IJC Heart & Vasculature 46 (2023) 101211

Contents lists available at ScienceDirect

IJC Heart & Vasculature

journal homepage: www.sciencedirect.com/journal/ijc-heart-and-vasculature

Identification of patients with dilated phase of hypertrophic cardiomyopathy using a convolutional neural network applied to multiple, dual, and single lead electrocardiograms☆

Naomi Hirota [a,*], Shinya Suzuki [a], Jun Motogi [b], Takuya Umemoto [b], Hiroshi Nakai [c], Wataru Matsuzawa [b], Tsuneo Takayanagi [b], Akira Hyodo [b], Keiichi Satoh [b], Takuto Arita [a], Naoharu Yagi [a], Mikio Kishi [a], Hiroaki Semba [a], Hiroto Kano [a], Shunsuke Matsuno [a], Yuko Kato [a], Takayuki Otsuka [a], Tokuhisa Uejima [a], Yuji Oikawa [a], Takayuki Hori [d], Minoru Matsuhama [d], Mitsuru Iida [d], Junji Yajima [a], Takeshi Yamashita [a]

[a] Department of Cardiovascular Medicine, The Cardiovascular Institute, Tokyo, Japan
[b] Nihon Kohden Corporation, Tokyo, Japan
[c] Information System Division, The Cardiovascular Institute, Tokyo, Japan
[d] Department of Cardiovascular Surgery, The Cardiovascular Institute, Tokyo, Japan

ARTICLE INFO

Keywords:
Dilated phase of hypertrophic cardiomyopathy
Artificial intelligence
Electrocardiography

ABSTRACT

Background: This study sought to develop an artificial intelligence-derived model to detect the dilated phase of hypertrophic cardiomyopathy (dHCM) on digital electrocardiography (ECG) and to evaluate the performance of the model applied to multiple-lead or single-lead ECG.
Methods: This is a retrospective analysis using a single-center prospective cohort study (Shinken Database 2010–2017, n = 19,170). After excluding those without a normal P wave on index ECG (n = 1,831) and adding dHCM patients registered before 2009 (n = 39), 17,378 digital ECGs were used. Totally 54 dHCM patients were identified of which 11 diagnosed at baseline, 4 developed during the time course, and 39 registered before 2009. The performance of the convolutional neural network (CNN) model for detecting dHCM was evaluated using eight-lead (I, II, and V1-6), single-lead, and double-lead (I, II) ECGs with the five-fold cross validation method.
Results: The area under the curve (AUC) of the CNN model to detect dHCM (n = 54) with eight-lead ECG was 0.929 (standard deviation [SD]: 0.025) and the odds ratio was 38.64 (SD 9.10). Among the single-lead and double-lead ECGs, the AUC was highest with the single lead of V5 (0.953 [SD: 0.038]), with an odds ratio of 58.89 (SD:68.56).
Conclusion: Compared with the performance of eight-lead ECG, the most similar performance was achieved with the model with a single V5 lead, suggesting that this single-lead ECG can be an alternative to eight-lead ECG for the screening of dHCM.

1. Introduction

Hypertrophic cardiomyopathy (HCM) is one of the most common genetic heart diseases. In most patients with HCM, their systolic function is normal and they are asymptomatic for a long time. However, a life-long process of left ventricular remodeling and progressive dysfunction can occur in a minority of HCM patients [1–3]. This stage of HCM, called dilated phase HCM (dHCM), was reported to be associated with a poor prognosis [4,5]. Because of the slowly evolving nature of HCM, the timely identification of patients with a risk of left ventricular dysfunction or heart failure might improve the poor prognosis [6]. Although it is useful to perform echocardiography, cardiac magnetic resonance, or genetic screening in the diagnosis of dHCM, these tests are rather costly for annual check-ups of all patients with HCM. Therefore, there is a need to develop readily available and cost-effective methods for screening patients with dHCM. Electrocardiogram (ECG), which is a non-invasive

☆ **Grant support:** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
* Corresponding author at: Department of Cardiovascular Medicine, The Cardiovascular Institute, 3-2-19 Nishiazabu, Minato-Ku, Tokyo 106-0031, Japan.
E-mail address: n-hirota@cvi.or.jp (N. Hirota).

https://doi.org/10.1016/j.ijcha.2023.101211
Received 26 January 2023; Received in revised form 27 March 2023; Accepted 13 April 2023

and readily available test, is widely performed for the detection and management of cardiac diseases. Recently, mobile and smartwatch ECG technologies using artificial intelligence (AI) were developed [7,8]. Mobile ECG system are generally a more simple and more available method for physicians in the clinic, regardless of whether or not they are a cardiologist.

In the present study, we developed CNN models to detect dHCM from multiple-lead or single-lead ECG, with the aim of increasing the knowledge on the difference in the performance of CNN models using different ECG leads. For this work, we used a single-center ECG database.

2. Methods

2.1. Ethics and informed consent

This study was performed in accordance with the Declaration of Helsinki (revised in 2013) and the Ethical Guidelines for Medical and Health Research Involving Human Subjects (Public Notice of the Ministry of Education, Culture, Sports, Science and Technology, and the Ministry of Health, Labour and Welfare, Japan, issued in 2017). Written informed consent was obtained from all participants. The study protocol was reviewed by the Institutional Review Board of the Cardiovascular Institute.

2.2. Study population

The Shinken database includes all patients who newly visited the Cardiovascular Institute, Tokyo, Japan, except for foreign travelers and patients with active cancer. This single-hospital database was established in June 2004, and details of this database have been described elsewhere [9]. In the present study, 19 170 subjects registered between February 2010 and March 2018 were extracted from the Shinken database, because it has included a computerized electrocardiogram database since February 2010. We excluded 1831 subjects without a normal P wave (P wave duration = 0 on the automatic measurement by the GE system) because we put priority on existence of the P wave in the disease prediction model.

Due to a too small number of patients who were diagnosed as dHCM at baseline in the original database (dHCM1), we added 43 patients with dHCM as follows: (1) four patients who were registered as having HCM at baseline in the original database and progressed to dHCM later during the time course (dHCM2; for these patients, the index ECG was chosen on the day of or on the nearest day after the ultrasound cardiogram in which left ventricular systolic dysfunction was firstly observed), and (2) 39 patients with a diagnosis of dHCM who were included in our database before 2009 and with available ECG recorded between February 2010 and March 2018 (dHCM3). Consequently, ECG data from 17 378 subjects (54 dHCM, 17 324 non-dHCM) were included in the present study.

2.3. Diagnosis of dHCM

The diagnostic definition of dHCM was left ventricular systolic dysfunction (defined as left ventricular ejection fraction [LVEF] $< 50\%$) with a diagnosis of HCM, which included (1) intraventricular septal thickness (IVST) ≥ 15 mm without other causes of left ventricular hypertrophy; (2) IVST ≥ 13 mm and a family history of HCM; and (3) hypertrophy in the apex of the left ventricle.

2.4. ECG recording

Twelve-lead ECG was recorded for 10 s in the supine position using an ECG machine (GE CardioSoft V6.71 and MAC 5500 HD; GE Healthcare, Chicago, IL, USA) at a sampling rate of 500 Hz, and raw data of the digital recordings were stored using the MUSE data management system.

2.5. Dataset management

In the present study, given the small number of the positive cases (dHCM), we employed the five-fold cross validation method to enable all data to be included in the testing dataset [10]. Management of the dataset with the five-fold cross validation method was shown in Supplementary Fig. 1 and was as follows. First, the dataset was randomly divided into five groups. Second, one of the five groups was set as the testing dataset, and the others were set as the training dataset in which 12.5% of the data were used as the internal-validation dataset. Third, the model was run five times using different combinations of training and testing datasets. Accordingly, model output was obtained from five testing datasets of five different models, in which all data were included in the testing dataset.

2.6. Convolutional neural network (CNN) modeling

We used a convolutional neural network (CNN) to develop an AI-enabled ECG analysis system to predict dHCM from sinus rhythm. The CNN was constructed using the Keras framework with a Tensorflow (Google; Mountain View, CA, USA) backend and Python.

Of the eight physical leads and four augmented leads with a 10-second duration on the 12-lead ECG recordings, we selected the eight independent leads (leads I, II, and V1–6) with a 10-second duration.

The CNN model in the present study was constructed based on the model by Attia et al. [11–13]. The conceptual architecture of the CNN model is shown in Fig. 1 and the detail architecture of the CNN model is shown in Supplementary Fig. 2. This model included layers for a temporal axis and a lead axis. The layers for the temporal axis were composed of two parts: a convolution part and a residual part. The convolution part included a convolution layer, a batch-normalization layer, a layer for non-linear Rectified Linear Unit (ReLU) activation, and a maximum pooling layer. The residual part included a combination of two residual blocks based on the Residual Network (ResNet) and average pooling, which was repeated N times, with the value of N being tuned to obtain the best performance (the method is outlined below). The layers for the lead axis were composed of a paired batch-normalization layer and a layer for non-linear ReLU activation, followed by a convolution layer. Thereafter, a second paired batch-normalization layer and a layer for non-linear ReLU activation were included. Finally, the data were fed to a dropout layer with global average pooling and to the final output layer activated by the softmax function, which generated the probability of dHCM.

The model was trained using the Keras software library on a computer with 128 GB RAM and a single Quadro P-2200 (NVIDIA) graphics processing unit. In order to avoid over-training, we applied the early stopping while model training. The training was stopped if the loss was not decreased in 50 epochs in the internal-validation dataset, and employed the model with the lowest loss. Considering the class imbalance between the positive and the negative cases, we set the weight of the loss function as 320 times higher (based on the ratio of the number of negative to positive data in the training dataset) for the positive class samples as for the negative class samples.

A receiver operating characteristics (ROC) curve was created and the area under the curve (AUC) was used to evaluate the ability of the CNN model with ECG to determine whether dHCM was present or not. Using the ROC curve in the internal-validation dataset, we tuned the number of repetitions for the combination of the two residual blocks and average pooling described above (N). We determined the probability threshold for dHCM as the point on the ROC curve closest to the (0,1) point [14] in the internal-validation dataset.

2.7. Outcome measurement and statistical analysis

First, the performance of the CNN model applied to eight-lead (I, II, V1-6), single-lead, and double-lead (I, II) ECGs was assessed using AUC,

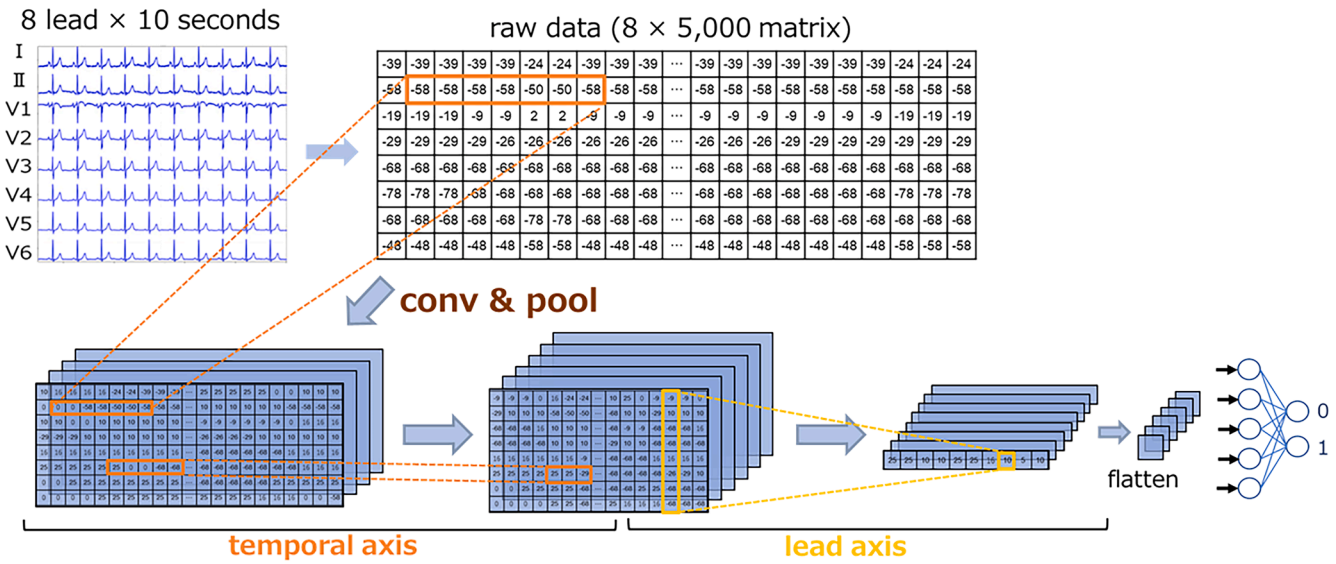

Fig. 1. Convolutional neural network (CNN) analysis.

sensitivity, specificity, accuracy, and F1 score. Second, the model output was described as the proportion of the patients in dHCM and non-dHCM group per probability, according to the levels of diagnostic probability yielded by the CNN model. The proportion of dHCM was separately described by the three subtypes according to the different time course (dHCM1, 2, and 3). Third, the odds ratios (=[the ratio of true/false positives] / [the ratio of false/true negatives]) were described according to eight-lead, single-lead, and double-lead ECGs. Fourth, we used the gradient-weighted class activation mapping (Grad-CAM) method [15] for the multi-input models.

The patient characteristics were summarized as mean (standard deviation [SD]) and n (%), and the differences of two groups were tested by unpaired *t*-test and Chi-squared test for continuous and categorical variables, respectively. The data of model performance and the odds ratios were presented as the mean (SD) of 5 model runs with five-fold cross validation. The statistical analyses were performed using R version 4.0.3 (The R Foundation, Vienna, Austria) and SPSS version 28.0. (IBM Corp., Armonk, NY, USA).

3. Results

3.1. Patient characteristics

The patient characteristics are listed in Table 1. Of a total of 17 378 included subjects, 54 were classified to the dHCM group and 17 324 to the non-dHCM group. The mean (SD) of ages were 69.1 (13.5) years for the dHCM group and 57.5 (15.6) years for the non-dHCM group ($p < 0.001$), with 42 men (77.8%) in the dHCM group and 10 120 men (58.4%) in the non-dHCM group ($p = 0.004$). The mean left ventricular ejection fractions in the dHCM and non-dHCM groups were 40.3% (SD 10.6%) and 66.2% (SD 9.7%), respectively ($p < 0.001$). The prevalences of congestive heart failure (admission within 90 days) in the dHCM and non-dHCM groups were 14.8% and 1.6%, respectively ($p < 0.001$).

3.2. Evaluation of the CNN model to detect dHCM using the test dataset

The performance of the CNN model to detect dHCM deriving from the training dataset was evaluated in the test dataset (n = 3477).

3.2.1. Basic performance of the CNN model

The basic performance of the CNN model for detecting dHCM on eight-lead, single-lead, and double-lead ECGs is shown in Table 2. The mean (SD) of AUCs of the 5 model runs were 0.929 (0.025) for eight-lead ECG and 0.897 (0.056) for double-lead ECG, with the AUCs for single-lead ECGs being around 0.9, and the highest for the single-lead of V5 for which the AUC (SD) was 0.953 (0.038).

The odds ratios for dHCM in the CNN models for eight-lead, single-lead, and double-lead ECGs are shown in Table 3. The CNN model with eight-lead ECG diagnosed dHCM with a mean odds ratio (SD) of 38.64 (9.10). The mean odds ratios for dHCM with single-lead and double-lead ECGs were lower than that of eight-lead ECG, except for V4 and V5 with the odds ratios (SD) of 45.79 (12.57) and 58.89 (65.56), respectively.

3.2.2. Distribution of the patients according to the CNN model outputs

The proportions of patients in the CNN model outputs are shown in Figs. 2 and 3. For the model with eight-lead ECG (Fig. 2), the proportion of patients in the dHCM group sharply increased with higher probability (model output > 0.9), and inversely, those in the non-dHCM group sharply increased with lower probability (model output < 0.1). A similar distribution was observed with the model with single-lead or double-lead ECGs, especially with I, V3, V4, and V5 leads (Fig. 3). The distribution was mostly similar among three dHCM subtypes with the different time course (dHCM1, 2, and 3; Figs. 2 and 3).

3.2.3. GradCAM for the diagnosis of dHCM in the CNN models

The locations on which the CNN model focused indicated by the GradCAM are displayed in Figs. 4 and 5. In Fig. 4, the images of GradCAM on the eight-lead ECG with 5 CNN model runs with a true positive result in a dHCM patient are displayed. In Fig. 5, the images of GradCAM on single-lead and double-lead ECGs with 5 CNN model runs in the same patient are displayed. The results of the GradCAM in Figs. 4 and 5 similarly indicated that the CNN model focused mainly on the QRS segment and partially on the ST-T segment and seemed to obtain the diagnostic information from any leads of the ECG.

4. Discussion

4.1. Major findings

We evaluated the performance of a CNN model for detecting dHCM applied to eight-lead, single-lead, and double-lead ECGs. Using eight-lead ECG, the diagnostic performance was summarized by an AUC (SD) of 0.929 (0.025) and an odds ratio (SD) of 38.64 (9.10). For the single-lead and double-lead ECGs, the AUC (SD) was the highest when using the single lead of V5 (0.953 [0.038] with an odds ratio (SD) of 58.89 (65.56).

Table 1
Patient characteristics.

	Total N = 17 378	dHCM N = 54	non-dHCM N = 17 324	P-value
Age, yeas	57.5 ± 15.7	69.1 ± 13.5	57.5 ± 15.6	<0.001
Male, n (%)	10 162 (58.5)	42 (77.8)	10 120 (58.4)	0.004
Height, cm	163.7 ± 9.4	164.0 ± 8.8	163.7 ± 9.4	0.390
Weight, kg	62.6 ± 13.7	65.0 ± 15.0	62.6 ± 13.7	0.104
BMI, kg/m^2	23.2 ± 3.9	24.0 ± 4.2	23.2 ± 3.9	0.083
Systolic BP, mmHg	127.8 ± 22.5	128.5 ± 30.5	127.8 ± 22.5	0.452
Diastolic BP, mmHg	75.5 ± 13.5	79.0 ± 20.8	75.5 ± 13.5	0.228
IVST, mm	9.6 ± 2.2	13.7 ± 3.6	9.6 ± 2.2	<0.001
PWT, mm	9.0 ± 1.6	10.8 ± 3.1	8.9 ± 1.5	<0.001
LVDd, mm	46.0 ± 5.6	54.0 ± 8.9	45.9 ± 5.6	<0.001
LVDs, mm	29.2 ± 6.2	42.0 ± 10.2	29.1 ± 6.2	<0.001
LVEF, %	66.1 ± 9.8	40.3 ± 10.6	66.2 ± 9.7	<0.001
LAD, mm	34.6 ± 6.3	47.6 ± 8.7	34.6 ± 6.2	<0.001
Congestive heart failure, n (%) (Heart failure admission within 90 days)	279 (1.6)	8 (14.8)	271 (1.6)	<0.001
Heart failure with reduced EF, n (%)	853 (4.9)	28 (51.9)	825 (4.8)	<0.001
Ischemic heart disease, n (%) (PCI within 90 days)	1788 (10.3)	2 (3.7)	1786 (10.3)	0.172
Asymptomatic ischemia, n (%)	407 (2.3)	1 (1.9)	406 (2.3)	1.000
Old myocardial infarction, n (%)	411 (2.4)	0 (0)	411 (2.4)	0.640
Acute coronary syndrome, n (%)	589 (3.4)	0 (0)	589 (3.4)	0.264
Aortic stenosis, n (%)	526 (3.0)	1 (1.9)	525 (3.0)	0.999
Aortic regurgitation, n (%)	369 (2.1)	2 (3.7)	367 (2.1)	0.318
Mitral regurgitation, n (%)	375 (2.2)	5 (9.3)	370 (2.1)	0.006
Mitral stenosis, n (%)	49 (0.3)	0 (0)	49 (0.3)	0.999
Tricuspid regurgitation, n (%)	207 (1.2)	6 (11.1)	201 (1.2)	<0.001
Dilated cardiomyopathy, n (%)	111 (0.6)	0 (0)	111 (0.6)	0.999
Hypertrophic cardiomyopathy, n (%)	152 (0.9)	54 (100)	98 (0.6)	<0.001
Hypertensive cardiomyopathy, n (%)	1905 (11.0)	0 (0)	1905 (11.0)	0.003
Ischemic cardiomyopathy, n (%)	211 (1.2)	0 (0)	211 (1.2)	1.000
Aortic aneurism, n (%)	271 (1.6)	3 (5.6)	268 (1.5)	0.051
Aortic dissection, n (%)	135 (0.8)	0 (0)	135 (0.8)	0.999
Hypertension, n (%)	7496 (43.1)	24 (44.4)	7472 (43.1)	0.893
Diabetes, n (%)	2057 (11.8)	6 (11.1)	2051 (11.8)	1.000
Smoking history, n (%)	7238 (41.7)	26 (48.1)	7212 (41.6)	0.337
Chronic kidney disease, n (%)	2695 (15.5)	29 (53.7)	2666 (15.4)	<0.001
Paroxysmal AF, n (%)	1459 (8.4)	30 (55.6)	1429 (8.2)	<0.001

Data are presented as mean ± SD unless otherwise stated.
dHCM, dilated phase of hypertrophic cardiomyopathy; BMI, body mass index; BP, blood pressure; IVST, intraventricular septum thickness; PWT, posterior left ventricular wall thickness; LVDd, left ventricular end-diastolic diameter; LVDs, left ventricular end-systolic diameter; LVEF, left ventricular ejection fraction; LAD, left atrial diameter; EF, ejection fraction; PCI, percutaneous coronary intervention; AF, atrial fibrillation.

4.2. Comparison with previous studies

Many studies have focused on using ECG for the detection of cardiomyopathy (including hypertrophic cardiomyopathy) [16–18], and recently, the number of such studies using AI applications has been growing [19,20]. However, we are only aware of one study evaluating the ECG characteristics of patients with dHCM, which was performed 1999, and which showed that a gradual decrease in the amplitude of the S wave in V1 plus the R wave in V5 during 20-year follow up periods was associated with ventricular tachycardia or a poor prognosis [21]. To the best of our knowledge, our study is the first to use an AI algorithm to detect dHCM from snapshot ECG.

In our study, the AUC for detecting dHCM using eight-lead ECG was 0.929 and the odds ratio was 38.64. A high AUC generally indicates that a model has a good differentiation threshold with high sensitivity and high specificity, while a high odds ratio generally indicates that the threshold shows a high positive predictive value and a low false-positive rate. However, caution is needed when interpreting the odds ratio of our CNN model, because in our results the ratio of true/false positives was very low (number of true positive = 7 + 7 + 8 + 5 + 8 = 35; number of false positive = [3465 –3313] + [3465 –3316] + [3465 –3132] + [3465 –3382] + [3464 –3291] = 890; ratio = 35 / 890 = 0.039: calculated from Table 3), while the ratio of false/true negatives was extremely low (number of false negative = [11 –7] + [11 –7] + [11 –8] + [10 –5] + [11 –8] = 19; number of true negative = 3313 + 3316 + 3132 + 3382 + 3291 = 16434; ratio = 19 / 16434 = 0.0011: calculated from Table 3), and the latter therefore contributed strongly to the high odds ratio.

4.3. Model performance according to differences in lead application

The ECG change in a specific lead might be affected by the electrical or structural condition of the myocardium around the lead position [22,23]. For example, in the diagnosis of ST-segment elevation myocardial infarction, the elevation of the ST segment on ECG in a specific lead can be useful for predicting the infarction area. In the case of occlusion in the proximal left anterior descending artery, the elevation of the ST segment on ECG is generally observed in V1, V2, V3, V4, I, and/or aVL, which is expanded to V3, V4, V5, V6, I, II, and/or aVF in the case of occlusion in the mid-to-distal left anterior descending artery [24]. As such, the lead-dependent changes generally reflect the structurally-damaged sites in heart. Based on such understanding, we evaluated the performance of the CNN model on eight-lead (I, II, V1-6), single-lead, and double-lead (I, II) ECGs, and this study is therefore the first study to evaluate the performance of CNN models according to different ECG leads for dHCM.

Compared with the performance of eight-lead ECG (AUC 0.929), the performance was mostly similar in any single-lead or double-lead ECGs (AUC was close to 0.9), suggesting that single-lead ECGs can be alternatives to eight-lead ECG for the screening of dHCM. According to the GradCAM, CNN models focused mainly on the QRS segment and partially on the ST-T segment and seemed to obtain the diagnostic information from any leads of the ECG. This suggests that the features of dHCM, generally accompanied by the left ventricular hypertrophy and the enlargement of left ventricle, would primarily be detected by the CNN-enabled ECG from any directions.

Among the models with single-lead or double-lead ECGs, the AUC was the highest in the model with a single V5 lead (AUC 0.953). Lead V5 is positioned on the left lateral side of the heart, and represents the myocardial condition in the mid to apical anterior to lateral wall of the left ventricle. The common left ventricular hypertrophic change in HCM patients is represented by a deep S wave in V1 or high R wave in V5 or V6 [25]. In the progression phase from HCM to dHCM, irreversible change such as systolic dysfunction and left ventricular dilatation (which resembles dilated cardiomyopathy) is observed, which can especially be identified in V5 to V6 leads.

Table 2
Performance of the AI-enabled CNN algorithm for detecting dHCM on eight-lead, single-lead, and double-lead ECGs.

Model pattern	Leads		AUC	Sensitivity	Specificity	F1 score	Accuracy
All leads (8 leads)		Mean	0.929	0.645	0.949	0.080	0.948
		SD	0.025	0.093	0.027	0.021	0.027
Single lead	I	Mean	0.914	0.682	0.908	0.053	0.907
		SD	0.037	0.164	0.053	0.022	0.052
Single lead	II	Mean	0.878	0.667	0.878	0.034	0.878
		SD	0.019	0.101	0.031	0.007	0.031
Single lead	V1	Mean	0.879	0.627	0.892	0.041	0.891
		SD	0.028	0.152	0.061	0.016	0.061
Single lead	V2	Mean	0.868	0.665	0.852	0.030	0.851
		SD	0.055	0.168	0.065	0.010	0.065
Single lead	V3	Mean	0.869	0.653	0.878	0.042	0.878
		SD	0.031	0.194	0.087	0.023	0.087
Single lead	V4	Mean	0.906	0.700	0.914	0.082	0.914
		SD	0.049	0.175	0.080	0.053	0.079
Single lead	V5	Mean	0.953	0.518	0.967	0.114	0.966
		SD	0.038	0.175	0.025	0.073	0.024
Single lead	V6	Mean	0.921	0.669	0.920	0.051	0.920
		SD	0.021	0.185	0.029	0.010	0.028
Double leads	I, II	Mean	0.897	0.718	0.901	0.053	0.901
		SD	0.056	0.145	0.072	0.020	0.072

The mean and SD of the performance values of 5 model runs by five-fold cross validation are displayed for models of all-lead ECG, single-lead ECGs, or double-lead ECG. Detail information of the 5 model runs are shown in Supplementary Table 1.
AUC, area under the curve; SD, standard deviation.

4.4. Clinical implications of our CNN model

As we mentioned above, our study is the first to use an AI algorithm to detect dHCM from snapshot ECG. However, its clinical implications should be carefully interpreted by two critical viewpoints: the diagnostic accuracy and the small number of positive cases.

In view of the diagnostic accuracy of our CNN model, we need to pay very close attention to that there were many patients with false positive results for dHCM diagnosis; the number of cases for whom the CNN model judged as dHCM was 925 in the model with all eight-lead ECG (In model 1, n = 7 + [3465 –3313] = 159; In total of 5 models, n = 159 + 156 + 341 + 88 + 181 = 925: calculated from the data in Table 3), while the number of true positive cases were 35 (=7 + 7 + 8 + 5 + 8; positive predictive rate, 4%). Meanwhile, as 35 cases were totally diagnosed by our CNN models out of 54 true dHCM cases (total of 5 models: calculated from Table 3), the total sensitivity can be calculated as 0.648. Therefore, the CNN model in our study can contribute as a screening tool for dHCM with a high odds ratio (38 times), a low positive predictive rate (4%), and a moderate sensitivity (65%).

In view of the small number of positive cases in our dataset, we have two discussions. First, considering the small number of positive cases in the large dataset, we employed in our CNN model the weighting of judgement for the characteristics of the positive cases. We applied loss function weighting when training the model, to address class imbalance in the dataset. This is generally used as a countermeasure against class imbalance in a dataset and is a method of learning models to pay more attention to the minority class by making the weight of the loss function of the minority class larger than that of the majority class. In this study, the weight of the loss function was set as 320 times higher (the number of training data / the number of positive classes in the training data ≈ 320) for the positive class samples as for the negative class samples. Second, the impact of our model can be roughly estimated based on the estimation of the number of dHCM patients in Japan. Miura et al. reported that the number of HCM patients in Japan was estimated to be 21 900 (95% confidence interval, 20 000 to 23 200)[26]. HCM patients has been reported to develop to dHCM by 5.3 per 1000 patient-years[4]. Then, simply talking, 116 HCM patients are estimated to newly progress to dHCM in Japan annually. Therefore, the number of dHCM in Japan is inherently small in total, and it is difficult to make a large-sized ECG dataset of dHCM patients. It is unclear whether our model deriving from 54 dHCM patients can be applied to a majority of dHCM patients in Japan or can cover only a minority of them. It is a future task to increase the number of positive cases in our model.

4.5. Future perspectives

Our study confirmed the high performance of a CNN algorithm for detecting dHCM on eight-lead, single-lead, and double-lead ECGs. The AUCs for single leads (I, V4, V5, and V6) and double leads (I, II) were comparable to that for all eight leads. Our model could be helpful for screening of patients with dHCM, especially in situations where echocardiography is unavailable, facilitating consultation to the cardiologist. The ECG with a smaller number of leads is generally regarded as using leads I or II, which can be recorded easily and would be useful especially when the patient has difficulty in undressing or ambulation to the bed in the clinic. Given the ease of recording, it would also be useful in mass screening. This would expand the possibility of using a mobile ECG with a single lead for screening dHCM. On the other hand, recording the lead V5 is somewhat inconvenient. Strictly speaking, V5 is not a single lead which requires the Wilson's central terminal (determined by three potentials) as a reference potential. Therefore, further studies investigating the use of lead V5 for screening dHCM in daily clinical practice will be necessary. These studies should, for example, assess whether a similar performance to the model using lead V5 can be achieved with CM5 or CC5, which are commonly used in ambulatory ECG recordings.

4.6. Limitations

There are several limitations to this study. First, this was a single-center study, and all participants were patients who visited a cardiovascular hospital in an urban area in Japan. Due to this selection bias, the application of our models might be limited only to Japanese population. Second, the number of entries in the training dataset was relatively small. Third, although we separated data for the entire cohort into a training dataset and test dataset to develop the models for internal validation, our model was not validated in an external cohort. Therefore, our findings should be validated in other populations from different hospitals or in the general population, or in large registries including Japanese patients with dHCM. Fourth, we chose patients with a sinus rhythm because we put priority on existence of the P wave in the disease prediction model. Actually, by the exclusion criteria in the present study, seven patients with dHCM were excluded from the analysis.

Table 3
The odds ratios for detecting dHCM with the AI-enabled CNN algorithm on eight-lead, single-lead, and double-lead ECGs.

Model pattern	Leads	5-fold models	Sensitivity	Specificity	Odds ratio
All leads (8 leads)		Model 1	64 (7/11)	96 (3313/3465)	37.96
		Model 2	64 (7/11)	96 (3316/3465)	38.89
		Model 3	73 (8/11)	90 (3132/3465)	25.08
		Model 4	50 (5/10)	98 (3382/3465)	40.67
		Model 5	73 (8/11)	95 (3291/3464)	50.60
		Mean			38.64
		SD			9.10
Single lead	I	Model 1	55 (6/11)	96 (3340/3465)	32.07
		Model 2	91 (10/11)	88 (3035/3465)	70.57
		Model 3	73 (8/11)	84 (2924/3465)	14.41
		Model 4	50 (5/10)	96 (3326/3465)	24.00
		Model 5	73 (8/11)	89 (3097/3464)	22.46
		Mean			32.70
		SD			22.08
Single lead	II	Model 1	64 (7/11)	84 (2921/3465)	9.38
		Model 2	64 (7/11)	91 (3160/3465)	18.11
		Model 3	82 (9/11)	86 (2966/3465)	26.72
		Model 4	70 (7/10)	87 (3021/3465)	15.90
		Model 5	55 (6/11)	91 (3149/3464)	11.96
		Mean			16.41
		SD			6.68
Single lead	V1	Model 1	46 (5/11)	96 (3330/3465)	20.57
		Model 2	73 (8/11)	81 (2821/3465)	11.65
		Model 3	64 (7/11)	90 (3112/3465)	15.38
		Model 4	50 (5/10)	94 (3247/3465)	14.87
		Model 5	82 (9/11)	85 (2937/3464)	25.07
		Mean			17.51
		SD			5.30
Single lead	V2	Model 1	91 (10/11)	79 (2734/3465)	37.35
		Model 2	46 (5/11)	91 (3160/3465)	8.65
		Model 3	64 (7/11)	78 (2685/3465)	6.02
		Model 4	60 (6/10)	88 (3049/3465)	11.00
		Model 5	73 (8/11)	90 (3131/3464)	25.08
		Mean			17.62
		SD			13.26
Single lead	V3	Model 1	36 (4/11)	95 (3285/3465)	10.43
		Model 2	64 (7/11)	95 (3278/3465)	30.61
		Model 3	73 (8/11)	75 (2602/3465)	8.03
		Model 4	90 (9/10)	92 (3195/3465)	106.38
		Model 5	64 (7/11)	82 (2854/3464)	8.18
		Mean			32.73
		SD			42.25
Single lead	V4	Model 1	73 (8/11)	94 (3257/3465)	41.72
		Model 2	55 (6/11)	98 (3392/3465)	55.84
		Model 3	91 (10/11)	80 (2758/3465)	38.98
		Model 4	50 (5/10)	98 (3410/3465)	61.50
		Model 5	82 (9/11)	87 (3024/3464)	30.90
		Mean			45.79
		SD			12.57
Single lead	V5	Model 1	64 (7/11)	93 (3216/3465)	22.52
		Model 2	73 (8/11)	99 (3413/3465)	174.87
		Model 3	46 (5/11)	97 (3347/3465)	23.72
		Model 4	50 (5/10)	96 (3340/3465)	26.78
		Model 5	27 (3/11)	99 (3436/3464)	46.56
		Mean			58.89
		SD			65.56
Single lead	V6	Model 1	46 (5/11)	96 (3337/3465)	21.73
		Model 2	91 (10/11)	89 (3080/3465)	80.00
		Model 3	55 (6/11)	93 (3236/3465)	16.95
		Model 4	80 (8/10)	90 (3122/3465)	36.40
		Model 5	64 (7/11)	92 (3170/3464)	18.81
		Mean			34.78
		SD			26.41
Double leads	I, II	Model 1	64 (7/11)	93 (3229/3465)	23.95
		Model 2	82 (9/11)	91 (3164/3465)	47.17
		Model 3	82 (9/11)	78 (2685/3465)	15.48
		Model 4	50 (5/10)	95 (3281/3465)	17.87
		Model 5	82 (9/11)	94 (3256/3464)	70.41
		Mean			34.98
		SD			23.44

SD, standard deviation.

Therefore, our model cannot be applied to patients with atrial fibrillation on ECG, which would be a topic for further analysis in a future study. Finally, our model was developed using ECG data only, and the patients' characteristics such as cardiac anatomical information, comorbidities, concomitant medications, and frailty, were not included.

4.7. Conclusions

We evaluated the performance of a CNN model for detecting dHCM on eight-lead, single-lead, and double-lead ECGs. Compared with the performance of eight-lead ECG, the performance of the model with a single V5 lead was largely similar, suggesting that this single-lead ECG can be an alternative to eight-lead ECG for the screening of dHCM.

Declarations

Consent for publication

Not applicable.

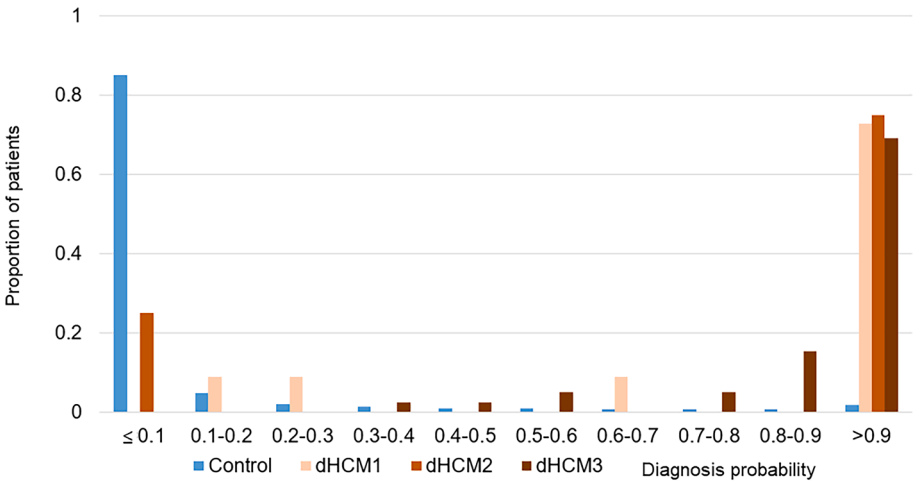

Fig. 2. The proportion of patients according to the model output in CNN-derived models using eight-lead ECGs The vertical scale indicates the proportion of the patients in dHCM group (orange) and non-dHCM group (blue). The horizontal scale indicates the diagnostic probability for dHCM yielded by CNN model. dHCM1 indicates patients diagnosed as dHCM at baseline in the original database (n = 11). dHCM2 indicates patients registered as having HCM at baseline in the original database and progressed to dHCM later during the time course (n = 4). dHCM3 indicates patients with a diagnosis of dHCM who were registered to our database before 2009 and with available ECG recorded between February 2010 and March 2018 (n = 39). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

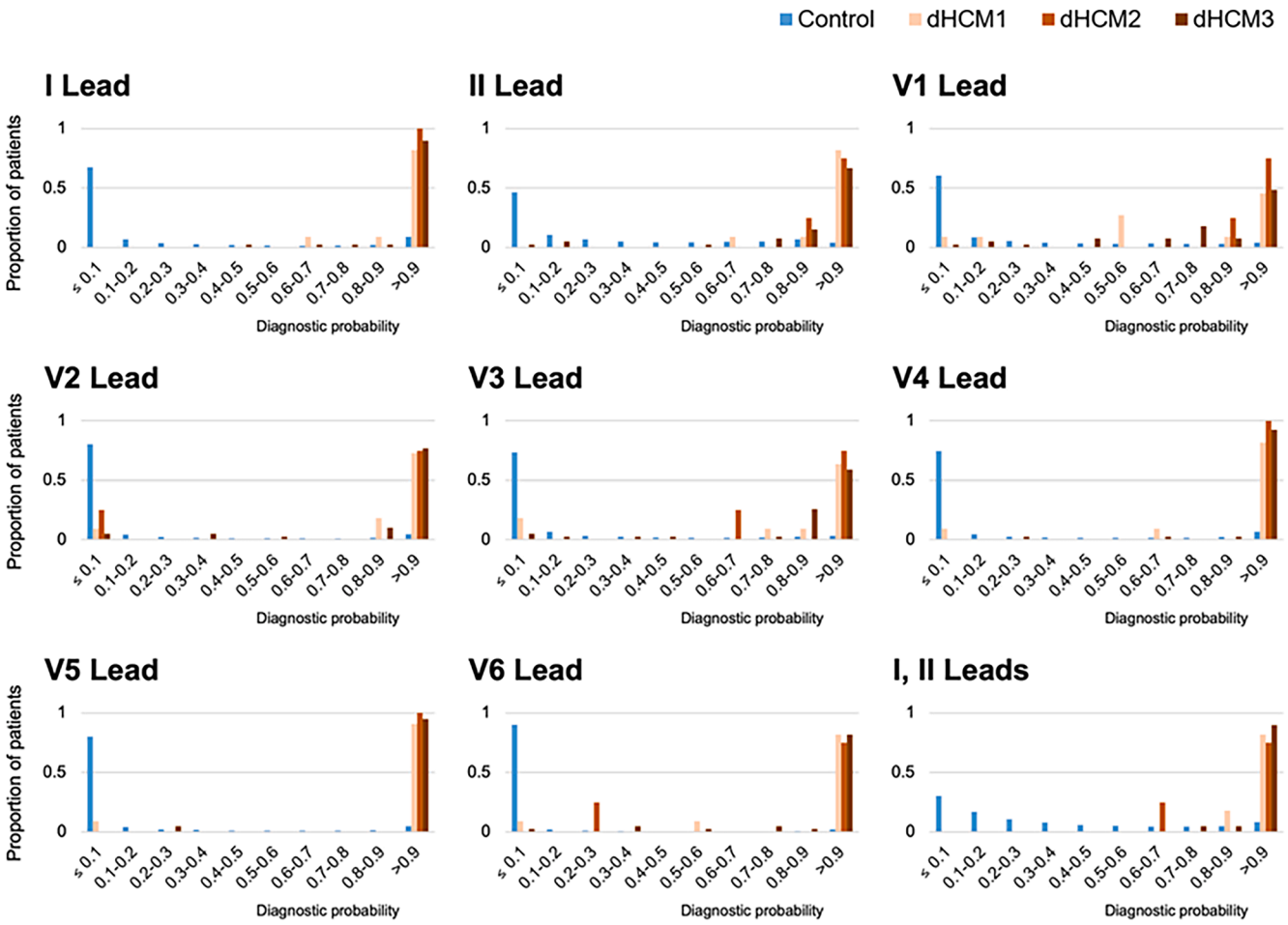

Fig. 3. The proportion of patients according to the model output in CNN-derived models using single-lead or double-lead ECGs The vertical scale indicates the proportion of the patients in dHCM group (orange) and non-dHCM group (blue). The horizontal scale indicates the diagnostic probability for dHCM yielded by each CNN model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Availability of data and materials

Data cannot be shared publicly because of lack of such a description in the study protocol and informed consent. Data are available from the Ethics Review Committee at the Cardiovascular Institute for researchers who meet the criteria for access to confidential data (contact via the corresponding author).

Authorship

All authors take responsibility for all aspects of the reliability and freedom from bias of the data presented and their discussed interpretation.

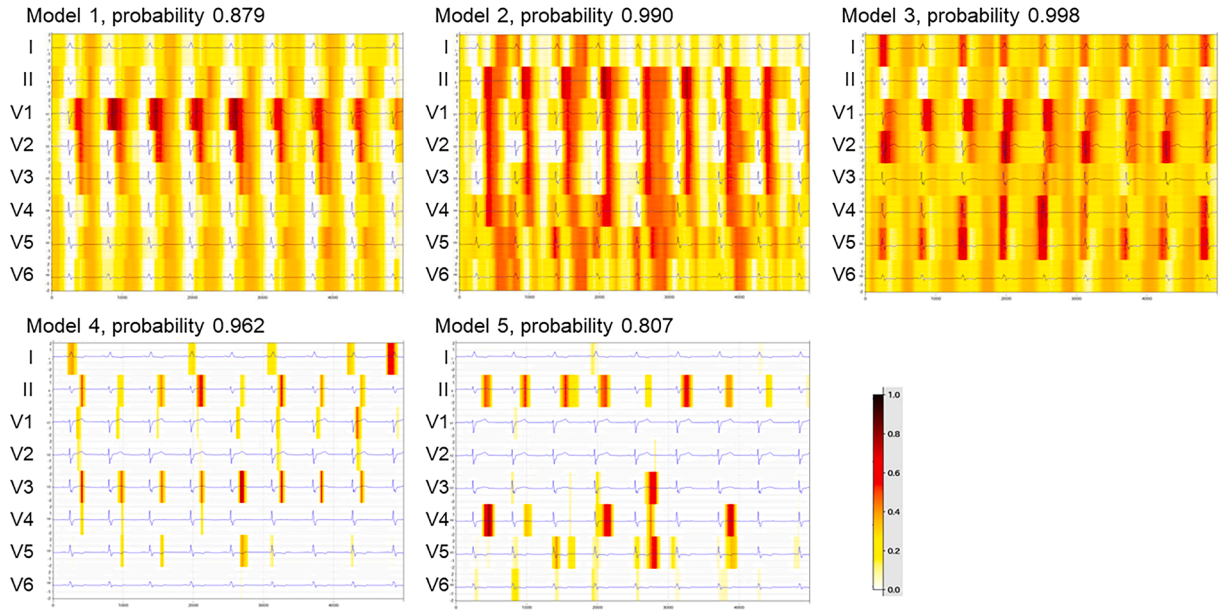

Fig. 4. Representative images of GradCAM with eight-lead ECGs The difference in color chart cannot be compared between the models or leads, because the color chart was determined within each model or lead,

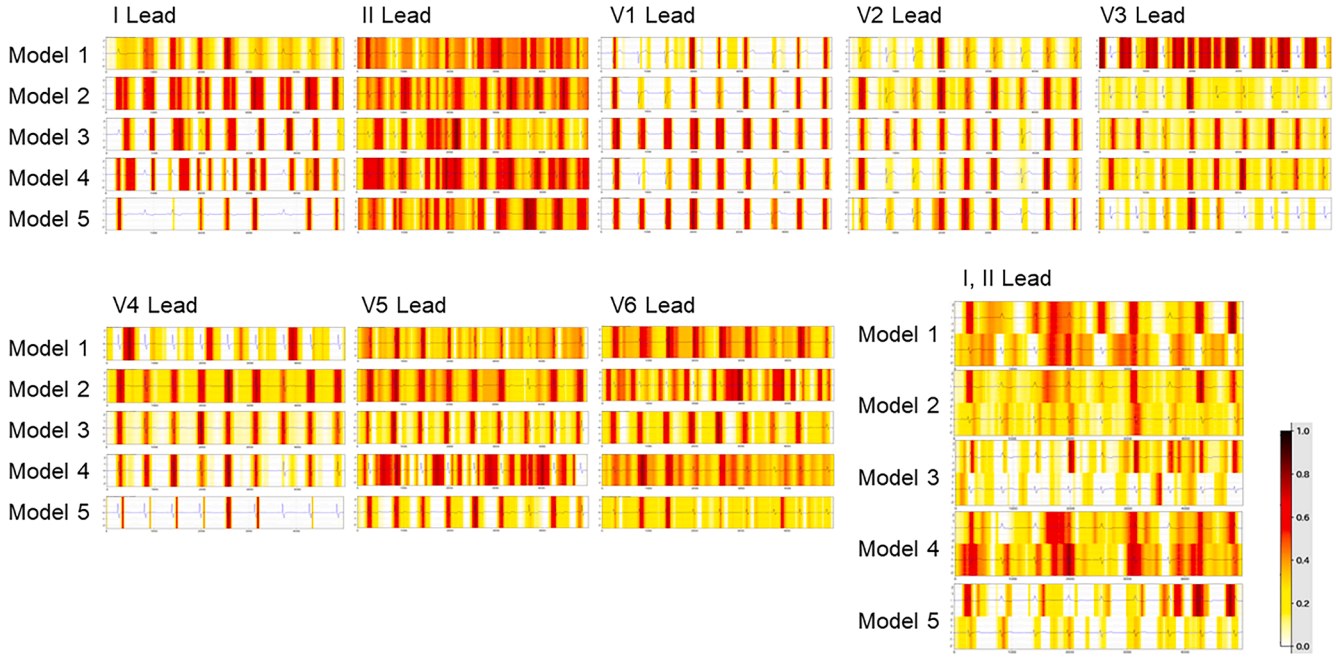

Fig. 5. Representative images of GradCAM with single-lead and double-lead ECGs The difference in color chart cannot be compared between the models or leads, because the color chart was determined within each model or lead,

Competing interests

Dr. Suzuki received lecture fees from Daiichi Sankyo and Bristol-Myers Squibb. Dr. Yamashita received research funds and/or lecture fees from Daiichi Sankyo, Bayer Yakuhin, Bristol-Myers Squibb, Pfizer, Nippon Boehringer Ingelheim, Eisai, Mitsubishi Tanabe Pharm, Ono Pharmaceutical, and Toa Eiyo. J Motogi, T Umemoto, W Matsuzawa, T Takayanagi, A Hyodo, and K Satoh are employee at Nihon Kohden Corporation.

Sources of funding

None.

Registration number of clinical studies

None.

CRediT authorship contribution statement

Naomi Hirota: Conceptualization, Validation, Writing – original draft. **Shinya Suzuki:** Conceptualization, Validation, Writing – review

& editing. **Jun Motogi:** Conceptualization, Validation. **Takuya Umemoto:** Data curation, Formal analysis, Writing – original draft, Writing – review & editing. **Hiroshi Nakai:** Conceptualization, Data curation, Validation. **Wataru Matsuzawa:** Data curation, Formal analysis. **Tsuneo Takayanagi:** Data curation, Formal analysis. **Akira Hyodo:** Formal analysis, Validation. **Keiichi Satoh:** Formal analysis, Validation. **Takuto Arita:** Writing – review & editing. **Naoharu Yagi:** Writing – review & editing. **Mikio Kishi:** Writing – review & editing. **Hiroaki Semba:** Writing – review & editing. **Hiroto Kano:** Writing – review & editing. **Shunsuke Matsuno:** Writing – review & editing. **Takayuki Otsuka:** Writing – review & editing. **Tokuhisa Uejima:** Writing – review & editing. **Yuji Oikawa:** Writing – review & editing. **Takayuki Hori:** Writing – review & editing. **Mitsuru Iida:** Writing – review & editing. **Junji Yajima:** Writing – review & editing. **Takeshi Yamashita:** Writing – review & editing.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Shinya Suzuki reports a relationship with Public Interest Incorporated Foundation The Cardiovascular Institute that includes: speaking and lecture fees. Takeshi Yamashita reports a relationship with Public Interest Incorporated Foundation The Cardiovascular Institute that includes: funding grants and speaking and lecture fees.

Acknowledgments

We thank Shiro Ueda and Nobuko Ueda at Medical Edge Company, Ltd., for assembling the database via the Clinical Study Supporting System, and Yurika Hashiguchi, Hiroaki Arai, and Takashi Osada for data management and system administration. We thank Karl Embleton, PhD, from Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijcha.2023.101211.

References

[1] I. Olivotto, F. Cecchi, C. Poggesi, M.H. Yacoub, Patterns of disease progression in hypertrophic cardiomyopathy: an individualized approach to clinical staging, Circ Heart Fail. 5 (2012) 535–546, https://doi.org/10.1161/CIRCHEARTFAILURE.112.967026.

[2] K.M. Harris, P. Spirito, M.S. Maron, et al., Prevalence, clinical profile, and significance of left ventricular remodeling in the end-stage phase of hypertrophic cardiomyopathy, Circulation. 114 (2006) 216–225, https://doi.org/10.1161/CIRCULATIONAHA.105.583500.

[3] T. Kubo, H. Kitaoka, M. Okawa, et al., Lifelong left ventricular remodeling of hypertrophic cardiomyopathy caused by a founder frameshift deletion mutation in the cardiac Myosin-binding protein C gene among Japanese, J Am Coll Cardiol. 46 (2005) 1737–1743, https://doi.org/10.1016/j.jacc.2005.05.087.

[4] E. Biagini, F. Coccolo, M. Ferlito, et al., Dilated-hypokinetic evolution of hypertrophic cardiomyopathy: prevalence, incidence, risk factors, and prognostic implications in pediatric and adult patients, J Am Coll Cardiol. 46 (2005) 1543–1550, https://doi.org/10.1016/j.jacc.2005.04.062.

[5] R. Thaman, J.R. Gimeno, R.T. Murphy, et al., Prevalence and clinical significance of systolic impairment in hypertrophic cardiomyopathy, Heart. 91 (2005) 920–925, https://doi.org/10.1136/hrt.2003.031161.

[6] M.S. Maron, I. Olivotto, B.J. Maron, et al., The case for myocardial ischemia in hypertrophic cardiomyopathy, J Am Coll Cardiol. 54 (2009) 866–875, https://doi.org/10.1016/j.jacc.2009.04.072.

[7] Z.I. Attia, D.M. Harmon, E.R. Behr, P.A. Friedman, Application of artificial intelligence to the electrocardiogram, Eur Heart J. 42 (2021) 4717–4730, https://doi.org/10.1093/eurheartj/ehab649.

[8] K.C. Siontis, P.A. Noseworthy, Z.I. Attia, P.A. Friedman, Artificial intelligence-enhanced electrocardiography in cardiovascular disease management, Nat Rev Cardiol. 18 (2021) 465–478, https://doi.org/10.1038/s41569-020-00503-2.

[9] S. Suzuki, T. Otsuka, K. Sagara, et al., Nine-Year Trend of Anticoagulation Use, Thromboembolic Events, and Major Bleeding in Patients With Non-Valvular Atrial Fibrillation- Shinken Database Analysis, Circ J. 80 (2016) 639–649, https://doi.org/10.1253/circj.CJ-15-1237.

[10] L. Baecker, R. Garcia-Dias, S. Vieira, C. Scarpazza, A. Mechelli, Machine learning for brain age prediction: Introduction to methods and clinical applications, EBioMedicine. 72 (2021), 103600, https://doi.org/10.1016/j.ebiom.2021.103600.

[11] Z.I. Attia, P.A. Noseworthy, F. Lopez-Jimenez, et al., An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: a retrospective analysis of outcome prediction, The Lancet. 394 (2019) 861–867, https://doi.org/10.1016/s0140-6736(19)31721-0.

[12] S. Suzuki, J. Motogi, H. Nakai, et al., Identifying patients with atrial fibrillation during sinus rhythm on ECG: Significance of the labeling in the artificial intelligence algorithm, Int J Cardiol Heart Vasc. 38 (2022), 100954, https://doi.org/10.1016/j.ijcha.2022.100954.

[13] N. Hirota, S. Suzuki, J. Motogi, et al., Cardiovascular events and artificial intelligence-predicted age using 12-lead electrocardiograms, Int J Cardiol Heart Vasc. 44 (2023), 101172, https://doi.org/10.1016/j.ijcha.2023.101172.

[14] M. Coffin, S. Sukhatme, Receiver operating characteristic studies and measurement errors, Biometrics. 53 (1997) 823–837.

[15] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. IEEE International Conference on Computer Vision. 2017; https://openaccess.thecvf.com/content_ICCV_2017/papers/Selvaraju_Grad-CAM_Visual_Explanations_ICCV_2017_paper.pdf:618-26. Doi: https://openaccess.thecvf.com/content_ICCV_2017/papers/Selvaraju_Grad-CAM_Visual_Explanations_ICCV_2017_paper.pdf.

[16] N. Sheikh, M. Papadakis, S. Ghani, et al., Comparison of electrocardiographic criteria for the detection of cardiac abnormalities in elite black and white athletes, Circulation. 129 (2014) 1637–1649, https://doi.org/10.1161/CIRCULATIONAHA.113.006179.

[17] D.D. Savage, S.F. Seides, C.E. Clark, et al., Electrocardiographic findings in patients with obstructive and nonobstructive hypertrophic cardiomyopathy, Circulation. 58 (1978) 402–408, https://doi.org/10.1161/01.cir.58.3.402.

[18] D. Corrado, A. Pelliccia, H.H. Bjornstad, et al., Cardiovascular pre-participation screening of young competitive athletes for prevention of sudden death: proposal for a common European protocol. Consensus Statement of the Study Group of Sport Cardiology of the Working Group of Cardiac Rehabilitation and Exercise Physiology and the Working Group of Myocardial and Pericardial Diseases of the European Society of Cardiology, Eur Heart J. 26 (2005) 516–524, https://doi.org/10.1093/eurheartj/ehi108.

[19] W.Y. Ko, K.C. Siontis, Z.I. Attia, et al., Detection of Hypertrophic Cardiomyopathy Using a Convolutional Neural Network-Enabled Electrocardiogram, J Am Coll Cardiol. 75 (2020) 722–733, https://doi.org/10.1016/j.jacc.2019.12.030.

[20] S. Goto, D. Solanki, J.E. John, et al., Multinational Federated Learning Approach to Train ECG and Echocardiogram Models for Hypertrophic Cardiomyopathy Detection, Circulation. 146 (2022) 755–769, https://doi.org/10.1161/CIRCULATIONAHA.121.058696.

[21] K. Doi, G. Toda, I.I. Iliev, M. Hayano, K. Yano, Clinical analysis of hypertrophic cardiomyopathy which evolved into dilated phase during long-term follow-up, Jpn Heart J. 40 (1999) 579–587, https://doi.org/10.1536/jhj.40.579.

[22] R.P. Grant, The relationship between the anatomic position of the heart and the electrocardiogram; a criticism of unipolar electrocardiography, Circulation. 7 (1953) 890–902, https://doi.org/10.1161/01.cir.7.6.890.

[23] F.D. Johnston, J.M. Ryan, J.M. Bryant, The electrocardiogram and the position of the heart, Am Heart J. 43 (1952) 306–315, https://doi.org/10.1016/0002-8703(52)90222-6.

[24] D.G. Strauss, D.D. Schocken, Marriott's Practical Electrocardiography, Thirteenth ed., Wolters Kluwer, 2021.

[25] E.W. Hancock, B.J. Deal, D.M. Mirvis, et al., AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram: part V: electrocardiogram changes associated with cardiac chamber hypertrophy: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society: endorsed by the International Society for Computerized Electrocardiology, Circulation. 119 (2009) e251–e261, https://doi.org/10.1161/CIRCULATIONAHA.108.191097.

[26] K. Miura, H. Nakagawa, Y. Morikawa, et al., Epidemiology of idiopathic cardiomyopathy in Japan: results from a nationwide survey, Heart. 87 (2002) 126–130, https://doi.org/10.1136/heart.87.2.126.