Prediction of Coronary Artery Calcium Using Deep Learning of Echocardiograms

Neal Yuan, M.D.[1], Alan C Kwan, M.D.[2,6], Grant Duffy [2], John Theurer [2], Jonathan H. Chen, M.D., Ph.D.[4], Koen Nieman, M.D., Ph.D.[4,5], Patrick Botting, M.S.P.H. [2], Damini Dey Ph.D.[6], Daniel S. Berman, M.D. [2], Susan Cheng, M.D., M.P.H., M.Sc.[2], David Ouyang, M.D.[2,3]

1. School of Medicine, University of California, San Francisco, CA;

Section of Cardiology, San Francisco Veterans Affairs Medical Center, San Francisco, CA

1. Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, CA
2. Department of Medicine, Division of Artificial Intelligence in Medicine, Cedars-Sinai Medical Center, Los Angeles, CA
3. Department of Medicine, Stanford University, Stanford, CA
4. Department of Radiology, Stanford University, Stanford, CA
5. Biomedical Imaging Research Institute, Cedars-Sinai Medical Center, Los Angeles, CA

Corresponding Authors

Neal Yuan

4150 Clement St

San Francisco, CA 94121

Neal.Yuan@ucsf.edu

David Ouyang

8700 Beverly Blvd

Los Angeles, CA 90048

David.Ouyang@cshs.org

Abstract

Background

Coronary artery calcification (CAC), often assessed by computed tomography (CT), is a powerful marker of coronary artery disease (CAD) that can guide preventive therapies. CTs, however, are not always accessible or serially obtainable. It remains unclear if other widespread tests such as transthoracic echocardiograms (TTEs) can be used to predict CAC.

Methods

Using a dataset of 2881 TTE videos paired with coronary calcium CTs, we trained a video-based artificial intelligence (AI) convolutional neural network to predict CAC scores from parasternal long

axis (PLAX) views. We evaluated the model's ability to classify patients from a held-out sample as well as an external site sample into zero CAC and high CAC (CAC ≥400 Agatston units) groups by receiver operating characteristic (ROC) and precision-recall curves. We also investigated whether such classifications prognosticated significant differences in 1-year mortality rates by log-rank test of Kaplan-Meier curves.

Results

TTE AI models had high discriminatory abilities in predicting zero CAC (ROC AUC 0.81 (95% CI 0.74-0.88), F1 0.95) and high CAC (AUC 0.74 (0.68-0.8), F1 0.74). This performance was confirmed in an external test dataset of 92 TTEs ((AUC 0.75 (0.65-0.85), F1 0.77), (AUC 0.85 (0.76-0.93), F1 0.59), respectively). Risk stratification by TTE-predicted CAC performed similarly to CT CAC scores in prognosticating significant differences in 1-year survival in high CAC patients (CT CAC≥400 vs. CT CAC<400 p=0.03, TTE-predicted CAC≥400 vs. TTE-predicted CAC<400 p=0.02).

Conclusions

A video-based deep learning model successfully used TTE videos to predict zero CAC and high CAC with high accuracy. TTE-predicted CAC prognosticated differences in 1-year survival similar to CT CAC. Deep learning of TTEs holds promise for future adjunctive CAD risk stratification to guide preventive therapies.

Keywords: coronary artery calcium, echocardiogram, deep learning, machine learning, convolutional neural network, preventive cardiology, risk stratification

Abbreviations

Coronary artery calcification (CAC)

Computed tomography (CT)

Transthoracic echocardiogram (TTE)

Digital Imaging and Communications in Medicine (DICOM)

Parasternal long axis (PLAX)

Apical 4 chamber (A4C)

Receiver operating characteristic (ROC)

Precision-recall (PR)

Area under the curve (AUC)

Left ventricular ejection fraction (LVEF)

Net reclassification index for non-events (NRIne)

Net reclassification index for events (NRIe)

Disclosures

The authors have no conflicts of interest to declare.

Koen Nieman reports unrestricted institutional research support from the National Institutes of Health, Siemens Healthineers, Bayer, HeartFlow Inc., and consulting for Siemens Medical Solutions, unrelated to the present work.

David Ouyang reports involvement in InVision, an artificial intelligence company that uses deep learning algorithms similar to those used in this paper. InVision currently has no commercial interests or pursuits relevant to the research in this paper.

Funding

This works was supported by the following grants:

NIH T32 5T32HL116273-07 (NY)

NIH K99 HL157421-01 (DO)

Doris Duke Charitable Foundation Grant 2020059 (AK)

The Erika J. Glazer Family Foundation

Funding sources were not involved in study design, data collection, or analysis

Contributions

NY, DO were responsible for study design. PB helped with data extraction and curation. NY, GD, JT conducted deep learning and survival analysis. NY wrote manuscript. ACK, HC, KN, DD, DSB, SC, DO contributed manuscript edits.

Introduction

Coronary artery calcification (CAC) is a highly specific marker of atherosclerosis that is the result of pathogenic inflammatory, metabolic, and developmental processes (1). Assessment of CAC now plays an increasingly important role in risk stratification for coronary artery disease and prognostication in asymptomatic patients. Guidelines from both the American College of Cardiology/ American Heart Association and European Society of Cardiology endorse using CAC scores for guiding decisions on the use of lipid-lowering therapy and aspirin as well as informing discussions around modification of cardiovascular risk factors (2,3). Individuals with no coronary calcium have an extremely low risk of cardiovascular disease and mortality over the next 15 years and may safely de-escalate certain medical therapies (4,5). Conversely, individuals with any coronary calcium have a greater risk of long-term cardiovascular events and mortality that increases with the amount of calcium and represents risk beyond those predicted by traditional clinical risk factors (6–8).

Although coronary artery calcification is visualizable by multiple x-ray based imaging modalities including chest radiography and fluoroscopy, CAC is most frequently assessed by the Agatston

scoring method using computed tomography (CT) imaging (9). Currently, all forms of imaging to identify coronary calcium requires ionizing radiation(14,15), a reason for hesitation for some patients and providers, and are paralleled by concerns about inappropriate use of testing, healthcare costs, and an increase in the detection of incidental non-cardiac findings (10–13). In light of these concerns, at least one guideline body, the US Preventive Task Force, concluded in 2018 that there was insufficient evidence to formally recommend CAC for cardiovascular risk stratification (16). These limitations have also tempered the use of serial CAC CTs for ongoing disease surveillance, even though the rate of CAC progression is known to provide additional prognostic insight (7,17).

Recent advances in computational techniques have revealed that machine learning when applied to information-dense medical images can identify disease phenotypes and prognosticate outcomes beyond what is possible by expert clinician observation alone (18–21). Given that CAC CTs may not always be accessible nor serially obtainable, we sought to understand whether applying deep learning to other widely used cardiovascular tests, such as transthoracic echocardiograms (TTE), could complement and extend the prognostic power of CAC assessment.

In this retrospective study of patients with both CAC CTs and TTE studies, we used a video-based deep learning architecture to predict 1.) the presence of any coronary artery calcium and 2.) the presence of high coronary artery calcium levels (CAC score ≥400) from standard TTE views. In addition to validating results in an external test dataset, we tested whether these model predictions could be used to successfully prognosticate differences in 1-year mortality and applied model attribution methods to better understand areas of focus by the deep learning model.

Methods

Dataset

We identified all patients with CT studies with CAC scoring performed at Cedars-Sinai Medical Center, a large multi-site urban health system, from 1/1/2015 to 12/31/2020 who also had a TTE completed within one year of the CAC study. While a CAC score could be paired with potentially multiple TTEs, each TTE could only be paired with a single CAC score. Using independent studies from the same patient to predict a single label has been established as a helpful method of data augmentation that can improve the performance of deep learning models (19), and was only used during model training. In both internal and external test cohorts, one echo-CT pair was identified for each patient.

Per institutional protocol, CT imaging was obtained on Siemens or GE scanners and CAC scoring was performed at the computer on non-contrast image volumes according to standard imaging acquisition and Agatston scoring methods (9).

Echocardiograms were acquired using Philips EPIQ 7 or iE33 ultrasound machines. Echocardiogram characteristics (left ventricular ejection fraction, presence of wall motion abnormalities, and presence of at least moderate valvular disease) were obtained from echocardiogram reports, based off of the

final clinician report and evaluation of disease severity using ASE criteria. To better characterize our cohort, we distinguished echocardiograms with “>= Moderate Valvular Disease“ as any TTE with moderate or worse regurgitant or stenotic disease in any valve. For deep learning, each TTE study was initially sourced in Digital Imaging and Communications in Medicine (DICOM) format and contained multiple video loops and still images. Videos corresponding to the standard Parasternal Long Axis (PLAX) view were extracted, masked, and downsampled by cubic interpolation to a resolution of 112 by 112 pixels per previously described methods (18). The PLAX view was chosen as it is a standard view readily attainable by even novice scanners, visualizes anatomical features adjacent to the proximal coronary arteries (i.e. the aorta and aortic valve) that can experience early calcification, and can be readily identified by automated view classifiers (22). We randomly split our dataset using 80% of TTEs for model training, 10% for model validation, and 10% for hold-out testing. For a supplemental analysis, we also extracted apical 4 chamber (A4C) views from the same TTEs. Since some studies did not have both high-quality PLAX and A4C views, the number of PLAX and A4C videos was different.

We obtained an additional limited external dataset from Stanford Healthcare containing all TTEs completed within one year of CAC studies performed from 7/1/2016 to 8/1/2018. A4C and PLAX videos were extracted from these TTEs and then masked and downsampled according to a workflow similar to what was used for processing TTEs from Cedars-Sinai.

This study was approved by IRBs at Cedars-Sinai Medical Center and Stanford Healthcare.

Assessment of patient characteristics and outcomes

Patient characteristics were derived from electronic health records data according to standard Elixhauser comorbidity definitions (23). The Atherosclerotic Cardiovascular Disease (ASCVD) risk score for 10-year risk of heart disease or stroke was calculated for a subgroup of 402 patients who had available cholesterol and blood pressure information (24). Mortality information was sourced from both deaths recorded in the EHR as well as the California Department of Public Health Death Index.

Deep learning model selection and training

We trained a convolutional neural network model with residual connection and spatiotemporal convolutions across frames to predict CAC scores. This model, based on the R2+1D architecture, has been previously used successfully to predict left ventricular ejection fraction from TTEs (18). Our model was initialized with pretrained weights from the EchoNet-Dynamic dataset (18). Models were trained to minimize the squared loss between the prediction and log (CAC score + 1) using an Adam optimizer with a learning rate of 0.001 and batch size of 10 across 40 epochs.

The model input used video clips of 32 frames which were created by sampling every other frame. The weights from the epoch with the lowest validation loss was used for final model testing on the held out dataset. Model training was done in Python using the publicly available PyTorch deep learning library.

Model Performance and Survival Analysis

We evaluated the ability of our models to correctly classify patients into coronary calcium versus no coronary calcium groups as well as CAC score ≥400 versus <400 groups. We tested our model on the held-out test set, an external test set from Stanford Healthcare, as well as a subset of the original held-out dataset containing only patients aged less than 70 years old. We displayed receiver operating characteristic (ROC) and precision-recall (PR) curves to demonstrate model performance across classification thresholds. As a measure of overall model performance, we reported the area under the curve (AUC) for the ROC and precision-recall curves as well as F1 scores. Confidence intervals were reported using 10,000 bootstrapped samples.

In order to compare our models' performance with the prognostic abilities of a CAC score of zero by CT, we used predicted calcium scores to stratify patients into low or high risk using the score cutoff that was most accurate for predicting presence of CAC (Youden's index or threshold with highest mean of sensitivity and specificity). We compared the risk stratification with CAC by whether they had a CAC score of 0 or > 0 by CT. We additionally compared whether there were differences in mortality rates between these groups by log-rank test of Kaplan-Meier curves over a 1-year period from the date of the TTE or CT. The same comparison was repeated with the predictive abilities of a CAC score of above or below 400. Lastly, we calculated net reclassification indices for 1-year mortality when going from risk stratification by CT CAC score of 0 to risk stratification by TTE-predicted CAC score of 0 (25).

Survival analysis was conducted using R software (version 3.4.1, Vienna, Austria) survival and survminer packages.

Model Interpretability

We attempted to visualize input features most influential in our deep learning models using the Integrated Gradients attribution method, which has several advantages to other attribution methods including its ability to better meet the tests of sensitivity and implementation invariance (26). We chose representative output images using this attribution method on PLAX models.

Results

We trained CNN models on a dataset of 2881 TTEs paired with coronary calcium scores from 1635 patients. Patients had a mean age of 72 years (SD 14.3), were 37.4% female, 31.4% non-white, and had a range of cardiovascular comorbidities (25.4% hypertension, 36.9% heart failure, 59.8% hyperlipidemia) (**Table 1**). Patients with available cholesterol and blood pressure information were on average at intermediate cardiovascular risk with a mean ASCVD risk score of 13.6% (SD 11.4%). Over three quarter of patients had a normal left ventricular ejection fraction (LVEF) and 69.2% of echocardiograms had no wall motion abnormalities. The mean absolute difference in time between TTE and CT was 68.6 days (SD 96.0) with 13.4% having a calcium score of 0 and 53.0% having a calcium score ≥ 400. An external dataset from Stanford Healthcare contained 92 TTEs paired with

coronary calcium scores from 92 patients. The mean patient age was 58.4 years (SD 11.7) and 43.4% were female. Almost all TTEs had a normal LVEF, with 47.8% of TTEs paired with a calcium score of 0 and 17.4% with a calcium score ≥ 400.

When tested on a held-out dataset not used for model training, the CNN model had high discriminatory abilities in predicting which patients had zero calcium (ROC AUC 0.81 (95% CI 0.74-0.88), F1 score 0.95) (**Figure 1A**). The model performed modestly well in predicting patients who had high calcium scores > 400 Agatston units (AUC 0.74 (0.68-0.8), F1 0.74) (**Figure 1B**). On the external dataset of TTEs from Stanford Healthcare, the model was able to again predict patients who had zero coronary calcium (AUC 0.75 (0.65-0.85), F1 0.77) as well as those with high coronary calcium (AUC 0.85 (0.76-0.93), F1 0.59) (**Figure 2**).

In supplemental analyses, we found that a model trained using A4C videos was able to also predict both zero CAC (AUC 0.73 (0.67-0.79), F1 0.95) as well as high CAC (AUC 0.72 (0.68-0.76), F1 0.73) (**Supplemental Figure 1**). However, performance was worse compared to the model using PLAX videos, especially when applied to the external dataset (AUC 0.55 (0.47-0.64), F1 0.67 for zero CAC, AUC 0.73 (0.62-0.84), F1 0.33 for CAC > 400).

Since our patient cohort was on average older than the typical target population for CAC screening, we additionally tested our model in a limited subset of patients younger than 70 years old who were not used in model training. This subset had a mean age of 57.7 years (SD 9.8) and mean ASCVD risk score of 10.1% (SD 9.2%) with 22% of patients having a calcium score of 0 (**Supplemental Table 1**). Prediction performance was similar when compared to testing in the entire cohort of patients in predicting zero coronary calcium (AUC 0.79 (0.73-0.85), F1 0.87) and high coronary calcium (AUC 0.74 (0.67-0.81), F1 0.64) (**Supplemental Figure 2**).

We compared the risk discrimination abilities of our TTE CAC prediction model to having a CT CAC score of zero in predicting 1-year survival. Both risk stratification by TTE-predicted CAC and CT CAC produced survival curves that separated by 1 year, although in both cases, the separation was not statistically significant (CT calcium vs. CT no calcium p=0.20, TTE-predicted calcium vs. TTE-predicted no calcium p=0.07) (**Figure 3A**). When comparing survival curves in patients with a TTE-predicted CAC score > 400 or a CT CAC > 400, both methods of CAC assessment were able to significantly predict 1-year survival (**Figure 3B**). When going from CT CAC score of 0 to TTE-predicted CAC score of 0, there was an improved net reclassification index for non-events (NRIne= 0.12 for PLAX) and similar net reclassification index for events (NRIe=-0.04 for PLAX) **(Supplemental Table 2)**.

In an exploratory interpretability analysis using the Integrated Gradients attribution method, we found that the model appeared to focus more heavily on the aortic valve anulus, aorta, left ventricular outflow tract (LVOT), and mitral valve (**Figure 4**).

Discussion

Assessing risk of coronary artery disease in asymptomatic individuals continues to be a cornerstone for ensuring that patients receive appropriate preventive care and therapies. Echocardiography is a widely available, non-invasive, and radiation free diagnostic tool that captures a large amount of physiological information that could be used for helping better assess CAD risk. In this retrospective study of patients with both CAC and TTE studies, we demonstrated for the first time that a video-based deep learning algorithm using standard TTE views successfully predicted with high accuracy whether individuals had developed coronary artery calcium as well as high levels of coronary calcium. These predictions additionally stratified patients into risk groups that performed similarly in predicting 1-year mortality when compared to CAC scores by CT.

With the application of machine learning techniques, it has become increasingly clear that diagnostics such as TTEs and ECGs contain patterns of information that can be used for predicting patient characteristics and disease states beyond what may be immediately discernible by the human expert (19–21,27). In this study, we showed that deep learning of TTEs successfully predicted the presence of both any coronary artery calcium as well as high levels of calcium. Using TTE images to predict the presence of CAC, as opposed to other endpoint such as hard cardiovascular events for example, is particularly valuable because the development of CAC is an early upstream process that may identify a key time window during which preventative therapies may be most effective. Indeed, there already exists a robust body of evidence linking the detection of CAC to actionable therapies such as the use of aspirin and statins (2,3). While we do not anticipate echocardiograms can replace CAC, our study shows the opportunity to provide additional information from already obtained cardiovascular imaging and opportunities for opportunistic screening and maximize the value of the imaging.

The connection between TTE images, obtained without radiation, CAC, cardiovascular outcomes, and mortality has a compelling physiological basis and warrant further investigation. Cardiac structural changes that are visualizable in the PLAX view, such as systolic dilation/dysfunction, aortic and mitral valve calcification, and aortic valve restriction/stenosis, are well known to be associated with CAD as well as cardiovascular outcomes (28–31). Deep learning of PLAX TTE images have similarly been able to predict aortic stenosis with high accuracy (32) and show similar regions of interest with attribution methods, suggesting such models also use calcium as a surrogate of disease severity. Attribution methods confirmed that our deep learning algorithm appeared to focus on the aorta, LVOT, aortic, and mitral valves, and further work is needed to further explore the relationship between calcificiation and cardiovascular outcomes.

We confirmed that the calcium predictions by TTE deep learning models appeared to stratify patients into risk groups that had 1-year all-cause mortality rates similar to those predicted using CT CAC scores. While our study was not powered to fully explore the relationship between CAC and cardiovascular risk factors, as an exploratory endpoint, we even observed a slight improvement in net reclassification of non-events, meaning that patients who were deemed high risk by CT-CAC were correctly reclassified into the low risk group when using TTE-predicted CAC instead. As a result, in

contrast to the TTE-predicted CAC survival curves, there was no statistical difference seen yet for the CT calcium vs. CT no calcium curves at 1 year given the lower difference in mortality rates between the two groups (although we would expect this difference to become significant with additional follow-up and with larger sample sizes). The reason for this improved performance using TTE-predicted CAC is not immediately clear and remains to be confirmed in additional studies. It could be that there are additional features present in a TTE video that may help distinguish when a subset of patients with coronary artery calcification are lower risk for future adverse events. It is known, for example, that statins, while reducing future cardiovascular events, can accelerate coronary artery calcification, meaning that calcification in certain patients may not in fact be as worrisome (33).

If confirmed in subsequent prospective testing, the use of deep learning of TTEs for CAD risk stratification could have several potentially intriguing applications. Given the widespread availability of TTE, deep learning of TTEs could conceivably be used as a substitute for CAC testing when CAC scoring by CT is deferred by the patient or inaccessible. There also may be a role for using these models to help patients and clinicians decide whether to pursue a CAC CT if the decision is unclear. When CAC scores by CT are obtained, such models could act as an adjunct to inform the interpretation of CAC scores to give an overall more accurate picture of disease. TTEs could also have a potential role in serial monitoring for the development of CAD years after a zero CAC score. Prediction of high CAC scores may have additional uses such as determining when coronary CT angiography might have an increased likelihood of being nondiagnostic due to the presence of high calcium levels.

Several study limitations warrant consideration. TTE and CAC assessment were not performed simultaneously, although we ensured that they occurred within one year of each other with the mean absolute difference in time being slightly over two months. Our internal training cohort study population was older and had a high proportion of positive CAC scores, which is likely due to the older population that is seen at our institution and possibly because reimbursement issues limiting the availability of this risk stratification test. Our study utilized clinically obtained imaging based on standard clinician referral patterns, however we do not have the indication of imaging and could not identify whether patients had anginal symptoms at the time of their TTE or CT. In the subgroup of patients with available cholesterol and blood pressure information, the mean ASCVD risk score fell within the intermediate cardiovascular risk range. Model performance remained similar when limiting the test dataset to only patients aged less than 70 years old. The model was additionally able to perform well in an external dataset of younger patients who had a lower proportion of positive CAC. We used all-cause death as our main outcome given its adjudication certainty, although cardiovascular-related outcomes would be more specific to CAD. While survival curves based on TTE-predicted CAC diverged at 1 year, longer term follow-up would also be informative. Future prospective studies involving multiple centers will be critical for further confirming our findings.

Conclusion

A video-based deep learning model successfully used TTE videos to predict the presence of coronary artery calcium as well as high levels of coronary calcium with a high degree of accuracy. CAC predictions were additionally able to successfully prognosticate differences in 1-year mortality. These results speak to the potential role of using deep learning of TTEs for future CAD risk stratification to guide preventive therapies.

References

1. Greenland P., Blaha MJ., Budoff MJ., Erbel R., Watson KE. Coronary Calcium Score and Cardiovascular Risk. J Am Coll Cardiol 2018;72(4):434–47. Doi: 10.1016/j.jacc.2018.05.027.

2. Arnett DK., Blumenthal RS., Albert MA., et al. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation 2019;140(11):e596–646. Doi: 10.1161/CIR.0000000000000678.

3. Piepoli MF., Hoes AW., Agewall S., et al. 2016 European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts)Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). Eur Heart J 2016;37(29):2315–81. Doi: 10.1093/eurheartj/ehw106.

4. Valenti V., Ó Hartaigh B., Heo R., et al. A 15-Year Warranty Period for Asymptomatic Individuals Without Coronary Artery Calcium: A Prospective Follow-Up of 9,715 Individuals. JACC Cardiovasc Imaging 2015;8(8):900–9. Doi: 10.1016/j.jcmg.2015.01.025.

5. Budoff MJ., Young R., Burke G., et al. Ten-year association of coronary artery calcium with atherosclerotic cardiovascular disease (ASCVD) events: the multi-ethnic study of atherosclerosis (MESA). Eur Heart J 2018;39(25):2401–8. Doi: 10.1093/eurheartj/ehy217.

6. Nakanishi R., Li D., Blaha MJ., et al. All-cause mortality by age and gender based on coronary artery calcium scores. Eur Heart J Cardiovasc Imaging 2016;17(11):1305–14. Doi: 10.1093/ehjci/jev328.

7. Lehmann N., Erbel R., Mahabadi AA., et al. Value of Progression of Coronary Artery Calcification for Risk Prediction of Coronary and Cardiovascular Events. Circulation 2018;137(7):665–79. Doi: 10.1161/CIRCULATIONAHA.116.027034.

8. Peng AW., Mirbolouk M., Orimoloye OA., et al. Long-Term All-Cause and Cause-Specific Mortality in Asymptomatic Patients With CAC ≥1,000: Results From the CAC Consortium. JACC Cardiovasc Imaging 2020;13(1 Pt 1):83–93. Doi: 10.1016/j.jcmg.2019.02.005.

9. 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(4):827–32. Doi: 10.1016/0735-1097(90)90282-t.

10. ACC News Story. Coronary Artery Calcium: Score? or No More? American College of Cardiology. Available at: https://www.acc.org/latest-in-cardiology/articles/2014/05/20/14/40/http%3a%2f%2fwww.acc.org%2flatest-in-cardiology%2farticles%2f2014%2f05%2f20%2f14%2f40%2fcover-story-coronary-artery-calcium. Accessed July 17, 2021.

11. Pletcher MJ., Pignone M., Earnshaw S., et al. Using the coronary artery calcium score to guide statin therapy: a cost-effectiveness analysis. Circ Cardiovasc Qual Outcomes 2014;7(2):276–84. Doi: 10.1161/CIRCOUTCOMES.113.000799.

12. Machaalany J., Yam Y., Ruddy TD., et al. Potential clinical and economic consequences of noncardiac incidental findings on cardiac computed tomography. J Am Coll Cardiol 2009;54(16): 1533–41. Doi: 10.1016/j.jacc.2009.06.026.

13. Onuma Y., Tanabe K., Nakazawa G., et al. Noncardiac findings in cardiac imaging with multidetector computed tomography. J Am Coll Cardiol 2006;48(2):402–6. Doi: 10.1016/j.jacc. 2006.04.071.

14. Messenger B., Li D., Nasir K., Carr JJ., Blankstein R., Budoff MJ. Coronary calcium scans and radiation exposure in the multi-ethnic study of atherosclerosis. Int J Cardiovasc Imaging 2016;32(3): 525–9. Doi: 10.1007/s10554-015-0799-3.

15. Mandrola J., Foy A. The Case Against Coronary Artery Calcium Scoring for Cardiovascular Disease Risk Assessment. Am Fam Physician 2019;100(12):734–5.

16. US Preventive Services Task Force, Curry SJ., Krist AH., et al. Risk Assessment for Cardiovascular Disease With Nontraditional Risk Factors: US Preventive Services Task Force Recommendation Statement. JAMA 2018;320(3):272–80. Doi: 10.1001/jama.2018.8359.

17. Hecht H., Blaha MJ., Berman DS., et al. Clinical indications for coronary artery calcium scoring in asymptomatic patients: Expert consensus statement from the Society of Cardiovascular Computed Tomography. J Cardiovasc Comput Tomogr 2017;11(2):157–68. Doi: 10.1016/j.jcct.2017.02.010.

18. Ouyang D., He B., Ghorbani A., et al. Video-based AI for beat-to-beat assessment of cardiac function. Nature 2020;580(7802):252–6. Doi: 10.1038/s41586-020-2145-8.

19. Attia ZI., Noseworthy PA., Lopez-Jimenez F., et al. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. Lancet 2019;394(10201):861–7. Doi: 10.1016/S0140-6736(19)31721-0.

20. Attia ZI., Kapa S., Lopez-Jimenez F., et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. Nat Med 2019;25(1):70–4. Doi: 10.1038/s41591-018-0240-2.

21. Ko W-Y., Siontis KC., Attia ZI., et al. Detection of Hypertrophic Cardiomyopathy Using a Convolutional Neural Network-Enabled Electrocardiogram. J Am Coll Cardiol 2020;75(7):722–33. Doi: 10.1016/j.jacc.2019.12.030.

22. Madani A., Arnaout R., Mofrad M., Arnaout R. Fast and accurate view classification of echocardiograms using deep learning. NPJ Digit Med 2018;1:6. Doi: 10.1038/s41746-017-0013-1.

23. Elixhauser A., Steiner C., Harris DR., Coffey RM. Comorbidity measures for use with administrative data. Med Care 1998;36(1):8–27. Doi: 10.1097/00005650-199801000-00004.

24. Goff DC., Lloyd-Jones DM., Bennett G., et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation 2014;129(25 Suppl 2):S49-73. Doi: 10.1161/01.cir.0000437741.48606.98.

25. Kerr KF., Wang Z., Janes H., McClelland RL., Psaty BM., Pepe MS. Net reclassification indices for evaluating risk prediction instruments: a critical review. Epidemiology 2014;25(1):114–21. Doi: 10.1097/EDE.0000000000000018.

26. Sundararajan M., Taly A., Yan Q. Axiomatic Attribution for Deep Networks. ArXiv:170301365 [Cs] 2017.

27. Ghorbani A., Ouyang D., Abid A., et al. Deep learning interpretation of echocardiograms. NPJ Digit Med 2020;3:10. Doi: 10.1038/s41746-019-0216-8.

28. Mansour MJ., Chammas E., Hamoui O., Honeine W., AlJaroudi W. Association between left ventricular diastolic dysfunction and subclinical coronary artery calcification. Echocardiography 2020;37(2):253–9. Doi: 10.1111/echo.14580.

29. Lin FY., Zemedkun M., Dunning A., et al. Extent and severity of coronary artery disease by coronary CT angiography is associated with elevated left ventricular diastolic pressures and worsening diastolic function. J Cardiovasc Comput Tomogr 2013;7(5):289-296.e1. Doi: 10.1016/j.jcct.2013.08.008.

30. Nasir K., Katz R., Al-Mallah M., et al. Relationship of aortic valve calcification with coronary artery calcium severity: the Multi-Ethnic Study of Atherosclerosis (MESA). J Cardiovasc Comput Tomogr 2010;4(1):41–6. Doi: 10.1016/j.jcct.2009.12.002.

31. Allison MA., Cheung P., Criqui MH., Langer RD., Wright CM. Mitral and Aortic Annular Calcification Are Highly Associated With Systemic Calcified Atherosclerosis. Circulation 2006;113(6):861–6. Doi: 10.1161/CIRCULATIONAHA.105.552844.

32. Holste G., Oikonomou EK., Mortazavi BJ., et al. Automated detection of severe aortic stenosis using single-view echocardiography: A self-supervised ensemble learning approach 2022:2022.08.30.22279413. Doi: 10.1101/2022.08.30.22279413.

33. Puri R., Nicholls SJ., Shao M., et al. Impact of statins on serial coronary calcification during atheroma progression and regression. J Am Coll Cardiol 2015;65(13):1273–82. Doi: 10.1016/j.jacc.2015.01.036.

Tables

Patients	**N (%)**	**Echocardiogram videos paired with CAC**	**N (%)**
	1635		2881
Age (mean (SD))	71.9 (14.3)	Normal LVEF	2166 (75.2)
Female (%)	612 (37.4)	Wall Motion Abnormality	886 (30.8)
Race/Ethnicity (%)		>= Moderate Valvular Disease	569 (19.8)
American Indian	5 (0.3)	Days from TTE to CT (mean (SD))	68.6 (96.0)
Asian	101 (6.2)	CAC = 0	385 (13.4)
Black	143 (8.7)	CAC = 0-99	518 (18.0)
Hispanic	119 (7.3)	CAC = 100-399	450 (15.6)
Non-Hispanic White	1121 (68.6)	CAC > 400	1528 (53.0)
Other	129 (7.9)		
Unknown	17 (1.0)	**External site echocardiogram videos**	
Hypertension	416 (25.4)		92
Heart Failure	603 (36.9)	Age (mean (SD))	58.4 (11.7)
Hyperlipidemia	977 (59.8)	Female (%)	40 (43.4)
Diabetes	345 (21.1)	Normal LVEF	89 (96.7)
Obesity	116 (7.1)	>= Moderate Valvular Disease	5 (5.4)
Chronic Lung Disease	622 (38.0)	Days from TTE to CT (mean (SD))	40 (169.9)
Chronic Kidney Disease	287 (17.6)	CAC = 0	44 (47.8)
Active Smoker	36 (2.2)	CAC = 0-99	16 (17.4)

On Hypertension Medication	817 (50)	CAC = 100-399	16 (17.4)
ASCVD Risk Score*	13.6 (11.4)	CAC > 400	16 (17.4)

Table 1. Baseline patient, echocardiography, and calcium score characteristics

Abbreviations: SD = standard deviation, TTE = transthoracic echocardiogram, CT = computed tomography, CAC = coronary artery calcium

*ASCVD Risk Scores were calculable for a subset of 402 patients that had complete data

Figure Legends

Figure 1. Performance characteristics of a deep learning model for predicting coronary artery calcium (CAC) using parasternal long axis (PLAX) transthoracic echocardiogram (TTE) videos when applied to a held-out test dataset. Receiver-Operating Characteristic and Precision-Recall Curves across different classification thresholds with Area Under the Curve (AUC) (95% confidence interval) and F1 score for predicting **A.** Presence versus absence of CAC **B.** CAC score < 400 versus > 400 Agatston units.

Figure 2. Performance characteristics of a deep learning model for predicting coronary artery calcium (CAC) using parasternal long axis (PLAX) transthoracic echocardiogram (TTE) videos when applied to an external site test dataset. Receiver-Operating Characteristic and Precision-Recall Curves across different classification thresholds with Area Under the Curve (AUC) (95% confidence interval) and F1 score for predicting **A.** Presence versus absence of CAC **B.** CAC score < 400 versus > 400 Agatston units.

Figure 3. Kaplan-Meier survival curves for all-cause mortality over 1 year for patients in a held-out test dataset stratified by **A.** CT CAC score 0 vs. > 0 or by TTE-predicted CAC score 0 vs. > 0 **B.** CT CAC score < 400 vs. > 400 or by TTE-predicted CAC score <400 vs. > 400

Figure 4. Areas of focus by the deep learning model (bright pixels) according to results from the Integrated Gradients attribution method.