Circulation: Cardiovascular Imaging

ORIGINAL ARTICLE

# Multimodal Artificial Intelligence for Cardiac Amyloidosis Diagnosis: Integrating Echocardiography With Clinical and Laboratory Data for Improved Detection

Jeremy A. Slivnick, MD; Sze Chi Lim, BS; Michael Randazzo, MD; Mathew S. Maurer, MD; Stephen Helmke, BS, MPH; Marielle Scherrer-Crosbie, MD, PhD; Azin Vakilpour, MD; Karolina M. Zareba, MD; Akash Goyal, MD; Richard Cheng, MD; Nicole Wakamatsu, MD; Tetsuji Kitano, MD, PhD; Masaaki Takeuchi, MD, PhD; Viviane Tiemi Hotta, MD, PhD; Marcelo Luiz Campos Vieira, MD, PhD; Pablo Elissamburu, MD; Ricardo E. Ronderos, MD, PhD; Aldo Prado, MD; Efstratios Koutroumpakis, MD; Anita Deswal, MD; Amit Pursnani, MD; Nitasha Sarswat, MD; Karima Addetia, MD; Juan Cotella, MD; Frederick L. Ruberg, MD; Matthew Frost, BE; Marianna Fontana, MD, PhD; Carolyn Lam Su Ping, MBBS, PhD; Roberto M. Lang, MD; Federico M. Asch, MD

**BACKGROUND:** Cardiac amyloidosis (CA) is an underdiagnosed yet treatable cause of heart failure in which timely diagnosis is essential to initiate life-prolonging therapies. While artificial intelligence (AI)-based tools using transthoracic echocardiography (TTE), electrocardiography, or electronic health records have demonstrated promise for CA detection, most rely on single data sources. We aimed to evaluate whether integrating clinical, laboratory, and TTE biomarkers improves the performance of an existing TTE-based AI model for CA detection.

**METHODS:** We developed and tested a combined AI echo-clinical model (AI-ECM) incorporating demographics, laboratory biomarkers, and TTE parameters into a previously validated TTE-only AI model (Us2.Ca). Model training and internal validation were performed using the Amyloidosis Imaging International Consortium, a global multiethnic registry comprised of 727 patients with CA and 316 controls, including 202 with suspected transthyretin-CA with negative diagnostic evaluation and 114 patients with biopsy-proven extracardiac light chain amyloidosis without cardiac involvement. Ground truth CA diagnosis was adjudicated per consensus criteria. AI-ECM and Us2.Ca performance was assessed using area under the curve, accuracy, sensitivity, and specificity.

**RESULTS:** In building the AI-ECM, feature importance analysis showed that having the Us2.Ca prediction scores, relative wall thickness, gender, and estimated glomerular filtration rate contributed most to performance. The AI-ECM demonstrated superior performance (area under the curve, 0.94; accuracy, 90%; sensitivity, 93%; specificity, 85%) compared with the Us2.Ca (area under the curve, 0.89; accuracy, 80%; sensitivity, 76%; specificity, 91%; $P$=0.006). While the Us2.Ca model classification was indeterminate in 9% of the cases, the AI-ECM allowed classification of all cases. The AI-ECM improved sensitivity for light chain-CA detection and maintained high accuracy across subtypes and control groups.

**CONCLUSIONS:** A multiparametric AI model integrating basic clinical, laboratory, and TTE data with the deep learning Us2.Ca improved performance for CA detection over Us2.Ca alone. This approach represents a step toward scalable, AI-guided precision diagnostics for CA in diverse populations.

**GRAPHICAL ABSTRACT:** A graphical abstract is available for this article.

**Key Words:** artificial intelligence ■ biomarkers ■ deep learning ■ heart failure ■ prealdumin

See Editorial by Author

Correspondence to: Federico M. Asch, MD, MedStar Health Research Institute, 100 Irving St, NW. Ste 5123., Washington, DC 20010. Email federico.asch@medstar.net
This manuscript was sent to Linda D. Gillam, Senior Guest Editor, for review by expert referees, editorial decision, and final disposition.
Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/CIRCIMAGING.126.019610.

*Circulation: Cardiovascular Imaging* is available at www.ahajournals.org/journal/circimaging

## CLINICAL PERSPECTIVE

Cardiac amyloidosis remains underrecognized despite increasing availability of disease-modifying therapies and growing awareness of the disease. In this international multicenter study, we developed and validated a fully automated multimodal artificial intelligence echo-clinical model integrating deep learning-based echocardiographic analysis with routinely available demographic, laboratory, and echocardiographic parameters. Compared with a transthoracic echocardiography-only artificial intelligence model and established multiparametric echocardiographic scores, the multimodal model improved diagnostic performance while eliminating indeterminate classifications. The addition of clinical and laboratory parameters particularly improved sensitivity for light-chain cardiac amyloidosis detection. These findings support the potential role of multimodal artificial intelligence as an accurate and scalable approach for cardiac amyloidosis screening with echocardiography.

## Nonstandard Abbreviations and Acronyms

| | |
|---|---|
| **AI** | artificial intelligence |
| **AI-ECM** | artificial intelligence echo-clinical model |
| **AL-CA** | light chain cardiac amyloidosis |
| **AUC** | area under the curve |
| **ATTR-CA** | transthyretin cardiac amyloidosis |
| **CA** | cardiac amyloidosis |
| **eGFR** | estimated glomerular filtration rate |
| **GLS** | global longitudinal strain |
| **IWT** | increased wall thickness |
| **LV** | left ventricular |
| **NT-proBNP** | N-terminal pro-B-type natriuretic peptide |
| **RWT** | relative wall thickness |
| **SAB** | septal apical-to-basal |
| **TTE** | transthoracic echocardiography |

Cardiac amyloidosis (CA) is an increasingly recognized and treatable cause of heart failure due to deposition of misfolded proteins within the myocardium. While effective life-prolonging therapies exist for the 2 most common subtypes, they are most impactful when initiated before onset of late-stage disease.[1–4] Timely diagnosis is therefore critical to ensure access to treatment.

In clinical practice, suspicion for CA is often raised by the presence of characteristic red flag features spanning demographics (older age, male sex), clinical history (bilateral carpal tunnel, lumbar spinal stenosis), laboratory biomarkers (elevated cardiac troponins and NT-proBNP [N-terminal pro-B-type natriuretic peptide]), and electrocardiographic and echocardiographic findings. Transthoracic echocardiography (TTE) plays a pivotal role in this workup due to its wide availability and low cost. Classical TTE features—including increased wall thickness (IWT), impaired global longitudinal strain (GLS) sparing apical segments, increased myocardial echogenicity, and biatrial enlargement—may prompt further confirmatory testing such as nuclear scintigraphy, cardiac magnetic resonance, and myocardial biopsy. However, despite improved awareness and the emergence of targeted therapies, diagnostic delays remain all too common. Even in the modern era, the median time from clinical suspicion to diagnosis can approach 1 year, likely underestimating true delays from symptom onset.[5] While provider awareness is 1 factor, another key barrier is that current frontline tests lack sufficient accuracy for broad screening. Even TTE has limitations: recent studies show that hallmark features such as apical sparing in GLS may have lower diagnostic performance in modern cohorts.[6]

To address this gap, recent advances in artificial intelligence (AI) and machine learning are being applied to various modalities—TTE, ECG, electronic health records, and laboratory data—to improve early CA detection.[7–15] These models have shown promising performance with the potential to facilitate scalable, automated CA screening. In particular, deep learning-based AI-driven algorithms for automated analysis of TTE images are showing potential as effective tools to screen for CA.[7–9] Our group has recently showed that a novel deep learning-based Us2.Ca model (recently FDA-cleared) outperformed screening scores currently being used in clinical practice, with high accuracy.[16] However, a key limitation of such models is their exclusion of clinical, laboratory, and additional imaging data that inform real-world decision-making, potentially overlooking important diagnostic cues.

Accordingly, we hypothesized that integration of clinical, laboratory, and TTE data—including outputs from an AI TTE model—into a fully automated, comprehensive machine learning TTE and AI echo-clinical model (AI-ECM) would improve performance over a previously developed deep learning-based AI TTE model (Us2.Ca) for CA detection.[16] We therefore aimed to evaluate the incremental value of adding clinical and laboratory features to the TTE-only AI model within a large, global, multiethnic registry.

# METHODS

## Combined Echo-Clinical Model AI Cohort

The AI-ECM, integrating clinical, laboratory, and imaging biomarkers into the previously developed deep learning-based model, Us2.Ca, was trained and validated using the Amyloidosis

Imaging International Consortium database. This global, multiethnic registry contains clinical, laboratory, and imaging data, as well as de-identified TTE digital imaging and communications in medicine. The entire Amyloidosis Imaging International Consortium cohort includes 1140 patients with CA and phenotypically relevant controls referred for clinically indicated TTE exam, from 9 academic medical centers spanning 4 countries: Columbia University Irving Medical Center (New York City), The Ohio State University (Columbus), University of Chicago (Chicago), University of Pennsylvania (Philadelphia), University of Washington (Seattle), Hospital of the University of the Occupational and Environmental Health (Kitakyushu, Japan), Instituto do Coração (Sao Paulo, Brazil), Instituto Cardiovascular de Buenos Aires (Buenos Aires, Argentina),and National University of Tucumán (Tucuman, Argentina). Each site obtained appropriate ethical permission with waiver of informed consent for this retrospective study. The cohort used in this analysis is independent of previously published or submitted validation cohorts for the TTE-only Us2.Ca model.[17] Because of the sensitive nature of the data collected for this study, requests to access the data set from qualified researchers trained in human subject confidentiality protocols may be sent to the corresponding author.

After exclusion of 92 (8.1%) patients in whom an apical 4-chamber view was not available for analysis or was unable to be detected by the AI (n=73), or was present but of insufficient quality for AI analysis (n=19), this study included 727 patients with proven CA and 316 phenotypic controls—comprised of 202 patients in which CA was suspected but excluded by technetium pyrophosphate scan and 114 patients with biopsy-proven AL amyloidosis without cardiac involvement as defined by negative cardiac magnetic resonance and negative NT-proBNP.[18] TTEs were acquired closest to the time of diagnosis in patients with CA, or at the initiation of CA workup in control patients (median time, 27 [IQR, 5–110] days between TTE and CA diagnosis). All studies were performed by trained sonographers using commercial ultrasound systems (Philips or GE). Clinical and laboratory biomarker data—including demographics, NT-proBNP, serum and urine immunofixation, and serum free light chains closest to the TTE—were collected through review of the electronic health record at each institution.

## CA Diagnosis

The presence of CA was independently adjudicated by a CA expert at each enrolling center, according to multisocietal guidelines: (1) positive endomyocardial biopsy, (2) positive extracardiac biopsy with typical cardiac imaging features and elevated NT-proBNP (for light chain-CA [AL-CA]), or (3) noninvasively by Perugini grade 2 or 3 cardiac uptake on 99mTc-DPD scintigraphy in the absence of an abnormal free light chain ratio or monoclonal gammopathy.[19,20] Technetium pyrophosphate scintigraphy was interpreted using the Perugini visual grading system; grade 2 to 3 uptake in the absence of a monoclonal protein was considered diagnostic of transthyretin-CA (ATTR-CA) in the absence of a monoclonal light chain, whereas grade 0 to 1 uptake was considered negative for ATTR-CA, and such patients were classified as controls in accordance with established diagnostic guidelines. Among patients with ATTR-CA, 4 had Perugini grade 0 and 7 had Perugini grade 1 uptake on technetium pyrophosphate scintigraphy; all of these 11 had endomyocardial biopsy-confirmed CA (Table S1).

## TTE Image Analysis and Multiparametric TTE Score Calculation

The IWT score and the systemic light chain amyloidosis (AL) score are established multiparametric echocardiographic scoring systems designed to aid in the diagnosis of CA, particularly its subtypes ATTR-CA and AL-CA amyloidosis.[21] In the original study, an IWT score threshold of ≥6 points and an AL score threshold of ≥3 points yielded the optimal balance of sensitivity and specificity for CA detection. In the present study, these scores were calculated using these established thresholds to serve as a conventional clinical benchmarks against which the performance of the TTE-only deep learning model (Us2.ai) and the proposed integrated clinical-AI-ECM were compared.

The IWT score incorporates 5 echocardiographic parameters: relative wall thickness (RWT >0.6, 3 points), E/e´ ratio (>11, 1 point), tricuspid annular plane systolic excursion (≤19 mm, 2 points), GLS (longitudinal strain ≤−13%, 3 points), and septal apical-to-basal (SAB) strain ratio (>2.9, 3 points), with scores ≥6 considered abnormal. The AL score includes 4 variables: RWT >0.52 (2 points), E/e´ >10 (2 points), tricuspid annular plane systolic excursion ≤19 mm (1 point), and longitudinal strain ≤−14% (1 point), with scores ≥3 considered abnormal. These parameters and their respective thresholds were applied in the derivation of IWT scores. All TTE measurements used to derive the multiparametric scores were obtained through automated analysis using a validated FDA and CE-marked AI software including GLS (Us2.ai, Singapore).

## Previous Echo AI Model Development

Us2.Ca is a regulatory-cleared (FDA-cleared and CE-marked) pattern recognition model for CA detection (Us2.ai, Singapore).[16] It was developed using apical 4-chamber echocardiographic videos from 4371 patients and 1526 controls. For each patient, a continuous CA probability score was generated by Us2.ai software, which employs an automated pipeline to identify appropriate views, segment cardiac cycles, and perform beat-by-beat inference. Classification thresholds in this study were applied as per the original model's specifications: scores ≥0.80 were considered positive for CA, scores ≤0.45 were considered negative, and scores between 0.45 and 0.80 were classified as indeterminate.

## Combined Echo-Clinical Model Development

The data set for developing the AI-ECM model was randomly split into training (70%), validation (15%), and test (15%) sets, with stratification to ensure proportional representation of ATTR-CA, AL-CA, and control cases across each split. The AI-ECM model is a supervised machine learning model developed using the XGBoost model, designed to integrate multimodal clinical, imaging, and biochemical data to improve the diagnosis of CA.[22] XGBoost is a high performance, tree-based ensemble method that builds decision trees sequentially, where each new tree is trained to correct the residual errors of the preceding ensemble. It optimizes a regularized objective function with a penalty term that discourages overly complex models. This regularization improves generalizability, particularly in

data sets with correlated or high-dimensional inputs. The final prediction is made by aggregating the outputs of all trees in the ensemble, weighted by their respective learning contributions.

The model incorporates probability scores from the Us2.Ca alongside selected demographic parameters, TTE measurements, and biochemical markers. Specifically, the input features can included RWT, global longitudinal left ventricular apex-to-basal strain ratio (LV SAB), as well as patient demographics such as age, gender, height, weight, and body mass index. Additional serum and urine biomarkers included serum kappa and lambda light chains, the presence of monoclonal proteins detected via immunofixation in serum and urine, and renal and cardiac biomarkers such as creatinine and estimated glomerular filtration rate (eGFR). NT-proBNP was not included as an input feature in the fusion models because it forms part of the diagnostic criteria used to adjudicate cardiac involvement in AL amyloidosis. Inclusion of this biomarker in the predictive model would therefore introduce incorporation bias when distinguishing CA from noncardiac controls.[20] In all instances, the laboratory values closest to the time of TTE acquisition were used.

The XGBoost model was developed in a stepwise manner to assess the incremental value of each group of features. The initial model, Us2.Ca+RWT, combined the pattern recognition CA score with RWT alone. The second iteration, Us2.Ca +RWT+SAB, added LV SAB to enhance strain-based myocardial deformation characterization. The third model, Us2.Ca+RWT+SAB+Clin, further incorporated demographic variables to account for patient-specific factors. Subsequently, Us2.Ca+RWT+SAB+Clin+Labs included protein biomarkers and finally, the AI-ECM model integrated all available features, adding the normalized biomarkers of renal function (creatinine and eGFR) to the aforementioned parameters and the Us2.Ca score. The final AI-ECM did not use indeterminate values.

Importantly, the CA score generated by the Us2.Ca served as a mandatory input for all subsequent XGBoost fusion models developed in this study. Each XGBoost model version was built upon the foundation of the Us2.Ca score, with progressively additional clinical, imaging, demographic, and biochemical features layered in to assess their incremental diagnostic value when paired with the Us2.Ca. To isolate the specific added value of the Us2.Ca pattern recognition score, a separate sensitivity analysis was also performed by training an XGBoost model using only demographic, biomarker, and TTE data (excluding the Us2.Ca score).

Hyperparameters of the XGBoost models were optimized via 5-fold cross-validation on the training and validation data splits. Notably, the same set of hyperparameters was consistently selected across all fusion models, indicating stable model tuning despite incremental feature additions. This uniformity in hyperparameter selection facilitates a fair comparison of model performance attributable to the varying feature sets. Early stopping with a patience of 10 rounds was used based on validation performance to prevent overfitting and enhance generalizability.

## Handling of Missing Data in Echo-Clinical Models

Missing data rates for all clinical, laboratory, and echocardiographic input features are reported by CA status in Table S2. XGBoost natively accommodates missing values through a learned sparsity-aware split strategy, in which the algorithm determines the optimal default branch direction for instances with missing values at each tree node during training based on which direction minimizes the loss function. This approach avoids the need for imputation and has been shown to perform well under both random and nonrandom missingness patterns.

However, as certain echocardiographic parameters (eg, tricuspid annular plane systolic excursion and strain ratios) may exhibit differential missingness between CA cases and controls—likely reflecting disease-related image quality limitations—we performed the following additional analyses to assess the impact of missing data on model performance. First, effective sample sizes (ie, the number of patients with nonmissing values) are reported for each input feature stratified by CA status and by model iteration (Table S3). Second, a complete-case sensitivity analysis was performed by restricting the test set to patients with no missing values across all AI-ECM input features. AI-ECM performance metrics were recalculated in this complete-case subset to assess whether differential missingness influenced model comparisons. Finally, to directly evaluate whether missingness patterns contributed to predictions, we trained an auxiliary XGBoost model using only binary missingness indicators (ie, missing versus Nonmissing) for each input feature. The discriminative performance of this missingness-only model was assessed to quantify the extent to which the pattern of data availability alone was predictive of CA status.

## Automated Quantification of Echocardiographic Features

All echocardiographic parameters used in the calculation of the IWT score, AL score, and as input features to the AI-ECM model were derived automatically from complete transthoracic echocardiographic studies using a fully automated, regulatory-cleared software pipeline (Us2.ai, Singapore). The software performs end-to-end image processing, including automated view identification, cardiac chamber segmentation, and quantitative measurement of RWT, LV SAB, tricuspid annular plane systolic excursion, longitudinal strain, and E/e′ parameters, without manual input or operator adjustment. These measurements are generated from the raw TTE study upon completion of automated analysis, ensuring reproducible quantification across large data sets. This fully automated workflow enables scalable application of the IWT score, AL score, and AI-ECM model, allowing advanced echocardiographic features such as RWT and SAB to be incorporated into routine clinical and research workflows with the same ease as conventional measurements.

## Statistical Analysis

Model performance for each of the staged combined clinical and TTE models was evaluated on the validation and test sets using metrics including the area under the receiver operating characteristic curve, sensitivity, specificity, and overall accuracy. Comparative assessment across the staged models allowed quantification of the added predictive value contributed by each additional set of features. For all combined clinical and TTE models, the classification threshold for distinguishing CA from control was selected such that specificities in the validation and test splits were closest to 85%, allowing for a consistent benchmark against the Us2.Ca model. Us2.Ca and IWT model performance were similarly evaluated using the area under the

curve (AUC), sensitivity, specificity, yield, and overall accuracy. Baseline characteristics were summarized using descriptive statistics; continuous variables were compared between CA and control groups using Student *t* test and categorical variables using the $\chi^2$ test (Table 1).

To quantify the incremental benefit of the IWT model and newer XGBoost-based staged models over the baseline Us2.Ca model, the categorical net reclassification improvement was calculated. Given the different architectures, the net reclassification improvement was determined using the model-specific optimal thresholds (fixed at 85% specificity for the staged models and cutoff score of 6 for the IWT scoring model) to identify net correct movement of patients into appropriate risk categories. Overall model calibration and probabilistic accuracy were assessed using the Brier score, representing the mean squared difference between predicted probabilities and actual outcomes. For the IWT scoring model, raw scores were normalized to a 0 to 1 scale to facilitate the Brier score calculation.

Ninety-five percent CIs for AUC were calculated using the DeLong method, while CIs for all other point estimates were computed using the Wilson method. All statistical tests were 2-sided, with a significance level set at $\alpha$ <0.05. To assess the relative importance of model inputs, feature importance scores were extracted from the XGBoost classifier using the gain metric, which quantifies the average improvement in model performance, or a reduction in loss, brought by a feature when it is used in a decision tree split. Gain-based importance reflects how useful each feature was in improving model purity and decision boundaries across all boosting rounds.

Given the lower prevalence of CA in clinical practice, we performed nonparametric stratified bootstrap analysis to simulate PPV and NPV across a range of hypothetical disease prevalences. The upper range of prevalence reflects clinical scenarios in which the pretest probability of CA is elevated, such as among patients referred for echocardiography with clinical suspicion of infiltrative cardiomyopathy. For each investigated model and each disease prevalence, we used a stratified bootstrap resampling approach (2000 iterations). Specifically, sensitivity and specificity were estimated by resampling the test results of gold-standard positive and negative cases independently with replacement. These estimates were then used to calculate the simulated PPV and NPV using Bayes' theorem. Point estimates for PPV and NPV were reported as the mean of the bootstrap distribution, with 95% CIs derived using the percentile method (2.5th and 97.5th percentiles). This simulation allowed for a robust comparison of the models across varying clinical contexts.

# RESULTS

## Patient Demographics

A comparison of clinical, laboratory, and TTE features between patients with CA and controls is presented in Table 1. Overall, patient demographics were similar between



**Table 1. Comparison of Clinical, Laboratory, and Transthoracic Echocardiography Parameters Between CA and Controls. Continuous Variables Are Compared Using Student *t* Test and Categorical Variables Using the $\chi^2$ Test Variables**

| Clinical characteristic | CA cases (N=727) | Controls (N=316) | *P* value |
|---|---|---|---|
| Age, y, mean±SD | 70.9±11.1 | 69.3±13.9 | 0.0864 |
| Weight, kg, mean±SD | 78.2±19.7 | 82.8±22.0 | 0.0014 |
| Height, cm, mean±SD | 170.6±11.6 | 169.1±11.4 | 0.0580 |
| BMI, kg/m$^2$, mean±SD | 26.8±6.0 | 29.0±7.5 | <0.0001 |
| Gender, male, n (%) | 535 (73.6%) | 182 (57.6%) | <0.0001 |
| Ethnicity (n%) | | | |
| Black | 210 (28.9%) | 105 (33.2%) | <0.0001 |
| Hispanic or Latino | 120 (16.5%) | 19 (6.0%) | |
| White | 280 (38.5%) | 163 (51.6%) | |
| Others/not reported | 117 (16.1%) | 29 (9.2%) | |
| NT-proBNP, pg/mL; median (IQR) | 2492.0 (693.2–6602.5) | 826.0 (182.0–3804.0) | 0.2129 |
| BNP, pg/mL, median (IQR) | 505.0 (237.5–930.5) | 87.0 (41.0–280.9) | <0.0001 |
| Creatinine, mg/dL; median (IQR) | 1.2 (1.0–1.7) | 1.2 (0.9–1.8) | 0.6654 |
| eGFR, mL/min per 1.73 m$^2$; median (IQR) | 56.4 (37.6–72.0) | 59.0 (34.2–80.0) | 0.2954 |
| RWT, mean±SD | 0.7±0.2 | 0.5±0.2 | <0.0001 |
| TAPSE, mm, mean±SD | 17.8±4.9 | 21.9±5.2 | <0.0001 |
| E/e′ mean, mean±SD | 16.5±6.7 | 12.2±6.0 | <0.0001 |
| A4C LV GLS, %, mean±SD | −12.5±4.6 | −15.6±5.1 | <0.0001 |
| A4C LV SAB, mean±SD | 1.8±0.9 | 1.3±0.6 | <0.0001 |
| LV SAB, mean±SD | 2.2±2.0 | 1.4±0.7 | <0.0001 |

A4C indicates apical 4-chamber; BMI, body mass index; CA, cardiac amyloidosis; eGFR, estimated glomerular filtration rate; GLS, global longitudinal strain; LV, left ventricular; NT-proBNP, N-terminal pro B-type natriuretic peptide; SAB, septal apical basal; and TAPSE, tricuspid annular planar systolic excursion.

patients with CA and controls, apart from higher male and Hispanic/Latino prevalence in the population with CA. NT-proBNP and BNP were significantly more abnormal in the CA group. The percentage of missing clinical, laboratory, and echo parameters is depicted in Table S2.

## Echo-Clinical Model Development and Validation

Although the XGBoost model is robust to collinearity, Spearman correlation analysis is depicted to show the relationship between different parameters (Figure 1). This demonstrated no significant collinearity between clinical, laboratory, TTE measurements, and the TTE-based Us2.Ca output (Figure 1). Feature importance of clinical, laboratory, and imaging biomarkers for CA detection using the AI-ECM is depicted in Table 2. The TTE-based Us2.Ca score was the most important parameter, followed by RWT, abnormal serum/urine immunofixation, gender, and LV SAB. Ultimately, the final AI-ECM model included all features listed in Table 2: the Us2.Ca model, RWT, LV SAB, demographics, light chain biomarkers, and renal function biomarkers. Within the internal validation data set, the AUC and accuracy of the AI-ECM model were 0.93 (95% CI, 0.89–0.97) and 85.1%, respectively (Table S4 and Figure S1).

## AI-ECM Performance

The AI-ECM machine learning model demonstrated high performance for CA detection in the internal test data set, with AUC, accuracy, sensitivity, and specificity of 0.94 (95% CI, 0.90–0.98), 90.1%, 92.5%, and 84.8%, respectively (Figure 2; Table 3). Reduction in the number of clinical, laboratory, and TTE parameters used resulted in numerically lower performance compared with the optimal AI-ECM model (Table 3; Figure S2). Compared with Us2.Ca, the fully integrated AI-ECM achieved a positive net reclassification improvement (categorical net reclassification improvement, +0.18), indicating a modest net increase in correct reclassification of cases and controls at the prespecified operating threshold. A subanalysis using demographics, biomarkers, and TTE parameters (including RWT and SAB) without the Us2.Ca score demonstrated lower diagnostic accuracy compared with the full AI-ECM (Table S5).

The AI-ECM demonstrated similarly high performance across both CA subtypes in the internal test data set (Tables S6 and S7). This included both relevant subgroups of ATTR-CA versus controls suspected to have ATTR-CA with negative technetium pyrophosphate (Table S6) and AL-CA versus controls with proven AL amyloidosis without cardiac involvement (Table S7). Patients with CA and unknown subtypes were included in the overall test data set analysis but excluded from subtype-specific analyses, accounting for the discrepancy between total CA cases and the sum of subtype subgroups.

Stratified subtype analyses demonstrated differences in model performance between ATTR-CA and AL-CA (Tables S6 and S7). For ATTR-CA, the echocardiography-based AI model (Us2.Ca) achieved strong discriminatory performance, with incremental improvement observed with the multimodal AI-ECM model. In contrast, in AL-CA, the addition of clinical and laboratory variables resulted in greater gains in sensitivity compared with the echocardiography-only model. These subtype-specific performance patterns resulted in overall improvements in pooled model performance for the multimodal model.

## Contribution of Renal Function to AI-ECM Performance

Renal function was similar across all groups: median eGFR was 56.0 (IQR, 40.0–68.9) mL/min per 1.73 $m^2$ in ATTR-CA, 59.2 (IQR, 34.8–80.0) in AL-CA, and 59.0 (IQR, 34.2–80.0) in controls, with no significant pairwise differences (all $P$>0.08). Creatinine was comparable between patients with CA and controls (ATTR-CA versus controls, $P$=0.052; AL-CA versus controls, $P$=0.56), though a modest difference was observed between CA subtypes (ATTR-CA versus AL-CA, $P$=0.010).

In ablation analyses, removal of renal function markers had minimal impact on AI-ECM performance. The model excluding both creatinine and eGFR achieved an AUC of 0.94, accuracy of 89.5%, sensitivity of 91.5%, and specificity of 84.8%, compared with the full AI-ECM (AUC, 0.94; accuracy, 90.1%; sensitivity, 92.5%; specificity, 84.8%). Models incorporating creatinine alone (AUC, 0.94; accuracy, 85.5%; sensitivity, 85.8%) or eGFR alone (AUC, 0.94; accuracy, 88.8%; sensitivity, 90.6%) yielded identical AUCs, confirming that neither renal marker provides meaningful incremental discriminatory value. Although eGFR ranked among the top features by gain-based importance during model training (Table 2), it did not materially affect performance, suggesting that XGBoost redistributes predictive signal across features when individual variables are excluded.

In patients with preserved renal function (eGFR ≥60 mL/min per 1.73 $m^2$; n=63, 41.4% of the test set), the AI-ECM maintained high performance (AUC, 0.94, accuracy 92.1%, sensitivity 93.8%, specificity 86.7%). Taken together, these findings demonstrate that the AI-ECM does not rely on renal impairment as a proxy for illness severity.

## Missing Data Analysis

In the complete-case sensitivity analysis, restricted to the 79 patients (52.0% of the test set) with nonmissing values across all AI-ECM input features, AI-ECM

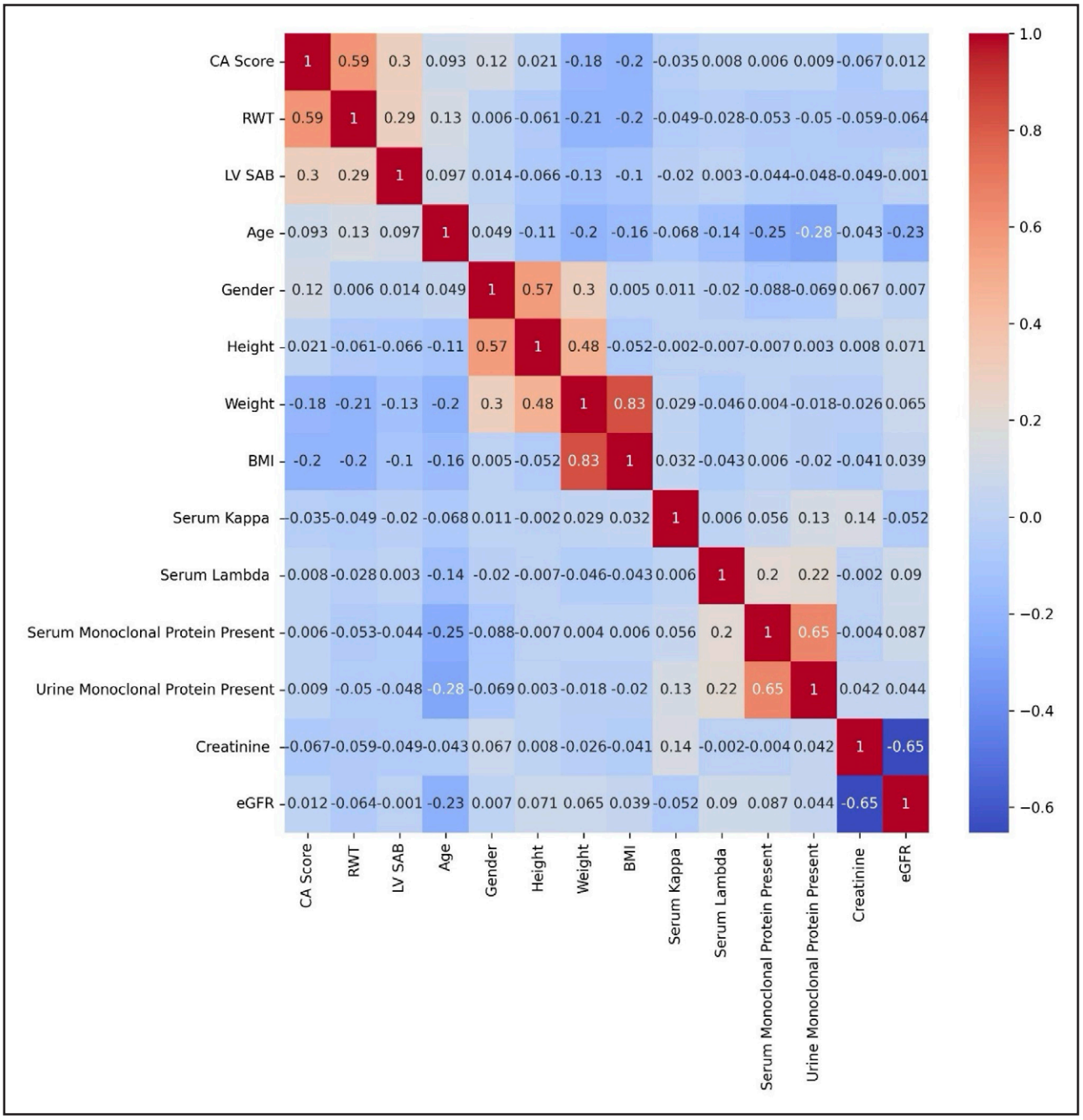


**Figure 1. Demonstrating Spearman correlation between parameters demonstrating minimal collinearity.**
Of note, the XGBoost model is robust to collinearity but is depicted to show the relationship between different variables. BMI indicates body mass index; eGFR, estimated glomerular filtration rate; RWT, relative wall thickness; and SAB, septal apical basal.

performance was preserved (AUC, 0.97 versus 0.94 in the full test set), arguing against the possibility that differential missingness artificially inflated model performance. The numerically higher performance in the complete-case subset likely reflects the higher overall data quality in patients with complete clinical, laboratory, and echocardiographic assessments. The reduced sample size of the complete-case subset limits statistical power for formal comparison, but the directional finding that performance improves rather than degrades when missingness is eliminated supports the conclusion that the AI-ECM's discriminatory ability is driven by measured feature values rather than patterns of data absence.

To evaluate whether missingness patterns themselves contributed to model predictions, an auxiliary XGBoost model trained exclusively on binary missingness indicators achieved an AUC of 0.50, indicating that data availability patterns were not predictive of CA status. This confirms that the AI-ECM's discriminatory performance

**Table 2. Feature Importance Rankings for Cardiac Amyloidosis Prediction Derived From the XGBoost Classifier. Importance Is Quantified Using the Gain Metric, Which Measures the Average Improvement in Model Performance Attributable to Each Feature Across All Decision Tree Splits During Training. Higher Values Indicate Greater Contribution to the Model's Discriminative Ability**

| Feature | Importance |
|---|---|
| Us2.Ca score | 12.799 |
| RWT | 3.181 |
| Abnormal urine monoclonal protein present (immunofixation) | 2.31 |
| Gender | 2.309 |
| Abnormal serum monoclonal protein present (immunofixation) | 1.912 |
| LV apex-to-basal strain ratio | 1.776 |
| eGFR | 1.631 |
| Serum kappa | 1.345 |
| Weight | 1.341 |
| Serum lambda | 1.318 |
| Age | 1.275 |
| Height | 1.166 |
| BMI | 1.159 |
| Creatinine | 1.15 |

BMI indicates body mass index; eGFR, estimated glomerular filtration rate; LV, left ventricular; and RWT, relative wall thickness.



is driven entirely by measured feature values rather than patterns of data absence.

### Us2.Ca Model and IWT Score Performance

The performance of the Us2.Ca model and IWT scores in the test data set in our cohort is displayed in Table 3 and Figure 2. The Us2.Ca model demonstrated high accuracy, although it was limited by 9% of patients being classified as indeterminate and therefore could not be classified as CA/no CA. The Us2.Ca model performance was similar across both CA subtypes except for numerically lower accuracy and sensitivity in AL-CA of 74.5% and 67.5%, respectively, when allowing for indeterminate Us2.Ca scores (Tables S6 and S7). The performance of the Us2.Ca model and IWT scores is depicted in Table S6.

### Comparison of AI-ECM, Us2.Ca, and IWT Score Performance

Within our internal test data set, the AI-ECM accuracy was significantly higher than both the Us2.Ca ($P$=0.006) and IWT scores ($P$≤0.001, Figure 1; Table 3). Notably, the AI-ECM—which does not use indeterminate predictions—outperformed the TTE-only Us2.Ca model even while the latter allowed for indeterminate AI predictions.

Nonparametric bootstrap simulation across the clinically relevant prevalence spectrum demonstrated that the AI-ECM provided consistently superior NPV compared with the TTE-only Us2.Ca model. However, this gain in sensitivity and rule-out capability was associated with a reduction in PPV (Figure 3). At a simulated disease prevalence of 15%, the AI-ECM yielded a higher NPV (98.4% versus 95.5%) but a slightly lower PPV (53.1% versus 61.0%) compared with the Us2.Ca score. These findings suggest that while both models are robust, the AI-ECM integration of clinical and biochemical data optimizes the detection of CA, making it particularly effective as a high-sensitivity screening tool in populations with a moderate-to-high clinical suspicion of amyloidosis.

The AI-ECM numerically outperformed the Us2.Ca score and IWT scores in subgroup analyses of ATTR-CA versus controls with suspected ATTR-CA but negative technetium pyrophosphate/light chains. The AI-ECM similarly outperformed the Us2.Ca score and the systemic AL score in the subgroup of AL-CA versus controls with extracardiac CA without cardiac involvement.

## DISCUSSION

In this study, we developed and validated a fully automated, multimodal, machine learning-based AI-ECM for diagnosis of CA that includes echocardiographic, clinical, and laboratory data. In addition, we proved that by doing a more comprehensive evaluation of the patient, model performance is significantly improved as compared with the TTE-only Us2.Ca model and other multiparametric TTE-based scores currently used in clinical practice, without need for indeterminate values.

To our knowledge, this is the first study to fuse deep learning-based AI models for TTE with multimodal clinical, laboratory, and TTE biomarker data to augment CA detection. Previous models integrating data across these domains have largely been derived from parameters without integration of raw TTE images.[23,24] Our work aligns with a growing body of evidence supporting the use of TTE-based AI models for CA detection and further expands upon this area by demonstrating the potential for integration of multimodal clinical, laboratory, and TTE biomarkers to augment AI model accuracy.[7]

In particular, TTE parameters—including RWT and SAB—light chains, renal function, and gender appeared to provide the most incremental value. Although increased RWT and apical sparing in GLS—as assessed by SAB—are well-established echocardiographic markers of CA, their explicit inclusion in the fusion model provided only modest incremental improvement over the TTE-only Us2.Ca model.[6,21,25] This likely reflects the fact that the deep learning-based Us2.Ca implicitly encodes myocardial geometric and deformation patterns from raw echocardiographic videos. Nevertheless, because RWT and SAB are rapidly and automatically quantified, their inclusion

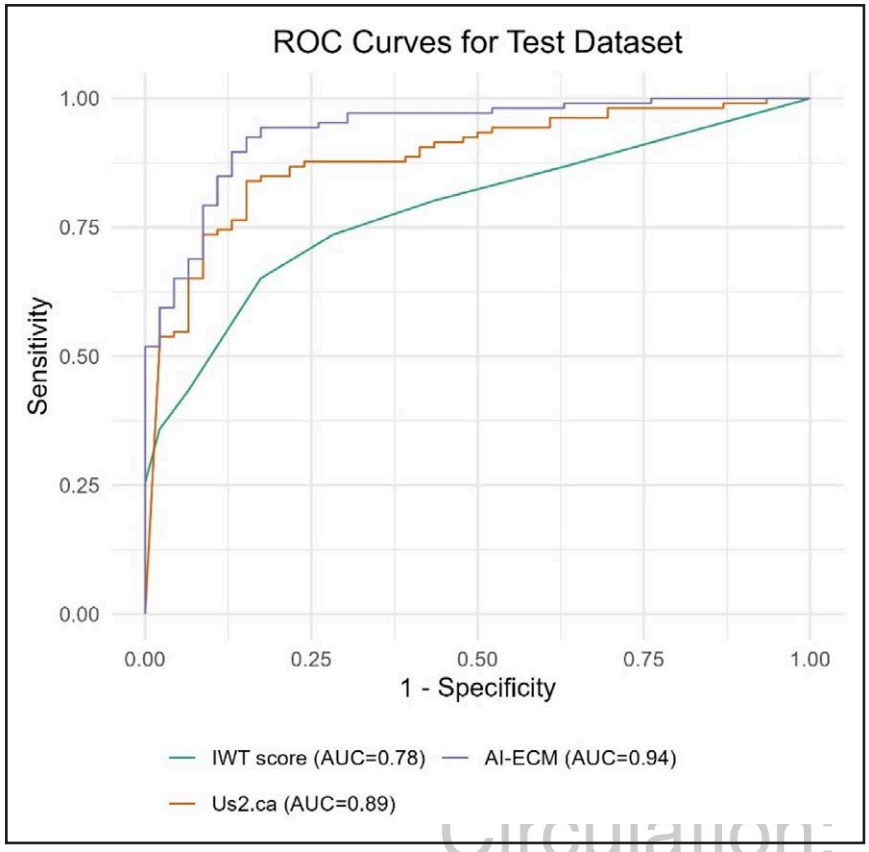


**Figure 2. Receiver operator curve (ROC) indicating performance of the increased wall thickness (IWT), Us2.Ca, and artificial intelligence echo-clinical models (AI-ECM) for cardiac amyloidosis (CA) detection.**



adds minimal complexity while providing complementary and interpretable diagnostic information within the multimodal AI-ECM framework.

In addition, light chain biomarkers are present in nearly all cases of AL-CA and therefore likely augment diagnosis in this subset.[26] Importantly, clinical and laboratory biomarkers demonstrated minimal collinearity with the TTE-only AI model, further supporting their additive value. Importantly, as with the TTE-only AI model, RWT and SAB were quantified automatically using FDA-approved AI software, further supporting their ease of use. While the addition of multiple parameters made the models more complex, the use of fully automated TTE measurements nature improves the simplicity, efficiency, and reproducibility of the model.

Importantly, the data elements incorporated into the AI-ECM are consistent with routine components of the early diagnostic evaluation for CA. In patients with echocardiographic findings suspicious for CA, current guidelines recommend obtaining serum and urine light chain testing before pursuing nuclear scintigraphy or endomyocardial biopsy.[19] Demographic data and renal function are typically available in all patients at the time of TTE. As such, the AI-ECM is designed to operate early within the diagnostic evaluation. Demographics, biomarkers, and conventional TTE parameters alone were

**Table 3. Comparing Performance of IWT Score, Cardiac Amyloidosis Score, and All Fused Models in the Test Cohort**

| | Accuracy (%) | Sensitivity, % (95% CI) | Specificity, % (95% CI) | AUC (95% CI) | TP | FN | TN | FP | Total N | Yield (%) | Brier score | NRI* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IWT score | 54.6 | 35.8 (27.4–45.3) | 97.8 (88.7–99.6) | 0.78 (0.71–0.86) | 38 | 68 | 45 | 1 | 152 | 100.0 | 0.30 | −0.26 |
| Us2.Ca | 80.4 | 75.8 (66.3–83.3) | 90.7 (78.4–96.3) | 0.89 (0.84–0.95) | 72 | 23 | 39 | 4 | 138 | 90.8 | 0.16 | … |
| Us2.Ca+RWT | 83.6 | 83.0 (74.7–89.0) | 84.8 (71.8–92.4) | 0.90 (0.85–0.95) | 88 | 18 | 39 | 7 | 152 | 100.0 | 0.13 | +0.09 |
| Us2.Ca+RWT+SAB | 81.6 | 80.2 (71.6–86.7) | 84.8 (71.8–92.4) | 0.91 (0.86–0.95) | 85 | 21 | 39 | 7 | 152 | 100.0 | 0.13 | +0.06 |
| Us2.Ca+RWT+ SAB+Clin | 82.9 | 82.1 (73.7–88.2) | 84.8 (71.8–92.4) | 0.92 (0.87–0.96) | 87 | 19 | 39 | 7 | 152 | 100.0 | 0.11 | +0.08 |
| Us2.Ca+RWT+SAB+ Clin+Lab | 89.5 | 91.5 (84.6–95.5) | 84.8 (71.8–92.4) | 0.94 (0.90–0.97) | 97 | 9 | 39 | 7 | 152 | 100.0 | 0.10 | +0.17 |
| AI-ECM | 90.1 | 92.5 (85.8–96.1) | 84.8 (71.8–92.4) | 0.94 (0.90–0.98) | 98 | 8 | 39 | 7 | 152 | 100.0 | 0.10 | +0.18 |

AI-ECM indicates artificial intelligence echo-clinical model; Clin, clinical; IWT, increased wall thickness; NRI, net reclassification ratio; RWT, relative wall thickness; and SAB, septal apical basal ratio.

*Using Us2.Ca as a benchmark.

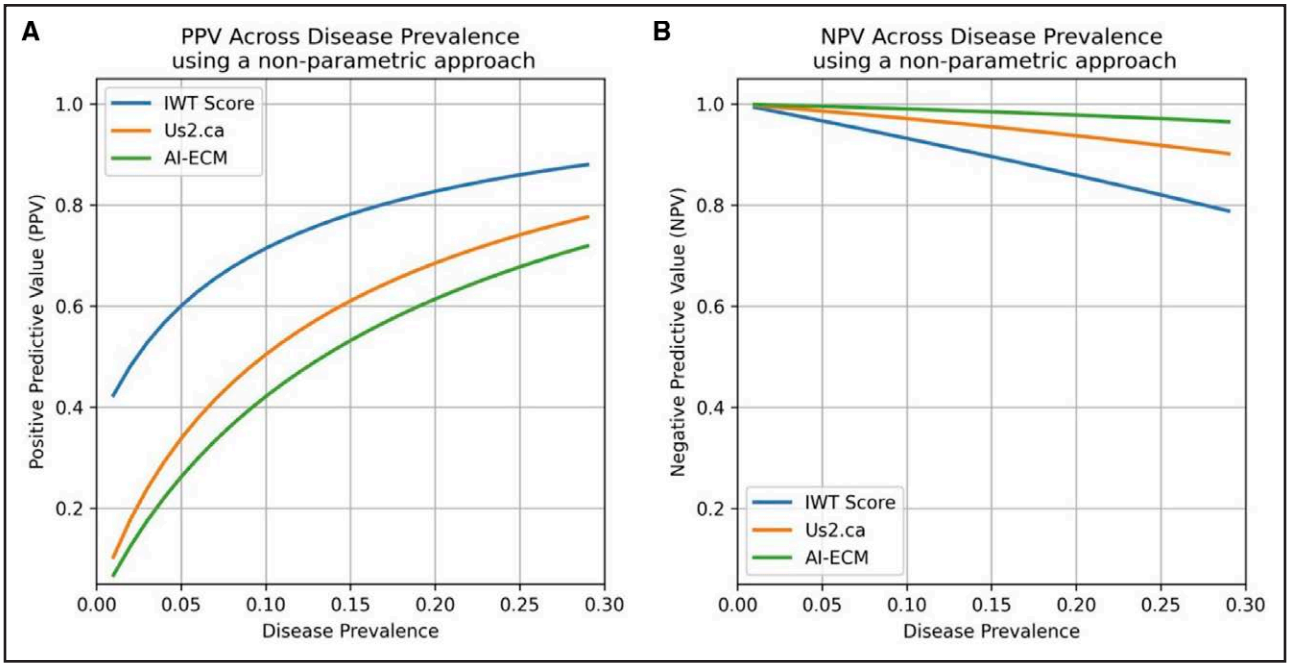


**Figure 3. Simulated positive predictive value (PPV) and negative predictive value (NPV) across a range of hypothetical cardiac amyloidosis prevalence for each model.**

Values on both axes are expressed as proportions (eg, 0.3=30%). AI-ECM indicates artificial intelligence echo clinical model; and IWT, increased wall thickness.



insufficient to achieve the performance of the full AI-ECM, highlighting the complementary value of the deep learning-derived Us2.Ca score.

The addition of clinical, laboratory, and imaging biomarkers had the greatest effect on augmenting sensitivity. This is critically important given the frontline screening role of TTE in clinical practice. At the time of referral for TTE, patients often have undifferentiated symptoms, and clinical suspicion for CA may not yet be high. Failure to detect CA with TTE may therefore cause significant delays in diagnosis and timely treatment, with potential prognostic implications. In our cohort, the augmentation of screening efficacy was particularly valuable for AL-CA, where AI model sensitivity was lowest. Not only did the incorporation of clinical and laboratory biomarkers improve accuracy, but it did so without the need for indeterminate AI model values. Given the significant harms of false predictions, indeterminate values are often used in TTE models—including AI-based—to overcome accuracy limitations. Our AI-ECM model has the potential to overcome these limitations by augmenting accuracy while eliminating the need for indeterminate values, potentially facilitating deployment in a greater number of patients.

Subtype analyses suggest that different components of the model contribute variably to detection of ATTR-CA and AL-CA. The echocardiographic AI component appears to provide strong discriminatory signal for ATTR-CA, whereas clinical and laboratory variables, including renal function and biomarkers relevant to plasma cell dyscrasia, contribute more prominently to AL-CA classification. These findings highlight the complementary roles of imaging-derived phenotypes and systemic biomarkers in amyloid detection and support the rationale for multimodal fusion approaches. Future work could explore subtype-specific model architectures, though such approaches would need to be weighed against increased complexity and the clinical reality that subtype is typically unknown at the time of initial echocardiographic screening.

Despite the transformative potential of existing AI models to rapidly and accurately detect underdiagnosed diseases such as CA, a limitation has largely been the use of only single-source data (eg, electronic health record, ECG, TTE). The integration of multimodal clinical, laboratory, and imaging data more closely aligns with clinical decision-making in which a wide array of heterogeneous data is integrated into generating a diagnosis. Although human decision-making is slower and less consistent than AI, it stands apart in its remarkable ability to synthesize complex, nuanced information, offering insights that set the benchmark for clinical intuition. In this way, multimodal data fusion may offer the best of both worlds, pairing the speed and precision of AI with the nuanced, integrative reasoning that has long defined expert human clinical judgment.

It should be noted that this study used only a limited number of additional clinical, laboratory, and TTE biomarkers. This was necessitated by the large size and multicenter nature of the registry, which did not?

allow for extensive data collection. Further studies are needed to better understand the incremental value of expanded multimodal biomarkers—perhaps including CA red flag signs and symptoms—for augmenting Us2.Ca performance. It is even conceivable that future studies could integrate multiple AI models spanning different domains—including electronic health record, ECG, and TTE—to create a screening pipeline to streamline CA detection.[27] This concept aligns with previous work by Goto et al,[7] which demonstrated promising potential for combining ECG- and TTE-based AI to boost accuracy for differentiating CA from other causes of left ventricular hypertrophy. Obviously, these potential advances must be balanced against the added complexity of incorporating increasing amounts of data.

## Limitations

Our study has several limitations. The model lacks external validation, although we attempted to address this with the use of both diverse, multiethnic, multivendor data sets and cross-validation techniques. In our cohort, comprehensive light chain testing was available in the majority of patients; however, in routine clinical practice, pyrophosphate imaging is sometimes performed before complete serum and urine light chain evaluation. As such, the high rate of light chain testing in our data set may not fully reflect real-world practice. In addition, although the initial TTE-only AI model was developed in a cohort where the prevalence of CA reflected clinical practice, this study was performed in an enriched, referred population with a high prevalence of CA. Although prevalence bootstrapping was performed to estimate performance across different disease prevalences, this approach does not address potential spectrum bias, and model performance in undifferentiated echocardiography populations will require prospective external validation in populations whose prevalence of CA more closely reflects real-world screening. In addition, given the heterogeneity in assays used, we were unable to incorporate cardiac biomarkers such as NT-proBNP or troponin into the model, which may have incrementally improved performance; further evaluation of these markers in more homogeneous data sets could be considered in future studies. Finally, given that the cohort enrolled patients across a wide date range, the severity of disease may have been higher. It is therefore unclear how well the AI-ECM would perform in early-stage disease.

## Conclusions

Our machine learning-based AI-ECM model combining clinical, laboratory, and TTE parameters with an Us2.Ca provided increased accuracy while eliminating the need for indeterminate values. The integration of multimodal data sources into existing CA screening tools has the potential to augment CA diagnosis by better identifying patients who would benefit most from confirmatory testing, advancing toward real-world AI-guided precision diagnostics in CA.

## ARTICLE INFORMATION

Received January 19, 2026; accepted May 6, 2026.

### Affiliations

University of Chicago, IL (J.A.S., N.S., K.A., R.M.L.). Us2.AI, Singapore (S.C.L., M. Frost, C.L.S.P.). Hospital of the University of Pennsylvania, Philadelphia (M.R., M.S.-C., A.V.). Columbia University Irving Medical Center, New York, NY (M.S.M., S.H.). The Ohio State University Wexner Medical Center, Columbus (K.M.Z., A.G.). University of Washington, Seattle (R.C., N.W.). Hospital of University of Occupational and Environmental Health, Japan (T.K., M.T.). Heart Institute (InCor), Sao Paolo, Brazil (V.T.H., M.L.C.V.). ICBA, Buenos Aires, Argentina (P.E., R.E.R.). Centro Privado de Cardiología, Tucuman, Argentina (A. Prado). University of Texas MD Anderson Cancer Center, Houston (E.K., A.D.). Endeavor NorthShore, Evanston, IL (A. Pursnani). Cardiólogo en Altos de Salta Swiss Medical Group, Tucuman, Argentina (J.C.). Boston University Chobanian and Avedisian School of Medicine, MA (F.L.R.). National Amyloidosis Center, London, United Kingdom (M. Fontana). MedStar Health Research Institute, Washington, DC (F.M.A.).

### Acknowledgments

The authors like to thank Christine Gouillard and Glenda Chin for their contributions to executing the presented work. Additional Information: Coauthor Roberto M. Lang, MD, died June 10, 2025.

### Sources of Funding

None.

### Disclosures

Dr Slivnick has received personal fees from Alnylam, BridgeBio, GE, and Pfizer. S.C. Lim and M. Frost are employees of Us2.ai. Dr Maurer has received grant support from NIH grants R01HL139671 and AG081582; has received institutional research funding from Alnylam Pharmaceuticals, Attralus, BridgeBio, Intellia, and Ionis; and has received personal fees from Akcea, Alnylam Pharmaceuticals, Bayer, AstraZeneca, Attralus, Intellia, and Novo Nordisk. Dr Scherrer-Crosbie receives grant support from the American Society of Echocardiography and institutional funding from Atman Health. Anita Deswal is a consultant for Bayer. Dr Lam Su Ping is supported by a Clinician Scientist Award from the National Medical Research Council of Singapore; has received research support from Novo Nordisk and Roche Diagnostics; has served as a consultant or on the Advisory Board/Steering Committee/Executive Committee for Alnylam Pharma, AnaCardio AB, Applied Therapeutics, AstraZeneca, Bayer, Biopeutics, Boehringer Ingelheim, Boston Scientific, Bristol Myers Squibb, Corteria, CPC Clinical Research, Cytokinetics, Eli Lilly, Impulse Dynamics, Intellia Therapeutics, Janssen Research & Development LLC, Medscape/WebMD Global LLC, Merck, Novartis, Novo Nordisk, Quidel Corporation, Radcliffe Group, Ltd, Roche, and Us2.ai; serves as cofounder and nonexecutive director of Us2.ai; and is an owner of US patent no 10 702 247 and has a patent pending (PCT/SG2016/050217). Dr Lang is on the advisory board and speaker's bureau for Philips. Dr Asch receives institutional (MedStar Health) research grants from Us2.ai, TOMTEC/Philips, Ultromics, GE, MyCardium, Caption Health, Egnite, Abbott, Edwards, BSC, Medtronic, Neovasc, Ancora Heart, InnovHeart, Polares Medical, Corflow, Aria CV, Laminar/Johnson & Johnson, Corcym, Xeltis, Tricares, VDyne, Croivalve, BMS, Sanofi, and Novartis. The other authors report no conflicts.

### Supplemental Material

Tables S1–S7
Figures S1–S2

## REFERENCES

1. Fontana M, Berk JL, Gillmore JD, Witteles RM, Grogan M, Drachman B, Damy T, Garcia-Pavia P, Taubel J, Solomon SD, et al; HELIOS-B Trial Investigators. Vutrisiran in patients with transthyretin amyloidosis with cardiomyopathy. *N Engl J Med.* 2025;392:33–44. doi: 10.1056/NEJMoa2409134

2. Maurer MS, Schwartz JH, Gundapaneni B, Elliott PM, Merlini G, Waddington-Cruz M, Kristen AV, Grogan M, Witteles R, Damy T, et al; ATTR-ACT Study Investigators. Tafamidis treatment for patients with transthyretin amyloid cardiomyopathy. *N Engl J Med*. 2018;379:1007–1016. doi: 10.1056/NEJMoa1805689
3. Gillmore JD, Judge DP, Cappelli F, Fontana M, Garcia-Pavia P, Gibbs S, Grogan M, Hanna M, Hoffman J, Masri A, et al; ATTRibute-CM Investigators. Efficacy and safety of acoramidis in transthyretin amyloid cardiomyopathy. *N Engl J Med*. 2024;390:132–142. doi: 10.1056/NEJMoa2305434
4. Kastritis E, Palladini G, Minnema MC, Wechalekar AD, Jaccard A, Lee HC, Sanchorawala V, Gibbs S, Mollee P, Venner CP, et al; ANDROMEDA Trial Investigators. Daratumumab-based treatment for immunoglobulin light-chain amyloidosis. *N Engl J Med*. 2021;385:46–58. doi: 10.1056/NEJMoa2028631
5. Vogel J, Jura S, Settelmeier S, Buehning F, Lerchner T, Carpinteiro A, Rassaf T, Michel L. Delays in diagnosis and treatment of ATTR cardiac amyloidosis: a real-world data analysis. *ESC Heart Fail*. 2025;12:2969–2975. doi: 10.1002/ehf2.15311
6. Cotella J, Randazzo M, Maurer MS, Helmke S, Scherrer-Crosbie M, Soltani M, Goyal A, Zareba K, Cheng R, Kirkpatrick JN, et al. Limitations of apical sparing pattern in cardiac amyloidosis: a multicentre echocardiographic study. *Eur Heart J Cardiovasc Imaging*. 2024;25:754–761. doi: 10.1093/ehjci/jeae021
7. Goto S, Mahara K, Beussink-Nelson L, Ikura H, Katsumata Y, Endo J, Gaggin HK, Shah SJ, Itabashi Y, MacRae CA, et al. Artificial intelligence-enabled fully automated detection of cardiac amyloidosis using electrocardiograms and echocardiograms. *Nat Commun*. 2021;12:2726. doi: 10.1038/s41467-021-22877-8
8. Duffy G, Cheng PP, Yuan N, He B, Kwan AC, Shun-Shin MJ, Alexander KM, Ebinger J, Lungren MP, Rader F, et al. High-throughput precision phenotyping of left ventricular hypertrophy with cardiovascular deep learning. *JAMA Cardiol*. 2022;7:386–395. doi: 10.1001/jamacardio.2021.6059
9. Slivnick JA, Hawkes W, Oliveira J, Woodward G, Akerman A, Gomez A, Hamza I, Desai VK, Barrett-O'Keefe Z, Grogan M, et al. Cardiac amyloidosis detection from a single echocardiographic video clip: a novel artificial intelligence-based screening tool. *Eur Heart J*. 2025;46:4090–4101. doi: 10.1093/eurheartj/ehaf387
10. Grogan M, Lopez-Jimenez F, Cohen-Shelly M, Dispenzieri A, Attia ZI, Abou Ezzedine OF, Lin G, Kapa S, Borgeson DD, Friedman PA, et al. Artificial intelligence-enhanced electrocardiogram for the early detection of cardiac amyloidosis. *Mayo Clin Proc*. 2021;96:2768–2778. doi: 10.1016/j.mayocp.2021.04.023
11. Haimovich JS, Diamant N, Khurshid S, Di Achille P, Reeder C, Friedman S, Singh P, Spurlock W, Ellinor PT, Philippakis A, et al. Artificial intelligence-enabled classification of hypertrophic heart diseases using electrocardiograms. *Cardiovasc Digit Health J*. 2023;4:48–59. doi: 10.1016/j.cvdhj.2023.03.001
12. Huda A, Castaño A, Niyogi A, Schumacher J, Stewart M, Bruno M, Hu M, Ahmad FS, Deo RC, Shah SJ. A machine learning model for identifying patients at risk for wild-type transthyretin amyloid cardiomyopathy. *Nat Commun*. 2021;12:2725. doi: 10.1038/s41467-021-22876-9
13. Agibetov A, Seirer B, Dachs TM, Koschutnik M, Dalos D, Rettl R, Duca F, Schrutka L, Agis H, Kain R, et al. Machine learning enables prediction of cardiac amyloidosis by routine laboratory parameters: a proof-of-concept study. *J Clin Med*. 2020;9:1334. doi: 10.3390/jcm9051334
14. Delbarre MA, Girardon F, Roquette L, Blanc-Durand P, Hubaut MA, Hachulla E, Semah F, Huglo D, Garcelon N, Marchal E, et al. Deep learning on bone scintigraphy to detect abnormal cardiac uptake at risk of cardiac amyloidosis. *JACC Cardiovasc Imaging*. 2023;16:1085–1095. doi: 10.1016/j.jcmg.2023.01.014
15. Spielvogel CP, Haberl D, Mascherbauer K, Ning J, Kluge K, Traub-Weidinger T, Davies RH, Pierce I, Patel K, Nakuz T, et al. Diagnosis and prognosis of abnormal cardiac scintigraphy uptake suggestive of cardiac amyloidosis using artificial intelligence: a retrospective, international, multicentre, cross-tracer development and validation study. *Lancet Digit Health*. 2024;6:e251–e260. doi: 10.1016/S2589-7500(23)00265-0
16. U.S. Food and Drug Administration, Center for Drug Evaluation and Research. *K250151 US2.Ca approval letter*. 2025;chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.accessdata.fda.gov/cdrh_docs/pdf25/K250151.pdf
17. Ioannou A, Khouri MG, Kitai T, Vemulapalli S, Hung C-L, Lim SC, Frost M, Tee WW, Mansell J, Sheikh A, et al. Diagnosis of cardiac amyloidosis on echocardiography using artificial intelligence. *Circ Cardiovasc Imaging*. 2026;19:e018991. doi: 10.1161/CIRCIMAGING.125.018991
18. Gertz MA, Comenzo R, Falk RH, Fermand JP, Hazenberg BP, Hawkins PN, Merlini G, Moreau P, Ronco P, Sanchorawala V, et al. Definition of organ involvement and treatment response in immunoglobulin light chain Amyloidosis (AL): a consensus opinion from the 10th International Symposium on Amyloid and Amyloidosis, Tours, France, 18-22 April 2004. *Am J Hematol*. 2005;79:319–328. doi: 10.1002/ajh.20381
19. Dorbala S, Ando Y, Bokhari S, Dispenzieri A, Falk RH, Ferrari VA, Fontana M, Gheysens O, Gillmore JD, Glaudemans A, et al. ASNC/AHA/ASE/EANM/HFSA/ISA/SCMR/SNMMI expert consensus recommendations for multimodality imaging in cardiac amyloidosis: part 1 of 2-evidence base and standardized methods of imaging. *Circ Cardiovasc Imaging*. 2021;14:e000029. doi: 10.1161/hci.0000000000000029
20. Grogan M, Dispenzieri A, Gertz MA. Light-chain cardiac amyloidosis: strategies to promote early diagnosis and cardiac response. *Heart*. 2017;103:1065–1072. doi: 10.1136/heartjnl-2016-310704
21. Boldrini M, Cappelli F, Chacko L, Restrepo-Cordoba MA, Lopez-Sainz A, Giannoni A, Aimo A, Baggiano A, Martinez-Naharro A, Whelan C, et al. Multiparametric echocardiography scores for the diagnosis of cardiac amyloidosis. *JACC Cardiovasc Imaging*. 2020;13:909–920. doi: 10.1016/j.jcmg.2019.10.011
22. Chen T, Guestrin C. *XGBoost: a scalable tree boosting system*. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016;785–794. doi: 10.1145/2939672.2939785
23. Arana-Achaga X, Goena-Vives C, Villanueva-Benito I, Solla-Ruiz I, Rengel Jimenez A, Gaspar TI, Urreta-Barallobre I, Barge-Caballero G, Seijas-Marcos S, Cabrera E, et al. Development and validation of a prediction model and score for transthyretin cardiac amyloidosis diagnosis: T-amylo. *JACC Cardiovasc Imaging*. 2023;16:1567–1580. doi: 10.1016/j.jcmg.2023.05.002
24. Davies DR, Redfield MM, Scott CG, Minamisawa M, Grogan M, Dispenzieri A, Chareonthaitawee P, Shah AM, Shah SJ, Wehbe RM, et al. A simple score to identify increased risk of transthyretin amyloid cardiomyopathy in heart failure with preserved ejection fraction. *JAMA Cardiol*. 2022;7:1036–1044. doi: 10.1001/jamacardio.2022.1781
25. Pagourelias ED, Mirea O, Duchenne J, Van Cleemput J, Delforge M, Bogaert J, Kuznetsova T, Voigt JU. Echo parameters for differential diagnosis in cardiac amyloidosis: a head-to-head comparison of deformation and nondeformation parameters. *Circ Cardiovasc Imaging*. 2017;10:e005588. doi: 10.1161/CIRCIMAGING.116.005588
26. Palladini G, Russo P, Bosoni T, Verga L, Sarais G, Lavatelli F, Nuvolone M, Obici L, Casarini S, Donadei S, et al. Identification of amyloidogenic light chains requires the combination of serum-free light chain assay with immunofixation of serum and urine. *Clin Chem*. 2009;55:499–504. doi: 10.1373/clinchem.2008.117143
27. Wardak R, Snipelisky D. The role of artificial intelligence in cardiac amyloidosis: a focus on diagnosis and clinical application. *J Cardiovasc Dev Dis*. 2025;12:221. doi: 10.3390/jcdd12060221