

ORIGINAL RESEARCH

HEART FAILURE AND CARDIOMYOPATHIES

International Validation of Echocardiographic Artificial Intelligence Amyloid Detection Algorithm



Grant Duffy, BS,^a Evangelos K. Oikonomou, MD, DPHIL,^{b,c} Nicholas Easton, BS,^d Hiroki Usuku, MD, PhD,^e Jigesh Patel, MD,^a Yoshinori Katsumata, MD, PhD,^f Daisuke Yamasawa, MD, PhD,^f Lily Stern, MD,^a Shinichi Goto, MD, PhD,^{e,f,g,h} Kenichi Tsujita, MD, PhD,^e Paul Cheng, MD, PhD,ⁱ Rohan Khera, MD, MS,^{b,c} Faraz S. Ahmad, MD,^d David Ouyang, MD^a

ABSTRACT

BACKGROUND Diagnosis of cardiac amyloidosis (CA) is often missed or delayed due to confusion with other causes of increased left ventricular wall thickness. Conventional transthoracic echocardiographic measurements like global longitudinal strain have shown promise in distinguishing CA, but with limited specificity.

OBJECTIVES We conducted a multisite retrospective case-control study to investigate the performance of a computer vision algorithm for CA identification across multiple international sites.

METHODS EchoNet-left ventricular hypertrophy (LVH) is a computer vision deep learning algorithm for the detection of CA based on parasternal long axis and apical-4-chamber view videos. We evaluated EchoNet-LVH's ability to distinguish between the echocardiogram studies of 574 CA patients and 979 controls. We reported discrimination performance with an area under the receiver operating characteristic curve and associated sensitivity, specificity, and positive predictive value at the prespecified threshold.

RESULTS EchoNet-LVH had an area under the receiver operating characteristic curve of 0.896 (95% CI: 0.875-0.916). At the prespecified model threshold optimizing for specificity, EchoNet-LVH had a sensitivity of 0.644 (95% CI: 0.601-0.685), specificity of 0.988 (95% CI: 0.978-0.994), positive predictive value of 0.968 (95% CI: 0.944-0.984), and negative predictive value of 0.828 (95% CI: 0.804-0.850). There was no evidence of heterogeneity in performance by site, race, sex, age, body mass index, CA subtype, or ultrasound manufacturer.

CONCLUSIONS EchoNet-LVH can assist with earlier and accurate diagnosis of CA. EchoNet-LVH achieved development goals to be highly specific to maximize positive predictive value of downstream confirmatory testing since CA is a rare disease. (JACC Adv. 2025;4:102067) © 2025 The Authors. Published by Elsevier on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

From the ^aSmidt Heart Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA; ^bSection of Cardiovascular Medicine, Department of Internal Medicine, Yale School of Medicine, New Haven, Connecticut, USA; ^cCardiovascular Data Science (CarDS) Lab, Yale School of Medicine, New Haven, Connecticut, USA; ^dDepartment of Medicine, Northwestern Medicine, Chicago, Illinois, USA; ^eDepartment of Cardiovascular Medicine, Graduate School of Medical Sciences, Kumamoto University, Chuo-ku, Kumamoto, Japan; ^fDepartment of Cardiology, Keio University School of Medicine, Shinjuku, Tokyo, Japan; ^gDepartment of Medicine, Tokai University School of Medicine, Isehara, Kanagawa, Japan; ^hOne Brave Idea, Division of Cardiovascular Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA; and the ⁱDivision of Cardiology, Department of Medicine, Stanford University, Palo Alto, California, USA.

**ABBREVIATIONS
AND ACRONYMS****AI** = artificial intelligence**AL** = light chain amyloidosis**ATTR** = transthyretin
amyloidosis**AUC** = area under the receiver
operating characteristic curve**CA** = cardiac amyloidosis**LVH** = left ventricular
hypertrophy**PPV** = positive predictive value

Cardiac amyloidosis (CA) is caused by deposition of misfolded proteins in the myocardium, including transthyretin or immunoglobulin light chains.^{1,2} Regardless of the etiology, CA leads to increased left ventricular wall thickness and heart failure; however, early symptoms can be nonspecific and not readily recognized.³⁻⁵ Common echocardiographic measurements are insufficient to precisely discriminate CA from other etiologies of left ventricular hypertrophy (LVH) or heart failure.⁶⁻⁸

There is concern that CA is underdiagnosed and diagnosed too late, which limit the opportunity to receive recent targeted therapies that improve quality of life and decrease mortality outcomes in CA patients.⁹⁻¹¹ Recent research has focused on methods that can assist with early identification of CA.¹²⁻¹⁴ Echocardiography is one of the most common initial tests when evaluating patients with heart failure symptoms, with typical CA features including increased left ventricular wall thickness, normal or small left ventricular cavity, biatrial enlargement, preserved left ventricular ejection fraction, and diastolic dysfunction.^{1,2} However, many of these features are also commonly found in other forms of heart failure with preserved ejection fraction and have limited specificity to identify CA,^{6,8} resulting hesitancy to highlight suspicion for CA.

Recent advances in computer vision and artificial intelligence (AI) have enabled precision phenotyping of structure and function in cardiac ultrasound as well as other forms of cardiovascular diagnostics.¹⁵⁻¹⁹ AI applied to echocardiography can precisely estimate wall thickness,¹³ assess mitral regurgitation severity,²⁰ and left ventricular ejection fraction²¹ as well as detect CA,¹³ hypertrophic cardiomyopathy, and diastolic dysfunction.²² EchoNet-LVH is one such computer vision model developed for the detection of CA.¹³ An automated pipeline, EchoNet-LVH identifies parasternal long-axis and apical-4-chamber views from transthoracic echocardiogram studies to precisely measure the wall thickness and identify texture and motion suggestive of CA. Trained against other forms of LVH as controls against CA cases, EchoNet-LVH was designed to be specific in

discriminating between CA and phenotypic mimics. This model was “frozen” and not further fine-tuned to perform further analyses and external validation.

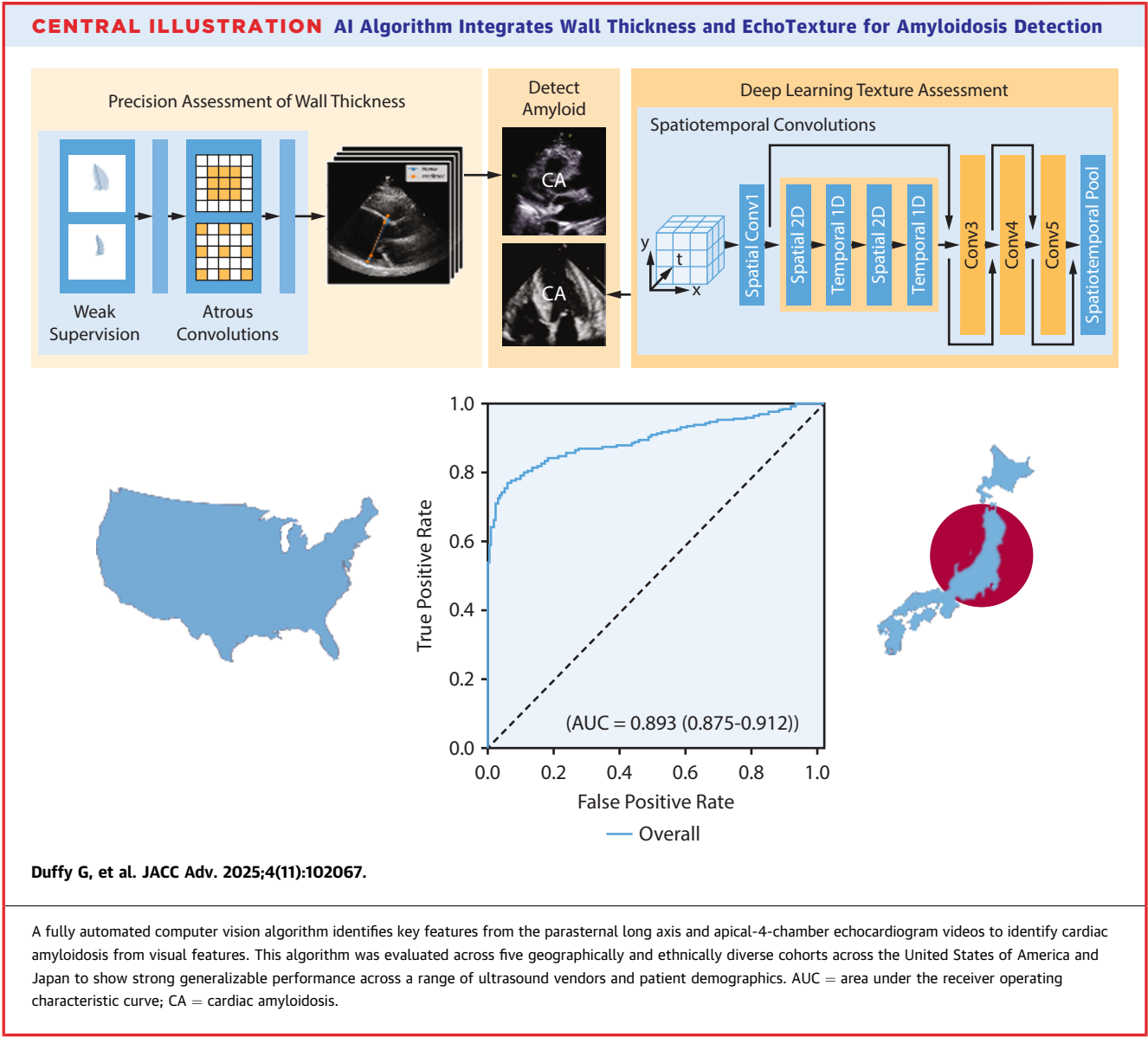
In this study, we evaluated the performance EchoNet-LVH across multiple new health care systems on videos that the model has never seen before (**Central Illustration**). We conducted a multisite international retrospective case-control study evaluating EchoNet-LVH’s ability to distinguish between the echocardiogram studies of CA patients and controls. We reported discrimination performance with area under the receiver operating characteristic curve (AUC) and associated sensitivity, specificity, and positive predictive value (PPV) at a prespecified threshold to identify CA.

METHODS

STUDY DESIGN. We conducted an international multicenter retrospective case-control cohort study with participants from multiple geographically distinct health care systems to evaluate EchoNet-LVH. EchoNet-LVH was previously developed using CA cases and controls from Stanford Health care,²⁰ so this study serves as geographically distinct external validation. Patients were sourced from Cedars-Sinai Medical Center in Los Angeles; California, Keio University in Tokyo, Japan; Northwestern Medicine in Chicago, Illinois; Kumamoto University in Kumamoto; and Yale-New Haven Hospital in New Haven, Connecticut. A total of 574 patients were identified as having CA and matched to 979 randomly selected controls from patients receiving echocardiography and at least 65 years of age. CA patients were diagnosed with transthyretin (ATTR), light chain (AL) amyloidosis, or other forms of CA using a combination of nuclear CA imaging (including single photon emission computed tomography with approved bone radiotracers, such as technetium-99m pyrophosphate), monoclonal gammopathy testing, genetic testing, and/or tissue biopsy. The controls either had negative testing for CA or testing was not performed. Given the low population prevalence of CA, the likelihood of undiagnosed CA in the controls was thought to be minor and unlikely to change the analysis. This study was approved by the Cedars-Sinai Institutional Review Board.

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

Manuscript received March 27, 2025; revised manuscript received July 8, 2025, accepted July 17, 2025.



COMPUTER VISION MODEL. EchoNet-LVH’s development approach and internal validation has been previously described.²⁰ In short, EchoNet-LVH is an automated machine learning pipeline that automatically identifies parasternal long axis and apical-4-chamber views from transthoracic echocardiogram studies, precisely measures wall thickness, and assesses texture and motion from the apical-4-chamber view echocardiogram videos to assess suspicion for CA. Information from the apical-4-chamber view is synthesized with a segmentation model’s assessment of wall thickness from the parasternal long-axis videos to come up with a suspicion for CA. Given the low population prevalence of CA, a

prespecified threshold (0.8) in the summative assessment was chosen to optimize and prioritize for specificity, which in turn, maximizes PPV. Echocardiogram videos were obtained in DICOM format and a fully automatic pipeline analyzed the study. All sites ran the model independently without transfer of imaging data through the sharing of a Docker container with a code for the full end-to-end pipeline.

STATISTICAL ANALYSIS. Continuous variables were reported using median with 25th to 75th percentiles (Q1-Q3), and categorical variables were reported with number (percentage). Performance of the model for discriminating CA was evaluated using AUC,

TABLE 1 Patient Characteristics

Age, y	78 (72-84)
Men	1,200 (77.3%)
Cardiac amyloidosis	574 (36.9%)
Intraventricular septum, cm	1.27 (1.02-1.48)
Subtype	
AL	169 (29.4%)
ATTR	389 (67.8%)
Other	16 (2.8%)
Age, y	
<74	515 (33.2%)
75-84	673 (43.3%)
>85	365 (23.5%)
BMI, kg/m ²	
<25	695 (44.8%)
25-30	538 (34.6%)
>30	313 (20.2%)
Ultrasound manufacturer	
Philips	1,065 (68.6%)
GE	187 (12.0%)
Siemens	130 (8.4%)
Vingmed	96 (6.2%)
Toshiba	66 (4.2%)
Fujifilm	9 (0.6%)

Values are median (Q1-Q3) or n (%).
AL = light chain amyloidosis; ATTR = transthyretin amyloidosis; BMI = body mass index; GE = general electric.

sensitivity, specificity, PPV, and negative predictive value with 95% CIs. AUC was calculated comparing ground truth labels and model output, both for overall AUC and for subgroup analysis. Statistical

analysis was performed in Python (Python Software Foundation).

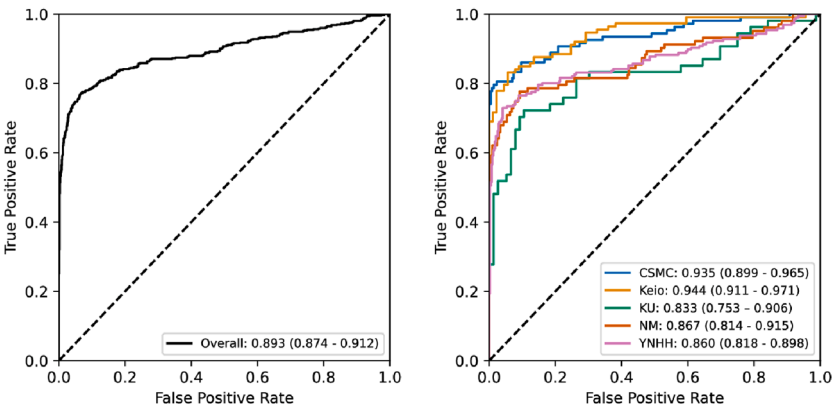
RESULTS

Demographics and clinical characteristics of the study cohort are shown in [Table 1](#). The median age of the cohort was 78 years (Q1-Q3: 72-84) and 77.3% male, with cases matched 2:1 with controls. In the cases, there was representation from both AL (29.4%) and ATTR (67.8%). There was a wide range of age, body mass index, and ultrasound manufacturers across the sites with subgroup analysis to evaluate for heterogeneity of EchoNet-LVH performance.

EchoNet-LVH had an overall AUC of 0.893 (95% CI: 0.874-0.912) with minimal site level variation in performance ([Figure 1](#), [Table 2](#)). The lowest site AUC was Kumamoto University with an AUC 0.833 (95% CI: 0.753-0.906) and the highest site AUC was Keio University with an AUC of 0.944 (95% CI: 0.911-0.971). The overall sensitivity was 0.648 (95% CI: 0.607-0.687) and the overall specificity was 0.982 (95% CI: 0.971-0.989). There was no significant heterogeneity in other performance characteristics across site. At a 2:1 ratio of controls to cases, EchoNet-LVH had a PPV of 0.954 (95% CI: 0.928-0.972) and a negative predictive value of 0.826 (95% CI: 0.804-0.848).

EchoNet-LVH also had similar performance across patient characteristics and ultrasound manufacturer ([Table 3](#)). EchoNet-LVH had an AUC of 0.907 (95% CI:

FIGURE 1 Receiver Operating Characteristic Curve for EchoNet-LVH



Overall performance of AI algorithm and subpopulation performance by site. CSMC = Cedars-Sinai Medical Center; KU = Kumamoto University; NU = Northwestern University; YNHH = Yale New Haven Hospital.

AUROC = area under the receiver operating characteristic curve; CSMC = Cedars-Sinai Medical Center; KU = Kumamoto University; NPV = negative predictive value; NU = Northwestern University; PPV = positive predictive value; YNH = Yale New Haven Hospital.

In this study, we evaluated the performance of a computer vision-driven AI workflow for the detection and screening of CA across a wide range of

^a7 patients did not have BMI available.
Abbreviations as in Table 1 and 2.

patients from an international cohort of geographically distinct health care systems. In this setting, we found EchoNet-LVH had strong performance in identifying patients with CA of all subtypes, and its performance was consistent across sites, ultrasound manufacturers, and patient characteristics.

A few things are worth considering in evaluating our algorithm. Because increased wall thickness is a hallmark of CA,^{1,23} our algorithm automated the approach to precisely measuring wall thickness as well as integrated this precise measurement with more “black box” features of motion and texture assessed in the apical-4-chamber view. Our approach was designed to minimize or exclude confounders,¹³ as our models were trained on controls matched on wall-thickness and limit the potential of AI models to shortcut on wall thickness alone.

Although EchoNet-LVH performed well across all sites, there was a site-level heterogeneity that might be the result of variations in imaging protocols across countries or patient populations. Further study is needed in larger populations to evaluate these impacts. Our approach sought to maximize PPV in the downstream testing of CA. In a rare disease, PPV is significantly impacted by the model specificity (as the number of false positives is likely to outweigh the number of potential true positives). In the algorithm development, we optimized for specificity as to minimize the number of false positives rejected in downstream testing.

In this study evaluating the performance of an AI algorithm on echocardiogram images, an end-to-end computer vision approach was able to detect ATTR and AL CA from an unselected population of patients 65 years of age or greater undergoing echocardiography. Performance was similar across patient characteristics and geographic sites, suggesting generalizability of a computer vision approach despite differences in practice patterns. If validated in future studies, incorporation of this easily obtainable software evaluation can assist with earlier diagnosis of CA and result in reduced morbidity and mortality.

STUDY LIMITATIONS. A few limitations are worth considering as further study is still warranted. Prospective trials of CA screening approaches that evaluate clinical outcomes have not yet completed.^{24,25} There are limited data on the true population prevalence of CA,^{3-5,26} which significantly impacts the PPV of any algorithm in true clinical deployment, and our prior work have shown that choice of controls in retrospective cohorts impact the descriptive model

performance.^{13,27} Other approaches have been proposed to screen for CA,²⁸⁻³⁰ however, EchoNet-LVH represent one of the few fully automated pipelines that can augment clinicians without additional data processing or feedback. Furthermore, there is significant observer variability in most echocardiographic measurements,³¹ which limit the precision of approaches based on routine echo measurements alone.⁸ Parallel work has assessed the performance of EchoNet-LVH in comparison to score-based decision aids and electronic health record-based algorithms and showed computer vision-based approaches having superior performance.³⁰

FUNDING SUPPORT AND AUTHOR DISCLOSURES

This work is funded by NIH NHLBI grants R00HL157421, R01HL173526, and R01HL173487 to D.O., and a grant from Alexion AstraZeneca Rare Disease. Grant Duffy is currently an employee of Meta. D.O. reports consulting fees and/or equity in Ultralytics, InVision, EchoIQ, and Pfizer. Evangelos K. Oikonomou is supported through NIH NHLBI (F32HL170592); he is a co-founder of Evidence2Health LLC, an ad hoc consultant for Caristo Diagnostics Ltd and Ensign-AI Inc; and has received royalty fees from technology licensed through the University of Oxford, outside this work. Rohan Khara reports support from NIH NHLBI (R01HL167858 and K23HL153775), NIH NIA (R01AG089981), and the Doris Duke Charitable Foundation (2022060); he is an Associate Editor of JAMA and receives research support, through Yale, from the Blavatnik Foundation, Bristol-Myers Squibb, Novo Nordisk, and BridgeBio; he is a coinventor of Pending Patent Applications WO2023230345A1, US20220336048A1, 63/346,610, 63/484,426, 63/508,315, 63/580,137, 63/606,203, 63/619,241, and 63/562,335; and a co-founder of Ensign-AI, Inc and Evidence2Health, LLC. Faraz S. Ahmad has received research support from Pfizer and Atman Health and consulting fees from Alnylam Pharmaceuticals. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr David Ouyang, 127 S. San Vicente Blvd., Suite A3600, Los Angeles, California 90048, USA. E-mail: david.ouyang@cshs.org.

PERSPECTIVES

COMPETENCY IN MEDICAL KNOWLEDGE:

Recently United States Food and Drug Administration cleared AI algorithm can assist clinicians in identifying patients >65 years of age with cardiac amyloidosis.

TRANSLATIONAL OUTLOOK: Additional research is needed to understand impact on clinical workflow in prospective evaluation.

REFERENCES

1. Writing Committee, Kittleson MM, Ruberg FL, et al. 2023 ACC expert consensus decision pathway on comprehensive multidisciplinary care for the patient with cardiac amyloidosis: a report of the American college of cardiology solution set oversight committee. *J Am Coll Cardiol*. 2023;81:1076-1126.
2. Kittleson MM, Maurer MS, Ambardekar AV, et al. Cardiac amyloidosis: evolving diagnosis and management: a scientific statement from the American heart association. *Circulation*. 2020;142:e7-e22.
3. Kim HM, Sohn D-W, Paeng JC. Prevalence of positive 99 mTc-DPD scintigraphy as an indicator of the prevalence of wild-type transthyretin amyloidosis in the elderly. *Int Heart J*. 2019;60:643-647.
4. Quock TP, Yan T, Chang E, Guthrie S, Broder MS. Epidemiology of AL amyloidosis: a real-world study using US claims data. *Blood Adv*. 2018;2:1046.
5. Treglia G, Martinello C, Dondi F, et al. Prevalence of incidental findings suspicious for transthyretin cardiac amyloidosis among patients undergoing bone scintigraphy: a systematic Review and a meta-analysis. *J Clin Med*. 2023;12:5698.
6. Lee GY, Kim K, Choi JO, et al. Cardiac amyloidosis without increased left ventricular wall thickness. *Mayo Clin Proc*. 2014;89:781-789.
7. Cipriani A, De Michieli L, Porcari A, et al. Low QRS voltages in cardiac amyloidosis: clinical correlates and prognostic value. *JACC CardioOncol*. 2022;4:458-470.
8. Cotella J, Randazzo M, Maurer MS, et al. Limitations of apical sparing pattern in cardiac amyloidosis: a multicenter echocardiographic study. *Eur Heart J Cardiovasc Imaging*. 2024;25:754-761. <https://doi.org/10.1093/ehjci/jeae021>
9. Ichikawa Y, Oota E, Odajima S, et al. Impact of tafamidis on echocardiographic cardiac function of patients with transthyretin cardiac amyloidosis. *Circ J*. 2023;87:508-516.
10. Maurer MS, Schwartz JH, Gundapaneni B, et al. Tafamidis treatment for patients with transthyretin amyloid cardiomyopathy. *N Engl J Med*. 2018;379:1007-1016.
11. Fontana M, Berk JL, Gillmore JD, et al. Vutrisiran in patients with transthyretin amyloidosis with cardiomyopathy. *N Engl J Med*. 2024;392:33-44.
12. Goto S, Mahara K, Beussink-Nelson L, et al. Artificial intelligence-enabled fully automated detection of cardiac amyloidosis using electrocardiograms and echocardiograms. *Nat Commun*. 2021;12:2726.
13. Duffy G, Cheng PP, Yuan N, et al. High-Throughput precision phenotyping of left ventricular hypertrophy with cardiovascular deep learning. *arXiv*. 2021.
14. Sangha V, Oikonomou EK, Khera R. Artificial intelligence applied to electrocardiographic images for scalable screening of transthyretin amyloid cardiomyopathy. *medRxiv*. 2024. <https://doi.org/10.1101/2024.09.30.24314651>
15. Omori H, Kawase Y, Mizukami T, et al. Diagnostic performance of artificial intelligence-based angiography-derived non-hyperemic pressure ratio using pressure wire as reference. *Circ J*. 2024;89:323-330.
16. Omori H, Kawase Y, Mizukami T, et al. Diagnostic accuracy of artificial intelligence-based angiography-derived fractional flow reserve using pressure wire-based fractional flow reserve as a reference. *Circ J*. 2023;87:783-790.
17. Yamamoto T, Kawase Y, Mizukami T, et al. Enhanced plaque stabilization effects of alirocumab - insights from artificial intelligence-aided optical coherence tomography analysis of the alirocumab for thin-cap fibroatheroma in patients with coronary artery disease estimated by optical coherence tomography (ALTAIR) study. *Circ J*. 2024;88:1809-1818.
18. Holste G, Oikonomou EK, Mortazavi BJ, et al. Severe aortic stenosis detection by deep learning applied to echocardiography. *Eur Heart J*. 2023;44:4592-4604.
19. Pirruccello JP, Bick A, Wang M, et al. Analysis of cardiac magnetic resonance imaging in 36,000 individuals yields genetic insights into dilated cardiomyopathy. *Nat Commun*. 2020;11:2254.
20. Vrudhula A, Duffy G, Vukadinovic M, et al. High throughput deep learning detection of mitral regurgitation. *bioRxiv*. 2024. <https://doi.org/10.1101/2024.02.08.24302547>
21. Vukadinovic M, Tang X, Yuan N, et al. EchoPrime: a multi-video view-informed vision-language model for comprehensive echocardiography interpretation. *arXiv*. 2024.
22. Akerman AP, Porumb M, Scott CG, et al. Automated echocardiographic detection of heart failure with preserved ejection fraction using artificial intelligence. *JACC Adv*. 2023;2:100452.
23. Pagourelas ED, Mirea O, Duchenne J, et al. Echo parameters for differential diagnosis in cardiac amyloidosis. *Circ Cardiovasc Imaging*. 2017;10:e005588.
24. Gouverneur C. BridgeBio partners with leading cardiovascular data science lab (CarDS lab) to improve transthyretin amyloid cardiomyopathy (ATTR-CM) diagnosis in diverse patient populations with multimodal artificial intelligence. BridgeBio. 2024. Accessed December 17, 2024. <https://bridgebio.com/news/bridgebio-partners-with-leading-cardiovascular-data-science-lab-cards-lab-to-improve-transthyretin-amyloid-cardio-myopathy-attr-cm-diagnosis-in-diverse-patient-populations-with-multimodal-artificial/>
25. National Library of Medicine. Artificial Intelligence Guided Echocardiographic Screening of Rare Diseases (EchoNet-Screening). Accessed December 17, 2024. <https://clinicaltrials.gov/study/NCT05139797>
26. Gilstrap LG, Dominici F, Wang Y, et al. Epidemiology of cardiac amyloidosis-associated heart failure hospitalizations among fee-for-service medicare beneficiaries in the United States. *Circ Heart Fail*. 2019;12:e005407.
27. Vrudhula A, Stern L, Cheng PC, et al. Impact of case and control selection on training artificial intelligence screening of cardiac amyloidosis. *JACC Adv*. 2024;3:100998.
28. Huda A, Castaño A, Niyogi A, et al. A machine learning model for identifying patients at risk for wild-type transthyretin amyloid cardiomyopathy. *Nat Commun*. 2021;12:2725.
29. Davies DR, Redfield MM, Scott CG, 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.
30. Hourmozdi J, Easton N, Benigeri S, et al. Evaluating the performance and potential bias of predictive models for the detection of transthyretin cardiac amyloidosis. *medRxiv*. 2024;10.09.24315202. <https://doi.org/10.1101/2024.10.09.24315202>
31. Pillai B, Salerno M, Schnittger I, Cheng S, Ouyang D. Precision of echocardiographic measurements. *J Am Soc Echocardiogr*. 2024;37:562-563.

KEY WORDS artificial intelligence, cardiac amyloidosis, computer vision, echocardiography