JACC: ADVANCES VOL. ■, NO. ■, 2026
© 2026 THE AUTHORS. PUBLISHED BY ELSEVIER ON BEHALF OF THE AMERICAN COLLEGE OF CARDIOLOGY FOUNDATION. THIS IS AN OPEN ACCESS ARTICLE UNDER THE CC BY LICENSE (http://creativecommons.org/licenses/by/4.0/).

ORIGINAL RESEARCH

Artificial Intelligence Detection of Aortic Valve Calcification on Routine Noncontrast Chest CT

A Real-World Cohort

Ofir Koren, MD, FESC,[a] Vivek Patel, MS,[a] Orit Wimpfheimer, MD,[b] Alexander Kukuev, MD,[a] Parker Chu, MPH,[a] Harsh Jain, MBBS,[a] Kazuki Suruga, MD,[a] Hidemasa Shitan, MD,[a] Shiri Hassid, PhD,[b] Daniel Ng, BS,[a] Shir Hofstetter, PhD,[b] Raj R. Makkar, MD[a]

ABSTRACT

BACKGROUND Routine noncontrast chest computed tomography (CT) may reveal aortic valve calcification (AVC) outside aortic stenosis evaluation. Its prevalence, burden, and documentation remain uncertain. Artificial intelligence (AI) may enable systematic detection.

OBJECTIVES The study aims to evaluate AVC prevalence and burden, clinical correlates, radiology documentation, and AI-assisted case finding.

METHODS This retrospective single-center study screened 511 consecutive noncontrast, non-electrocardiographically gated chest CT examinations in adults aged ≥50 years. After 15 protocol exclusions and 56 software-designated non-evaluable examinations, 440 underwent quantification by 2 blinded readers. AVC was defined as >10 AU using a 130-Hounsfield-unit threshold adjusted for blood-pool attenuation and image noise when needed. Locked HealthAVC V1.0 was a secondary comparator.

RESULTS Manual quantification identified AVC in 172 of 440 scans (39.1%; 95% CI: 34.6%-43.7%); 97 of 440 (22.0%) had ≥100 AU. Inter-reader agreement at >10 AU was 99.5% (kappa = 0.990). Age (adjusted OR: 1.92 per 5 years; 95% CI: 1.63-2.26) and male sex (adjusted OR: 2.09; 95% CI: 1.32-3.30) were associated with AVC. Reports documented 27 of 172 positive scans (15.7%; 95% CI: 11.0%-21.9%). AI detected 169 of 172 manually confirmed cases (98.3%) among 440 evaluable scans, including 142 of 145 undocumented cases. Of 496 submitted scans, 56 (11.3%) were nonevaluable. End-to-end sensitivity counting positive rejected scans as missed was 89.9%. Among 86 AI-positive patients with echocardiography, 10 (11.6%) had mild-or-greater aortic stenosis.

CONCLUSIONS AVC was common but infrequently documented on routine chest CT. AI detected most cases among evaluable scans; 11.3% were nonevaluable, yielding 89.9% end-to-end sensitivity. Prospective studies should assess clinical pathways. (JACC Adv. 2026;■:103446) © 2026 The Authors. Published by Elsevier on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

From the Smidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA; and the NanoxAI Ltd, Neve Ilan, Israel.

The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the Author Center.

Manuscript received August 25, 2026; revised manuscript received October 6, 2026, accepted October 7, 2026.

ISSN 2772-963X https://doi.org/10.1016/j.jacadv.2026.103446

ABBREVIATIONS AND ACRONYMS

AI = artificial intelligence

AS = aortic stenosis

AU = Agatston units

AUPRC = area under the precision-recall curve

AUROC = area under the receiver operating characteristic curve

AVA = aortic valve area

AVC = aortic valve calcification

CT = computed tomography

ECG = electrocardiography

ICC = intraclass correlation coefficient

SAVR = surgical aortic valve replacement

TAVR = transcatheter aortic valve replacement

Aortic valve calcification (AVC) is the anatomical substrate of calcific aortic valve disease and an early phenotype along the continuum from valve sclerosis to hemodynamically important aortic stenosis (AS). Progressive leaflet mineralization and fibrosis increase cusp stiffness and cause obstruction. Quantitative computed tomography (CT) AVC is a marker of disease burden and a guideline-supported adjunct when echocardiographic AS severity is discordant or uncertain.[1-4]

Community studies extend its relevance beyond AS referrals. In the MESA (Multi-Ethnic Study of Atherosclerosis), AVC predicted graded long-term risk of severe AS among participants without a known cardiovascular disease, whereas a score of 0 identified an extremely low-risk group.[5] Longitudinal analyses identified factors associated with incident calcification.[6] Lower burdens may mark an earlier disease; surveillance remains uncertain.

Routine noncontrast chest CT for noncardiac indications frequently includes the valve, allowing assessment without additional imaging. Consensus recommendations support recognition and visual grading of incidental AVC on nongated thoracic CT,[7] including lung cancer-screening cohorts.[8] Incidental valve calcification is associated with cardiovascular events, AS, mortality, and management changes.[9,10]

Systematic recognition is technically difficult. The oblique annular plane requires separating leaflet from adjacent aortic root, coronary, and outflow-tract calcium by multiplanar reconstruction. Nongated acquisition introduces motion, while slice thickness, reconstruction kernel, noise, and attenuation threshold affect scoring.[7,11,12] These constraints may contribute to incomplete documentation.

A downstream pathway remains undefined. Sex-specific severe AVC thresholds support selected diagnostic settings, and nongated CT thresholds can identify significant AS.[13] Lower incidental burdens lack validated referral or surveillance pathways. MESA and sex-specific Tastet/Ali categories organize burden but do not establish management.[4,5] Artificial intelligence (AI) prediction of progression requires prospective validation.[14]

We evaluated manually adjudicated AVC prevalence, clinical correlates, report documentation, and exploratory echocardiography in consecutive noncontrast, non-electrocardiography (ECG)-gated chest CT examinations from adults not selected for AS. Expert manual quantification defined the reference standard, and HealthAVC V1.0 (NanoxAI Ltd), a pretrained and locked AI model, was a secondary comparator for identifying manually confirmed AVC absent from original reports.

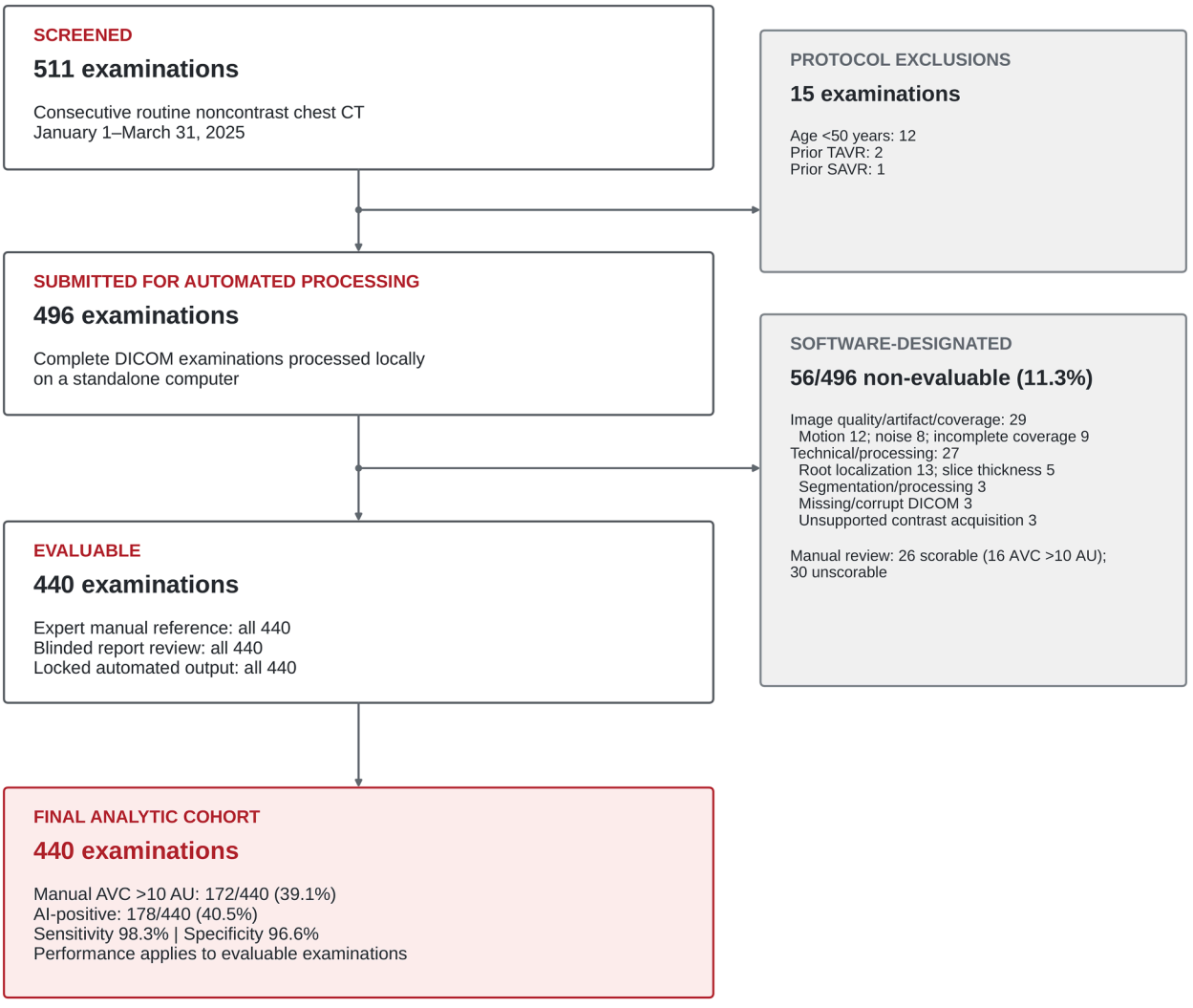
METHODS

STUDY DESIGN AND POPULATION. We retrospectively analyzed consecutive noncontrast, non-ECG-gated chest CT examinations at the Cedars-Sinai Medical Center from January 1 through March 31, 2025. Of 511 screened, 15 met protocol exclusions: age <50 years (n = 12), prior transcatheter aortic valve replacement (n = 2), or prior surgical valve replacement (n = 1). Of 496 submitted for AI processing, 56 were nonevaluable, leaving 440 (**Figure 1**). Three software-identified contrast-enhanced examinations were recognized only after submission and remained processing failures. Rejected scans were not reuploaded or reprocessed; no manual review was used to rescue AI output.

CT ACQUISITION AND RECONSTRUCTION. Routine multidetector CT examinations included the heart. We sampled DICOM headers of 440 evaluable examinations for reconstruction thickness, manufacturer/model, and tube voltage. Eligible original nonscout CT series had ≥30 indexed images and a recorded thickness >0 to ≤3 mm. If multiple thicknesses qualified, we selected one matching a derived Detect Index series. Metadata linked unambiguously to a diagnostic chest CT series in 401 examinations; 39 were incomplete or ambiguous. Manufacturers were GE (n = 181), Canon/Toshiba (n = 192), Siemens (n = 27), and Philips (n = 1); 373 of 401 (93.0%) were 120 kVp. Kernel and CTDIvol were inconsistently recorded and are not reported (Supplemental Table 14).

MANUAL AVC REFERENCE STANDARD. Two experienced structural-heart CT readers (V.P. and O.K.) independently quantified AVC in all 440 evaluable scans using 3mensio v9.0 (Pie Medical Imaging) at the Cedars-Sinai core laboratory. Both disclosed NanoxAI relationships related to a companion study. They were blinded to AI output, each other's scores, radiology reports, echocardiography, and clinical information; AI output was withheld until manual scores were final. Neither handled upload or processing. Starting at 130 HU, readers adjusted thresholds within 120 to 350 HU when needed for blood-pool attenuation or noise. A multiplanar review was performed including leaflet calcium and excluded aortic sinus, outflow-tract, aorto-mitral, and other nonleaflet calcium. Uncertain or discordant scans underwent joint consensus without a third reader. AVC >10 AU was prespecified as positive.

FIGURE 1 Study Design and Cohort Derivation



Of 511 consecutive noncontrast, non–ECG-gated chest CT examinations screened, 15 met protocol exclusions before automated processing. Of 496 examinations uploaded to the standalone system, 56 were software-designated nonevaluable, and 440 comprised the analytic cohort. A manual review scored 26 of the 56 nonevaluable examinations; 30 were unscorable (Supplemental Table 12). AVC = aortic valve calcification; CT = computed tomography; ECG = electrocardiography; SAVR = surgical aortic valve replacement; TAVR = transcatheter aortic valve replacement.

CLINICAL VARIABLES. Demographics, CT indication, anthropometrics, comorbidities, medication, prior AS, referrals, valve evaluation, and intervention were abstracted from records. The principal comparison was manual AVC >10 vs ≤10 AU; age was continuous in the primary model.

ECHOCARDIOGRAPHY AND MISSING-DATA HANDLING. Clinically obtained echocardiograms within ±365 days of CT were identified in AI-positive and AI-negative patients. AS severity was harmonized as none, mild, moderate, severe, or not assessed using report findings and peak-velocity categories (<2.0, 2.0-<3.0, 3.0-<4.0, ≥4.0 m/s); the more severe supported category was retained. Mean gradient, valve area, dimensionless index, and stroke-volume index were analyzed separately. Missing measurements were neither imputed nor considered normal. Analyses used variable-specific available pairs, with AI-negative findings summarized separately. Echocardiography was clinically selected, and its indication was not systematically captured, limiting inference.

ORIGINAL RADIOLOGY REPORT REVIEW. An investigator (AK) blinded to manual and AI scores reviewed original radiology reports. Only explicit

AVC wording counted as documented; nonspecific aortic, atherosclerotic, or coronary calcification did not. Documentation was assessed among 172 manually positive scans.

AUTOMATED ANALYSIS AS A SECONDARY COMPARATOR. HealthAVC V1.0 is a pretrained, locked postprocessing model for AVC detection and quantification on routine noncontrast, non-ECG-gated chest CT.

HealthAVC V1.0 uses U-Net-based convolutional neural networks for anatomical localization and calcium detection, producing an automated Agatston score and visual localization. According to the manufacturer, supervised training used noncontrast, non-ECG-gated CT from varied sources, scanner manufacturers, acquisition parameters, ages, and sexes, with annotations validated by American Board of Radiology-certified radiologists. Development, hyperparameter selection, and testing used separate data partitions; no Cedars-Sinai data contributed to development or calibration. The locked version was not trained or tuned on this cohort. Intended for assisted AVC detection and quantification in adults aged ≥40 years, its output requires clinical interpretation and does not replace CT images or the radiology report. Implementation details and weights remain proprietary.

Investigators supplied deidentified DICOM examinations on physical media. A NanoxAI employee uploaded complete examinations to an offline Cedars-Sinai computer and returned printed reports; no imaging data left Cedars-Sinai. The employee lacked access to clinical data, reports, echocardiography, and manual scores and did not analyze results. No manual series selection or preprocessing occurred. Outputs remained separate until analysis and provided an AVC score, burden category, and localization; positivity was prespecified as >10 AU. Five examinations were rejected for "slice thickness outside supported range." Outputs were not corrected or adjudicated. HealthAVC V1.0 was investigational and not Food and Drug Administration-cleared at study time; results did not affect care.

The companion technical-validation manuscript under review uses this cohort, manual reference measurements, and AI outputs. Its overlapping analyses are binary >10-AU diagnostic performance (see Results) and AI-manual agreement across burden frameworks (Supplemental Table 5), derived from the same data rather than independent replication. Paired scores are displayed descriptively in Results; a statistical analysis of continuous AI-manual agreement appears only in the companion manuscript. AVC prevalence, explicit report documentation, clinical correlates, and echocardiographic associations are reported only here.

NanoxAI supplied software access and onsite operational support. Company-affiliated authors provided model information and reviewed aggregate results and the manuscript. NanoxAI had no role in reference measurements, report review, clinical data abstraction, statistical analysis, interpretation, or the decision to submit; academic investigators retained final authority over content and submission.

AVC BURDEN FRAMEWORKS. Manual AVC >10 AU was the primary definition. Three complementary frameworks described burden: the primary 3-tier scheme (≤10 AU, detectable nonsevere, or sex-specific severe at ≥1,300 AU in women or ≥2,000 AU in men); modified MESA (≤10, 11-99, 100-299, and ≥300 AU); and modified Tastet/Ali, retaining ≤10 AU as no AVC with sex-specific higher tiers. These categories are descriptive, not interchangeable management thresholds.[1,4,5]

STATISTICAL ANALYSIS. Continuous variables are provided as median (IQR); categorical variables as n (%). Table 1 is descriptive; exploratory comparisons used Mann-Whitney *U* or Fisher exact tests. Preconsensus reader reproducibility used Pearson and Spearman correlations, Lin concordance, intraclass correlation coefficient (ICC)(A,1), Bland-Altman bias and 95% limits, mean absolute error, and categorical agreement. Fixed 130-HU manual remeasurement used the same series, valve region, and Agatston method and was compared with final consensus adaptive-threshold scores, not individual reader scores. Continuous agreement intervals used 2,000 bootstrap resamples; Wilson intervals described prevalence, reporting, and diagnostic proportions. The prespecified complete-case logistic model included age, sex, hypertension, diabetes, dyslipidemia, chronic kidney disease, coronary artery disease, heart failure, atrial fibrillation, and statin use. All 440 had complete model data; no imputation was used. Age odds ratios are per 5 years. A quadratic age term and a sensitivity model adding smoking and CT indication were examined. Prevalence and reporting were primary descriptive findings; other analyses were exploratory. No multiplicity adjustment was applied, and 2-sided *P* values are nominal. Spearman intervals used 2,000 paired bootstrap resamples; likelihood-ratio intervals used the log-normal approximation. Reader and fixed-versus-adaptive score plots appear in Supplemental Figures 1 and 2, with the three-tier cross-tabulation in Supplemental Table 4.

TABLE 1 Baseline Characteristics According to Manually Quantified Aortic Valve Calcium Burden

	Overall (N = 440)	AVC >10 AU (n = 172)	AVC ≤10 AU (n = 268)
Age, y, median [Q1-Q3]	71.0 [64.0-75.0]	73.0 [70.0-80.0]	68.0 [61.0-73.0]
Male, n (%)	194 (44.1)	89 (51.7)	105 (39.2)
Body mass index, kg/m^2, median [Q1-Q3]	25.5 [22.2-29.3]	25.2 [22.1-28.8]	25.7 [22.3-29.5]
CT indication group			
Pulmonary, n (%)	322 (73.2)	111 (64.5)	211 (78.7)
Cardiovascular/aortic, n (%)	60 (13.6)	35 (20.3)	25 (9.3)
Other/mixed/unknown, n (%)	58 (13.2)	26 (15.1)	32 (11.9)
Smoking history, n (%)	61 (13.9)	27 (15.7)	34 (12.7)
Hypertension, n (%)	181 (41.1)	89 (51.7)	92 (34.3)
Diabetes mellitus, n (%)	80 (18.2)	37 (21.5)	43 (16.0)
Dyslipidemia, n (%)	164 (37.3)	74 (43.0)	90 (33.6)
Chronic kidney disease, n (%)	41 (9.3)	24 (14.0)	17 (6.3)
Coronary artery disease, n (%)	72 (16.4)	41 (23.8)	31 (11.6)
Prior myocardial infarction, n (%)	15 (3.4)	9 (5.2)	6 (2.2)
Prior percutaneous coronary intervention, n (%)	17 (3.9)	12 (7.0)	5 (1.9)
Heart failure, n (%)	28 (6.4)	17 (9.9)	11 (4.1)
Atrial fibrillation, n (%)	28 (6.4)	17 (9.9)	11 (4.1)
Chronic obstructive pulmonary disease, n (%)	35 (8.0)	18 (10.5)	17 (6.3)
Cancer history, n (%)	110 (25.0)	40 (23.3)	70 (26.1)
Peripheral vascular disease, n (%)	12 (2.7)	6 (3.5)	6 (2.2)
Prior sternotomy, n (%)	11 (2.5)	5 (2.9)	6 (2.2)
Statin use, n (%)	80 (18.2)	37 (21.5)	43 (16.0)

Values are median [Q1-Q3] or n (%). All variables shown were available for all 440 evaluable examinations. Each row identifies its summary statistic; no baseline hypothesis tests are presented.

AU = Agatston units; AVC = aortic valve calcification; CT = computed tomography.

The sensitivity model added smoking and grouped CT indication, retaining unknown indications in the other/mixed/unknown category. Its apparent area under the curve (AUC) was calculated in the fitting cohort without validation and is distinct from AI-detection AUC (Supplemental Table 6).

Receiver operating characteristic and precision-recall curves used continuous AI scores against manual AVC >10 AU in the 440 evaluable scans. Area under the receiver operating characteristic curve (AUROC) and trapezoidal area under the precision-recall curve (AUPRC) confidence intervals used 2,000

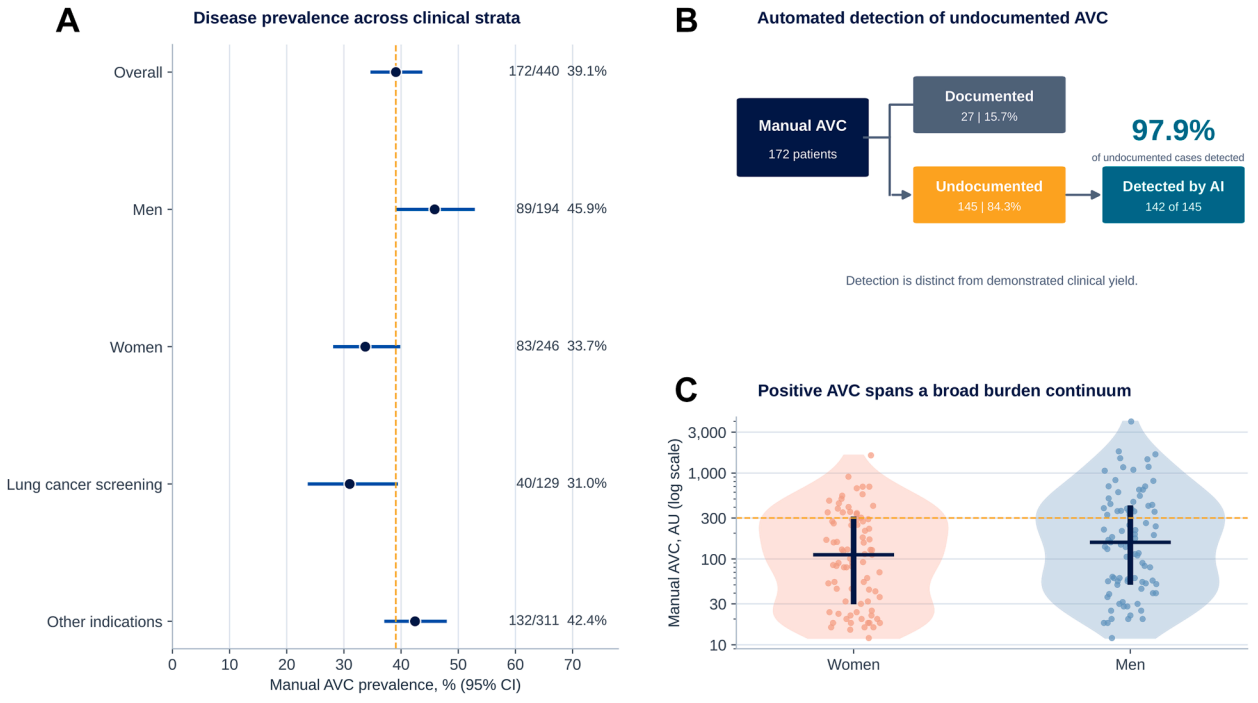
TABLE 2 Original-Report Documentation and Automated Detection of Manually Confirmed AVC >10 AU

Measure	Result	95% CI or Detail
Explicitly documented among manual AVC-positive examinations	27/172 (15.7%)	11.0%-21.9%; confirmed adjudicated report review
Not explicitly documented among manual AVC-positive examinations	145/172 (84.3%)	Explicit wording criterion
AI sensitivity	169/172 (98.3%)	95.0%-99.4%
AI specificity	259/268 (96.6%)	93.7%-98.2%
AI positive predictive value	169/178 (94.9%)	90.7%-97.3%
AI negative predictive value	259/262 (98.9%)	96.7%-99.6%
AI accuracy	428/440 (97.3%)	95.3%-98.4%
AI-identified, manually confirmed cases not explicitly documented	142/145 (97.9%)	Exploratory documentation comparison
AI AUROC	0.9989	0.9974-0.9999
AI AUPRC (trapezoidal)	0.9984	0.9962-0.9998
AI positive likelihood ratio (LR+)	29.26	15.39-55.63
AI negative likelihood ratio (LR−)	0.018	0.006-0.055

Documentation results use the 172 manually positive examinations. Automated diagnostic proportions use all 440 evaluable examinations. AUROC and AUPRC use continuous AI scores against manual AVC >10 AU; AUPRC uses trapezoidal integration. Curve confidence intervals use 2,000 class-stratified bootstrap resamples. See Supplemental Figure 3. LR+ and LR− confidence intervals use a log-normal approximation and apply to evaluable scans at >10 AU.

AI = artificial intelligence; AUPRC = area under the precision-recall curve; AUROC = area under the receiver operating characteristic curve; other abbreviations as in Table 1.

FIGURE 2 Aortic Valve Calcification Prevalence, Documentation, and Automated Detection



Manual AVC prevalence by clinical stratum with Wilson 95% CIs (A); explicit original-report documentation and AI detection of undocumented manual-positive cases (B); and the distribution of positive manual AVC scores in women and men, with median and interquartile range (C). The 300-AU line denotes the modified MESA high-burden threshold. Detection does not establish clinical yield. AI = artificial intelligence; MESA = Multi-Ethnic Study of Atherosclerosis; other abbreviations as in Figure 1.

class-stratified bootstrap resamples. The prespecified AI >10-AU cutoff was marked without optimization (Supplemental Figure 3).

Exploratory performance analyses among 401 scans with assignable DICOM metadata used thickness groups ≤1.0, >1.0 to 1.5, and >1.5 to 3.0 mm and scanner manufacturer data. Sensitivity and specificity used Wilson intervals; 39 others remained in primary analyses.

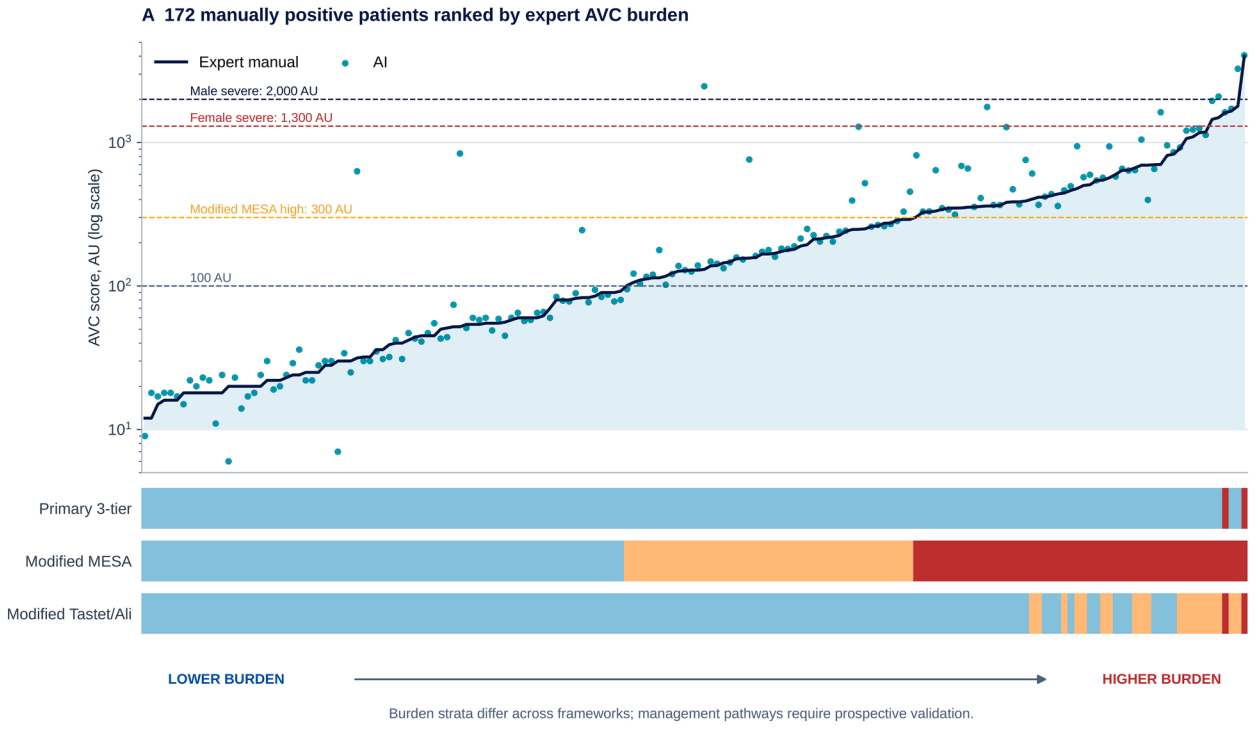
During revision, the same 2 readers manually reviewed all 56 rejected examinations using the 3mensio protocol, blinded to clinical data and reports but aware of rejection. Scans were unscorable if the valve/root was incompletely imaged, noise or motion obscured leaflet calcium, DICOM data were missing or corrupt, the acquisition was contrast-enhanced, or thickness exceeded 3 mm. Prevalence among manually scorable examinations used Wilson intervals; bounds treated the 30 unscorable examinations as all negative or all positive. End-to-end sensitivity counted manually positive rejected scans as missed detections because no AI result was returned (Supplemental Table 12).

ETHICS. The Cedars-Sinai Medical Center Institutional Review Board (STUDY00004566) initially approved the study effective November 7, 2025. The final protocol amendment applicable to this cohort was approved in April 2026, together with waivers of informed consent and Health Insurance Portability and Accountability Act (HIPAA) authorization.

RESULTS

COHORT AND PREVALENCE. The evaluable cohort comprised 440 examinations: median age 71.0 years (Q1-Q3: 64.0-75.0), with 194 men (44.1%). Manual AVC >10 AU occurred in 172 of 440 scans (39.1%; 95% CI: 34.6%-43.7%) and AVC ≥100 AU in 97 of 440 scans (22.0%; 95% CI: 18.4%-26.2%) (Supplemental Table 13). Prevalence was 89 of 194 (45.9%) in men and 83 of 246 (33.7%) in women. By CT indication, it was 111 of 322 (34.5%) for pulmonary, 35 of 60 (58.3%)

FIGURE 3 Manual and Automated AVC Burden Across Classification Frameworks



The 172 manually positive patients are ranked by manual AVC score with paired AI scores overlaid. Aligned stripes show primary 3-tier, modified MESA, and modified Tastet/Ali manual categories. Blue, orange, and red indicate increasing burden within each framework. Lines mark 100 AU, 300 AU, and female (1,300 AU) and male (2,000 AU) severe thresholds. These are not management thresholds on nongated CT. Full-cohort counts are in Supplemental Table 5. Abbreviations as in **Figures 1 and 2**.

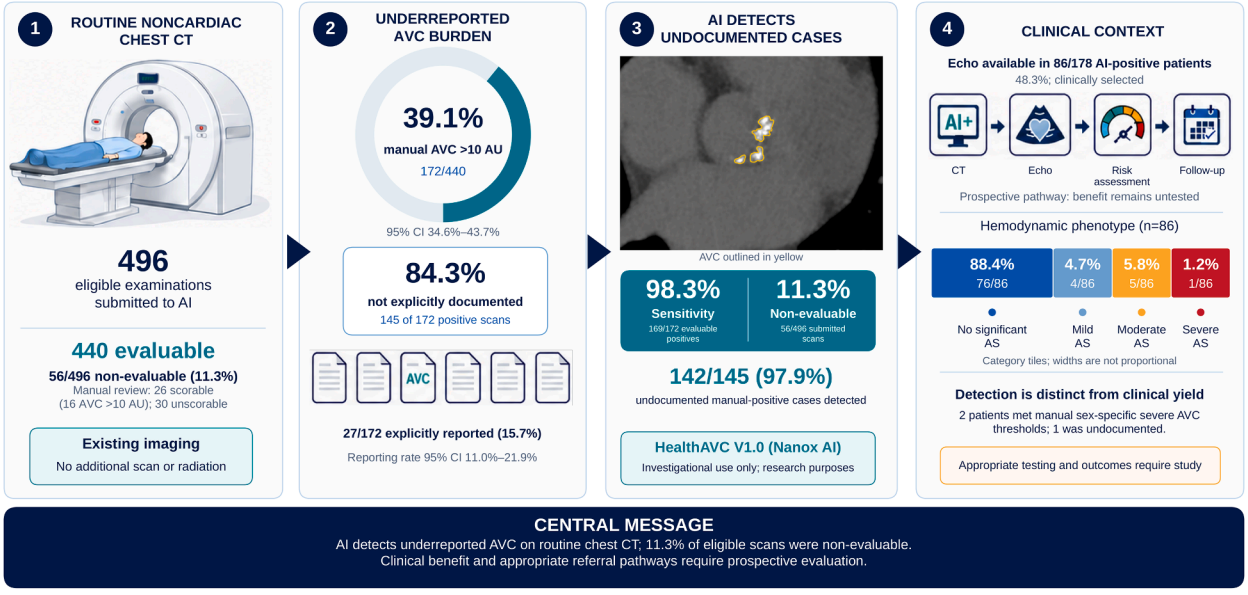
for cardiovascular/aortic, and 26 of 58 (44.8%) for other/mixed/unknown indications (**Table 1**, Supplemental Table 1A). By lung cancer screening status, prevalence was 40 of 129 (31.0%) in screening examinations and 132 of 311 (42.4%) in other indications (**Figure 2A**, Supplemental Table 1B).

Manual review scored 26 of the 56 nonevaluable examinations; 30 were unscorable because of incomplete aortic root coverage (n = 14), slice thickness >3 mm (n = 5), image noise (n = 4), missing or corrupt DICOM data (n = 3), contrast enhancement (n = 3), or motion (n = 1). AVC >10 AU was present in 16 of 26 scorable rejected examinations (61.5%; 95% CI: 42.5%-77.6%). Among all 466 manually scored examinations, AVC prevalence was 40.3% (188/466; 95% CI: 36.0%-44.9%); assigning all 30 unscorable examinations as negative- or positive-bounded prevalence across the 496 eligible examinations at 37.9% to 44.0% (Supplemental Table 12).

Among 56 software-designated nonevaluable examinations, software-reported reasons included motion artifact (n = 12), image noise (n = 8), incomplete valve/root coverage (n = 9), aortic-root localization failure (n = 13), unsupported slice thickness (n = 5), segmentation/processing failure (n = 3), missing/corrupt DICOM (n = 3), and contrast-enhanced acquisition not supported by the software (n = 3). Compared with the evaluable cohort, nonevaluable examinations involved younger patients (61.0 [59.0-63.3] vs 71.0 [64.0-75.0] years; $P < 0.001$), more men (60.7% vs 44.1%; $P = 0.022$), and lower prevalences of hypertension, diabetes, and dyslipidemia; broad CT-indication distribution did not differ ($P = 0.394$) (Supplemental Tables 2A and 2B).

READER REPRODUCIBILITY AND FIXED-THRESHOLD SENSITIVITY. Both readers independently scored all 440 examinations. Inter-reader ICC(A,1) was 0.991 (95% CI: 0.990-0.992); agreement at >10 AU was 438 of 440 (99.5%; kappa = 0.990). Thirteen examinations (3.0%) required consensus review: 7 for reader differences and 6 for reader uncertainty. Fixed 130-HU scoring produced identical AVC prevalence (172/

CENTRAL ILLUSTRATION AI Detection of Aortic Valve Calcification on Routine Noncontrast Chest CT



Koren O, et al. JACC Adv. 2026;■(■):103446.

Of 496 submitted scans, 56 (11.3%) were nonevaluable; manual review scored 26, including 16 with AVC >10 AU, while 30 were unscorable. Among 440 evaluable scans, 172 had manual AVC >10 AU (39.1%; 95% CI: 34.6%-43.7%). Explicit reporting occurred in 27 of 172 (15.7%; 95% CI: 11.0%-21.9%); AI detected 142 of 145 undocumented cases. Sensitivity of 98.3% applies to evaluable scans. Echocardiography was clinically selected in 86 of 178 AI-positive patients. Two met manual sex-specific severe AVC thresholds; one undocumented. Scanner/pathway artwork is illustrative; panel 3 shows the supplied CT image with yellow outlines. The proposed pathway was not tested. AI = artificial intelligence; AS = aortic stenosis; AU = Agatston unit; AVC = aortic valve calcification; CT = computed tomography.

440 [39.1%]) and primary 3-tier classification (440/440 agreement; kappa = 1.000); continuous-score ICC(A,1) was 0.999 (Supplemental Tables 3 and 4).

RADIOLOGY REPORTING GAP. Among 172 manually positive examinations, AVC was explicitly documented in 27 original radiology reports (15.7%; 95% CI: 11.0%-21.9%) and was not explicitly documented in 145 (84.3%). Explicit documentation increased with manual burden: 0 of 75 (0%) at >10 to <100 AU, 6 of 45 (13.3%) at 100 to <300 AU, and 21 of 52 (40.4%) at ≥300 AU. This endpoint describes explicit report documentation and does not establish whether the interpreting radiologist recognized but chose not to report AVC (Table 2).

AUTOMATED ANALYSIS AS A SECONDARY DETECTION COMPARATOR. The automated analysis identified 169 of 172 manually positive examinations (sensitivity: 98.3%; 95% CI: 95.0%-99.4%) and correctly classified 259 of 268 manually negative examinations (specificity: 96.6%; 95% CI: 93.7%-98.2%). It identified 142 of the 145 manually confirmed cases not explicitly documented in the original report. Overall accuracy was 97.3% (95% CI: 95.3%-98.4%). Performance remained high in lung cancer-screening examinations (n = 129; sensitivity: 100.0%, specificity: 98.9%, accuracy: 99.2%) and other indications (n = 311; sensitivity: 97.7%, specificity: 95.5%, accuracy: 96.5%) (Table 2, Figure 2B, Central Illustration, Supplemental Table 1B). The positive likelihood ratio was 29.26 (95% CI: 15.39-55.63), and the negative likelihood ratio was 0.018 (95% CI: 0.006-0.055). Counting the 16 manually positive nonevaluable examinations as missed detections, end-to-end sensitivity among all 188 manually positive examinations was 89.9% (169/188; 95% CI: 84.8%-93.4%) (Supplemental Table 12).

AUROC was 0.9989 (95% CI: 0.9974-0.9999), and AUPRC was 0.9984 (95% CI: 0.9962-0.9998); curves are shown in Supplemental Figure 3.

Among the 145 manually positive examinations without explicit report documentation, 75 had >10 to <100 AU, 39 had 100 to <300 AU, and 31 had ≥300 AU. AI detected 72, 39, and 31 cases in these strata,

respectively (142/145, 97.9%). One undocumented case met the manual sex-specific severe AVC threshold; 2 met this threshold in the full evaluable cohort. These counts describe additional detection and do not establish a need for further testing (Supplemental Table 9).

DISTRIBUTION ACROSS COMPLEMENTARY AVC FRAMEWORKS. Applying the primary 3-tier framework yielded 268 patients with no AVC, 170 with detectable nonsevere AVC, and 2 with severe AVC by manual scoring; the automated analysis assigned 262, 173, and 5 patients, respectively. Modified MESA manual counts were 268 no AVC, 75 mild, 45 moderate, and 52 high, compared with automated counts of 262, 79, 38, and 61, respectively. Modified Tastet/Ali manual counts were 268, 151, 19, and 2 across ascending tiers, compared with 262, 147, 26, and 5 (**Figure 3**, Supplemental Table 5).

CLINICAL CORRELATES OF MANUALLY DEFINED AVC. Baseline characteristics are described in **Table 1**. Median age was 73.0 (70.0-80.0) years with AVC and 68.0 (61.0-73.0) without AVC; cardiovascular comorbidities were generally more frequent in the AVC group.

In the prespecified complete-case model (n = 440; 172 events), age (adjusted OR per 5 years: 1.92; 95% CI: 1.63-2.26; nominal $P < 0.001$) and male (OR: 2.09; 95% CI: 1.32-3.30; nominal $P = 0.002$) were associated with AVC. Apparent model AUC was 0.783; collinearity was low, and a quadratic age term was not significant (nominal $P = 0.587$) (**Table 3**, **Figure 4**). The exploratory sensitivity model adding smoking and CT indication had apparent AUC 0.799 (Supplemental Table 6).

ECHOCARDIOGRAPHIC FINDINGS AND VALVE PATHWAY. Among 178 AI-positive patients, 86 (48.3%) had echocardiography within ±365 days of CT; median timing was 4 days (Q1-Q3: −8 to 93). Mild-or-greater AS was present in 6 of 53 patients (11.3%) within ±90 days, 7 of 63 (11.1%) within ±180 days, and 10 of 86 (11.6%) within ±365 days. At ±365 days, 76 had no significant AS, 4 mild AS, 5 moderate AS, and 1 severe AS. Among 262 AI-negative patients, 99 had echocardiography within ±365 days; none had moderate or severe AS (**Table 4**). Within the AI-positive group, patients with vs without echocardiography had a similar age and calcium burden but were more often male and more frequently had heart failure and atrial fibrillation (Supplemental Table 7).

In the AI-positive ±365-day subset, manual AVC correlated with the mean gradient (rho = 0.306; 95% CI: 0.034-0.529; n = 56) and inversely with valve area (rho = −0.301; 95% CI: −0.546 to −0.026; n = 52). All 8 correlations, paired bootstrap intervals, and nominal *P* values are in **Table 4**. Estimates were modest and uncertain, with selected samples of 49 to 71 patients (**Figure 5**). Documentation and echo severity are summarized in **Figure 6**.

Four AI-positive patients had a prior AS diagnosis; none had valve replacement documented during available follow-up. One had concordantly high automated and manual AVC (4,051 and 3,991 AU, respectively), above the male severe CT-AVC threshold, but echocardiography showed moderate AS (Vmax: 2.88 m/s; mean gradient: 24 mm Hg; aortic valve area: 1.05 cm^2; dimensionless index: 0.34), and surveillance continued. Among the remaining 3, one had manual AVC 416 AU without significant AS, one had 18 AU with limited hemodynamics, and one was manual-negative at 10 AU but AI-positive at 12 AU (Supplemental Table 8).

TABLE 3 Multivariable Logistic Regression of Clinical Factors Associated With Manually Confirmed AVC >10 AU

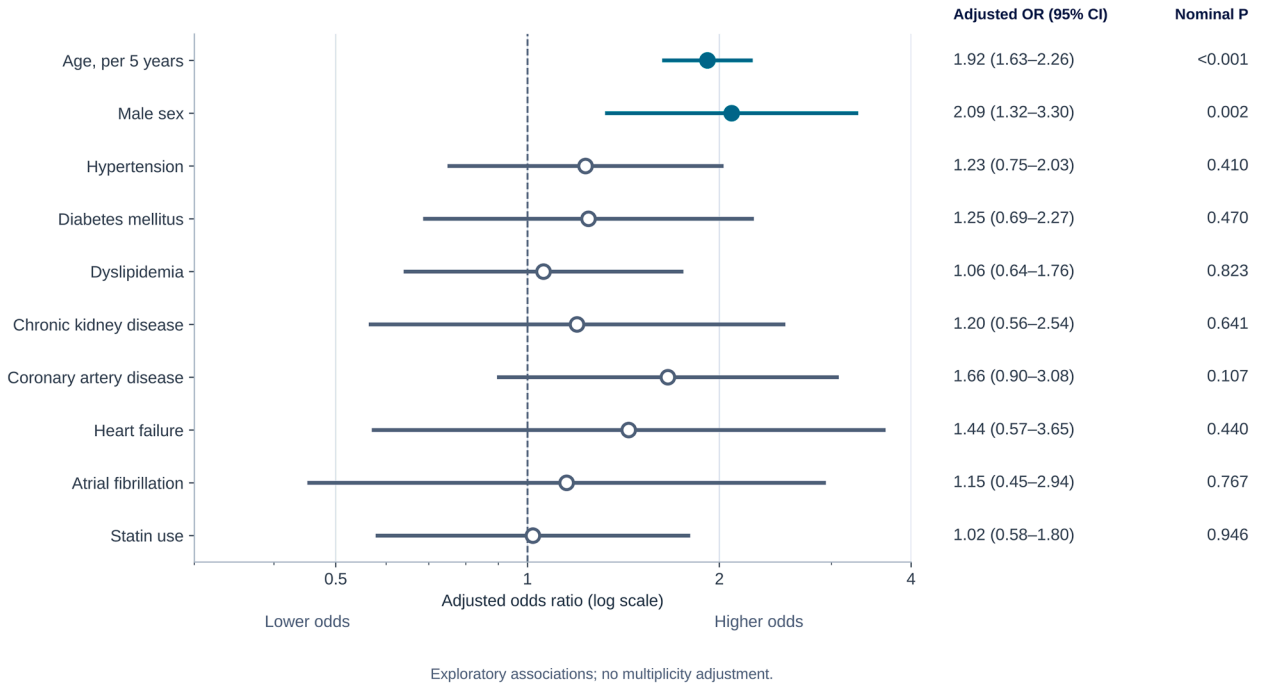
	Adjusted OR	95% CI	Nominal *P* Value
Age, per 5 y	1.92	1.63-2.26	<0.001
Male	2.09	1.32-3.30	0.002
Hypertension	1.23	0.75-2.03	0.410
Diabetes	1.25	0.69-2.27	0.470
Dyslipidemia	1.06	0.64-1.76	0.823
CKD	1.20	0.56-2.54	0.641
CAD	1.66	0.90-3.08	0.107
Heart failure	1.44	0.57-3.65	0.440
Atrial fibrillation	1.15	0.45-2.94	0.767
Statin use	1.02	0.58-1.80	0.946

Complete-case model: n = 440, 172 events, no imputation; apparent AUC 0.783.
AUC = area under the curve; CAD = coronary artery disease; CKD = chronic kidney disease; other abbreviations as in Table 1.

DISCUSSION

In this consecutive cohort, 172 of 440 (39.1%) had manually confirmed AVC, yet only 27 of 172 (15.7%) positive reports explicitly documented it. Documentation increased with burden, from 0/75 below 100 AU to 21/52 (40.4%) at ≥300 AU. AI detected 169 of 172 manual-positive cases, including 142 of 145 undocumented cases. Reader and fixed-threshold analyses supported the prevalence estimate, and including scorable rejected scans yielded 40.3%. These findings support automated case finding; additional detection does not itself establish any clinical benefit.

FIGURE 4 Independent Clinical Correlates of Manually Confirmed Calcification



Adjusted odds ratios and 95% confidence intervals from the complete-case multivariable model (n = 440; 172 events; area under the curve: 0.783). Age and male sex were independently associated with manually confirmed aortic valve calcification. Age is expressed per 5-year increment; association analyses and *P* values are exploratory.

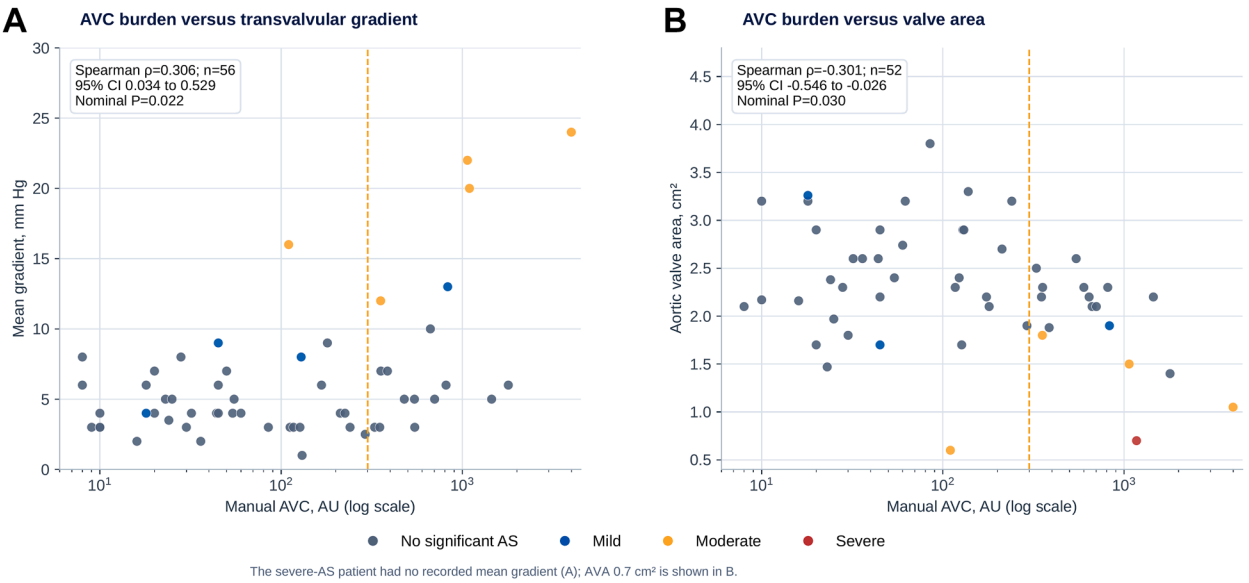
TABLE 4 Exploratory Echocardiographic Sensitivity Analyses

Analysis	Result	Denominator or Note
AI-positive with echo within ±365 d	86/178 (48.3%)	Median CT-to-echo 4 d [−8 to 93]
Echo timing	26 before; 7 same day; 53 after CT	AI-positive ±365-d subset
Mild-or-greater AS within ±90 d	6/53 (11.3%)	3 mild; 2 moderate; 1 severe
Mild-or-greater AS within ±180 d	7/63 (11.1%)	3 mild; 3 moderate; 1 severe
Mild-or-greater AS within ±365 d	10/86 (11.6%)	4 mild; 5 moderate; 1 severe
AI negative with echo within ±365 d	99/262 (37.8%)	92 no significant AS; 2 mild; 5 not assessed; no moderate/severe
Available hemodynamics in combined ±365-d subset	n = 185	Vmax 148; mean gradient 120; AVA 101; DI 107; SVI 95
AI AVC vs peak velocity	rho = 0.086 (95% CI: −0.186 to 0.344); nominal *P* = 0.476	n = 71
AI AVC vs mean gradient	rho = 0.240 (95% CI: −0.045 to 0.493); nominal *P* = 0.075	n = 56
AI AVC vs AVA	rho = −0.280 (95% CI: −0.529 to −0.010); nominal *P* = 0.045	n = 52
AI AVC vs dimensionless index	rho = −0.141 (95% CI: −0.425 to 0.154); nominal *P* = 0.333	n = 49
Manual AVC vs peak velocity	rho = 0.142 (95% CI: −0.112 to 0.381); nominal *P* = 0.238	n = 71
Manual AVC vs mean gradient	rho = 0.306 (95% CI: 0.034-0.529); nominal *P* = 0.022	n = 56
Manual AVC vs AVA	rho = −0.301 (95% CI: −0.546 to −0.026); nominal *P* = 0.030	n = 52
Manual AVC vs dimensionless index	rho = −0.184 (95% CI: −0.454 to 0.125); nominal *P* = 0.207	n = 49

Echocardiography was clinically obtained rather than protocol mandated. Analyses used variable-specific denominators; missing hemodynamic measures were not imputed. Correlation confidence intervals are paired-bootstrap percentile intervals (2,000 resamples); *P* values are nominal without multiplicity adjustment.

AVA = aortic valve area; DI = dimensionless index; SVI = stroke-volume index; other abbreviations as in Tables 1-3.

FIGURE 5 Calcium Burden vs Echocardiographic Hemodynamics



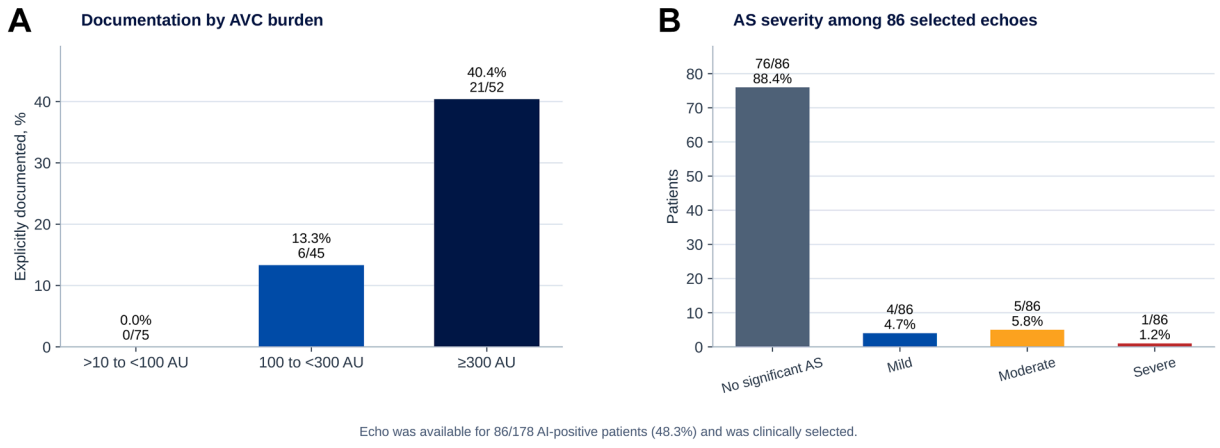
Manual AVC score is plotted against available mean gradient (A; n = 56) and aortic valve area (B; n = 52) among AI-positive patients with clinically obtained echocardiography within ±365 days. Point colors indicate recorded AS severity; the dashed line marks 300 AU. The patient with severe AS had no recorded mean gradient and appears only in panel B. Annotations show Spearman correlations, paired-bootstrap 95% confidence intervals, and nominal *P* values. Available paired measurements were used without imputation; analyses were exploratory. All manual and AI correlation estimates are reported in Table 4. AS = aortic stenosis; other abbreviations as in Figures 1 and 2.

The significance of incidental AVC extends beyond patients who already have severe AS. Among 6,814 MESA participants free of known cardiovascular disease at baseline, 13.4% had detectable AVC; increasing burden was associated with graded long-term risk of incident severe AS, while a score of zero identified a very low-risk group.[5] Longitudinal MESA work also documented incident calcification and its clinical correlates.[6] That community cohort and its CT protocol differ from our older, clinically imaged population. Only 2 patients here met sex-specific severe thresholds, whereas 170 had detectable nonsevere AVC, and 52 met the modified MESA high-burden definition. Severe thresholds help adjudicate suspected AS; modified MESA and Tastet/Ali categories describe burden. They are not interchangeable or validated referral thresholds for incidental nongated CT. Prospective studies should establish which burdens predict echocardiographic disease and progression.

The documentation gap should not be equated with radiologist error. Routine chest CT addresses other clinical indications. The oblique annular plane may require a multiplanar review to separate leaflet calcium from deposits in the aortic sinuses, sinotubular junction, coronary ostia, left ventricular outflow tract, fibrous trigones, or aorto-mitral continuity. Motion and partial-volume effects on nongated imaging, together with noise, slice thickness, reconstruction, and attenuation differences, complicate interpretation and Agatston scoring.[7,11-13] When calcium is subtle and no standardized reporting pathway exists, it may reasonably remain undocumented. Automated localization may improve consistency, but whether alerts improve appropriate evaluation or outcomes requires prospective testing. A numerical AVC score cannot diagnose hemodynamic AS.

Echocardiography was clinically selected: Among 86 AI-positive patients tested within ±365 days, 10 had mild-or-greater AS, and 1 had severe AS. Testing was more frequent among men and patients with heart failure or atrial fibrillation; selection and verification bias preclude treating these proportions as clinical yield. In the EARLY TAVR (Evaluation of Transcatheter Aortic Valve Replacement Compared to Surveillance for Patients With Asymptomatic Severe Aortic Stenosis) trial, 901 patients with echocardiographically defined asymptomatic severe AS

FIGURE 6 Radiology Documentation and Echocardiographic Context



Explicit original-report documentation by manual AVC burden was 0/75 (>10 to <100 AU), 6/45 (100 to <300 AU), and 21/52 (≥300 AU) (A). Among 86 of 178 AI-positive patients with clinically obtained echocardiography within ±365 days, 76 had no significant AS, 4 mild, 5 moderate, and 1 severe AS (B). This selected subset does not establish AS prevalence in the entire cohort or downstream benefit from AI detection. Abbreviations as in Figures 1, 2, and 5.

were randomized, and an early intervention reduced death, stroke, or unplanned cardiovascular hospitalization vs surveillance (26.8% vs 45.3%; HR: 0.50).[15] Those findings do not apply to the predominantly nonsevere incidental AVC observed here. Trials in moderate AS with cardiac damage may inform future pathways.[16] For now, incidental AVC warrants assessment in the context of symptoms, examination, prior disease, and available echocardiography rather than automatic referral.

Implementation requires a plan for rejected scans. The 11.3% rejection rate may differ in routine use; rejection accounted for 16 of 19 manually positive scans without an automated positive result. No AI rescue or AI-directed care occurred here. A future workflow should return an explicit nonevaluable status and retain a standard radiologist review. Human confirmation, clinical context, and a defined follow-up pathway would be needed before acting on an AI result. Referral thresholds, surveillance intervals, outcomes, and costs require prospective evaluation. Prediction models could eventually combine AVC burden, age, sex, renal and metabolic factors, valve morphology, and echocardiography,[14] informed by longitudinal and biological studies of progression.[6,17,18] This retrospective study cannot prescribe management.

STUDY LIMITATIONS. Several factors limit interpretation. The primary cohort excluded 56 AI-rejected examinations, and those patients were younger and had fewer selected risk factors. Readers subsequently scored 26 rejected examinations while being aware of the rejection status; 30 remained unscorable. Their unknown AVC status leaves potential selection bias despite the prevalence bounds. The absence of explicit wording in a report does not establish that a radiologist failed to recognize AVC. Echocardiography was clinically selected and available in 86 of 178 AI-positive and 99 of 262 AI-negative patients; hemodynamic measurements were incomplete and analyzed without imputation. Burden categories are descriptive, and clinical benefit from automated detection was not tested.

Acquisition data also constrain generalizability. DICOM metadata were assignable in 401 of 440 scans; the median reconstruction thickness was 1.0 mm (Q1-Q3: 0.625-1.0; range: 0.6-3.0). Across thickness groups, sensitivity ranged from 97.8% to 100%, and specificity from 93.3% to 100%, but groups thicker than 1.0 mm were small (Supplemental Tables 10 and 11). By manufacturer, sensitivity ranged from 95.9% to 100%, and specificity from 96.3% to 100% (Supplemental Table 14). Most assignable studies used 120 kVp on GE or Canon/Toshiba systems; kernel and dose information were inconsistently recorded. Performance with other protocols, especially near the >10-AU cutoff, remains uncertain. Exploratory *P* values were nominal, without an adjustment for multiple comparisons, and

echocardiographic correlations came from small selected samples.

CONCLUSIONS

In 440 AI-evaluable routine noncontrast chest CT examinations, manual AVC was present in 39% but explicitly reported in only 16% of positive examinations. AI detected 98.3% of manual-positive cases among evaluable scans; 56 of 496 (11.3%) submitted scans were rejected, and end-to-end sensitivity was 89.9% when manually positive rejected scans were counted as missed. Prospective studies should establish how incidental AVC and nonevaluable AI results are managed and whether AI-assisted detection improves care.

ETHICAL STATEMENT. The Cedars-Sinai Medical Center Institutional Review Board (STUDY00004566) initially approved the study effective November 7, 2025. The final protocol amendment applicable to this cohort was approved in April 2026 with waivers of informed consent and HIPAA authorization.

DATA AVAILABILITY Patient-level data are not publicly available because of institutional review board and privacy restrictions. Requests for deidentified analytic data or statistical code may be directed to the corresponding author, subject to institutional approval and a data-use agreement.

DECLARATION OF AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS During preparation of this revision, AI-assisted tools were used for language editing, document formatting, and figure preparation. All statistical analyses were performed by the authors, who reviewed all AI-assisted output and take responsibility for the accuracy and integrity of the manuscript.

ACKNOWLEDGMENTS The authors thank a NanoxAI Ltd employee who uploaded the deidentified examinations to the onsite study computer and returned the automated output reports. This individual had no access to clinical data or manual scores and did not participate in analysis, interpretation, or manuscript preparation.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

NanoxAI Ltd provided research access to HealthAVC V1.0 and onsite support for software operation but no financial support for this study. Dr Koren has received honoraria from NanoxAI Ltd for image-labeling consultation conducted before and separately from this study and for work related to a companion technical-validation study using the same cohort and manual reference measurements (under review); no compensation was received specifically for the present analysis, and no data from the preliminary image-labeling work were used. Mr Patel received compensation from NanoxAI Ltd for manual reconstruction work for the companion study; no payment had been received at the time of submission. Dr Koren and Mr Patel performed the manual reference measurements for the present study while blinded to the automated output. Drs Wimpfheimer, Hassid, and Hofstetter are employees of and hold restricted stock units in NanoxAI Ltd. Dr Makkar has received grant support from Edwards Lifesciences; serves as a consultant to Abbott Vascular, Cordis, and Medtronic; and holds equity in Entourage Medical. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr Ofir Koren, The Karsh Center of Interventional Cardiology, Smidt Heart Institute, Cedars-Sinai Medical Center, 127 S. San Vicente Boulevard, Los Angeles, California 90048, USA. E-mail: Ofir.Koren@csmc.edu.

PERSPECTIVES

COMPETENCY IN MEDICAL KNOWLEDGE: Aortic valve calcification was present in approximately 4 of every 10 evaluable routine noncontrast chest CT examinations in adults aged 50 years or older but was explicitly documented in fewer than 1 in 6 positive reports. Automated detection may assist case finding in evaluable scans (about 1 in 9 submitted scans could not be analyzed) but does not diagnose hemodynamic stenosis or determine management.

TRANSLATIONAL OUTLOOK: Automated analysis can surface most undocumented calcification, but the downstream pathway remains undefined. Prospective studies are needed to establish which incidental burden categories predict echocardiographic disease, calcium progression, and incident severe aortic stenosis; to define referral and surveillance triggers; and to determine whether structured opportunistic reporting improves time to diagnosis, appropriate referral, and clinical outcomes.

REFERENCES

1. Otto CM, Nishimura RA, Bonow RO, et al. 2020 ACC/AHA guideline for the management of patients with valvular heart disease: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice guidelines. *J Am Coll Cardiol*. 2021;77:e25-e197.

2. Praz F, Borger MA, Lanz J, et al. 2025 ESC/EACTS guidelines for the management of valvular heart disease. *Eur Heart J*. 2025;46:4635-4736.

3. Pawade T, Clavel MA, Tribouilloy C, et al. Computed tomography aortic valve calcium scoring in patients with aortic stenosis. *Circ Cardiovasc Imaging*. 2018;11:e007146.

4. Tastet L, Ali M, Pibarot P, et al. Grading of aortic valve calcification severity and risk stratification in aortic stenosis. *J Am Heart Assoc*. 2024;13:e035605.

5. Whelton SP, Jha K, Dardari Z, et al. Prevalence of aortic valve calcium and the long-term risk of incident severe aortic stenosis. *JACC Cardiovasc Imaging*. 2024;17:31-42.

6. Katz R, Budoff MJ, O'Brien KD, et al. Relationship of metabolic syndrome with incident aortic valve calcium and aortic valve calcium progression: the multi-ethnic study of atherosclerosis. *Diabetes*. 2009;58:813-819.

7. Williams MC, Abbas A, Tirr E, et al. Reporting incidental coronary, aortic valve and cardiac calcification on non-gated thoracic computed tomography: a consensus statement from the BSCI/BSCCT and BSTI. *Br J Radiol*. 2021;94:20200894.

8. Zhu Y, Wang Y, Gioia WE, et al. Visual scoring of aortic valve calcifications on low-dose CT in lung cancer screening. *Eur Radiol*. 2020;30:2658-2668.

9. Gunning SGS, Graby J, Mody Y, et al. Visual ordinal grading of aortic valve calcification on routine non-gated chest CT predicts prognosis and alters management. *Eur Radiol*. 2025;35:6291-6301.

10. Gondrie MJA, van der Graaf Y, Jacobs PC, Oen AL, Mali WPTM. The association of incidentally detected heart valve calcification with future cardiovascular events. *Eur Radiol*. 2011;21:963-973.

11. Lessmann N, van Ginneken B, Zreik M, et al. Automatic calcium scoring in low-dose chest CT using deep neural networks with dilated convolutions. *IEEE Trans Med Imaging*. 2018;37:615-625.

12. Ouchi K, Sakuma T, Nojiri A, et al. Accuracy of aortic valve calcification volume score for identification of significant aortic stenosis on non-electrocardiographic-gated computed tomography compared with the agatston scoring system. *J Cardiovasc Comput Tomogr*. 2024;18:352-362.

13. Ouchi K, Sakuma T, Nojiri A, et al. Optimal threshold score of aortic valve calcification for identification of significant aortic stenosis on non-electrocardiographic-gated computed tomography. *Eur Radiol*. 2023;33:1243-1253.

14. Sanabria M, Tastet L, Pelletier S, et al. AI-enhanced prediction of aortic stenosis progression: insights from the PROGRESSA study. *JACC Adv*. 2024;3:101234.

15. Généreux P, Schwartz A, Oldemeyer JB, et al. Transcatheter aortic-valve replacement for asymptomatic severe aortic stenosis. *N Engl J Med*. 2025;392:217-227.

16. Mihatov N, Pibarot P. Moderate aortic stenosis with cardiac damage: a new type of severe aortic stenosis. *Struct Heart*. 2024;8:100336.

17. Zheng KH, Tsimikas S, Pawade T, et al. Lipoprotein(a) and oxidized phospholipids promote valve calcification in patients with aortic stenosis. *J Am Coll Cardiol*. 2019;73:2150-2162.

18. Dweck MR, Jenkins WSA, Vesey AT, et al. 18F-sodium fluoride uptake is a marker of active calcification and disease progression in patients with aortic stenosis. *Circ Cardiovasc Imaging*. 2014;7:371-378.

KEY WORDS aortic stenosis, aortic valve calcification, artificial intelligence, computed tomography, incidental findings, opportunistic screening

APPENDIX For supplemental figures and tables, please see the online version of this paper.