

Artificial Intelligence–Enabled Acquisition and Interpretation for Screening Aortic Stenosis

Eunjung Lee, PhD; Jwan A. Naser, MBBS; Conor J. Kane; Jude L. Kovac; Christie Greason, BA, MS; Méabh M. Killalea; Mayari A. Gulati, BS; John I. Jackson, PhD; Jordan Borgeson, BS; Daniel A. Schonfeld; Jeffrey G. Malins, PhD; D. M. Anisuzzaman, PhD; Maria M. Crestanello, BA; Jessica Zacher, BA; Seda Camalan, PhD; Jeremy J. Thaden, MD; Vidhu Anand, MBBS; Vuyisile T. Nkomo, MD, MPH; Ratnasari Padang, MBBS, PhD; Timothy J. Poterucha, MD; Chieh-Ju Chao, MD; Patricia A. Pellikka, MD; Francisco Lopez-Jimenez, MD, MS; Paul A. Friedman, MD; Jae K. Oh, MD; Garvan C. Kane, MD, PhD; Zach I. Attia, PhD; Sorin V. Pislaru, MD, PhD; Jared G. Bird, MD; Gal Tsaban, MD, PhD

IMPORTANCE Timely identification of aortic stenosis (AS) is essential for appropriate clinical management, yet screening remains limited by dependence on comprehensive echocardiography and trained imaging personnel.

OBJECTIVE To develop and validate a deep learning algorithm for detection of moderate or greater AS and prospectively evaluate its performance using artificial intelligence (AI)–guided focused cardiac ultrasound (FoCUS) acquired by novice operators.

DESIGN, SETTING, AND PARTICIPANTS This diagnostic study included retrospective algorithm development and validation and prospective evaluation of AI-guided FoCUS across Mayo Clinic sites in the Midwest, Arizona, and Florida. The model was developed using 6753 patients and evaluated in internal validation (n = 852), internal test (n = 844), and validation (n = 1912) cohorts. Performance was assessed on FoCUS acquired by experienced sonographers (n = 602) and prospectively by novice operators (n = 1302). The retrospective model development and validation cohorts comprised studies performed from January 2005 through September 2022. Prospective study was conducted in 2 enrollment periods from June to August 2024 and from June to September 2025. Participants from both periods were combined to comprise the final prospective cohort.

EXPOSURE AI-guided FoCUS acquisition and automated deep learning–based assessment for detection of moderate or greater AS.

MAIN OUTCOMES AND MEASURES Detection of moderate or greater AS. Performance was assessed using area under the receiver operating characteristic curve (AUROC), sensitivity, specificity, and positive predictive value.

RESULTS The model demonstrated excellent discrimination in the internal test cohort (AUROC, 0.99; 95% CI, 0.98–1.00) and geographically distinct validation cohorts in Arizona (AUROC, 0.99; 95% CI, 0.97–1.00) and Florida (AUROC, 0.99; 95% CI, 0.96–1.00). Among FoCUS examinations acquired by experienced sonographers, sensitivity was 95% (95% CI, 82–99) and specificity was 97% (95% CI, 95–98). In the prospective novice-operator cohort, 1258 of 1302 examinations (96.6%) were suitable for automated analysis. Sensitivity was 93% (95% CI, 82–99) and specificity was 96% (95% CI, 95–97). Expert review of AI-positive and uninterpretable examinations increased the positive predictive value from 49.4% to 91.1%, with sensitivity of 85.4%.

CONCLUSIONS AND RELEVANCE A deep learning algorithm accurately detected moderate or greater AS across validation cohorts in this study. In prospective evaluation, novice operators were able to acquire AI-guided FoCUS examinations that enabled accurate detection of moderate or greater AS. These findings support a scalable strategy that may expand access to AS detection in settings with limited echocardiography resources.

JAMA Cardiol. doi:10.1001/jamacardio.2026.3829
Published online August 28, 2026.

 Editorial

 Multimedia

 Supplemental content

Author Affiliations: Department of Cardiovascular Medicine, Mayo Clinic, Rochester, Minnesota.

Corresponding Authors: Jared G. Bird, MD (bird.jared@mayo.edu), and Gal Tsaban, MD, PhD (tsaban.gal@mayo.edu), Department of Cardiovascular Medicine, Mayo Clinic, 200 First St SW, Rochester, MN 55905.

Aortic valve stenosis (AS) affects approximately 12% of individuals aged 75 years and older¹ and is the most common indication for valve intervention worldwide, yet it remains frequently underdiagnosed and undertreated.^{2,3} Delayed recognition is associated with adverse left ventricular remodeling, heart failure, and mortality.⁴⁻⁶ Because many patients remain asymptomatic until advanced disease, strategies for earlier identification of clinically significant AS are needed.

Transthoracic echocardiography (TTE) is the diagnostic reference standard for AS; however, routine echocardiographic screening of asymptomatic populations is not recommended because of cost, equipment, and workforce constraints. Physical examination may identify characteristic murmurs, but diagnostic accuracy varies widely outside cardiology settings, limiting its reliability as a screening tool.^{7,8} More recently, artificial intelligence (AI)-enabled electrocardiography has shown promise for AS detection, but modest accuracy (approximately 70% to 75%) limits its utility for guiding downstream echocardiography, particularly in lower-prevalence screening populations.⁹⁻¹¹

Handheld focused cardiac ultrasound (FoCUS) is increasingly used at the point of care.¹² However, FoCUS examinations are often performed without Doppler imaging and by operators with limited echocardiography training, constraining reliable identification of significant AS. An AI approach capable of accurately classifying AS from 2-dimensional FoCUS images acquired by novice operators could overcome these limitations and enable scalable, low-resource testing strategies across diverse clinical environments.

In this study, we sought to (1) develop and validate an AI model trained on 2-dimensional echocardiographic images to identify moderate or greater AS across geographically distinct cohorts; (2) assess model generalizability to images acquired using handheld devices; and (3) evaluate diagnostic performance in a prospective cohort of patients scanned by novice operators (those without any clinical or imaging experience) guided by AI-assisted image acquisition (concomitant Doppler echocardiography acting as the reference standard).

Methods

This study included retrospective development and multicenter validation of a deep learning model (January 2005 through September 2022), followed by prospective validation using AI-guided FoCUS acquired by novice operators (June to August 2024 and June to September 2025). Reporting followed the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) and Proposed Requirements for Cardiovascular Imaging-Related Machine Learning Evaluation (PRIME) reporting guidelines. The Mayo Clinic Institutional Review Board approved all study components. Prospective participants provided written informed consent; retrospective analyses included patients who had authorized use of their medical records for research.

Key Points

Question Can a deep learning echocardiography algorithm for detection of moderate or greater aortic stenosis be accurately applied to artificial intelligence (AI)-guided focused cardiac ultrasound examinations acquired by novice operators?

Findings In this diagnostic study, a deep learning echocardiography algorithm demonstrated high diagnostic performance across validation cohorts and maintained high diagnostic accuracy during prospective evaluation using AI-guided focused cardiac ultrasound examinations acquired by novice operators. A 2-step workflow incorporating expert review of AI-positive and uninterpretable examinations substantially improved positive predictive value while requiring review of only approximately 10% of examinations.

Meaning AI-guided focused cardiac ultrasound combined with automated interpretation may provide a scalable triage and possibly even screening strategy for moderate or greater aortic stenosis in settings with limited access to comprehensive echocardiography.

Model Development Cohort and Validation of AI Models

The development cohort included patients with a clinically indicated TTE performed at Mayo Clinic sites in Minnesota and Wisconsin between January 2005 and September 2022. Patients with prior aortic valve replacement were excluded, and only the earliest examination per patient was included. Moderate or greater AS was defined according to consensus criteria, as previously described.⁹ The dataset was randomly divided into training ($n = 6753$), internal validation ($n = 852$), and internal test ($n = 844$) cohorts. The model was validated in randomly sampled patients undergoing clinical TTE at Mayo Clinic sites in Arizona ($n = 959$) and Florida ($n = 953$).

FoCUS Using Handheld Devices by Experienced Sonographers

To assess performance on handheld images, 602 randomly selected patients underwent FoCUS performed by experienced sonographers using a Philips Lumify device between November 2022 and March 2023. FoCUS was performed within 15 minutes of a clinically indicated comprehensive TTE.

Prospective AI-Guided FoCUS Study by Novice Operators

We prospectively enrolled 1302 consenting adults undergoing clinically indicated comprehensive TTE at a single center. Participants were randomly recruited and excluded only for complex congenital heart disease or prior aortic valve prosthesis. Nine research staff without prior clinical or sonography experience performed FoCUS using a Philips Lumify handheld device with US Food and Drug Administration-approved real-time AI acquisition guidance (UltraSight) after 4 hours of standardized training. Operators attempted acquisition of 6 standard 2-dimensional views; the AS model used the parasternal long-axis, parasternal short-axis at the aortic valve level, and apical 3-chamber views, without Doppler imaging or quantitative measurements. Reference labels were derived from clinically indicated comprehensive TTE using a GE E95 system, performed by trained sonographers and interpreted by level III

echocardiographers according to American Society of Echocardiography guidelines.¹³ This examination served as the reference standard. Clinical echocardiography personnel were blinded to FoCUS findings. An expert echocardiographer graded FoCUS image quality while blinded to clinical and AI results (eMethods 1 in [Supplement 1](#)).

Furthermore, a subsequent 2-step hybrid screening approach was evaluated: studies flagged by the AI algorithm as positive for moderate or greater AS, as well as studies deemed uninterpretable by the algorithm were reviewed by expert echocardiographers and classified as negative, positive for at least moderate AS, or uninterpretable. The latter 2 categories would identify patients requiring comprehensive echocardiography.

Echocardiography Reference Standard

All patients underwent clinically indicated TTE in accordance with established guidelines.¹³ Aortic valve area was derived using the continuity equation from the left ventricular outflow tract diameter, left ventricular outflow tract time-velocity integral, and transaortic time-velocity integral, with Doppler signals obtained from multiple acoustic windows to ensure optimal alignment. The dimensionless velocity index was calculated as the ratio of left ventricular outflow tract time-velocity integral to aortic time-velocity integral. Measurements were averaged across 3 cardiac cycles in patients with sinus rhythm and 5 cycles in those with atrial fibrillation. Moderate or greater AS was defined by at least 1 of the following: aortic valve area ≤ 1.5 cm², mean transaortic gradient ≥ 20 mm Hg, peak velocity ≥ 3.0 m/s, or dimensionless velocity index ≤ 0.35 , as previously described.^{9,14}

Deep Learning Framework

The deep learning framework incorporated automated view classification followed by prediction of moderate or greater AS from standard 2-dimensional echocardiographic video clips. It generated a continuous probability score that was thresholded for binary classification. Model architecture, preprocessing, training, hyperparameter optimization, and model selection are described in eMethods 2 in [Supplement 1](#). Studies were considered interpretable when the automated pipeline detected at least 1 eligible view and generated a prediction. Studies were considered AI-uninterpretable when none of the 3 targeted views was identified. These studies were excluded from standalone model-performance analyses but included in evaluation of the 2-step workflow.

Statistical Analysis

Baseline characteristics were summarized using descriptive statistics and reported as means and standard deviations or medians and interquartile ranges. Model performance was evaluated using area under the receiver operating characteristic curve (AUROC), sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall accuracy. The Wilson score method was used to compute 95% confidence intervals. The optimal probability threshold was determined using the Youden index on the internal validation set

and subsequently applied unchanged to the internal test, geographically distinct validation, handheld, and prospective cohort. Discrimination in the prospective FoCUS evaluation was examined using receiver operating characteristic and precision-recall curves, while calibration was assessed using a calibration plot comparing predicted probabilities with observed event rates; decision curve analysis explored net clinical benefit across a range of threshold probabilities relative to treat-all and treat-none strategies. Selected graphics were generated using BioRender. Statistical analyses were performed using R version 4.5.2 (R Foundation) and JMP Pro 19.0.

Results

Baseline Characteristics of the Development and Validation Models

Baseline characteristics of the development and testing cohorts are summarized in [Table 1](#). The Midwest development cohort included 8449 patients undergoing clinically indicated TTE, 927 of whom (10.97%) had at least moderate AS. The median (IQR) age was 66.7 (53.8-76.9) years, and 4164 (49.3%) were female. The prevalence of at least moderate AS was lower in the geographically distinct validation cohorts (3.65%-3.88%). Consistent with this difference, patients in the testing cohorts had a lower burden of cardiovascular and non-cardiovascular comorbidities.

Model Validation

Model performance is summarized in [Table 2](#). The AS classification algorithm demonstrated excellent discrimination in the internal validation cohort (AUROC, 0.99; 95% CI, 0.99-1.00; sensitivity, 96%; 95% CI, 90-98; specificity, 97%; 95% CI, 95-98) and the internal Midwest test set (AUROC, 0.99; 95% CI, 0.98-1.00; sensitivity, 95%; 95% CI, 87-98; specificity, 96%; 95% CI, 94-97).

Geographically distinct validation cohorts from Arizona (n = 959) and Florida (n = 953) had AUROCs of 0.99 (95% CI, 0.97-1.00) and 0.99 (95% CI, 0.96-1.00), sensitivities of 89% (95% CI, 74-95) and 97% (95% CI, 86-100), specificities of 97% (95% CI, 96-98) and 95% (95% CI, 93-96), and accuracies of 97% and 95%, respectively. Performance remained robust across sites ([Table 2](#); [eFigure 1](#) in [Supplement 1](#)). Incorporating multiple imaging windows improved accuracy and reduced variability across sites ([eFigure 2](#) in [Supplement 1](#)). Performance was also consistent across sinus or atrial fibrillation rhythms ([eFigure 3](#) in [Supplement 1](#)). Discrimination remained consistently high across acquisition periods and ultrasound systems ([eFigure 4](#) in [Supplement 1](#)).

Model Validation on Images Acquired Using Handheld Ultrasound Devices

We next evaluated the model in 602 patients scanned by professional sonographers using a handheld device. In this cohort with an AS prevalence of 6.2%, the model achieved AUROC of 0.99 (95% CI, 0.96-1.00), a sensitivity of 95% (95% CI, 82-99), and a specificity of 97% (95% CI, 95-98) with an overall accuracy of 97%.

Table 1. Patient Characteristics for the Derivation and Validation Cohorts

Characteristic	Midwest (n = 8449)	Arizona (n = 959)	Florida (n = 953)	FoCUS by sonographers (n = 602)	AI-guided FoCUS by novices (n = 1302)
Age, median (IQR), y	67 (54-77)	66 (54-76)	67 (56-77)	66.3 (55-75)	67 (56-75)
Sex, No. (%)					
Male	4285 (50.7)	496 (51.7)	518 (54.3)	363 (60.3)	762 (58.5)
Female	4164 (49.3)	463 (48.3)	435 (45.7)	239 (39.7)	540 (41.5)
Race and ethnicity, No. (%) ^a					
African American	176 (2.08)	25 (2.61)	100 (10.49)	15 (2.49)	31 (2.40)
Asian	134 (1.59)	18 (1.88)	13 (1.36)	14 (2.33)	26 (2.00)
Hispanic	47 (0.56)	17 (1.77)	7 (0.73)	2 (0.33)	27 (2.10)
Native American	36 (0.43)	16 (1.67)	2 (0.21)	4 (0.66)	6 (0.46)
White	7813 (92.47)	867 (90.41)	798 (83.74)	558 (92.69)	1181 (90.70)
Other (unspecified), choose not to disclose, or unknown	243 (2.88)	16 (1.67)	33 (3.46)	9 (1.50)	31 (2.40)
BMI, median (IQR)	28.5 (24.7-33.1)	27.2 (23.9-31.4)	28.4 (24.9-32.9)	28.1 (24.7-31.5)	27.4 (24.5-31.2)
≥Moderate aortic stenosis, No. (%)	927 (11.0)	35 (3.7)	37 (3.9)	37 (6.2)	48 (3.7)
≥Moderate aortic regurgitation, No. (%)	137 (1.6)	4 (0.4)	15 (1.6)	25 (4.2)	44 (3.4)
≥Moderate mitral regurgitation, No. (%)	541 (6.4)	35 (3.7)	41 (4.3)	60 (10.0)	90 (6.9)
≥Moderate tricuspid regurgitation, No. (%)	915 (10.8)	44 (4.6)	54 (5.7)	57 (9.5)	104 (8.0)
History of MI, No. (%)	1251 (14.81)	76 (7.92)	95 (9.97)	73 (12.13)	138 (10.60)
History of CHF, No. (%)	2351 (27.83)	85 (8.86)	161 (16.89)	241 (40.03)	529 (40.63)
History of CAD, No. (%)	3020 (35.74)	189 (19.71)	291 (30.54)	227 (37.71)	487 (37.40)
History of PVD, No. (%)	2599 (30.76)	105 (10.95)	197 (20.67)	271 (45.02)	541 (41.55)
History of stroke, No. (%)	1591 (18.83)	117 (12.20)	144 (15.11)	88 (14.62)	178 (13.67)
History of AF, No. (%)	1914 (22.65)	139 (14.49)	210 (22.04)	190 (31.56)	424 (32.57)
History of chronic pulmonary disease, No. (%)	2882 (34.11)	159 (16.58)	183 (19.20)	144 (23.92)	330 (25.35)
History of chronic kidney disease, No. (%)	1687 (19.97)	152 (15.85)	187 (19.62)	145 (24.09)	320 (24.58)
History of cancer, No. (%)	1413 (16.72)	154 (16.06)	167 (17.52)	134 (22.26)	264 (20.28)
History of diabetes, No. (%)	1956 (23.15)	169 (17.62)	204 (21.41)	115 (19.10)	223 (17.13)
History of hypertension, No. (%)	4526 (53.57)	474 (49.43)	532 (55.82)	338 (56.15)	691 (53.07)

Abbreviations: AF, atrial fibrillation; AI, artificial intelligence; BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); CAD, coronary artery disease; CHF, congestive heart failure; FoCUS, focused cardiac ultrasound; MI, myocardial infarction; PVD, peripheral vascular disease.

^a Race and ethnicity data were collected via electronic health records and reported to provide a more complete description of the study population.

Table 2. Performance of the Deep Learning Aortic Stenosis Model Across Retrospective Internal Test and Geographically Distinct Validation Cohorts Including Focused Cardiac Ultrasound (FoCUS) Scans by Experienced Sonographers

Model	No.	AUROC (95% CI)	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, %	NPV, %	Accuracy, %	TN, No. (%)	FP, No. (%)	FN, No. (%)	TP, No. (%)
Midwest	844	0.99 (0.98-1.00)	95 (87-98)	96 (94-97)	68	99	95	734 (86.97)	34 (4.03)	4 (0.47)	72 (8.53)
Arizona	959	0.99 (0.97-1.00)	89 (74-95)	97 (96-98)	54	100	97	898 (93.64)	26 (2.71)	4 (0.42)	31 (3.23)
Florida	953	0.99 (0.96-1.00)	97 (86-100)	95 (93-96)	44	100	95	871 (91.40)	45 (4.72)	1 (0.1)	36 (3.78)
FoCUS by sonographers	602	0.99 (0.96-1.00)	95 (82-99)	97 (95-98)	65	100	97	546 (90.7)	19 (3.16)	2 (0.33)	35 (5.81)

Abbreviations: AUROC, area under the receiver operating characteristic curve; FN, false negative; FP, false positive; NPV, negative predictive value; PPV, positive predictive value; TN, true negative; TP, true positive.

Use of the Model in a Prospective Study of AI-Guided FoCUS Acquired by Novice Operators

The model was then prospectively evaluated in 1302 patients who underwent AI-guided FoCUS within 15 minutes of a comprehensive TTE. The mean (SD) age was 64 (16) years, and 540 (41.5%) were women; additional characteristics are provided in Table 1 and the eTable in [Supplement 1](#). The prevalence of

moderate or greater AS was 3.7%, and the median (IQR) scan time was 4 (3-5) minutes.

Across the 3 conventional imaging views of interest, feasibility and image quality varied. Parasternal long-axis and parasternal short-axis at the aortic valve level views were the most feasible, whereas the apical 3-chamber view was the least feasible (eFigure 5 in [Supplement 1](#)). Complete, good, or ad-

Table 3. Diagnostic Performance of Artificial Intelligence (AI) Only, Expert Review, and 2-Step Hybrid Aortic Stenosis Classifications in the Prospective AI-Guided Focused Cardiac Ultrasound (FoCUS) Study

Measure	AI model alone	Expert review ^a	2-Step hybrid ^b
Feasibility, No. (%)	1258 (96.6)	1268 (97.4)	1285 (98.7)
Sensitivity, % (95% CI)	93.5 (82.1-98.6)	91.7 (80.0-97.7)	85.4 (72.2-93.9)
Specificity, % (95% CI)	96.4 (95.2-97.4)	99.5 (98.9-99.8)	99.7 (99.2-99.9)
PPV, % (95% CI)	49.4 (38.5-60.4)	88.0 (75.7-95.5)	91.1 (78.8-97.5)
NPV, % (95% CI)	99.7 (99.3-100.0)	99.7 (99.2-99.9)	99.4 (98.8-99.8)
Accuracy, % (95% CI)	96.3 (95.1-97.2)	99.2 (98.6-99.6)	99.1 (98.5-99.6)

Abbreviations: NPV, negative predictive value; PPV, positive predictive value.

^a Expert refers to the independent blinded review of the 2-dimensional images acquired by novices from FoCUS examinations only. These reviewers interpreted these limited 2-dimensional FoCUS images in isolation (blinded to

the comprehensive echocardiographic findings).

^b Two-step hybrid evaluation based on first pass by AI followed by expert overread of AI aortic stenosis-positive and uninterpretable cases.

equate quality images were obtained in 80% to 90% of examinations for parasternal views and in 65% of examinations for the apical view. Although not all views were obtained in every participant, adequate images for diagnostic assessment were achieved in 1268 patients (97.4%). Representative class activation mapping-family heatmaps layered over 2-dimensional representative video clips of participants with and without AS (Video) showed spatial activation centered predominantly around the aortic valve region, in both participants with and without AS.

Diagnostic performance for AI classification alone and for the 2-step evaluation incorporating physician overread of AI-positive and uninterpretable cases is summarized in **Table 3**. Confusion matrices are provided in **Figure 1A** and **B**, and model performance metrics, including receiver operating characteristic curves, precision-recall analyses, calibration, and net benefit, are presented in **Figure 1C-F**. AI classification alone was feasible in 96.6% of cases and demonstrated a sensitivity of 93.5% and a specificity of 96.4%, with a PPV of 49.4% (Table 3). The 2-step evaluation, in which expert review classified some AI-uninterpretable studies, increased overall classification feasibility from 96.6% to 98.7% (confusion matrix in **Figure 1A** and **B**) and yielded a sensitivity of 85.4% and a specificity of 99.7%, with a marked improvement in PPV to 91.1% (Table 3). NPV remained high for both approaches (99.7% for AI-only and 99.4% for the 2-step evaluation). Mean (SD) expert review time was 2 (1) minutes per examination. A proposed screening process using a hybrid approach with AI-only classification, followed by expert overread of positive or uninterpretable studies only (10% of all studies) is illustrated in **Figure 2**. Compared with AI alone, the hybrid strategy reduced false-positive referrals from 44 to 4 but introduced 4 additional false-negative classifications during expert review; overall, 7 of 48 patients with moderate or greater AS were missed. Details on false-positive and false-negative cases are outlined in eResults 1 in **Supplement 1**. Sensitivity analyses demonstrated no significant differences in diagnostic accuracy between patients with and without concomitant valvular disease, between bicuspid and trileaflet valve morphology, and between patients younger than 75 years and those aged 75 years or older (eResults 2 in **Supplement 1**). In an exploratory nested substudy, 91 participants underwent repeated AI-guided FoCUS later the same day; screening classifications were concordant in all 91 examinations.

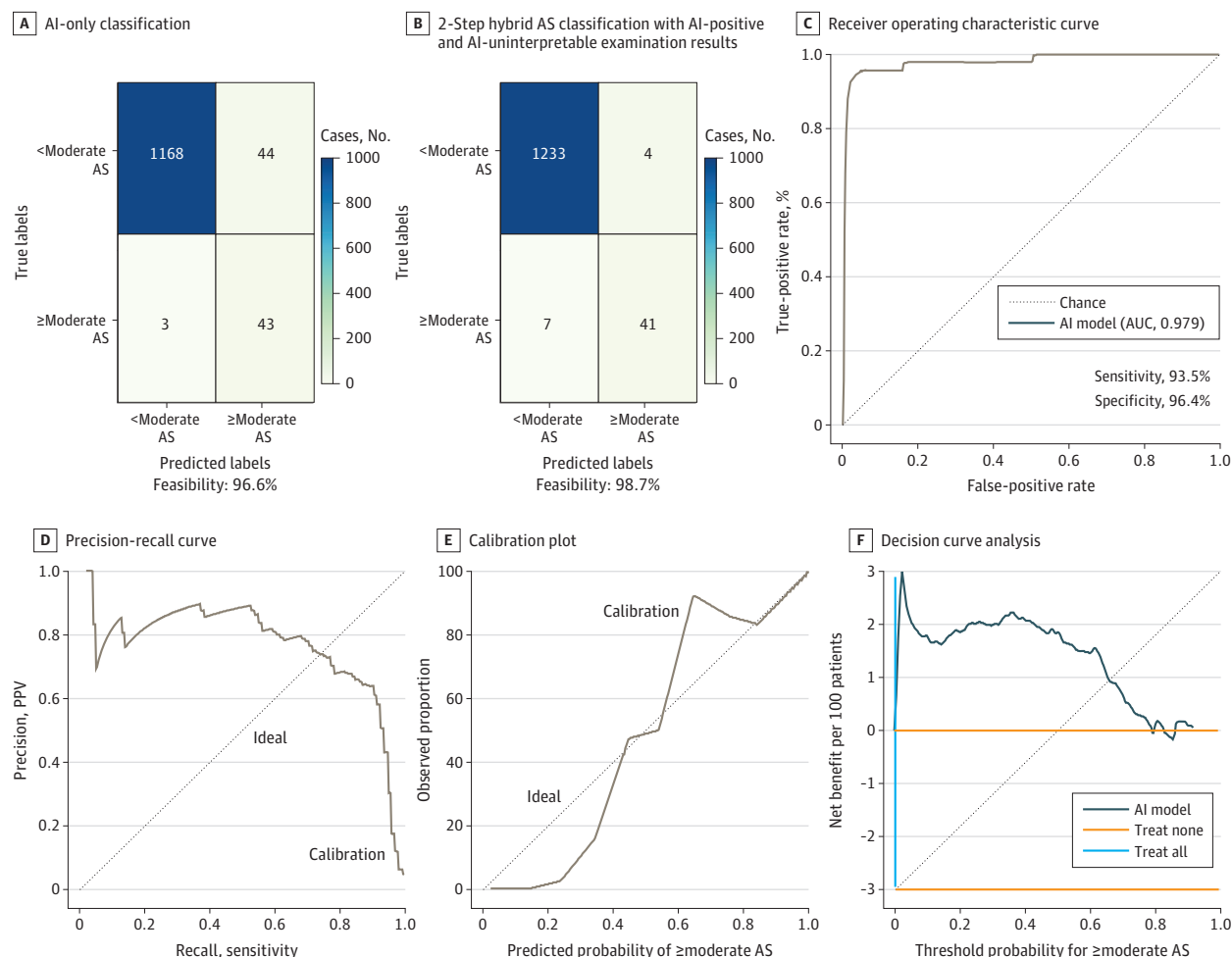
Discussion

In this multistage evaluation spanning conventional TTE cohorts, sonographer-acquired handheld studies, and prospective novice-performed, AI-guided FoCUS examinations, a deep learning model applied to 2-dimensional ultrasound video clips accurately identified patients with moderate or greater AS. Performance remained robust across internal, geographically distinct, and handheld-ultrasound cohorts. In the prospective cohort, novice operators acquired examinations suitable for automated analysis in most participants. A 2-step workflow in which only AI-positive or uninterpretable studies, approximately 10% of examinations, underwent expert review increased the positive predictive value to 91% and resulted in referral of fewer than 5% of participants for comprehensive TTE. These findings support AI-guided acquisition with automated interpretation for extending structural heart disease screening beyond the echocardiography laboratory.

Automated detection of AS using machine learning has been increasingly explored in echocardiography. Prior studies have shown that deep learning can identify clinically significant AS from expert-acquired TTE, including Doppler-independent approaches and models incorporating multiple echocardiographic views.¹⁵⁻¹⁷ These studies established the feasibility of AI-enabled AS detection. Two features distinguish the present approach. First, the model accepts multiple standard echocardiographic views rather than relying on a single imaging window,¹⁶⁻¹⁸ which may improve robustness when image acquisition is variable. Second, it relies exclusively on 2-dimensional imaging without requiring spectral or color Doppler, thereby simplifying acquisition and reducing the technical expertise required. These characteristics make the approach particularly suited to handheld ultrasound.

Unlike previous studies that primarily evaluated retrospective datasets of expert-acquired examinations,¹⁵⁻²⁰ we prospectively evaluated AI-guided handheld FoCUS acquired by operators without prior clinical or sonography experience for AI-assisted assessment of moderate or greater AS. Validation against an independently interpreted comprehensive Doppler echocardiographic examination is an important step be-

Figure 1. Confusion Matrices and Line Graphs of Model Performance in the Prospective Artificial Intelligence (AI)-Guided Focused Cardiac Ultrasound (FoCUS) Study



A and B, Confusion matrices of AI-only classification (A) and the 2-step hybrid aortic stenosis (AS) classification incorporating expert review of AI-positive and AI-uninterpretable examinations (B) in the prospective novice-acquired FoCUS study. Details of false-positive and false-negative cases are provided in eResults 1 in [Supplement 1](#). C, Area under the receiver operating characteristic curve (AUC) showing excellent discrimination (0.979), with high true-positive rates achieved at low false-positive rates. D, Precision-recall curve demonstrating

sustained precision across a broad range of recall, supporting favorable positive predictive value at clinically relevant operating points. E, Calibration plot using fixed-width probability bins (0.1). Points show the mean predicted probability vs observed event rate per bin; the diagonal line represents perfect calibration. F, Decision curve analysis indicating net clinical benefit for the AI model across a range of threshold probabilities compared with treat-all and treat-none strategies.

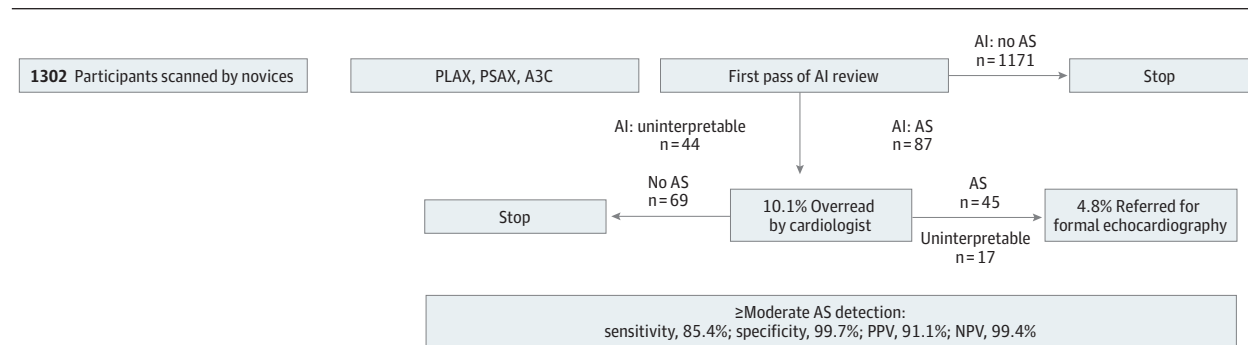
fore evaluation in a population-based testing program. Participants were recruited from a clinically referred echocardiography population rather than an unselected screening cohort. Nevertheless, the prevalence of moderate or greater AS was 3.7%, like that reported in a community study²¹ of adults who may represent a target population for future screening initiatives.

The intended role of this approach is triage rather than definitive diagnosis. It is not intended to replace physical examination or comprehensive echocardiography, but rather to provide an objective, imaging-based assessment that identifies individuals who should undergo comprehensive Doppler echocardiography. The model does not directly measure valve hemodynamics or determine AS severity; instead, it identifies a 2-dimensional morphological phenotype associated with mod-

erate or greater AS. Positive examinations require comprehensive Doppler echocardiography for confirmation. By relying only on brief B-mode cine loops, the workflow may enable image acquisition by nonexpert operators while preserving comprehensive TTE for patients with a higher likelihood of disease.

This study provides prospective evidence supporting AI-guided FoCUS screening for AS by novice operators. These findings confirm that individuals with minimal training can obtain interpretable cardiac images when supported by guided acquisition tools²²⁻²⁴ and demonstrate such images are suitable for automated AS assessment. AI-guided acquisition and interpretation do not eliminate operator dependence but reduce the expertise required to obtain diagnostically useful images. Image quality remains the principal determinant of per-

Figure 2. Artificial Intelligence (AI)–Guided 2-Step Screening Workflow for Detection of at Least Moderate Aortic Stenosis (AS) Using Novice-Operated Focused Cardiac Ultrasound (FoCUS)



A total of 1302 participants underwent FoCUS performed by novice operators using AI-guided handheld devices. All studies were first evaluated by the validated AI algorithm. Examinations classified as negative exited the screening pathway and were deemed negative for at least moderate AS, whereas AI-positive and uninterpretable cases underwent secondary review by an expert cardiologist (10% of studies) and were subsequently deemed positive, negative, or uninterpretable. This 2-step approach would result in only those with positive studies or uninterpretable studies (4.8% of total individuals)

needing referral for comprehensive transthoracic echocardiography. Overall diagnostic performance for detection of at least moderate AS demonstrated high sensitivity (85.4%), specificity (99.7%), positive predictive value (PPV; 91.1%), and negative predictive value (NPV; 99.4%), supporting efficient triage and scalable screening using novice-acquired FoCUS. A3C indicates apical 3-chamber view; PLAX, parasternal long-axis view; PSAX, parasternal short-axis view.

formance, while the algorithm performs most first-pass review and limits the need for expert interpretation. This workflow is intended to stratify patients by their likelihood of moderate or greater AS, not to confirm its presence or severity. This workflow would require physician review of only 10% of AI-FoCUS examinations and limit downstream referral to TTE to only 4.8% of patients. The small number of missed cases should be interpreted in the context of a screening strategy whose purpose is to efficiently identify patients requiring definitive comprehensive Doppler echocardiography rather than replace expert diagnostic evaluation. Comprehensive Doppler echocardiography remains the diagnostic reference standard for definitive assessment.

The prevalence of moderate or greater AS in the prospective cohort was 3.7%, consistent with contemporary epidemiologic estimates in community-dwelling older adults.^{21,25} In this low-prevalence setting, AI alone achieved a positive predictive value of 49%, which increased to 91% using the proposed 2-step workflow incorporating expert review of AI-positive and uninterpretable AI-FoCUS examinations using 2-dimensional imaging only. Although the development cohort was intentionally enriched for moderate or greater AS to support model training, the decision threshold was prespecified using the internal validation cohort and applied unchanged across all validation and prospective cohorts, with consistently robust discrimination.

Our findings complement prior work using AI-based electrocardiography to identify AS. Although AI-based electrocardiography is a low-cost population-scale prescreening tool, its modest accuracy limits its value in low-prevalence populations.⁹ A sequential strategy in which AI-based electrocardiography serves as an initial prescreen followed by AI-assisted FoCUS may improve efficiency while reducing unnecessary referrals for comprehensive TTE.²⁶

AS may be particularly suitable for a morphology-based screening strategy because it is common, progressive, frequently underrecognized, and characterized predominantly by leaflet thickening, fibrosis, and calcification.^{25,27-30} The present workflow requires only brief acquisition of several 2-dimensional views and selective expert review. Nevertheless, a disease-specific examination may be less efficient than a multipathology cardiac screening approach. The same novice-acquired FoCUS examination may ultimately support detection of additional structural abnormalities, and future studies should determine whether broader automated interpretation improves clinical yield and cost-effectiveness.^{22,23,26,31}

The proposed workflow is a triage strategy rather than a replacement for expert interpretation. The primary role of the AI system is to identify examinations that are likely normal while directing a relatively small subset of AI-positive or uninterpretable studies for expert review. In this cohort, this approach limited physician review to approximately 10% of studies while substantially increasing positive predictive value and maintaining a very high negative predictive value. Such a framework may be particularly attractive in resource-constrained environments where expert echocardiography interpretation capacity is limited.

Limitations

Several limitations merit consideration. First, although model performance was robust across geographically distinct institutions within one health care system and translated to handheld imaging, including novice-acquired studies, prospective validation across independent health care systems, ultrasound vendors, and heterogeneous acquisition protocols is needed. Second, participants were recruited from a clinically referred echocardiography population rather than an unselected screening population and symptom status was not available. Third, although this study

demonstrated diagnostic accuracy and operational feasibility, its effects on referral patterns, resource utilization, and patient outcomes remain unknown. Fourth, comprehensive Doppler TTE interpreted according to standard clinical practice served as the reference standard. Additional testing, including stress echocardiography, computed tomographic calcium scoring, or invasive hemodynamic assessment, may be required in selected patients with discordant findings. Fifth, the underlying cause of AS may influence its 2-dimensional appearance. The development and validation cohorts predominantly reflected calcific and fibrotic valve disease.²⁹ Although performance was consistent across bicuspid and trileaflet valve morphology and across age groups, further validation is needed in populations enriched for less common etiologies, including rheumatic valve disease.

Conclusions

In conclusion, AI-assisted 2-dimensional focused cardiac ultrasound accurately detected moderate or greater AS across diverse acquisition settings and operator experience, including novice-performed, AI-guided handheld FoCUS at the point of care. Prospective validation in a clinically referred cohort with relatively low disease prevalence, coupled with demonstration of both guided image acquisition and automated interpretation, supports further evaluation of this approach as a scalable triage strategy, and possibly even screening strategy, in primary care, community, and resource-limited settings.

ARTICLE INFORMATION

Accepted for Publication: August 7, 2026.

Published Online: August 28, 2026.
doi:10.1001/jamacardio.2026.3829

Author Contributions: Dr G. C. Kane and Dr Tsaban had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Drs Lee and Naser and Mr Kane contributed equally.

Concept and design: Lee, Naser, C. J. Kane, Killalea, Thaden, Lopez-Jimenez, Oh, G. C. Kane, Pislaru, Bird, Tsaban.

Acquisition, analysis, or interpretation of data: Lee, Naser, C. J. Kane, Kovac, Greason, Killalea, Gulati, Jackson, Borgeson, Malins, Anisuzzaman, Crestanello, Zacher, Camalan, Anand, Nkomo, Padang, Poterucha, Chao, Pelliikka, Friedman, G. C. Kane, Attia, Pislaru, Bird.

Drafting of the manuscript: Lee, Naser, C. J. Kane, Greason, Killalea, Jackson, Tsaban.

Critical review of the manuscript for important intellectual content: Lee, Naser, C. J. Kane, Kovac, Killalea, Gulati, Borgeson, Malins, Anisuzzaman, Crestanello, Zacher, Camalan, Thaden, Anand, Nkomo, Padang, Poterucha, Chao, Pelliikka, Lopez-Jimenez, Friedman, Oh, G. C. Kane, Attia, Pislaru, Bird.

Statistical analysis: Lee, C. J. Kane, Killalea, Borgeson, Zacher, Camalan, G. C. Kane.

Administrative, technical, or material support: Greason, Jackson, Malins, Camalan, Anand, Poterucha, Pelliikka, Friedman, G. C. Kane, Attia, Tsaban.

Supervision: Thaden, Nkomo, Lopez-Jimenez, G. C. Kane, Pislaru, Bird, Tsaban.

Conflict of Interest Disclosures: Drs Lee, Malins, Naser, Pislaru, G. C. Kane, and Attia invented algorithms licensed to UltraSight and may benefit from their commercialization through Mayo Clinic. Drs Friedman, Lopez-Jimenez, and Attia are members of the scientific advisory board for Anumana, an artificial intelligence (AI) company commercializing AI-based electrocardiography. Dr Lopez-Jimenez is also a coinventor of several AI-based electrocardiography algorithms licensed to Anumana and an artificial heart valve licensed to Colibri. Dr Poterucha reported research grants paid to affiliated institutions from Pfizer, Eidos Therapeutics, Janssen, the American Heart Association, and the Tianqiao and Chrissy Chen Foundation; stock ownership in Baxter and Abbott Laboratories; and a patent for the use of

electrocardiographic analysis to detect structural heart disease. No other disclosures were reported.

Funding/Support: Mr C. J. Kane, Mr Kovac, Ms Gulati, and Mr Schonfeld received support from the Mayo Clinic Program for Undergraduate Learning in Structural Heart Disease and Echocardiography, and Ms Killalea received support through the University College Dublin Gaensler Scholarship and the Ad Astra Academy Bursary.

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Meeting Presentation: This paper was presented at ESC Congress 2026; August 28, 2026; Munich, Germany.

Additional Information: ChatGPT-5.6 Sol was used on July 28, 2026, to improve the wording and readability of portions of the manuscript. It was not used to generate scientific content, analyze or interpret data, or create or modify tables, figures, or other data presentations. Dr Tsaban takes responsibility for the integrity of the content.

Data Sharing Statement: See [Supplement 2](#).

REFERENCES

- Osnabrugge RL, Mylotte D, Head SJ, et al. Aortic stenosis in the elderly: disease prevalence and number of candidates for transcatheter aortic valve replacement: a meta-analysis and modeling study. *J Am Coll Cardiol*. 2013;62(11):1002-1012. doi:10.1016/j.jacc.2013.05.015
- Généreux P, Sharma RP, Cubeddu RJ, et al. The mortality burden of untreated aortic stenosis. *J Am Coll Cardiol*. 2023;82(22):2101-2109. doi:10.1016/j.jacc.2023.09.796
- Thoenes M, Bramlage P, Zamorano P, et al. Patient screening for early detection of aortic stenosis (AS)-review of current practice and future perspectives. *J Thorac Dis*. 2018;10(9):5584-5594. doi:10.21037/jtd.2018.09.02
- Lancellotti P, Magne J, Dulgheru R, et al. Outcomes of patients with asymptomatic aortic stenosis followed up in heart valve clinics. *JAMA Cardiol*. 2018;3(11):1060-1068. doi:10.1001/jamacardio.2018.3152
- Gahl B, Celik M, Head SJ, et al. Natural history of asymptomatic severe aortic stenosis and the association of early intervention with outcomes: a systematic review and meta-analysis. *JAMA Cardiol*. 2020;5(10):1102-1112. doi:10.1001/jamacardio.2020.2497
- Praz F, Borger MA, Lanz J, et al; ESC/EACTS Scientific Document Group. 2025 ESC/EACTS guidelines for the management of valvular heart disease. *Eur Heart J*. 2025;46(44):4635-4736. doi:10.1093/eurheartj/ehaf194
- Fink JC, Schmid CH, Selker HP. A decision aid for referring patients with systolic murmurs for echocardiography. *J Gen Intern Med*. 1994;9(9):479-484. doi:10.1007/BF02599215
- Mangione S, Nieman LZ. Cardiac auscultatory skills of internal medicine and family practice trainees: a comparison of diagnostic proficiency. *JAMA*. 1997;278(9):717-722. doi:10.1001/jama.1997.03550090041030
- Cohen-Shelly M, Attia ZI, Friedman PA, et al. Electrocardiogram screening for aortic valve stenosis using artificial intelligence. *Eur Heart J*. 2021;42(30):2885-2896. doi:10.1093/eurheartj/ehab153
- Kwon JM, Lee SY, Jeon KH, et al. Deep learning-based algorithm for detecting aortic stenosis using electrocardiography. *J Am Heart Assoc*. 2020;9(7):e014717. doi:10.1161/JAHA.119.014717
- Elias P, Poterucha TJ, Rajaram V, et al. Deep learning electrocardiographic analysis for detection of left-sided valvular heart disease. *J Am Coll Cardiol*. 2022;80(6):613-626. doi:10.1016/j.jacc.2022.05.029
- Spencer KT, Flachskampf FA. Focused cardiac ultrasonography. *JACC Cardiovasc Imaging*. 2019;12(7 pt 1):1243-1253. doi:10.1016/j.jcmg.2018.12.036
- Mitchell C, Rahko PS, Blauwet LA, et al. Guidelines for performing a comprehensive transthoracic echocardiographic examination in adults: recommendations from the American Society of Echocardiography. *J Am Soc Echocardiogr*. 2019;32(1):1-64. doi:10.1016/j.echo.2018.06.004
- Baumgartner H, Hung J, Bermejo J, et al. Recommendations on the echocardiographic assessment of aortic valve stenosis: a focused update from the European Association of Cardiovascular Imaging and the American Society of Echocardiography. *J Am Soc Echocardiogr*. 2017;30(4):372-392. doi:10.1016/j.echo.2017.02.009
- Yang F, Zhang Y, Gao Y, et al. Deep learning-assisted aortic stenosis detection and grading based on multiview versus single-view

- echocardiography. *Quant Imaging Med Surg*. 2025; 15(11):11192-11204. doi:10.21037/qims-2025-415
16. Wessler BS, Huang Z, Long GM Jr, et al. Automated detection of aortic stenosis using machine learning. *J Am Soc Echocardiogr*. 2023;36(4):411-420. doi:10.1016/j.echo.2023.01.006
 17. Holste G, Oikonomou EK, Mortazavi BJ, et al. Severe aortic stenosis detection by deep learning applied to echocardiography. *Eur Heart J*. 2023;44(43):4592-4604. doi:10.1093/eurheartj/ehad456
 18. Krishna H, Desai K, Slostad B, et al. Fully automated artificial intelligence assessment of aortic stenosis by echocardiography. *J Am Soc Echocardiogr*. 2023;36(7):769-777. doi:10.1016/j.echo.2023.03.008
 19. Holste G, Oikonomou EK, Tokodi M, Kovács A, Wang Z, Khera R. Complete AI-enabled echocardiography interpretation with multitask deep learning. *JAMA*. 2025;334(4):306-318. doi:10.1001/jama.2025.8731
 20. Vukadinovic M, Chiu IM, Tang X, et al. Comprehensive echocardiogram evaluation with view primed vision language AI. *Nature*. 2026;650(8103):970-977. doi:10.1038/s41586-025-09850-x
 21. Brenner MI, Chuang ML, Nishimura R, et al. Population prevalence of valvular heart disease in the United States: the PREVUE-VALVE study. *J Am Coll Cardiol*. 2026;87(23):3227-3239. doi:10.1016/j.jacc.2026.02.5137
 22. Mor-Avi V, Khandheria B, Klempfner R, et al. Real-time artificial intelligence-based guidance of echocardiographic imaging by novices: image quality and suitability for diagnostic interpretation and quantitative analysis. *Circ Cardiovasc Imaging*. 2023;16(11):e015569. doi:10.1161/CIRCIMAGING.123.015569
 23. Kane CJ, Borgeson J, Greason C, et al. Smarter FoCUS: AI-guided focused cardiac ultrasound enables novice detection of left ventricular dysfunction. *Eur Heart J Digit Health*. 2026;7(5):ztag066. doi:10.1093/ehjdh/ztag066
 24. Killalea MM, Borgeson J, Kane CJ, et al. Improving structural heart disease screening: AI-ECG and novice AI-guided focused cardiac ultrasound. *NPJ Digit Med*. Published online June 9, 2026. doi:10.1038/s41746-026-02856-3
 25. Nkomo VT, Gardin JM, Skelton TN, Gottdiener JS, Scott CG, Enriquez-Sarano M. Burden of valvular heart diseases: a population-based study. *Lancet*. 2006;368(9540):1005-1011. doi:10.1016/S0140-6736(06)69208-8
 26. Alexandrino FB, Schlesinger R, Bird J, et al. Integrating AI-ECG and point-of-care cardiac ultrasound for screening structural heart disease: a proof-of-concept study. *Am Heart J*. 2026;294:107337. doi:10.1016/j.ahj.2025.107337
 27. Coffey S, Harper AR, Cairns BJ, Roberts IS, Prendergast BD. Clinical information has low sensitivity for postmortem diagnosis of heart valve disease. *Heart*. 2017;103(13):1031-1035. doi:10.1136/heartjnl-2016-310718
 28. Lindman BR, Sukul D, Dweck MR, et al. Evaluating medical therapy for calcific aortic stenosis: JACC State-of-the-Art Review. *J Am Coll Cardiol*. 2021;78(23):2354-2376. doi:10.1016/j.jacc.2021.09.1367
 29. Otto CM, Newby DE, Hillis GS. Calcific aortic stenosis: a review. *JAMA*. 2024;332(23):2014-2026. doi:10.1001/jama.2024.16477
 30. Praz F, Beyersdorf F, Haugaa K, Prendergast B. Valvular heart disease: from mechanisms to management. *Lancet*. 2024;403(10436):1576-1589. doi:10.1016/S0140-6736(23)02755-1
 31. Schlesinger RP, Alexandrino FB, Lee E, et al. Structural heart disease screening using artificial intelligence-enabled electrocardiogram and novice handheld cardiac ultrasound: a cost analysis. *Eur Heart J Open*. 2026;6(2):oeag049. doi:10.1093/ehjopen/oeag049