Screening performance of a deep learning model versus an automated wall thickness score in biopsy-proven cardiac amyloidosis

Authors

Y Nomura[1] , Y Hirata[1] , S Nishio[1] , Y Saijo[2] , R Nonaka[1] , N Yamaguchi[1] , R Zheng[2] , Y Okushi[2] , T Takahashi[2] , K Kusunose[3] , H Yamada[4] , M Sata[2] , [1]Tokushima University Hospital, Ultrasound Examination Center - Tokushima - Japan , [2]Tokushima University Hospital , Department of Cardiovascular Medicine - Tokushima - Japan , [3]University of the Ryukyus, Graduate School of Medicine, Department of Cardiovascular Medicine, Nephrology, and Neurology - Okinawa - Japan , [4]Tokushima University Graduate School of Biomedical Sciences, Community Medicine for Cardiology - Tokushima - Japan ,

Abstract

Citation: N/A

Background: Cardiac amyloidosis (CA) is a progressive cardiomyopathy associated with poor prognosis. The recent emergence of disease-modifying therapies has underscored the importance of early diagnosis. Because definitive diagnosis requires multiple tests, an accessible screening approach is essential. Echocardiography plays a pivotal role in screening. An artificial intelligence (AI)-based fully automated echocardiographic analysis system has been developed to detect CA from an apical four-chamber (A4C) cine clip using a deep learning model.

Purpose: To compare the screening performance of a deep learning model (AI-CNN) with a model that determines CA by calculating an increased wall thickness score from AI-derived automated measurements (AI-IWTS), using endomyocardial biopsy as the reference standard.

Methods: We retrospectively screened 171 consecutive patients who underwent endomyocardial biopsy at our institution between 2016 and 2025. Patients with AI-measured left ventricular (LV) wall thickness ≥ 12 mm were included in the analysis because this threshold is required for both AI-CNN and AI-IWTS. CA was defined as amyloid protein positivity on pathological examination. Diagnostic performance was evaluated by sensitivity, specificity, predictive values, and c-index with 95% confidence intervals. Paired classifications were compared using McNemar's test.

Results: Among the 92 patients with AI-measured LV wall thickness ≥ 12 mm, 58 had biopsy-confirmed CA, including 48 with transthyretin amyloidosis (ATTR), 9 with light-chain amyloidosis (AL), and 1 with apolipoprotein A-I amyloidosis (ApoA-I). AI-CNN demonstrated a c-index of 0.82 (95% CI 0.72–0.89), with a sensitivity of 98% and a specificity of 65%. AI-IWTS demonstrated a c-index of 0.59 (95% CI 0.48–0.69), with a sensitivity of 26% and a specificity of 91%. Positive predictive value (PPV) was 83% for both methods, while negative predictive value (NPV) was higher for AI-CNN than for AI-IWTS (96% vs. 43%). McNemar's test showed a significant difference in classifications between the two methods (P<0.001).

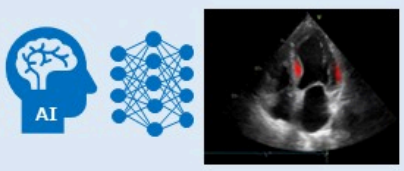
Conclusions: In a biopsy-referenced cohort selected by AI-measured LV wall thickness ≥ 12 mm, AI-CNN showed higher sensitivity and discrimination than AI-IWTS. These findings suggest that AI-CNN may serve as a practical echocardiographic screening tool to support early detection of CA in routine clinical practice.

Patients undergoing endomyocardial biopsy with AI-measured LV wall thickness ≥12 mm (n=92)

Cardiac amyloidosis (n=58)	
ATTR	n=48
AL	n=9
ApoA-1	n=1

Non-amyloidosis (n=34)	
Hypertrophic cardiomyopathy	n=5
Hypertensive heart disease	n=8
Dilated cardiomyopathy	n=7
Cardiac sarcoidosis	n=2
Myocarditis	n=8
Others	n=4

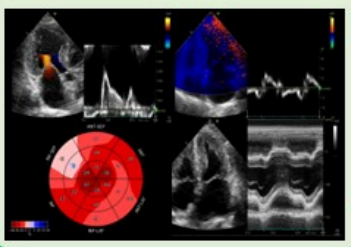
AI-CNN
Deep learning model



A4C cine clip analysis

Sensitivity	98%
Specificity	65%
PPV	83%
NPV	96%
c-index (95% CI)	0.82 (0.72–0.89)

AI-IWTS
Automated measurement–based IWT score model



Fully automated measurements

↓

IWT score (Positive if ≥8 points)

Sensitivity	26%
Specificity	91%
PPV	83%
NPV	43%
c-index (95% CI)	0.59 (0.48-0.69)