The International Journal of Cardiovascular Imaging (2023) 39:339–348
https://doi.org/10.1007/s10554-022-02738-1

ORIGINAL PAPER

Machine learning algorithms to automate differentiating cardiac amyloidosis from hypertrophic cardiomyopathy

Zi-Wen Wu[1] · Jin-Lei Zheng[1] · Lin Kuang[1] · Hui Yan[1]

Received: 14 April 2022 / Accepted: 27 September 2022 / Published online: 19 October 2022

Abstract
Cardiac amyloidosis has a poor prognosis, and high mortality and is often misdiagnosed as hypertrophic cardiomyopathy, leading to delayed diagnosis. Machine learning combined with speckle tracking echocardiography was proposed to automate differentiating two conditions. A total of 74 patients with pathologically confirmed monoclonal immunoglobulin light chain cardiac amyloidosis and 64 patients with hypertrophic cardiomyopathy were enrolled from June 2015 to November 2018. Machine learning models utilizing traditional and advanced algorithms were established and determined the most significant predictors. The performance was evaluated by the receiver operating characteristic curve (ROC) and the area under the curve (AUC). With clinical and echocardiography data, all models showed great discriminative performance (AUC > 0.9). Compared with logistic regression (AUC 0.91), machine learning such as support vector machine (AUC 0.95, p = 0.477), random forest (AUC 0.97, p = 0.301) and gradient boosting machine (AUC 0.98, p = 0.230) demonstrated similar capability to distinguish cardiac amyloidosis and hypertrophic cardiomyopathy. With speckle tracking echocardiography, the predictive performance of the voting model was similar to that of LightGBM (AUC was 0.86 for both), while the AUC of XGBoost was slightly lower (AUC 0.84). In fivefold cross-validation, the voting model was more robust globally and superior to the single model in some test sets. Data-driven machine learning had shown admirable performance in differentiating two conditions and could automatically integrate abundant variables to identify the most discriminating predictors without making preassumptions. In the era of big data, automated machine learning will help to identify patients with cardiac amyloidosis and timely and effectively intervene, thus improving the outcome.

Keywords Cardiac amyloidosis · Hypertrophic cardiomyopathy · Machine learning

Introduction

Cardiac amyloidosis (CA) is a part of systemic amyloidosis, in which misfolded amyloid proteins are deposited outside cardiomyocytes and lead to restrictive pathology of the heart, often denoting a poor outcome [1, 2]. In recent years, several new therapies that significantly improve the prognosis of patients with CA have been developed, including bortezomib-based induction and consolidation strategies, autologous stem cell transplantation, immunomodulatory drugs, etc. [3]. Unfortunately, for patients with advanced cardiac involvement, current treatments are still limited. Moreover, patients with CA could be easily misdiagnosed as hypertrophic cardiomyopathy (HCM) who have similar phenotypes that are difficult to distinguish on routine echocardiography, often leading to delayed diagnosis. However, CA has high mortality and poor prognosis, which makes early detection and differential diagnosis quite important.

Because of the advantages of wide application and superior diastolic function assessment, echocardiography has become the preferred screening method for CA. Advanced two-dimensional speckle tracking echocardiography (2D-STE) and strain, and strain rate imaging have been proven to differentiate CA from other causes of concentric cardiac hypertrophy [4]. Since supersonic inspection always produces lots of imaging data and the variables interact with each other to varying degrees, it is difficult to identify the most discriminative predictors through ordinary statistical analysis. Therefore, more powerful data processing

✉ Hui Yan
michelleyan@zju.edu.cn

[1] Department of Cardiology, The First Affiliated Hospital, School of Medicine, Zhejiang University, No.79 qingchun Road, Hangzhou 310003, Zhejiang, China

approaches are urgently needed to extract and analyze imaging data.

Machine learning (ML) utilizes computer algorithms to seek inherent patterns in datasets with massive variables without making preassumptions. It can learn from established datasets and facilitate the prediction of risk models on new data. In recent years, ML has become an effective means for prediction and intelligent decision-making [5–7] and has achieved commendable success in cardiovascular medicine, such as differentiation of constrictive pericarditis from restrictive cardiomyopathy [8], risk prediction of readmission of patients with heart failure [9], diagnosing different arrhythmias [10, 11], etc. Given this, we proposed an intelligent identification study of CA and HCM based on ML.

Methods

A case–control study of 138 subjects, including 74 verified CA cases and 64 patients with verified HCM cases were referred to the First Affiliated Hospital of Zhejiang University School of Medicine from June 2015 to November 2018. The type of amyloid of all patients with CA assessed by immune histology was the light chain and patients were eligible for inclusion if they met any of the following criteria: (1) an endomyocardial biopsy confirmed amyloid deposits; (2) a positive non-cardiac biopsy for amyloidosis combined with cardiac magnetic resonance or non-strain-based echocardiography which presented typical characteristic of CA, with relevant clinical history and laboratory findings. The characteristics of CA are consistent with the Expert Consensus Recommendations for Multimodality Imaging [12]. Cardiac involvement of CA was assessed by imaging scans, of which forty patients involved the left ventricle, twenty-seven patients involved the left and right ventricles, one patient implicated two ventricles and the left atrium, and six patients involved all ventricles and atria. We further incorporated 64 patients with HCM as comparator groups whose diagnose were created according to recently published guidelines from the American College of Cardiology/the American Heart Association [13], and they underwent both echocardiography and cardiac magnetic resonance imaging to further assess HCM and exclude other pathologies. Three of them had also genetic analysis and all showed heterozygous mutation. An echocardiographic examination was performed in all patients with HCM who presented unexplained left ventricular asymmetrical hypertrophy with septal wall thickness ≥ 15 mm. In the case of positive family history (such as sudden death, cardiac hypertrophy, etc.), interventricular septal thickness ≥ 13 mm was also enrolled. Subjects with left ventricular ejection fraction $< 45\%$, secondary cardiac hypertrophy caused by severe aortic valve disease, long-term uncontrolled hypertension, or thyroid disease were excluded from the study. Patients were also excluded if the relevant data were not available. The local institutional ethics committee approved the study.

Echocardiographic examination

All echocardiographic studies were conducted on GE Vivid E9 Color Doppler Ultrasound system (GE Medical, Milwaukee, Wisconsin, USA) equipped with a 2-dimension probe M5S with a frequency of 2.0–4.5 MHz and a frame rate of 50–70 frames per second. The grayscale dynamic images of the 4-chamber views, the long axis view of the left ventricle, the 2-chamber views, and the short axis section with 3 consecutive cardiac cycles were obtained and stored on the hard disk. M-mode and tissue Doppler ultrasound were used to collect ultrasonic parameters, which included: the left atrial volume index using an ellipse formula, end-diastolic left ventricular diameter, end-systolic left ventricular diameter, end-diastolic left ventricular volume, end-systolic left ventricular volume, ejection fraction using the biplane Simpson's method in 4-chamber views and 2-chamber views, septal wall thickness, posterior wall thickness. The eccentricity index was calculated as septal wall thickness divided by posterior wall thickness. Relative wall thickness was calculated as the ratio of 2 septal wall thickness divided by end-diastolic left ventricular diameter. The left ventricular mass index was calculated based on the Cube formula. Concentric hypertrophy was diagnosed in patients with relative wall thickness > 0.42 and a left ventricular mass index > 115 g/m^2. Diastolic parameters, including peak early (E) and late (A) diastolic mitral inflow velocity and the ratio of E/A, e′, and the ratio of E/e′ ratio were also measured (Fig. 1).

2D-STE acquisition and analysis

Offline analysis of the video clips was based on Echo PAC Version 201 software (GE Company, Fairfield, Connecticut, USA), running on Windows 10 Version 1709 (Microsoft Corporation, Washington State, USA). Selecting clear dynamic images and using the 4-chamber views, the long axis view of the left ventricle, and the 2-chamber views, the left ventricular endocardial and epicardial myocardium were automatically tracked combined with manually adjusted frame by frame throughout the cardiac cycle, and divided into 16 segments to generate a 'bull's-eye' plot.

The strain data are gathered by time and space parameters. Each cardiac cycle was divided into 17 equal segments (T1–T17), and Tj represented the corresponding time points ($j = 1,2\ldots17$); Strain measurements were included as follows: longitudinal strain (LS), global longitudinal strain,

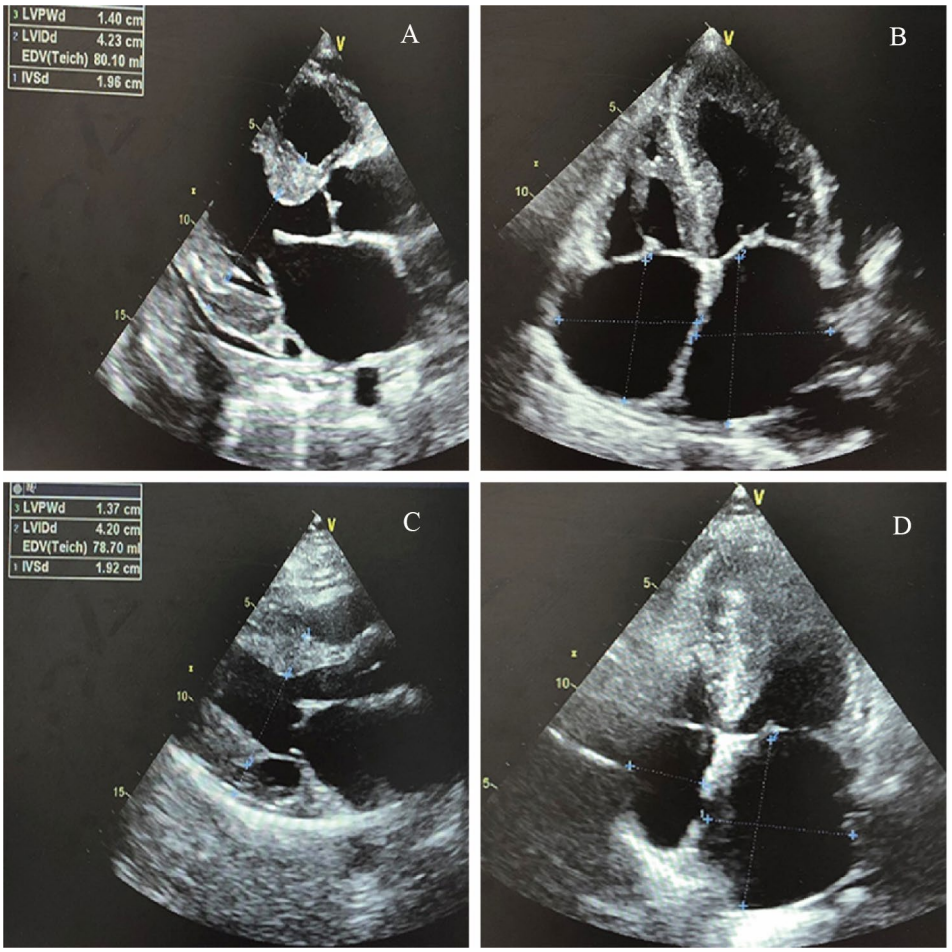

Fig. 1 Cardiac imaging. This example showed the morphological similarities between cardiac amyloidosis (**A**, **B**) and hypertrophic cardiomyopathy (**C**, **D**)

longitudinal strain velocity, longitudinal strain rate, longitudinal displacement, circumferential strain, global circumferential strain, circumferential strain rate, radial strain, global radial strain, radial strain rate, rotational rate, left ventricular twist, left ventricular twist rate. According to the above methods, 3791 (223×17) variables were systematically extracted for each patient (223 are strain-derived variables and 17 are time points). Average time strain-derived variables (223 variables) were used to train the models. Relative apical sparing was calculated as average apical LS divided by the sum of the average basal and mid-LS, septal apical to base ratio as apical septal LS divided by basal septal LS, and ejection fraction strain ratio as ejection fraction divided by global longitudinal strain.

Establishment and assessment of prediction models

Two ML-based prediction models were established respectively: one model was built using clinical characteristics, conventional echocardiography, and 2D-STE data; the other was to build models using only 2D-STE data.

Prediction models using clinical characteristics, conventional echocardiography, and 2D-STE data

We developed prediction models using four approaches: logistic regression, support vector machine, random forest, and XGBoost. These represent the comprehensive analysis from traditional logistic regression to classic ML algorithms (support vector machine, random forest), and then to advanced gradient boosting (XGBoost). To assess the validity of the models, we performed tenfold (or fivefold) cross-validation by randomly dividing the entire data into 10 (or 5) parts for 10 (or 5) iterations. In each iteration, we selected 7 parts as training data and 3 parts as test sets. We reported average results for each model on 30% of unseen test sets.

Logistic regression

Logistic regression is the most commonly used risk prediction model. First, univariate logistic regression was used to screen out the variables that were meaningful to predict CA. Then, variables with $p < 0.1$ were enrolled in the multivariate

regression analysis for modeling according to the previous research, clinical experience, and the multiple requirements between variables and outcome. In addition, Spearman correlation was used to exclude the influence of collinearity among variables.

Support vector machine

Support vector machine converts data into complex high-dimensional space to look for the largest difference margin to realize the differentiation of diseases [14]. We applied linear basis kernel and cost function to build the model and tuning parameters to minimize the error classification.

Random forest

Random forest is a tree-based method, the essence of which is to continuously split variables at discrete cutting points, usually presenting in the form of a tree graph [14]. A separate tree is built from bootstrapped data and variables, and the final model is a collection of many trees.

Gradient boosting

The core idea of gradient boosting is to set up a series of initial models based on the decision tree, which is called base classifiers [15, 16]. Subsequently, weaker base classifiers are iterated and adjusted the weights to create a single stronger classifier. Information gain (IG), a technique of feature selection, is defined as a metric of effective classification. It is measured in terms of the entropy reduction of the class, which reflects additional information about the class provided by the variables.

Prediction models of using 2D-STE data

Boosting-based algorithms are increasingly used because they involve the sequential creation of models, with each iteration attempting to correct errors in the previous models. LightGBM and XGBoost are two widely used algorithms. We developed predictive classifiers using 2D-STE data: (1) LightGBM; (2) XGBoost; (3) voting model based on LightGBM and XGBoost. To evaluate the validity of models, a fivefold cross-validation was performed. We split the dataset into the training set and test set in a 4:1 ratio and reported the performance on the test data.

Statistical analysis

Categorical variables were expressed as the number of cases and percentages and were compared using the chi-square test or Fisher's test. Continuous variables were expressed as mean ± SD. Kolmogorov–Smirnov test was used to determine whether the data were normally distributed. If the data conform to the normal distribution, the independent sample T-test was used for comparison; otherwise, the Mann–Whitney U test was suitable. $p < 0.05$ was considered statistically significant. Sensitivity, specificity, positive predictive value, negative predictive value, accuracy, receiver operating characteristic curve, and area under the curve (AUC) were used to evaluate the performance of models. DeLong test was used to evaluate whether the AUC in different models was statistically significant. The data analysis was implemented on SPSS 23.0 (Version 23.0), R (Version 4.0.3), and Python (Version 3.7).

Table 1 Baseline patient characteristics

Variable	CA (n = 74)	HCM (n = 64)	p value*
Age (years)	60.9 ± 9.7	50.3 ± 15.3	0.000
Sex (% male)	45 (60.8)	42 (65.6)	0.559
BMI (kg/m^2)	21.5 ± 4.6	24.3 ± 3.6	0.000
BSA (m^2)	1.6 ± 0.2	1.7 ± 0.2	0.000
SBP (mmHg)	104 ± 17	124 ± 20	0.000
DBP (mmHg)	67 ± 13	74 ± 14	0.004
HR (bpm)	82 ± 14	75 ± 13	0.003
NYHA (> 2 class)	16 (21.7)	10 (15.6)	0.369
Atrial fibrillation	8 (10.8)	12 (18.8)	0.186
Coronary heart disease	4 (5.4)	10 (15.6)	0.047
Hypertension	12 (16.2)	27 (42.2)	0.001
Hyperlipemia	29 (39.2)	14 (21.9)	0.029
Diabetes mellites	5 (6.8)	8 (12.5)	0.249
Chronic kidney disease	47 (63.5)	5 (7.8)	0.000
WBC (10^9/l)	7.1 ± 2.6	6.6 ± 2.3	0.423
HB (g/l)	122 ± 23	139 ± 23	0.000
Platelet (10^9/l)	193 ± 85	196 ± 64	0.795
NT-proBNP (pg/ml)	2984 ± 3213	1947 ± 2574	0.107
Troponin I (mmo/l)	0.46 ± 1.66	0.28 ± 1.18	0.001
creatinine (umol/l)	111 ± 94	103 ± 199	0.057
eGFR (ml/min)	72.82 ± 25.57	83.28 ± 28.22	0.007

Values are mean ± SD

CA cardiac amyloidosis; *HCM* hypertrophic cardiomyopathy; *BMI* body mass index; *BSA* body surface area; *SBP* systolic blood pressure; *DBP* diastolic blood pressure; *HR* heart rate; *HYHA* New York Heart Association; *WBC* White blood cell count; *HB* hemoglobin; *NT-proBNP* N-terminal pro-brain natriuretic peptide; *eGFR* estimated glomerular filtration rate

*Comparison performed between those with CA and HCM

Results

Study population

The clinical characteristics of both groups are summarized in Table 1. The age (60.9 ± 9.7 vs 50.3 ± 15.3, $p < 0.001$),

heart rate (82 ± 14 vs 75 ± 13, p = 0.003) and troponin I (0.46 ± 1.66 vs 0.28 ± 1.18, p = 0.001) level of patients with CA were higher, while the body mass index (21.5 ± 4.6 vs 24.3 ± 3.6, p < 0.001), body surface area (1.6 ± 0.2 vs, p = 0.000), systolic (104 ± 17 vs 124 ± 20, p < 0.001) and diastolic (67 ± 13 vs 74 ± 14, p = 0.004) blood pressure, and glomerular filtration rate (72.82 ± 25.57 vs 83.28 ± 28.22, p = 0.007) were lower than those with HCM, and the differences were statistically significant.

Conventional echocardiography

End-systolic left ventricular diameter (28.4 ± 4.0 vs 25.2 ± 4.9, p < 0.001), end-systolic left ventricular volume (31.8 ± 11.2 vs 24.1 ± 10.4, p < 0.001), left ventricular posterior wall thickness (14.0 ± 2.8 vs 12.3 ± 3.6, p < 0.001), relative wall thickness (0.70 ± 0.16 vs 0.60 ± 0.23, p < 0.001), E (0.85 ± 0.30 vs 0.70 ± 0.24, p = 0.001), E/A (1.81 ± 1.11 vs 1.19 ± 0.69, p = 0.001), E/e′ (18.1 ± 10.1 vs 15.0 ± 15.7, p = 0.010) and relative apical sparing (0.93 ± 0.31 vs 0.68 ± 0.24, p < 0.001) of patients with CA were higher, while end-diastolic left ventricular diameter (40.6 ± 4.3 vs 42.5 ± 5.7, p = 0.045), end-diastolic left ventricular volume (73.6 ± 18.5 vs 81.4 ± 24.6, p = 0.041), ejection fraction (56.7 ± 10.7 vs 70.7 ± 8.5, p < 0.001), left ventricular septum thickness (15.5 ± 3.6 vs 23.1 ± 6.7, p < 0.001), left ventricular mass index (159.6 ± 53.1 vs 190.8 ± 58.7, p = 0.001), eccentricity index (1.1 ± 0.3 vs 2.0 ± 0.7, p < 0.001) and A (0.59 ± 0.25 vs 0.69 ± 0.26, p = 0.026) were lower than those with HCM. There was no statistical difference in left atrial volume index, e′, global longitudinal strain, septal apical to base ratio, and ejection fraction strain ratio between the two groups, as shown in Table 2.

Clinical predictive models

The models based on clinical characteristics, conventional echocardiography, and 2D-STE data

ML algorithms all have good predictive performance (AUC > 0.9), among which the XGBoost has the highest AUC of 0.98 (Table 3). Compared with the logistic regression (sensitivity 92%, specificity 94%, AUC 0.91), ML such as support vector machine (sensitivity 89%, specificity 100%, AUC 0.95, p = 0.477), random forest (sensitivity 96%, specificity 100%, AUC 0.97, p = 0.301) and XGBoost (sensitivity 88%, specificity 95%, AUC 0.98, p = 0.230) presented similar capability to predict CA.

Relative apical sparing, age, left atrial volume index, and eccentric index were found to be significantly predictive of CA using multivariable logistic regression analysis (p < 0.05). The ejection fraction was ruled out due to the effect of multicollinearity. Information gain of XGBoost to feature selection showed that chronic kidney disease (IG = 0.26) was the most important predictor, followed by left ventricular septum thickness (IG = 0.17), ejection fraction (IG = 0.13), relative apical sparing (IG = 0.11), systolic blood pressure (IG = 0.07) and eccentricity index (IG = 0.05) (Fig. 2).

The models based on 2D-STE data

After training with the tuned hyperparameters, the feature importance of the voting model integrated LightGBM and XGBoost was obtained and ranked. The results indicated that RadStrain3 (the radial strain of the middle ventricular septum) was the most important predictor, followed by LongStrainEpi7 (the longitudinal strain of the anterior wall of the epicardial basement segment) and CirStrainR4 (circumferential strain rate of the posterior wall of the basement segment) (Table 4).

Among all the three ML algorithms (XGBoost, LightGBM, and voting model), the discriminant ability of the voting model was similar to LightGBM (AUC of both was 0.86), while the AUC of XGBoost was slightly lower, which was 0.84 (Fig. 3).

In the fivefold cross-validation, the mean AUC of LightGBM, XGBoost, and voting models were 0.89 ± 0.19, 0.85 ± 0.43, and 0.87 ± 0.30, respectively. The voting model was globally more robust and outperformed to individual model on test sets (Fig. 4).

Discussions

ML combined with 2D-STE to carry out intelligent identification on CA and HCM discovered: that ML had a great performance in the differential diagnosis and could automatically integrate plentiful variables to identify the most discriminative predictors without preassumption.

CA is a rare and complex disease with high mortality and poor prognosis. Although new treatments which significantly improve the outcomes have been developed, the available management with advanced cardiac involvement is still very limited. In addition, clinical confusion about cardiac hypertrophy caused by other causes (e.g., HCM, hypertension, and aortic stenosis) often leads to delayed diagnosis, which prevents patients with CA from receiving an early and effective intervention. Echocardiography has become the preferred screening approach for patients with CA due to its wide application, low risk, low cost, convenience, and superior diastolic function assessment.

Table 2 Echocardiographic characteristics

Variable	CA (n = 74)	HCM (n = 64)	p value*
LAVI (ml/m^2)	34.4 ± 13.5	39.0 ± 14.9	0.085
LVDd (mm)	40.6 ± 4.3	42.5 ± 5.7	0.045
LVDs (mm)	28.4 ± 4.0	25.2 ± 4.9	0.000
EDV (ml)	73.6 ± 18.5	81.4 ± 24.6	0.041
ESV (ml)	31.8 ± 11.2	24.1 ± 10.4	0.000
LVEF (%)	56.7 ± 10.7	70.7 ± 8.5	0.000
LVSd (mm)	15.5 ± 3.6	23.1 ± 6.7	0.000
LVPWd (mm)	14.0 ± 2.8	12.3 ± 3.6	0.000
LVMI (g/m^2)	159.6 ± 53.1	190.8 ± 58.7	0.001
RWT	0.70 ± 0.16	0.60 ± 0.23	0.000
Eccentric index	1.1 ± 0.3	2.0 ± 0.7	0.000
E (m/s)	0.85 ± 0.30	0.70 ± 0.24	0.001
A (m/s)	0.59 ± 0.25	0.69 ± 0.26	0.026
E/A	1.81 ± 1.11	1.19 ± 0.69	0.001
e′ (m/s)	0.06 ± 0.03	0.06 ± 0.02	0.263
E/e′	18.1 ± 10.1	15.0 ± 15.7	0.010
GLS	11.9 ± 4.6	13.1 ± 4.1	0.121
SAB	2.63 ± 1.45	2.81 ± 2.38	0.322
EFSR	5.4 ± 2.1	6.0 ± 2.7	0.098
RELAPS	0.93 ± 0.31	0.68 ± 0.24	0.000

Values are mean ± SD

CA cardiac amyloidosis; *HCM* hypertrophic cardiomyopathy; *LAVI* left atrial volume index; *LVDd* end-diastolic left ventricular diameter; *LVDs* end-systolic left ventricular diameter; *EDV* end-diastolic left ventricular volume; *ESV* end-systolic left ventricular volume; *LVEF* left ventricular ejection fraction; *LVSd* left ventricular septum thickness; *LVPWd* left ventricular posterior wall thickness; *LVMI* left ventricular mass index; *RWT* relative wall thickness; *E* peak early mitral diastolic flow velocity; *A* peak late mitral diastolic flow velocity; *e′* early mitral annular diastolic tissue velocity; *GLS* global longitudinal strain; *SAB* septal apical to base ratio; *EFSR* ejectionfraction strain ratio; *RELAPS* relative apical sparing

*Comparison performed between those with CA and HCM

Therefore, many related studies had been done to distinguish CA from other causes of cardiac hypertrophy.

Cardiac deformation analysis of 2D-STE could reveal early systolic abnormalities. Early reports by Sun et al. [17] suggest that global longitudinal strain, global circumferential strain, and the global radial strain were significantly reduced in patients with advanced CA compared with HCM and hypertensive heart disease, and although there was some overlap between the groups, the three causes of cardiac hypertrophy could be distinguished to a certain extent. Di Bella et al. [18] displayed that the epicardial strain in patients with amyloid transthyretin was significantly lower than that in patients with HCM. Subsequently, Baccouche et al. [19] made use of 3-dimension speckle-tracking echocardiography to identify CA from HCM, presenting that most of the functional parameters in both groups were lower, while those in the CA group were the lowest. The radial strain of CA patients demonstrated the "reverse pattern" from base to apex, suggesting that two conditions could be distinguished based on functional patterns. Similarly, Phelan et al. [4] proposed that the "relative apical sparing" of longitudinal strain could well identify CA. Liu et al. [20] showed that septal apical to base ratio > 2.1 joint with deceleration time < 200 ms helps to differentiate CA from other causes of ventricular hypertrophy. In recent years, Pagourelias et al. [21] proposed that the ejection fraction strain ratio has the best CA differentiation effect (AUC 0.95; 95% CI 0.89–0.98). In the challenging subgroups (maximum wall thickness ≤ 16 mm and LVEF > 55%), ejection fraction strain ratio is still the best predictor of CA. Furthermore, Boldrini et al. [22] developed a scoring-based CA diagnostic model by analyzing morphological, functional, and strain-derived parameters. The results show that centripetal reconstruction and strain-derived parameters have the best diagnostic performance. The multivariate logistic regression model, which included relative wall thickness, E/e′, LS and tricuspid annular plane contraction deviation, had the best diagnostic effect on monoclonal immunoglobulin light chain amyloidosis (AUC 0.90; 95% CI 0.87–0.92).

The complexity of CA assessment has increased in terms of a large amount of data generated by supersonic inspection and increasing clinical variables. Traditional statistical analysis could only explore the relationships among limited variables and achieve a certain degree of predictive performance. However, in the era of big data, it is usually necessary to integrate abundant variables, which is a great challenge for clinicians. Therefore, we presented the study of differentiating CA and HCM based on ML. Combining

Table 3 Comparing machine learning models with Logistic regression

	AUC	Sensitivity	Specificity	True positive	False negative	Accuracy	p value*
LR	0.91	0.92	0.94	0.96	0.89	0.93	–
SVM	0.95	0.89	1	1	0.91	0.95	0.477
RF	0.97	0.96	1	1	0.94	0.98	0.301
XGBoost	0.98	0.88	0.95	0.93	0.91	0.92	0.230

LR logistic regression; *SVM* support vector machine; *RF* Random Forest; *XGBoost* extreme gradient boosting

*Comparison performed between machine learning models with traditional logistic regression

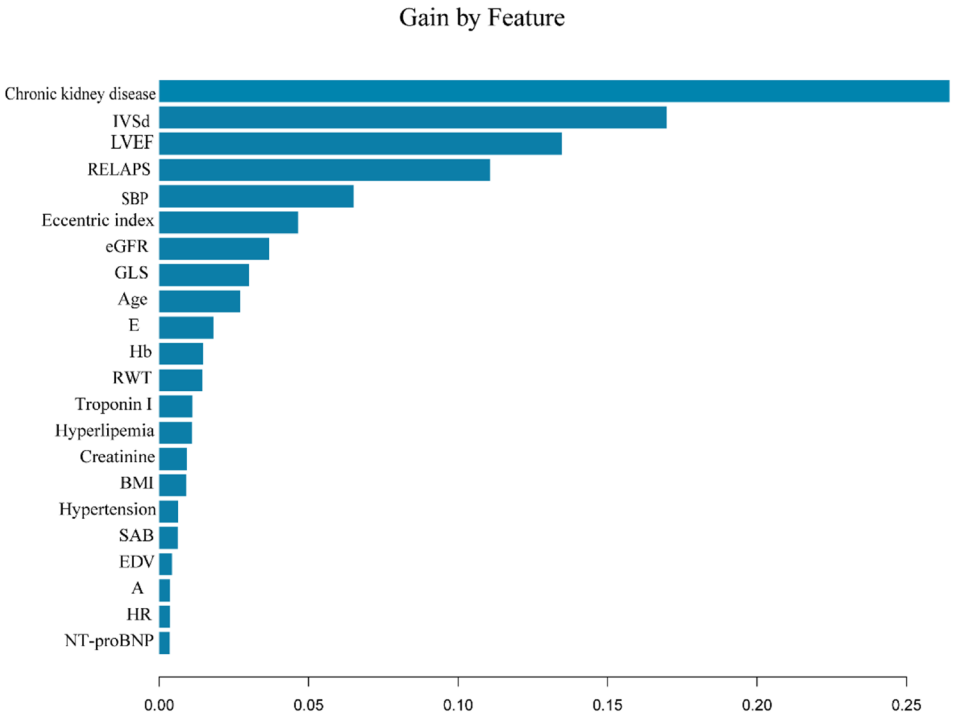

Fig. 2 Feature selection of XGBoost. Information gain (IG), a technique of feature selection, is defined as a metric of effective classification. It is measured in terms of the entropy reduction of the class, which reflects additional information about the class provided by the variables. Information gain of XGBoost combined clinical and echocardiography data, showing that chronic kidney disease (IG = 0.26) was the most important predictor, followed by Left ventricular septum thickness (IG = 0.17), ejection fraction (IG = 0.13), relative apical sparing (IG = 0.11), systolic blood pressure (IG = 0.07) and eccentricity index (IG = 0.05)

Table 4 The top 50 important features of the voting model

Feature	Rank	Feature	Rank	Feature	Rank	Feature	Rank	Feature	Rank
RadStrain3	1	BasalRotation3	11	LonStrainR10	21	CirStrainR8	31	CirStrainR11	41
LonStrainEpi7	2	LonStrainV8	12	PapiRotation6	22	PapiRotation2	32	GRSPAPI	42
CirStrainR4	3	LonStrainV13	13	LonStrainR1	23	CirStrainR16	33	LonStrainD2	43
LonStrain7	4	LonStrainEndo7	14	RadStrain4	24	GRSAPICAL	34	LonStrainEndo18	44
RadStrain2	5	LonStrainR7	15	PapiRotation3	25	CirStrain16	35	LonStrain13	45
RadStrain7	6	RadStrain8	16	ApicalRotationR4	26	CirStrainR13	36	CirStrainR17	46
LonStrainEpi16	7	LonStrainD9	17	LonStrainV14	27	CirStrainR5	37	LonStrain3	47
CirStrain10	8	RadStrainR12	18	LonStrainD10	28	RadStrain11	38	LonStrainR3	48
BasalRotation1	9	CirStrain4	19	LonStrainEpi15	29	CirStrain15	39	BasalRotation4	49
BasalRotation2	10	LonStrainEpi13	20	LonStrainV7	30	LonStrainD13	40	BasalRotationR3	50

The voting model that integrated the two algorithms of LightGBM and XGBoost used 2D-STE data to obtain the top 50 important features. The results indicated that RadStrain3 (the radial strain of the middle ventricular septum) was the most important predictor, followed by LongStrainEpi7 (the longitudinal strain of the anterior wall of the epicardial basement segment) and CirStrainR4 (the circumferential strain rate of the posterior wall of the basement segment)

clinical characteristics, routine echocardiography and 2D-STE data, support vector machine, random forest, and XGBoost manifested a favorable discriminative performance (AUC > 0.9). When based on 2D-STE data solely, the different gradient boosting models still performed well in the identification of CA patients. The voting model was more robust globally and superior to a single algorithm on some test sets. Although the difference in the AUC of ML was not statistically significant compared with the traditional logistic regression model ($p > 0.05$), it should be pointed out that this study was based on small sample data, and the performance of ML needs to be further discussed on larger data. Previously, Zhang et al. [23] employed ML to achieve automatic echocardiography interpretation. The algorithms can not only implement view recognition, image segmentation, structure, and function quantification but also realize automatic detection of CA, HCM, and pulmonary hypertension, which further reflects the effectiveness of ML. ML, a form of artificial intelligence that eliminates preassumptions, explores the unknown pattern with all useful data to avoid neglecting some important but not yet recognized predictors. Interestingly, ML also automatically identified traditional variables, such as ejection fraction, eccentricity index, and relative apical sparing, which further validates the potential

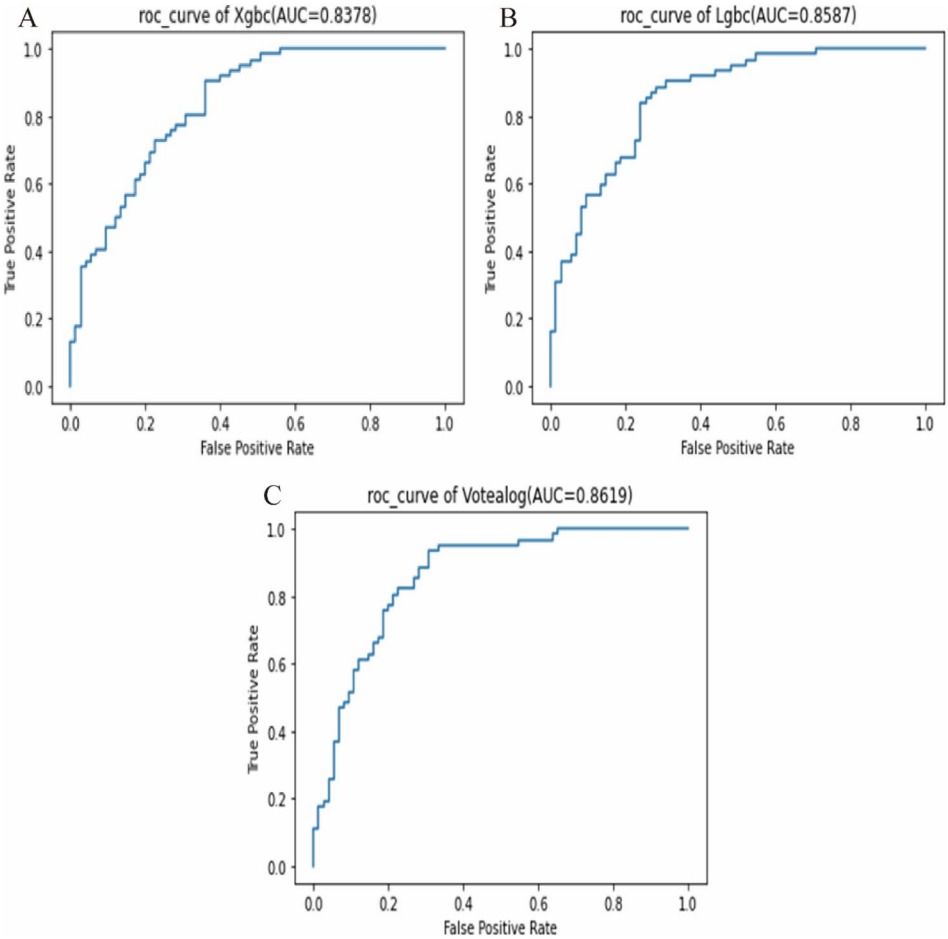

Fig. 3 The ROC curve of different gradient boosting models. Among all the three ML algorithms (XGBoost, LightGBM and voting model), the discriminant ability of voting model was similar to LightGBM (AUC of both were 0.86), while the AUC of XGBoost was slightly lower, which was 0.84. **A** XGBoost; **B** LightGBM; **C** Voting model

scalability and practicability. In addition, unexpected interactions between several weaker predictors would not be overlooked.

ML will not replace traditional statistical analysis, conversely, it provides a supplement and extension [24]. For rapidly growing data, ML explores non-linear patterns and automatically extracts important variables, thus simplifying feature selection and improving prediction, and facilitating the differentiation of similar phenotypic diseases. Also, ML seamlessly incorporates new data to continually update models and promote performance over time. Beyond that, ML possesses efficiency as it runs complex mathematical algorithms, such as gradient boosting, in a few seconds and produces easy-to-understand results with low variability and high accuracy.

Limitations of the study

There are some limitations to this study. First of all, the establishment of ML models was carried out in a small number of samples, and further validation needs to be conducted in a larger dataset. In addition, considering the imbalance of data among different areas, our study was single-center with certain specific population characteristics, further research needs to be trained and verified in multiple centers and regions to improve the generality of the models. Finally, our model was only evaluated in two-dimensional echocardiography with limited time and space, and further studies could include more ultrasonic sections or implement the models by other imaging methods. With the increment of samples, deep learning may improve the prediction of the models.

Conclusions

CA has a poor prognosis, and high mortality and is often misdiagnosed as HCM, leading to delayed diagnosis. If CA can be identified early and provided effective intervention in time, it is beneficial to improve the outcome for patients. We proposed intelligent identification of CA from HCM based on ML using 2D-STE data. The results indicated that the ML models had great discriminative performance, and could automatically integrate vast variables without making any preassumption, so as to identify the most important predictors. In the era of big data, automated ML will help

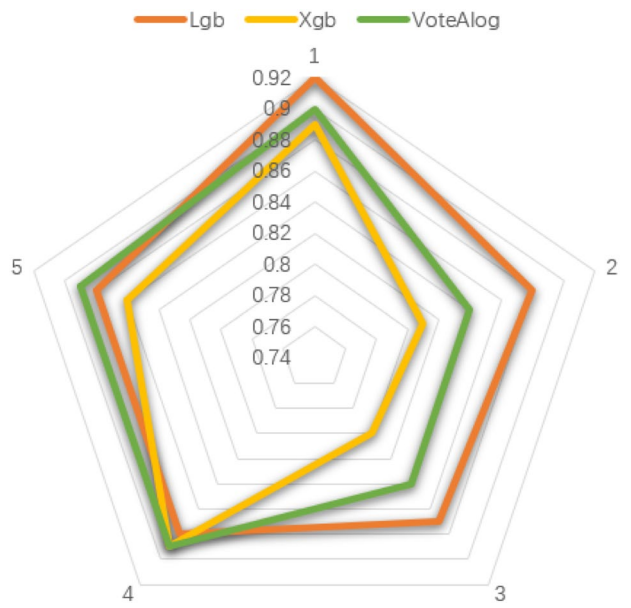

Fig. 4 Radar image of fivefold cross-validation of different gradient boosting. The mean AUC of LightGBM, XGBoost and voting models were 0.89 ± 0.19, 0.85 ± 0.43 and 0.87 ± 0.30, respectively. The voting model was globally more robust and outperformed to individual model on some test sets. Lgb (red): lightGBM; Xgb (yellow): XGBoost; Votealog (green): voting model

to identify patients with CA, so that timely and effective intervention can be carried out to improve the prognosis.

Acknowledgements The authors thank Xiao-Jun Chen for providing purely technical help.

Author contributions ZW as the first author collected data, drafted, and revised the manuscript. JZ and LK contributed to the acquisition, analysis, and interpretation of data. HY is the corresponding author, who performed the conception and design of the study, and the final approval of the submitted manuscript.

Funding This research did not receive any specific grant from funding: agencies in the public, commercial, or not-for-profit sectors.

Declarations

Competing interests The authors declare that they have no competing interests.

Ethical approval All of the procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The study was approved by the Clinical Research Ethics Committee of the First Affiliated Hospital, College of Medicine, Zhejiang University (No. IIT20200654A).

Consent for publication Consent for publication was obtained for every individual person's data included in the study.

Informed consent Informed consent was obtained from all of the individual participants included in the study.

References

1. Falk RH, Alexander KM, Liao R et al (2016) AL (light-chain) cardiac amyloidosis: a review of diagnosis and therapy. J Am Coll Cardiol 68(12):1323–1341. https://doi.org/10.1016/j.jacc.2016.06.053
2. Ruberg FL, Grogan M, Hanna M et al (2019) Transthyretin amyloid cardiomyopathy: JACC state-of-the-art review. J Am Coll Cardiol 73(22):2872–2891. https://doi.org/10.1016/j.jacc.2019.04.003
3. Wechalekar AD, Gillmore JD, Hawkins PN (2016) Systemic amyloidosis. Lancet (London, England) 387(10038):2641–2654. https://doi.org/10.1016/s0140-6736(15)01274-x
4. Phelan D, Collier P, Thavendiranathan P et al (2012) Relative apical sparing of longitudinal strain using two-dimensional speckle-tracking echocardiography is both sensitive and specific for the diagnosis of cardiac amyloidosis. Heart 98(19):1442–1448. https://doi.org/10.1136/heartjnl-2012-302353
5. Motwani M, Dey D, Berman DS et al (2017) Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. Eur Heart J 38(7):500–507. https://doi.org/10.1093/eurheartj/ehw188
6. Sanchez-Martinez S, Duchateau N, Erdei T et al (2018) Machine learning analysis of left ventricular function to characterize heart failure with preserved ejection fraction. Circ Cardiovasc Imaging 11(4):e007138. https://doi.org/10.1161/circimaging.117.007138
7. Narula S, Shameer K, Salem Omar AM et al (2016) Machine-learning algorithms to automate morphological and functional assessments in 2D echocardiography. J Am Coll Cardiol 68(21):2287–2295. https://doi.org/10.1016/j.jacc.2016.08.062
8. Sengupta PP, Huang YM, Bansal M et al (2016) Cognitive machine-learning algorithm for cardiac imaging: a pilot study for differentiating constrictive pericarditis from restrictive cardiomyopathy. Circ Cardiovasc Imaging. https://doi.org/10.1161/circimaging.115.004330
9. Awan SE, Bennamoun M, Sohel F et al (2019) Machine learning-based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics. ESC Heart Fail 6(2):428–435. https://doi.org/10.1002/ehf2.12419
10. Attia ZI, Noseworthy PA, Lopez-Jimenez F et al (2019) An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction. Lancet (London, England) 394(10201):861–867. https://doi.org/10.1016/s0140-6736(19)31721-0
11. Raj S, Ray KC (2018) A personalized arrhythmia monitoring platform. Sci Rep 8(1):11395. https://doi.org/10.1038/s41598-018-29690-2
12. Dorbala S, Ando Y, Bokhari S et al (2019) 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. J Card Fail 25(11):e1–e39. https://doi.org/10.1016/j.cardfail.2019.08.001
13. Ommen SR, Mital S, Burke MA et al (2020) 2020 AHA/ACC guideline for the diagnosis and treatment of patients with hypertrophic cardiomyopathy: executive summary: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice guidelines. Circulation 142(25):e533–e557. https://doi.org/10.1161/cir.0000000000000938
14. Churpek MM, Yuen TC, Winslow C et al (2016) Multicenter comparison of machine learning methods and conventional regression for predicting clinical deterioration on the wards. Crit Care Med 44(2):368–374. https://doi.org/10.1097/ccm.0000000000001571

15. Mortazavi BJ, Bucholz EM, Desai NR et al (2019) Comparison of machine learning methods with national cardiovascular data registry models for prediction of risk of bleeding after percutaneous coronary intervention. JAMA Netw Open 2(7):e196835. https://doi.org/10.1001/jamanetworkopen.2019.6835
16. Al'Aref SJ, Singh G, van Rosendael AR et al (2019) Determinants of in-hospital mortality after percutaneous coronary intervention: a machine learning approach. J Am Heart Assoc 8(5):e011160. https://doi.org/10.1161/jaha.118.011160
17. Sun JP, Stewart WJ, Yang XS et al (2009) Differentiation of hypertrophic cardiomyopathy and cardiac amyloidosis from other causes of ventricular wall thickening by two-dimensional strain imaging echocardiography. Am J Cardiol 103(3):411–415. https://doi.org/10.1016/j.amjcard.2008.09.102
18. Di Bella G, Minutoli F, Pingitore A et al (2011) Endocardial and epicardial deformations in cardiac amyloidosis and hypertrophic cardiomyopathy. Circ J 75(5):1200–1208. https://doi.org/10.1253/circj.cj-10-0844
19. Baccouche H, Maunz M, Beck T et al (2012) Differentiating cardiac amyloidosis and hypertrophic cardiomyopathy by use of three-dimensional speckle tracking echocardiography. Echocardiography 29(6):668–677. https://doi.org/10.1111/j.1540-8175.2012.01680.x
20. Liu D, Hu K, Niemann M et al (2013) Effect of combined systolic and diastolic functional parameter assessment for differentiation of cardiac amyloidosis from other causes of concentric left ventricular hypertrophy. Circ Cardiovasc Imaging 6(6):1066–1072. https://doi.org/10.1161/circimaging.113.000683
21. Pagourelias ED, Mirea O, Duchenne J et al (2017) Echo parameters for differential diagnosis in cardiac amyloidosis: a head-to-head comparison of deformation and nondeformation parameters. Circ Cardiovasc Imaging 10(3):e005588. https://doi.org/10.1161/circimaging.116.005588
22. Boldrini M, Cappelli F, Chacko L et al (2020) Multiparametric echocardiography scores for the diagnosis of cardiac amyloidosis. JACC Cardiovasc Imaging 13(4):909–920. https://doi.org/10.1016/j.jcmg.2019.10.011
23. Zhang J, Gajjala S, Agrawal P et al (2018) Fully automated echocardiogram interpretation in clinical practice. Circulation 138(16):1623–1635. https://doi.org/10.1161/circulationaha.118.034338
24. Shameer K, Johnson KW, Glicksberg BS et al (2018) Machine learning in cardiovascular medicine: are we there yet? Heart 104(14):1156–1164. https://doi.org/10.1136/heartjnl-2017-311198