

UCLA

UCLA Previously Published Works

Title

Artificial intelligence applied to coronary artery calcium scans (AI-CAC) significantly improves cardiovascular events prediction

Permalink

<https://escholarship.org/uc/item/1fq6z4kq>

Journal

npj Digital Medicine, 7(1)

ISSN

2398-6352

Authors

Naghavi, Morteza
Reeves, Anthony P
Atlas, Kyle
et al.

Publication Date

2024

DOI




10.1038/s41746-024-01308-0

Peer reviewed



Artificial intelligence applied to coronary artery calcium scans (AI-CAC) significantly improves cardiovascular events prediction



Morteza Naghavi¹ ✉, Anthony P. Reeves², Kyle Atlas¹, Chenyu Zhang³ , Thomas Atlas³, Claudia I. Henschke⁴, David F. Yankelevitz⁴, Matthew J. Budoff⁵ , Dong Li⁵, Sion K. Roy⁵, Khurram Nasir⁶, Sabee Molloy⁷, Zahi Fayad⁶, Michael V. McConnell⁸, Ioannis Kakadiaris⁹, David J. Maron⁸ , Jagat Narula⁹, Kim Williams¹⁰, Prediman K. Shah¹¹, Daniel Levy¹² & Nathan D. Wong¹³

Coronary artery calcium (CAC) scans contain valuable information beyond the Agatston Score which is currently reported for predicting coronary heart disease (CHD) only. We examined whether new artificial intelligence (AI) applied to CAC scans can predict non-CHD events, including heart failure, atrial fibrillation, and stroke. We applied AI-enabled automated cardiac chambers volumetry and calcified plaque characterization to CAC scans (AI-CAC) of 5830 asymptomatic individuals (52.2% women, age 61.7 ± 10.2 years) in the multi-ethnic study of atherosclerosis during 15 years of follow-up, 1773 CVD events accrued. The AUC at 1-, 5-, 10-, and 15-year follow-up for AI-CAC vs. Agatston score was (0.784 vs. 0.701), (0.771 vs. 0.709), (0.789 vs. 0.712) and (0.816 vs. 0.729) ($p < 0.0001$ for all), respectively. AI-CAC plaque characteristics, including number, location, density, plus number of vessels, significantly improved CHD prediction in the CAC 1–100 cohort vs. Agatston Score. AI-CAC significantly improved the Agatston score for predicting all CVD events.

Coronary artery calcium (CAC) scoring is the strongest predictor of risk for atherosclerotic cardiovascular disease (ASCVD) in asymptomatic individuals¹. Although CAC scoring is used for prediction of coronary heart disease events, it is not used for prediction of other cardiovascular disease (CVD) events such as stroke, heart failure (HF) and atrial fibrillation (AF). Beyond risk factor assessment, screening tools for overall CVD event prediction are limited due to cost-effectiveness and feasibility barriers.

The usage of CAC scans has increased significantly since the ACC/AHA Guideline on the Management of Blood Cholesterol in 2018² included CAC score in the algorithm for consideration of statin therapy, among those at borderline and intermediate risk for ASCVD. It is estimated that 45–50% of the US population aged 40–80 would fall in these groups defined as 5–20% risk of ASCVD events over 10 years^{3,4}. The possibility of applying artificial

intelligence (AI) to predict CVD has been previously published by some of our team members using the support vector machine algorithms in MESA⁵. We have sought to further enrich the value of CAC scans by applying AI that automatically measures all cardiac chamber volumes and left ventricular (LV) mass without using any contrast agent. For this manuscript, we refer to AI-enabled automated cardiac chambers volumetry from CAC scans as AI-CAC, and the AI-CAC model incorporates Agatston CAC Score, left atrial (LA), right ventricular (RV), left ventricular (LV) volume and mass.

We have recently shown that AI-CAC volumetry alone enabled the prediction of HF in the Multi-Ethnic Study of Atherosclerosis (MESA)^{6,7}. Additionally, we have demonstrated that AI-CAC LA volume alone improved the predictive value of CHARGE-AF Risk Score and NT-proBNP for the detection of individuals at high risk of AF^{8,9}. Such an add-on

¹HeartLung.AI, Houston, TX, 77021, USA. ²Department of Electrical and Computer Engineering, Cornell University, Ithaca, NY, 14853, USA. ³Tustin Teleradiology, Tustin, CA, 92780, USA. ⁴Mount Sinai Hospital, New York, NY, 10029, USA. ⁵The Lundquist Institute, Torrance, CA, 90502, USA. ⁶Houston Methodist Hospital, Houston, TX, 77030, USA. ⁷Department of Radiology, University of California Irvine, Irvine, CA, 92697, USA. ⁸Cardiovascular Medicine, Stanford School of Medicine, Stanford, CA, 94305, USA. ⁹The University of Texas Health Science Center at Houston, Houston, TX, 77030, USA. ¹⁰University of Louisville, Louisville, KY, 40292, USA. ¹¹Cedars-Sinai Medical Center, Los Angeles, CA, 90048, USA. ¹²Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, 20824, USA. ¹³Heart Disease Prevention Program, Division of Cardiology, University of California Irvine, Irvine, CA, 92697, USA. ✉ e-mail: mn@vp.org; morteza.naghavi@americanhearttechnologies.com

measurement can offer valuable insights into a patient’s overall CVD risk beyond the CAC score. In this study of MESA participants, we compared the performance of AI-CAC over the traditional Agatston CAC Score for the prediction of all CVD events (stroke, myocardial infarction, angina, resuscitated cardiac arrest, all cardiovascular disease-related deaths, HF, and AF) plus all-cause mortality. This approach broadens the scope and clinical significance of comparing AI-CAC vs. the Agatston CAC score

Results

The mean (SD) age of our subjects was 62 ± 10 years, 52% were female, 40% were non-Hispanic White, 26% non-Hispanic Black, 22% Hispanic, and 12% Chinese. Table 1 shows the baseline characteristics of MESA participants who experienced a CVD event vs. those who did not over the 15 years of follow-up, during which 1773 CVD events accrued. In univariate comparisons, participants experiencing CVD events were older, more likely male, and more likely non-Hispanic White. The cases that experienced a CVD event had higher cardiac chamber volumes for LA, LV, RA, and LV mass.

Figure 1 shows examples of three participants with enlarged LA and LV volumes with CAC score 0 and low risk (< 5%) ASCVD risk score who experienced CVD events. A significant number of low-risk participants with CAC 0 have enlarged cardiac chambers. With a higher CAC score category, there was a higher proportion of patients with LA and LV volumes in the highest quartile (*p*-trend = 0.0001). 17.7% of cases with CAC 0 who are considered low risk have enlarged LA volume that puts them at high risk for AF and stroke (Fig. 2a). Similarly, 22.7% of cases with CAC 0 have enlarged LV volume that puts them at risk of HF (Fig. 2b).

The median *C*-statistic (95% CI) for all CVD events over 15 years for pooled sexes between AI-CAC vs. Agatston CAC score was 0.742 (CI: 0.723–0.761) vs. 0.709 (CI: 0.688–0.728) (*p* < 0.0001). For females, the *C*-statistic between AI-CAC volumetry vs. Agatston CAC Score was 0.751 (0.738–0.778) vs. 0.705 (0.683–0.720) (*p* < 0.0001), respectively, and 0.701 (0.674–0.723) vs. 0.672 (0.651–0.693) (*p* = 0.0012), respectively, for males. AI-CAC had significantly higher discrimination than Agatston CAC Score for CVD events prediction across 1-, 5-, 10-, and 15-year follow-up (Fig. 3), including AF, HF, stroke, hard CVD, and All-cause mortality prediction

Table 1 | Baseline characteristics of the multi-ethnic study of atherosclerosis (MESA) participants, including cases with and without cardiovascular events at 15 years of follow-up

15-year follow-up outcome data	Overall (N = 5830)	Cardiovascular events		p-value
		No (N = 4057)	Yes (N = 1773)	
Age (years)	62.2 ± 10.3	60.1 ± 10.0	67.1 ± 9.1	<0.0001
Female sex (%)	52.2%	55.7%	43.8%	<0.0001
Body Surface Area	1.90 ± 0.24	1.89 ± 0.24	1.93 ± 0.24	<0.0001
Race (%)				
Non-Hispanic White	39.7%	37.7%	43.8%	0.0321
Chinese	12.1%	12.2%	11.0%	0.2054
Non-Hispanic Black	26.1%	27.2%	24.5%	0.0935
Hispanic	22.0%	22.9%	20.7%	0.8569
AI-CAC volumetry				
LA volume (cc)	102.5 ± 25.4	58.8 ± 14.5	67.3 ± 18.1	<0.0001
LV volume (cc)	61.4 ± 16.1	100.8 ± 24.4	106.6 ± 27.0	<0.0001
RA volume (cc)	134.3 ± 34.4	69.4 ± 17.1	74.4 ± 19.7	<0.0001
RV volume (cc)	77.0 ± 18.9	133.2 ± 34.1	136.7 ± 35.0	0.1254
LV mass (g)	107.8 ± 26.4	103.2 ± 25.0	110.9 ± 27.4	<0.0001
Total heart volume (cc)	482.3 ± 108.7	465.4 ± 104.6	495.9 ± 113.2	<0.0001
Coronary artery calcium (CAC)				
Agatston score	0.93 (0-90.66)	0 (0-36.9)	51.04 (0-280.4)	<0.0001
Number of plaques	0 (0-5)	0 (0-3)	5 (0-15)	<0.0001
Number of affected vessels	0 (0-2)	0 (0-1)	2 (0-3)	<0.0001
Mean CAC density	0 (0-79)	0 (0-48)	89 (0-399)	<0.0001
Risk factors				
Diabetes	12.7%	10.2%	18.7%	<0.0001
Hypertension	44.7%	38.5%	59.5%	<0.0001
Current smoking	13.0%	12.2%	13.3%	0.1765
Current alcohol usage	68.6%	69.7%	65.9%	<0.0001
Family history of coronary heart disease (%)	42.7%	40.5%	48.5%	<0.0001
LDL cholesterol (mg/dL)	117.2 ± 31	117.8 ± 31.1	115.5 ± 32.2	0.3017
HDL cholesterol (mg/dL)	51.0 ± 14	51.2 ± 14.9	50.1 ± 15.1	0.1087
Total cholesterol (mg/dL)	194.2 ± 35	195.0 ± 35.4	192.5 ± 36.8	<0.0001
Systolic blood pressure (mmHg)	126.5 ± 21.4	123.7 ± 20.4	133.4 ± 22.0	<0.0001
Diastolic blood pressure (mmHg)	71.9 ± 10.2	71.4 ± 10.1	73.0 ± 10.3	<0.0001
Blood pressure lowering Rx	37.0%	31.8%	49.7%	<0.0001
Lipid-lowering Rx	16.6%	14.6%	21.3%	<0.0001

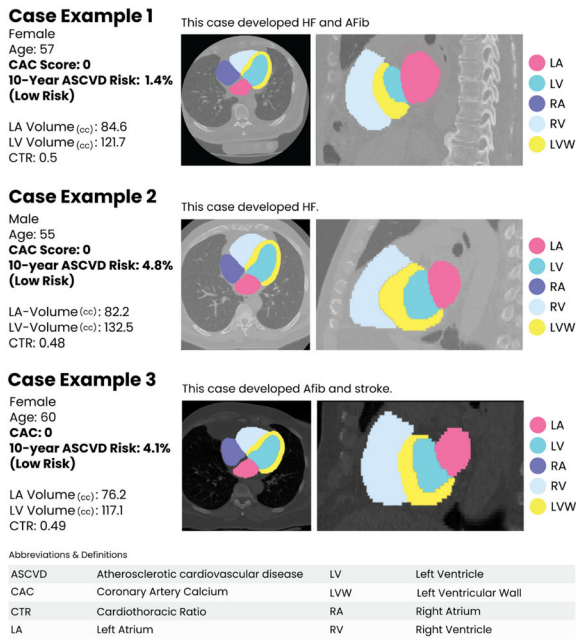


Fig. 1 | AI-enabled automated cardiac chambers volumetry and calcified plaque characterization to CAC scans (AI-CAC) definition and case examples. AI-CAC component diagram derived from coronary artery calcium (CAC) scan and

examples of AI-CAC volumetry detection of high-risk individuals with enlarged cardiac chambers in coronary artery calcium (CAC) scans with a calcium score of zero and low ASCVD risk.

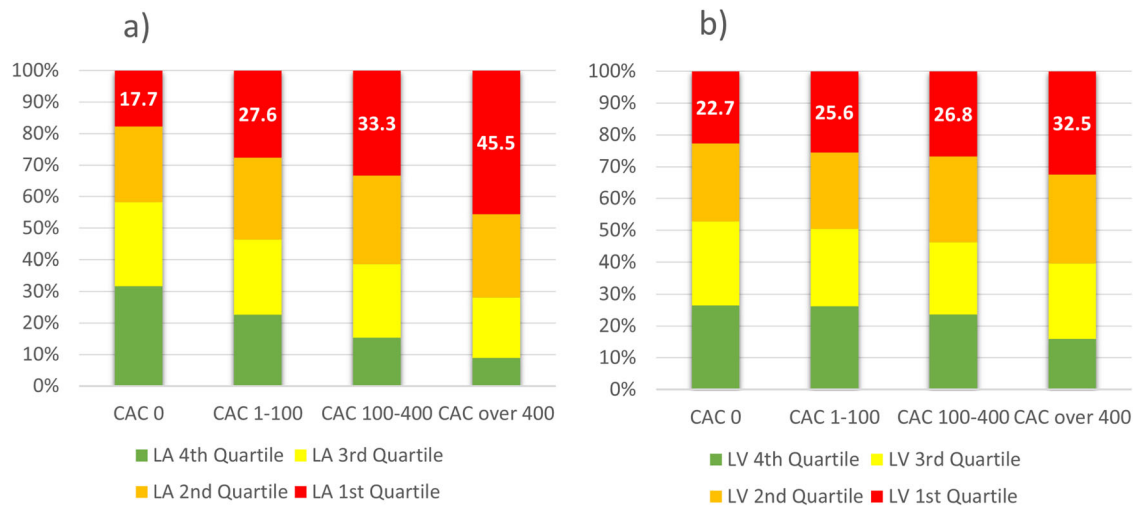


Fig. 2 | Quartiles of AI-enabled automated cardiac chambers volumetry and calcified plaque characterization to CAC scans (AI-CAC) Left Atrial (LA) and Left ventricular (LV) Volume by Agatston Coronary Artery Calcium (CAC) Score Quartiles. a AI-CAC LA volume vs. CAC score. Stacked bar chart of quartiles of AI-CAC LA volume by CAC score categories (0, 1–100, 101–400, over 400).

Despite the correlation, 17.7% of cases with CAC 0 who are considered low risk have enlarged LA volume that puts them at high risk for atrial fibrillation (AF) and stroke. **b** AI-CAC LV volume vs. CAC score. Stacked bar chart of quartiles of AI-CAC LV volume by CAC score categories (0, 1–100, 101–400, over 400). 22.7% of cases with CAC 0 have enlarged LV volume that puts them at risk of heart failure (HF).

(Table 2). Category-free NRI showed improvement across all follow-up periods for AF, HF, stroke, hard CVD, and All-Cause Mortality.

AI-CAC volumetry significantly added to a CVD risk factors-based model for all CVD event prediction (Supplementary Table 1). The AI-CAC biomarker model coefficients and hazard ratios for CVD, AF, and HF have been provided in Supplementary Tables 2, 3, 4. A significant increase in C-statistic for All CVD events was observed when adding AI-CAC measurements to basic risk factor models (0.745 (0.655–0.836) vs. 0.774 (0.693–0.852)) (Supplementary Table 5). A notable increase in discrimination for CVD prediction was demonstrated over 1-, 5-, 10-, and 15-year follow-ups. AI-CAC measurements demonstrated significant incremental value when added to CVD risk factors for 1-year CVD prediction

(0.803 vs. 0.749, $p < 0.0001$) (Supplementary Fig. 6a), 5-year CVD prediction (0.786 vs. 0.752, $p < 0.0001$) (Supplementary Fig. 6b), 10-year CVD prediction (0.801 vs. 0.774), $p < 0.0001$ (Supplementary Fig. 6c), and 15-year CVD prediction (0.823 vs. 0.816, $p < 0.0001$) (Supplementary Fig. 6d).

AI-CAC plaque characterization significantly improved CHD prediction in the CAC 1–100 cohort. AI-CAC plaque characteristics included the number of plaques, location, density, plus number of vessels affected. The addition of AI-CAC RV volume, LV volume, and LV mass further improved discrimination for CHD in this cohort. The AI-CAC composite model included LA volume, RV volume, LV volume, LV mass, AI-CAC derived plaque characterization, and Agatston CAC Score. Over 5- and 10-year follow-up, the time-dependent AUC for the AI-CAC composite model

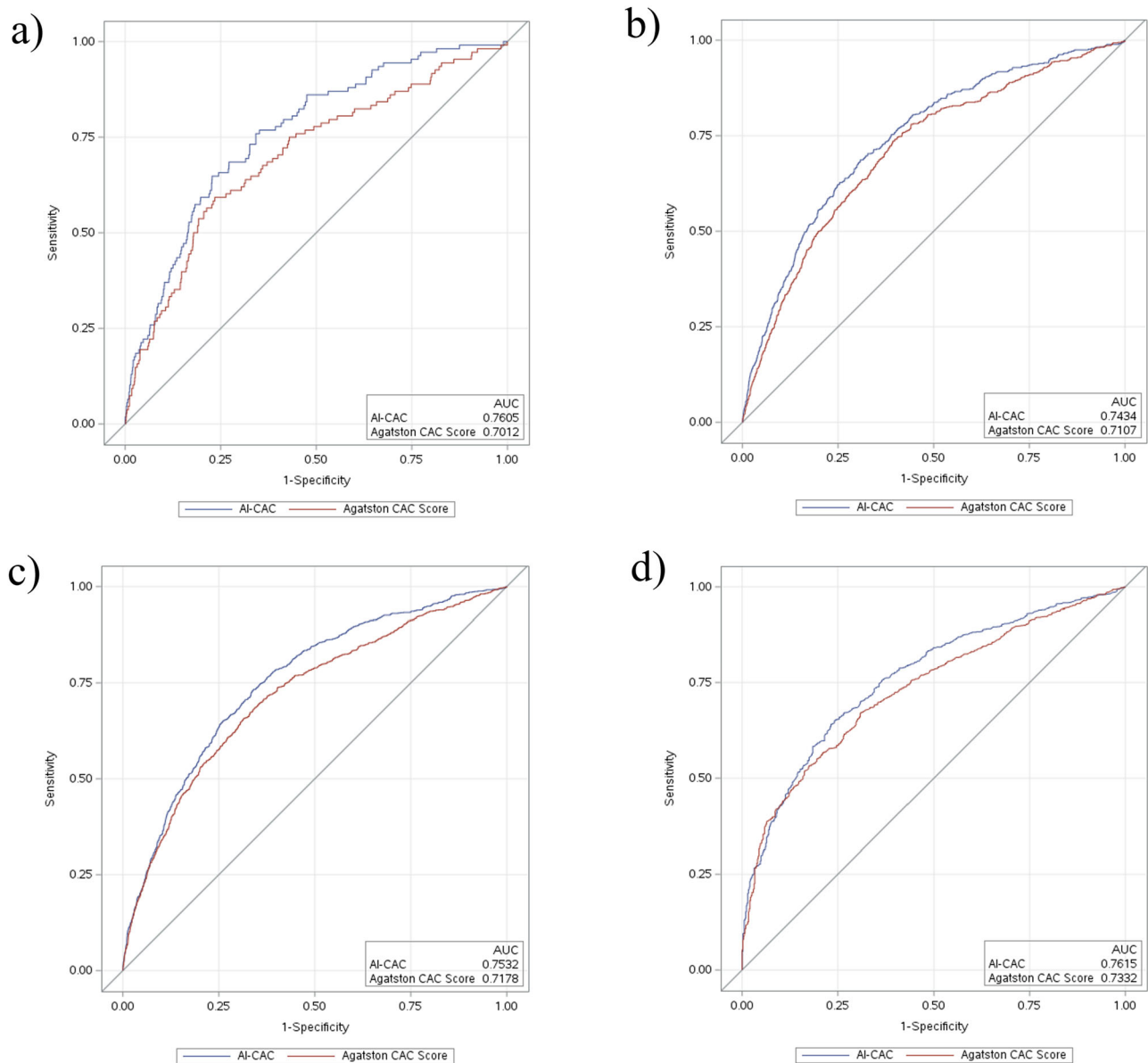


Fig. 3 | Time-dependent receiver operating curve (ROC) area under curve (AUC) for all cardiovascular events between AI-enabled automated cardiac chambers volumetry and calcified plaque characterization to CAC scans (AI-CAC) vs. Agatston coronary artery calcium (CAC) Score over 15 years. a Time-dependent AUC for AI-CAC vs. CAC score at 1-year follow-up. AI-CAC had significantly higher discrimination than Agatston CAC score for CVD events prediction over 1-year follow-up. The AUC at 1-year follow-up for AI-CAC vs. Agatston Score was 0.784 vs. 0.701 ($p < 0.0001$). **b** Time-dependent AUC for AICAC vs. CAC score over

5-years follow-up. At a 5-year follow-up, AI-CAC continued to demonstrate superior discrimination compared to the Agatston CAC score. The AUC for AI-CAC vs. Agatston score was 0.771 vs. 0.709 ($p < 0.0001$). **c** Time-dependent AUC for AI-CAC vs. CAC score over 10-years follow-up. For a 10-year follow-up, AI-CAC maintained a higher AUC compared to the Agatston score (0.789 vs. 0.712, $p < 0.0001$). **d** Time-dependent AUC for AI-CAC vs. CAC score over 15-years follow-up. At the 15-year follow-up, AI-CAC achieved the highest discrimination, with an AUC of 0.816 vs. 0.729 for the Agatston score ($p < 0.0001$).

vs. Agatston CAC Score was 0.654 vs. 0.557 ($p < 0.0001$) and 0.688 vs. 0.556 ($p < 0.0001$), respectively (Supplementary Fig. 8a, b).

Discussion

Our study primarily demonstrates the utility of applying AI to CAC scans to extract more actionable information than currently reported, which is the Agatston CAC score only. We found that AI volumetry significantly improves upon traditional CAC scoring for the prediction of risk for total CVD events as well as the prediction of individual CVD events of HF, stroke, AF, and all-cause mortality in a large multi-ethnic cohort. The plaque characterization component of AI-CAC specifically improved the predictive value of the Agatston score for CAC scores 1–100. Moreover, we show the value of this technique not only for longer-term event prediction (10–15

years) but also for nearer-term events (1 to 5-year follow-up). This is the first multi-ethnic outcome study of an easily implemented AI technology that can be applied to non-contrast CAC scans without additional radiation exposure to identify patients at risk of such events who would otherwise not be identified by Agatston CAC score. The potential utility of non-coronary findings in CAC scans has been reported previously using manual 2D measurements of LV^{10–13} and LA sizes^{14–17}. Our study corroborates findings from the Heinz Nixdorf Recall Study and others and further brings to light the value of non-coronary findings in CAC scans for a comprehensive CVD risk assessment beyond CHD^{14–18}. Kizer et al. showed that LA size was an independent predictor of CVD events¹⁹. Mahabadi et al.¹⁵ showed in the longitudinal Heinz Nixdorf Recall Study that two-dimensional LA size and epicardial adipose tissue from non-contrast CT were strongly associated

Table 2 | Time-dependent area under the curve (AUC) and category-free net reclassification index (NRI) at 1-, 5-, 10-, and 15-year follow-up between AI-CAC and Agatston CAC score (CAC) for individual and all cardiovascular event and mortality prediction

MESA outcomes 1-year follow up	CAC AUC (95% CI)	AI-CAC ^a AUC (95% CI)	AUC <i>p</i> -value	AI-CAC NRI over CAC	NRI <i>p</i> -value
AF	0.67 (0.58,0.75)	0.80 (0.71,0.88)	<0.0001	0.77	<0.0001
HF	0.65 (0.52,0.77)	0.91 (0.84,0.97)	<0.0001	1.17	<0.0001
CHD ^b	0.78 (0.72,0.85)	0.80 (0.72,0.87)	0.41	0.19	0.05
Stroke	0.67 (0.56,0.79)	0.79 (0.66,0.87)	<0.0001	0.65	<0.0001
Hard CVD ^c	0.68 (0.59,0.79)	0.77 (0.68,0.85)	<0.0001	0.34	<0.0001
All CVD events ^d	0.69 (0.65,0.75)	0.77 (0.73,0.83)	<0.0001	0.40	<0.0001
All-cause mortality	0.64 (0.52,0.68)	0.71 (0.58,0.82)	<0.0001	0.61	<0.0001
5-year follow up					
AF	0.67 (0.64,0.70)	0.74 (0.70,0.77)	<0.0001	0.36	<0.0001
HF	0.71 (0.65,0.75)	0.83 (0.79,0.88)	<0.0001	0.64	<0.0001
CHD ^b	0.79 (0.76,0.82)	0.81 (0.78,0.84)	0.18	0.18	0.006
Stroke	0.66 (0.60,0.72)	0.76 (0.70,0.81)	<0.0001	0.58	<0.0001
Hard CVD ^c	0.73 (0.70,0.77)	0.78 (0.75,0.80)	<0.0001	0.33	<0.0001
All CVD events ^d	0.71 (0.69,0.74)	0.75 (0.74,0.78)	<0.0001	0.28	<0.0001
All-cause mortality	0.67 (0.63,0.70)	0.70 (0.66,0.73)	<0.0001	0.32	<0.0001
10-year follow up					
AF	0.69 (0.67,0.71)	0.76 (0.74,0.78)	<0.0001	0.43	<0.0001
HF	0.71 (0.68,0.75)	0.81 (0.78,0.84)	<0.0001	0.52	<0.0001
CHD ^b	0.79 (0.77,0.81)	0.80 (0.78,0.83)	0.09	0.16	0.002
Stroke	0.66 (0.63,0.70)	0.75 (0.71,0.79)	<0.0001	0.45	<0.0001
Hard CVD ^c	0.72 (0.70,0.75)	0.77 (0.74,0.79)	<0.0001	0.33	<0.0001
All CVD events ^d	0.71 (0.70,0.73)	0.76 (0.75,0.78)	<0.0001	0.29	<0.0001
All-cause mortality	0.68 (0.65,0.70)	0.71 (0.69,0.73)	<0.0001	0.33	<0.0001
15-year follow up					
AF	0.69 (0.67,0.71)	0.75 (0.74,0.77)	<0.0001	0.33	<0.0001
HF	0.75 (0.71,0.78)	0.83 (0.79,0.86)	<0.0001	0.50	<0.0001
CHD ^b	0.81 (0.78,0.83)	0.82 (0.80,0.85)	0.73	0.17	0.002
Stroke	0.69 (0.66,0.73)	0.75 (0.70,0.79)	<0.0001	0.30	<0.0001
Hard CVD ^c	0.75 (0.72,0.78)	0.79 (0.77,0.82)	<0.0001	0.24	<0.0001
All CVD events ^d	0.72 (0.70,0.74)	0.76 (0.74,0.78)	<0.0001	0.23	<0.0001
All-cause mortality	0.68 (0.66,0.71)	0.72 (0.70,0.74)	<0.0001	0.29	<0.0001

AF atrial fibrillation, HF heart failure.

^aAI-CAC model: LA indexed by BSA, RV indexed by BSA, LV volume and mass indexed by BSA, log-transformed CAC.

^bCHD: myocardial infarction, definite angina, probable angina, resuscitated cardiac arrest, CHD death.

^cHard CVD: myocardial infarction, resuscitated cardiac arrest, stroke, CHD death, stroke death.

^dAll cardiovascular events: stroke, myocardial infarction, angina, resuscitated cardiac arrest, all cardiovascular disease-related deaths, heart failure, and atrial fibrillation.

with prevalent and incident AF and that LA size diminished the link of epicardial adipose tissue with AF, and was also associated with incident major CVD events independent of risk factors and CAC-score¹⁶.

Although there are multiple automated CAC scoring tools available, currently, there is no clinically available tool to clinicians for automated cardiac chamber volumetry in CAC scans that is validated against outcomes. Here, we provide evidence of the feasibility of using AI for automated 3D volumetry of cardiac chambers that takes on average 20 s. Currently, such measurements are only possible on contrast-enhanced CT scans, which require more radiation plus injection of an X-ray contrast agent that is burdensome²⁰. In contrast, AI-CAC volumetry can be applied to any new or existing non-contrast CAC scan for automated cardiac chamber measurement. Standalone cardiac MRI and echocardiography are not comparable to our solution, which is an opportunistic add-on to chest CT scans. While echocardiography and cardiac MRI provide valuable information on cardiac chamber volume, they are not indicated for the asymptomatic population and are usually performed in cardiovascular clinics. However, AI-

enabled cardiac chamber volumetry can be done on any chest CT scan, including lung cancer screening scans. This approach opens the door to identifying high-risk asymptomatic patients in non-cardiovascular clinics.

AI-CAC volumetry not only works on ECG-gated CAC scans but also non-gated lung CT scans²¹. Non-contrast chest CT scans are prime candidates for opportunistic AI-enabled cardiac chamber volumetry for the identification of patients at increased risk for AF²² and HF. The AI approach can enable automatic screening of the over 10 million chest CT scans done each year in the US alone²³. Such an AI tool can run in the background of radiology picture archiving and communication systems (PACS) and alert providers to cases with enlarged cardiac chambers. Unfortunately, many high-risk patients with enlarged cardiac chambers are currently undetected and, therefore, untreated. Early detection of these cases can allow for close monitoring of progression to AF for stroke prevention and guideline-directed medical therapy for HF prevention. In our study, we have found the unadjusted correlations between Agatston CAC score and LA and LV volumes to be low ($R = 0.20$ and $R = 0.10$, respectively) (Supplementary

Figs. 1 and 2), hence a substantial portion of the population with enlarged LA and LV chambers are found in low-risk CAC categories. The combination of the automated cardiac chambers volumetry component of AI-CAC plus automated AI-CAC plaque characterization showed a greater incremental AUC value over the Agatston score vs. each alone.

Finally, the lack of coverage for CAC scans by Medicare and healthcare insurance carriers has contributed to healthcare inequity in the US. Ikram and Williams²⁴ have shown that low-income people in the Chicago area are less likely to get CAC scans compared to people in higher-income zip codes. We hope that by applying AI to CAC scans and providing incremental value, the payers will be more likely to cover CAC scans.

Our study has several strengths and limitations. The multi-ethnic nature of MESA recruited from six field centers around the US provides for greater generalization of our findings than single-center studies. MESA included standardized methods of data collection, laboratory measurement, follow-up, as well as adjudication of CVD events. Internal validation was not performed with hold-out due to the low sample size of several events.

One limitation is that the MESA Exam 1 baseline CT scans were performed between 2000 and 2002 using electron-beam computed tomography (EBCT) or earlier generation multidetector CT scanners, and current CAC scanning utilizes more advanced multidetector CT scanning. However, since our AI training was done completely outside of MESA and used a modern multidetector (256 slices) scanner, we do not anticipate this to affect the generalizability of our findings.

Another limitation is the potential impact of different ECG gating methods (RR-interval) used in MESA for multidetector CT (50%) and EBCT (80%). This discrepancy resulted in significant differences in LA volume between participants scanned with EBCT vs. MDCT (57.4 cc vs. 65.4 cc, respectively, $p < 0.0001$). However, LV volume, LV mass, and RV volume measurements were not affected by scanner type (Supplementary Table 7). The cumulative incidence of AF between LA volumes measured by different scanners showed similar results (Supplementary Figs. 10 and 11). Although interaction terms between LA volume and scanner type were tested and found to be non-significant for outcome prediction, questions remain on the extent of the impact of scanner type on our findings.

Since MESA does not distinguish between HF subtypes (heart failure with reduced ejection fraction (HFrEF) vs. heart failure with preserved ejection fraction (HFpEF)), we were unable to compare the prediction of HF subtypes. However, in a preliminary study, we obtained data from 75 patients who underwent both a cardiac CT scan and echocardiography at Harbor UCLA medical center²⁵. AI-CAC LV volume index (LVVI) defined as LV volume divided by BSA was able to distinguish HFrEF vs. HFpEF comparably to echocardiography LVVI (Supplementary Fig. 9).

Finally, our study excluded 771 cases who did not consent to the use of their data by commercial entities. However, the baseline characteristics of these cases did not differ systematically with respect to the remaining participants, and we do not anticipate this to affect our findings.

In this study, we presented AI-CAC data on cardiac chambers volumetry and calcified plaque characterization obtained from existing CAC scans in a large multi-ethnic prospective study and compared it to Agatston CAC Score alone for prediction of all cardiovascular events (stroke, myocardial infarction, angina, resuscitated cardiac arrest, all CVD-related deaths, HF, and AF), over 15 years. AI-CAC significantly improved upon the Agatston CAC score for all cardiovascular events prediction (including all CHD in CAC 1–100 cohort), as well as total mortality. Moreover, significant improvement in risk prediction and reclassification of events was not only seen for longer-term (e.g., 10- and 15-year) events but also for nearer-term (e.g., 1- and 5-year) events, advancing the status quo to help identify individuals at risk of near-term CVD events and death.

The projected impact of our study lies in AI's ability to provide opportunistic screening of enlarged cardiac chambers in all types of chest CT scans, including lung cancer screening and non-cardiac thoracic diagnostic CT scans. This manuscript primarily demonstrates the utility of applying AI to CAC scans to extract more actionable information than the Agatston CAC score that is currently reported. Additionally, AI-CAC can

measure chamber volume much faster and cheaper and is operator-independent compared to manual methods.

Methods

Study population

MESA is a prospective, population-based, observational cohort study of 6814 men and women without clinical CVD at the time of recruitment. Six field centers in the United States participated in the study: Baltimore, Maryland; Los Angeles, California; Chicago, Illinois; Forsyth County, North Carolina; New York City, New York; and St. Paul, Minnesota. As part of the initial evaluation (2000–2002), participants received a comprehensive medical history, clinic examination, and laboratory tests. Demographic information, medical history, and medication use at baseline were obtained by self-report. An ECG-gated non-contrast CT was performed at the baseline examination to measure CAC (see below).

Outcomes

The primary outcome was a composite of all CVD events comprised of stroke, myocardial infarction, angina, HF, AF, resuscitated cardiac arrest, and all CVD-related deaths. Participants were contacted by telephone every 9–12 months during follow-up and asked to report all new CVD diagnoses. International Classification of Disease (ICD) codes were obtained. For participant reports of HF, coronary heart disease, stroke, and CVD mortality, detailed medical records were obtained, and diagnoses were adjudicated by the MESA Morbidity and Mortality Committee. Incident AF was identified by ICD codes 427.3x (version 9) or I48.x (version 10) from inpatient stays and, for participants enrolled in fee-for-service Medicare, from Medicare claims for outpatient and provider services. Hard CVD was defined as myocardial infarction, resuscitated cardiac arrest, stroke, CHD death, and stroke death. Angina was classified, except in the setting of MI, as definite, probable, or absent. Definite or probable angina requires symptoms of typical chest pain or atypical symptoms. Probable angina requires, in addition to symptoms, a physician's diagnosis of angina, and medical treatment for it. Definite angina required one or more additional criteria, including CABG surgery or other revascularization procedures; 70% or greater obstruction on coronary angiography; or evidence of ischemia by stress tests or by resting ECG. A detailed study design for MESA has been published elsewhere²⁶. MESA participants have been followed since the year 2000. Incident AF has been identified through December 2018. 70 cases with AF diagnosed prior to MESA enrollment were removed from the analysis.

From the 6814 MESA participants, we excluded 771 who did not consent to the use of their data by commercial entities, leaving 6043 participants at baseline. Among the remaining participants, 125 participants with missing slices in CAC scans and 88 participants with missing event or time follow-up data were excluded, resulting in data from 5830 participants for final analysis. Of the 125 cases with missing slices in CAC scans were, 49.8% male and 50.2% female, with age 60.8 ± 10.1 . These errors were random, and our investigations did not reveal any association between cases with missing slices and any of the dependent or independent variables in our study.

The AI tool for automated cardiac chamber volumetry

The automated cardiac chambers volumetry tool in AI-CAC referred to in this study is called AutoChamber™ (HeartLung.AI, Houston, TX). The deep learning model used TotalSegmentator²⁷ as the base input and was further developed to segment each of the four cardiac chambers: LA, LV, RA, RV also in addition to several other components such as automated CAC score and plaque characterization, which are not presented here (Fig. 1). The core machine learning component of AI-CAC is adapted from TotalSegmentator which is a widely used anatomical model published and validated by investigators independent from our group. The source code for the TotalSegmentator base model is available publicly. The base architecture of the TotalSegmentator model was trained on 1139 whole-body CT cases with 447 cases of coronary CT angiography (CCTA) independent from MESA using nnU-Net, a self-configuring method for deep learning-based biomedical image segmentation²⁸. The initial input training data were

matched to non-contrast and contrast-enhanced ECG-gated cardiac CT scans with 1.5 mm slice thickness. Because the images were taken from the same patients in the same session, registration was done with good alignment. Following this transfer of segmentations, a nnU-Net deep learning tool was used for training the model. Additionally, iterative training was implemented whereby human supervisors corrected errors made by the model, and the corrected data were used to further train the model, leading to improved accuracy. To standardize the comparison in MESA, cardiac chambers were indexed by body surface area (BSA).

We developed a post-processing pipeline to identify instances of poor-quality segmentations where the region of interest was absent. Furthermore, we employed connected components analysis to eliminate ectopic segmentation islands. These methods were implemented to ensure quality control and optimize the model's performance. Expert rules built into the AI model excluded 125 cases due to missing slices in image reconstruction, which occurred with some of the electron-beam CT scanners used in MESA at baseline.

Agatston CAC score measurement

Three study sites used cardiac-gated electron-beam CT scanners, whereas the other three sites used multidetector CT scanners. Each participant was scanned twice at baseline examination, with a mean Agatston score used for analysis²⁹. All scans were phantom-adjusted and read by two trained CT image analysts at a central MESA CT reading center, with high reproducibility and comparability between electron-beam CT and multidetector CT scanning^{30,31}. Detailed information on CT scan methods and interpretation has been provided previously³⁰.

CAC area and density were derived from total Agatston and volume scores, which were provided in the original MESA data set. The methods for this derivation are elsewhere³².

AI-CAC plaque characterization beyond Agatston CAC score

In addition to AI-CAC cardiac chamber volumetry, AI-CAC enables calcified plaque characterization that currently is not reported by the Agatston CAC Score. These characteristics include the number of plaques, the number of vessels with plaques, plaque density, and location. In this study, we have only used these characteristics for calcified plaques, however, efforts are underway to characterize non-calcified (soft plaques) in non-contrast CAC scans using AI-CAC.

In MESA-1, human experts generated reports on plaque characteristics for each patient. For these reports, each expert manually identified plaques by clicking on them, extracting x, y, and z coordinates for each point, and measuring surface area using a connected components algorithm. This algorithm identified connected pixels that were adjacent side-by-side but excluded those connected diagonally. From these reports, we were able to extract information on plaque location, number of plaques, plaque density, and the number of vessels affected by plaques. We have chosen to use MESA's plaque characterization due to excessive noise in the coronary arteries in MESA-1 CT scans.

Statistical analysis

We used SAS (SAS Institute Inc., Cary, NC) and Python 3.10 for statistical analyses. All values are reported as means \pm SD except for CAC and plaque characteristics, which did not show normal distribution and are presented as median with interquartile range (IQR). All tests of significance were two-tailed, and significance was defined at Type I error (α) = 0.05 and Type II error (β) = 0.20. All analyses met the appropriate sample size and power considerations. Instances where these requirements were not met have been excluded and noted.

Survival analysis was performed using Cox proportional hazards regression. Model assumptions were tested using Schoenfeld and Martingale residuals and no violations of proportional hazards or non-linearity were detected in any variables. Discrimination was assessed using the time-dependent receiver operator characteristic (ROC) area under the curve (AUC)³³ and Uno's C-statistic³⁴. The time-dependent AUC was calculated

using the inverse probability of censoring weighting (IPCW) estimator without competing risks to determine discrimination at specific follow-up times. AUC confidence intervals were obtained using 1000 bootstrapped samples. Significance in the AUC difference between predictors was calculated based on the variance of the difference using the independent and identically distributed (iid)-representation of the AUC estimator. Uno's C-statistic was calculated to account for significant right censoring over 15 years of follow-up for all CVD predictions (70%). Significance in concordance discrimination was determined using 1000 bootstrapped samples.

Category-free (continuous) net reclassification index (NRI) was calculated using the sum of the differences between the proportions of upward reclassifications and downward reclassifications events and non-events, respectively. $P(\text{up}|\text{event})$ and $P(\text{down}|\text{nonevent})$ form the positive components of the NRI in expression, while events that move down and non-events that move up are mistakes introduced by the new marker. NRI was developed as a statistical measure to evaluate the improvement in risk prediction models when additional variables are incorporated into a base model³⁵.

The AI-CAC model, as presented, is comprised of the LA volume index, RV volume index, LV volume index, LV mass index, plaque characterization, and MESA-reported phantom-adjusted Agatston CAC score. Cardiac chamber volumetry was indexed by body surface area to standardize measurements. Because MESA participants were entirely asymptomatic without overt HF, LV volume and LV mass index demonstrated high collinearity and were combined into a composite variable. Correlation and variance inflation factor analysis showed low multicollinearity among the remaining predictors. Agatston CAC score was natural logarithm-transformed (ln-transformed + 1) to improve the interpretability of hazard ratios and avoid undue influence of large values. All predictors were modeled continuously and exhibited a linear relationship with outcomes.

The focus of this manuscript is comparing AI-CAC over Agatston CAC Score alone; therefore, no risk factors or other covariates were included in either model presented in the figures of this manuscript. However, incremental value of AI-CAC measurements over CVD risk factors have been provided in the Supplementary Information.

Data availability

No datasets were generated or analysed during the current study.

Code availability

SAS and Python codes used for statistical analysis in this study are available within a reasonable time from the publication date.

Received: 16 May 2024; Accepted: 22 October 2024;

Published online: 05 November 2024

References

- Greenland, P. & Lloyd-Jones, D. M. Role of coronary artery calcium testing for risk assessment in primary prevention of atherosclerotic cardiovascular disease: a review. *JAMA Cardiol.* **7**, 219–224 (2022).
- Grundy, S. M. et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA guideline on the management of blood cholesterol. *J. Am. Coll. Cardiol.* **73**, e285–e350 (2019).
- Goff, D. C. 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* **129**, S49–S73 (2014).
- Vega, G. L., Wang, J. & Grundy, S. M. Prevalence and significance of risk enhancing biomarkers in the United States population at intermediate risk for atherosclerotic disease. *J. Clin. Lipido.* **16**, 66–74 (2022).
- Kakadiaris, I. A. et al. Machine learning outperforms ACC / AHA CVD risk calculator in MESA. *J. Am. Heart Assoc.* **7**, e009476 (2018).
- Naghavi, M. et al. AI-enabled cardiac chambers volumetry in coronary artery calcium scans (AI-CACTM) predicts heart failure and

- outperforms NT-proBNP: the multi-ethnic study of atherosclerosis. *J. Cardiovasc. Comput. Tomogr.* **18**, 392–400 (2024).
7. Naghavi, M. et al. Automated left ventricular volumetry using artificial intelligence in coronary calcium scans (AI-CAC) predicts heart failure comparably to cardiac MRI and outperforms NT-proBNP: the multi-ethnic study of atherosclerosis (MESA). *J Am Coll Cardiol Img* [Under review]. (2024).
 8. Naghavi, M. et al. AI-enabled left atrial volumetry in coronary artery calcium scans (AI-CACTM) predicts atrial fibrillation as early as one year, improves CHARGE-AF, and outperforms NT-proBNP: the multi-ethnic study of atherosclerosis. *J. Cardiovasc. Comput. Tomogr.* **18**, 383–391 (2024).
 9. Reeves, A. et al. Artificial intelligence-enabled automated left atrial volumetry in coronary calcium scans predicts atrial fibrillation as early as one year: multi-ethnic study of Atherosclerosis. *J Cardiovasc Comput Tomogr. Society of Cardiovascular Computed Tomography.* **17**, S1–S96 (2023).
 10. Daniel, K. R. et al. Comparison of methods to measure heart size using noncontrast-enhanced computed tomography: correlation with left ventricular mass. *J. Comput. Assist. Tomogr.* **32**, 934–941 (2008).
 11. Bittencourt, M. S. et al. Left ventricular area on non-contrast cardiac computed tomography as a predictor of incident heart failure—the multi-ethnic study of atherosclerosis. *J. Cardiovasc. Comput. Tomogr.* **10**, 500–506 (2016).
 12. Qureshi, W. T. et al. Determination and distribution of left ventricular size as measured by noncontrast CT in the multi-ethnic study of atherosclerosis. *J. Cardiovasc. Comput. Tomogr.* **9**, 113–119 (2015).
 13. Dykun, I. et al. Left ventricle size quantification using non-contrast-enhanced cardiac computed tomography—association with cardiovascular risk factors and coronary artery calcium score in the general population: the Heinz Nixdorf Recall Study. *Acta Radiol.* **56**, 933–942 (2015).
 14. Mahabadi, A. A. et al. Left atrial size quantification using non-contrast-enhanced cardiac computed tomography—association with cardiovascular risk factors and gender-specific distribution in the general population: the Heinz Nixdorf Recall study. *Acta Radiol.* **55**, 917–925 (2014).
 15. Mahabadi, A. A. et al. Association of epicardial adipose tissue and left atrial size on non-contrast CT with atrial fibrillation: the Heinz Nixdorf Recall Study. *Eur. Heart J. Cardiovasc. Imaging* **15**, 863–869 (2014).
 16. Mahabadi, A. A. et al. Association of computed tomography-derived left atrial size with major cardiovascular events in the general population: the Heinz Nixdorf Recall Study. *Int. J. Cardiol.* **174**, 318–323 (2014).
 17. Mahabadi, A. A. et al. Noncoronary measures enhance the predictive value of cardiac CT above traditional risk factors and CAC score in the general population. *JACC Cardiovasc. Imaging* **9**, 1177–1185 (2016).
 18. Miller, R. J. H. et al. AI-defined cardiac anatomy improves risk stratification of hybrid perfusion imaging. *JACC Cardiovasc. Imaging* <https://doi.org/10.1016/j.jcmg.2024.01.006> (2024).
 19. Kizer JR, Bella JN, Palmieri V, et al. Left atrial diameter as an independent predictor of first clinical cardiovascular events in middle-aged and elderly adults: the Strong Heart Study (SHS). *Am Heart J.* 2006;151:412–418.
 20. Power, S. P. et al. Computed tomography and patient risk: facts, perceptions and uncertainties. *World J. Radiol.* **8**, 902–915 (2016).
 21. Reeves, A. et al. AI-enabled automated cardiac chambers volumetry in non-contrast ECG-gated cardiac scans vs. non-contrast non-gated lung scans. *J Cardiovasc Comput Tomogr. Soc. Cardiovasc. Comput. Tomogr.* **17**, S1–S96 (2023).
 22. Naghavi M, Yankelevitz D, Reeves AP, et al. AI-enabled left atrial volumetry in coronary artery calcium scans (AI-CACTM) predicts atrial fibrillation as early as one year, improves CHARGE-AF, and outperforms NT-proBNP: The multi-ethnic study of atherosclerosis. *J Cardiovasc Comput Tomogr.* 2024;18:383–391.
 23. Mahesh, M., Ansari, A. J. & Mettler, F. A. Patient exposure from radiologic and nuclear medicine procedures in the United States and worldwide: 2009–2018. *Radiology* **307**, e221263 (2023).
 24. Ikram, M. & Williams, K. A. Socioeconomics of coronary artery calcium: is it scored or ignored? *J. Cardiovasc. Comput. Tomogr.* **16**, 182–185 (2022).
 25. Naghavi, M. et al. *AI-Enabled Cardiac Chambers Volumetry in Non-contrast Cardiac CT scans Detects HFrEF vs. HFpEF.* (American Heart Association Scientific Session, 2024).
 26. Bild, D. E. et al. Multi-ethnic study of atherosclerosis: objectives and design. *Am. J. Epidemiol.* **156**, 871–881 (2002).
 27. Wasserthal J, Breit HC, Meyer MT, et al. TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. *Radiol Artif Intell.* 2023;5:e230024.
 28. Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J. & Maier-Hein, K. H. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**, 203–211 (2021).
 29. Agatston, A. S. et al. Quantification of coronary artery calcium using ultrafast computed tomography. *J. Am. Coll. Cardiol.* **15**, 827–832 (1990).
 30. Carr, J. J. et al. Calcified coronary artery plaque measurement with cardiac CT in population-based studies: standardized protocol of Multi-Ethnic Study of Atherosclerosis (MESA) and Coronary Artery Risk Development in Young Adults (CARDIA) study. *Radiology* **234**, 35–43 (2005).
 31. Detrano, R. C. et al. Coronary calcium measurements: effect of CT scanner type and calcium measure on rescan reproducibility—MESA study. *Radiology* **236**, 477–484 (2005).
 32. Criqui, M. H. et al. Calcium density of coronary artery plaque and risk of incident cardiovascular events. *JAMA* **311**, 271–278 (2014).
 33. Kamarudin, A. N., Cox, T. & Kolamunnage-Dona, R. Time-dependent ROC curve analysis in medical research: current methods and applications. *BMC Med. Res. Methodol.* **17**, 53 (2017).
 34. Uno, H., Cai, T., Pencina, M. J., D’Agostino, R. B. & Wei, L. J. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat. Med.* **30**, 1105–1117 (2011).
 35. Pencina, M. J., D’Agostino, R. B., D’Agostino, R. B. & Vasan, R. S. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat. Med.* **27**, 157–172 (2008).

Acknowledgements

This research was supported by 2R42AR070713 and R01HL146666 and MESA was supported by contracts 75N92020D00001, HHSN268201500003I, N01-HC-95159, 75N92020D00005, N01-HC-95160, 75N92020D00002, N01-HC-95161, 75N92020D00003, N01-HC-95162, 75N92020D00006, N01-HC-95163, 75N92020D00004, N01-HC-95164, 75N92020D00007, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168 and N01-HC-95169 from the National Heart, Lung, and Blood Institute, and by grants UL1-TR-000040, UL1-TR-001079, and UL1-TR-001420 from the National Center for Advancing Translational Sciences (NCATS). The authors thank the other investigators, the staff, and the participants of the MESA study for their valuable contributions. The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood Institute; the National Institutes of Health; or the U.S. Department of Health and Human Services. A full list of participating MESA investigators and institutions can be found at <http://www.mesa-nhlbi.org>.

Author contributions

M.N.: conceptualization of the study, research design, proposal to MESA, project administration, writing the original draft, and revision and editing. A.P.R. and C.Z.: AI development, interpretation. K.A.: AI development, data analysis, interpretation, and manuscript revision and editing. T.A., C.L.H.,

D.F.Y., D.L., M.J.B., S.K.R., K.N., J.N., I.K., S.M., Z.F., D.J.M., M.V.M., K.W., D.L. P.K.S.: research design, proposal to MESA, manuscript brainstorming, reviewing data analysis, revisions, and editing. N.D.W.: project administration, research design, proposal to MESA, manuscript brainstorming, reviewing data analysis, revisions, and editing.

Competing interests

Several members of the writing group are inventors of the AI tool mentioned in this paper. M.N. is the founder of HeartLung.AI. A.P.R., T.A., D.F.Y., N.D.W., and D.L. are advisors to HeartLung.AI. C.Z. is a software engineer for HeartLung.AI. K.A. is a graduate research associate of HeartLung.AI. The remaining authors declare no competing interests.

Ethical approval

As a longitudinal population-based study sponsored by the National Institute of Health (NIH), MESA has received proper ethical oversight. The MESA protocol was approved by the Institutional Review Board (IRB) of the 6 field centers (Columbia University IRB, Johns Hopkins Medicine IRB, Northwestern University IRB, UCLA Office of the Human Research Protection Program (OHRPP), University of Minnesota Human Research Protection Program, Wake Forest Baptist Health IRB) and the National Heart, Lung, and Blood Institute. Data from participants who did not consent to commercial use were removed from our study.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41746-024-01308-0>.

Correspondence and requests for materials should be addressed to Morteza Naghavi.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024