A method using deep learning to discover new predictors from left-ventricular mechanical dyssynchrony for CRT response

Zhuo He, BS,[a] Xinwei Zhang, MD,[b] Chen Zhao, MS,[a] Xing Ling, MS,[c] Saurabh Malhotra, MD,[d,e] Zhiyong Qian, MD, PhD,[b] Yao Wang, MD, PhD,[b] Xiaofeng Hou, MD,[b] Jiangang Zou, MD, PhD, FHRS,[b] and Weihua Zhou, PhD[a,f]

[a] College of Computing, Michigan Technological University, Houghton, MI
[b] Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University, Nanjing, Jiangsu, China
[c] Department of Mathematical Sciences, Michigan Technological University, Houghton, MI
[d] Division of Cardiology, Cook County Health and Hospitals System, Chicago, IL
[e] Division of Cardiology, Rush Medical College, Chicago, IL
[f] Center for Biocomputing and Digital Health, Institute of Computing and Cybersystems, Health Research Institute, Michigan Technological University, Houghton, MI

Received Oct 28, 2021; accepted Jun 22, 2022
doi:10.1007/s12350-022-03067-5

Background. **Studies have shown that the conventional parameters characterizing left ventricular mechanical dyssynchrony (LVMD) measured on gated SPECT myocardial perfusion imaging (MPI) have their own statistical limitations in predicting cardiac resynchronization therapy (CRT) response. The purpose of this study is to discover new predictors from the polarmaps of LVMD by deep learning to help select heart failure patients with a high likelihood of response to CRT.**

Methods. **One hundred and fifty-seven patients who underwent rest gated SPECT MPI were enrolled in this study. CRT response was defined as an increase in left ventricular ejection fraction (LVEF) > 5% at 6 ± 1 month follow up. The autoencoder (AE) technique, an unsupervised deep learning method, was applied to the polarmaps of LVMD to extract new predictors characterizing LVMD. Pearson correlation analysis was used to explain the relationships between new predictors and existing clinical parameters. Patients from the IAEA VISION-CRT trial were used for an external validation. Heatmaps were used to interpret the AE-extracted feature.**

Results. **Complete data were obtained in 130 patients, and 68.5% of them were classified as CRT responders. After variable selection by feature importance ranking and correlation analysis, one AE-extracted LVMD predictor was included in the statistical analysis. This new**

Zhuo He and Xinwei Zhang have contributed equally to this work.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12350-022-03067-5.

The authors of this article have provided a PowerPoint file, available for download at SpringerLink, which summarises the contents of the paper and is free for re-use at meetings and presentations. Search for the article DOI on SpringerLink.com.

The authors have also provided an audio summary of the article, which is available to download as ESM, or to listen to via the JNC/ASNC Podcast.

Reprint requests: Weihua Zhou, PhD, College of Computing, Michigan Technological University, 1400 Townsend Drive, Houghton, MI; *whzhou@mtu.edu*; Jiangang Zou, MD, PhD, FHRS, Department of Cardiology, The First Affiliated Hospital of Nanjing Medical University, Guangzhou Road 300, Nanjing 210029, Jiangsu, China; *jgzou@njmu.edu.cn*

J Nucl Cardiol
1071-3581/$34.00

Published online: 01 August 2022

AE-extracted LVMD predictor showed statistical significance in the univariate (OR 2.00, $P = .026$) and multivariate (OR 1.11, $P = .021$) analyses, respectively. Moreover, the new AE-extracted LVMD predictor not only had incremental value over PBW and significant clinical variables, including QRS duration and left ventricular end-systolic volume (AUC 0.74 vs 0.72, LH 7.33, $P = .007$), but also showed encouraging predictive value in the 165 patients from the IAEA VISION-CRT trial ($P < .1$). The heatmaps for calculation of the AE-extracted predictor showed higher weights on the anterior, lateral, and inferior myocardial walls, which are recommended as LV pacing sites in clinical practice.

Conclusions. **AE techniques have significant value in the discovery of new clinical predictors. The new AE-extracted LVMD predictor extracted from the baseline gated SPECT MPI has the potential to improve the prediction of CRT response. (J Nucl Cardiol 2022)**

Key Words: CRT • SPECT myocardial perfusion imaging • left ventricular mechanical dyssynchrony • deep learning • autoencoder

Abbreviations	
AE	Autoencoder
CRT	Cardiac resynchronization therapy
LVEDV	Left ventricular end diastolic volume
LVEF	Left ventricular ejection fraction
LVESV	Left ventricular end systolic volume
LVMD	Left ventricular mechanical dyssynchrony
MPI	Myocardial perfusion imaging
PSD	Phase histogram standard deviation
PBW	Phase histogram bandwidth
SPECT	Single-photon emission computed tomography

INTRODUCTION

Current guidelines recommend cardiac resynchronization therapy (CRT) for patients with low left ventricular ejection fraction (LVEF) (typically $\leq 35\%$), sinus rhythm, left bundle-branch block (LBBB) with a QRS duration greater than or equal to 150 ms, and New York Heart Association (NYHA) class II, III, or ambulatory IV symptoms on guideline-directed medical therapy.[1,2] However, 30%-40% of the selected patients fail to respond.[3–5]

Left ventricular mechanical dyssynchrony (LVMD) of gated single-photon emission computed tomography (SPECT) myocardial perfusion imaging (MPI) has shown significant value for CRT patient selection and prognosis.[6–10] However, existing statistical predictors characterizing LVMD such as phase standard deviation (PSD) and phase bandwidth (PBW) are easily affected by the outliers of phase measurement and have been reported to have significant limitations, which may influence the assessment of clinical values of LVMD,[11–13] and therefore could not predict the response to CRT in several studies.[14–16]

Machine learning enables computers to learn and develop rules without having to be instructed by human programmers at every step of the way.[17] Deep learning as a powerful new tool combines many linear and nonlinear transformations to obtain a more comprehensive and useful representation of data.[18] It has achieved breakthrough applications in lesion detection and disease classification by absorbing the image measurement engineering directly into a learning step while processing the data in its natural form.[19–21] Betancur et al[22] presented the effectiveness of using deep neural networks for feature extraction in gated SPECT MPI and prediction of obstructive CAD. Xu et al[23] demonstrated that unsupervised feature extraction by deep learning (single-layer network of K-means centroids, accuracy 93.65%) was as effective as a supervised method (fully convolution neural network, accuracy 94.52%) in the classification of histopathology images. In this study, we aimed to discover new predictors by deep learning from LVMD measured on gated SPECT MPI for CRT patient selection.

METHODS

Patient population

One hundred and fifty-seven CRT patients with gated resting SPECT MPI were enrolled at nine medical centers in China from July 2008 to July 2020. The enrollment criteria included (1) LVEF measured by echocardiography $\leq 35\%$; (2) QRS duration ≥ 120 ms; (3) New York Heart Association (NYHA) functional class from II to IV; and (4) optimal medical therapy for at least 3 months before CRT implantation. Patients with atrial fibrillation or right bundle branch block were excluded.

All the patients had baseline characteristics, pre-CRT echocardiography, pre-CRT resting gated SPECT MPI, and 6 months follow-up echocardiography. The study was approved by the Institutional Ethical

Committee of the First Affiliated Hospital of Nanjing Medical University, and informed consent was obtained from all patients.

Evaluation of LV function by echocardiography

Echocardiography data of all patients were assessed by experienced ultrasound experts blinded to any clinical data and MPI data before and 6 months after CRT. LVEF was measured by the 2-dimensional modified biplane Simpson method. An increase in LVEF $> 5\%$ at 6-month follow-up echocardiography is considered as a positive mechanical response to CRT.[24,25]

Gated SPECT MPI acquisition and quantification

Gated SPECT MPI was performed by SPECT systems with a low-energy, general-purpose collimator within 7 days before CRT implantation. Resting gated SPECT scan was performed 60-90 minute after injection of 25-30 mCi of 99mTc-MIBI. Images were acquired by 1-day resting gated SPECT MPI protocol with a dual-headed or triple-head camera by 180° orbits with a complimentary 8 frames ECG-gating, according to the current guideline.[26] The photoelectric window of 99mTc was set to a 20% energy window centered over 140 keV when the gated SPECT MPI was acquired. All the images were reconstructed by the OSEM method from Emory Reconstruction Toolbox (ERTb2, Atlanta, GA) with 3 iterations, 10 subsets, power 10, and a cutoff frequency of 0.3 cycles/mm. Short-axis images were generated by Emory Cardiac Toolbox (ECTb4, Atlanta, GA) for automated measurement of LV function and LVMD.

LVMD is measured by a phase analysis technique with the following three steps: (1) extracting three-dimensional maximal count myocardial perfusion from the short axis slices by sampling the myocardial wall; (2) calculating the systolic phase angles by the 1-harmonic Fourier approximation, which measures the change of counts in LV myocardium in an R-R cardiac cycle; and (3) generating polarmaps based on the LV phase angles in the LV myocardium.[27,28]

Extraction of new parameters by autoencoder techniques

Autoencoder (AE) is an unsupervised learning technique in which we use neural networks to accomplish the task of representation learning.[29] Specifically, the AE architecture forces the compressed knowledge to represent the original by copying the input to the output, compressing the input to a latent space representation, and then reconstructing the output from this representation.[30] It has an input layer, an output layer, and one or more hidden layers connecting them. The input layer and the output layer have the same number of nodes. During the training process, AE sets the target value equal to the input value and adjusts the model parameters by applying backpropagation.[31] The network is arranged in two stages, the encoder and the decoder. The encoder connects directly to the pixels of inputted images through a linear layer and compresses the high dimensional input images into the hidden layer, which has a lower dimension than the input images and can be used as features of the input images. The decoder attempts to reconstruct the input images from the generated features. The AE model learns to recognize key image features and reconstruct the input by minimizing the error between the reconstructed output and input image.

In this study, a multi-layer AE with one input layer, five fully connected hidden layers and one output layer was used to train the model, and the middle-most hidden layer was used to represent the features of LVMD polarmaps. The overall process of the AE model training, including the detailed architecture of our AE model, and the process of predicting CRT response with AE-extracted parameters is illustrated in Fig. 1. The gradient is calculated through a mean squared error loss function, and the parameters are updated by an Adam optimizer. Early stopping is used to avoid overfitting when the loss stops decreasing. The AE model was implemented in the Python programming language by the PyTorch deep learning toolkit (version 1.3.1).[32] Model training was performed on graphical processor units (TITAN Xp, NVIDIA, Santa Clara, California).

Statistical analysis

Discrete variables of the baseline characteristics were expressed in number and percentage and tested with the Chi-square test. Continuous variables were expressed as mean $\pm$ standard deviation and tested with the Student t-test. The univariate logistic regression analysis was applied to estimate the predictive values of baseline clinical variables, conventional LVMD parameters (PSD and PBW), and 32 AE-extracted LVMD parameters for CRT response. To overcome the curse of dimensionality and collinearity, only significant features in the univariate analysis were selected, and the Pearson correlation coefficient (PCC) analysis was further applied to exclude the variables that were highly correlated with each other. Subsequent multivariate analysis was performed to identify the predictive values for CRT response. Variables with $P < .05$ were

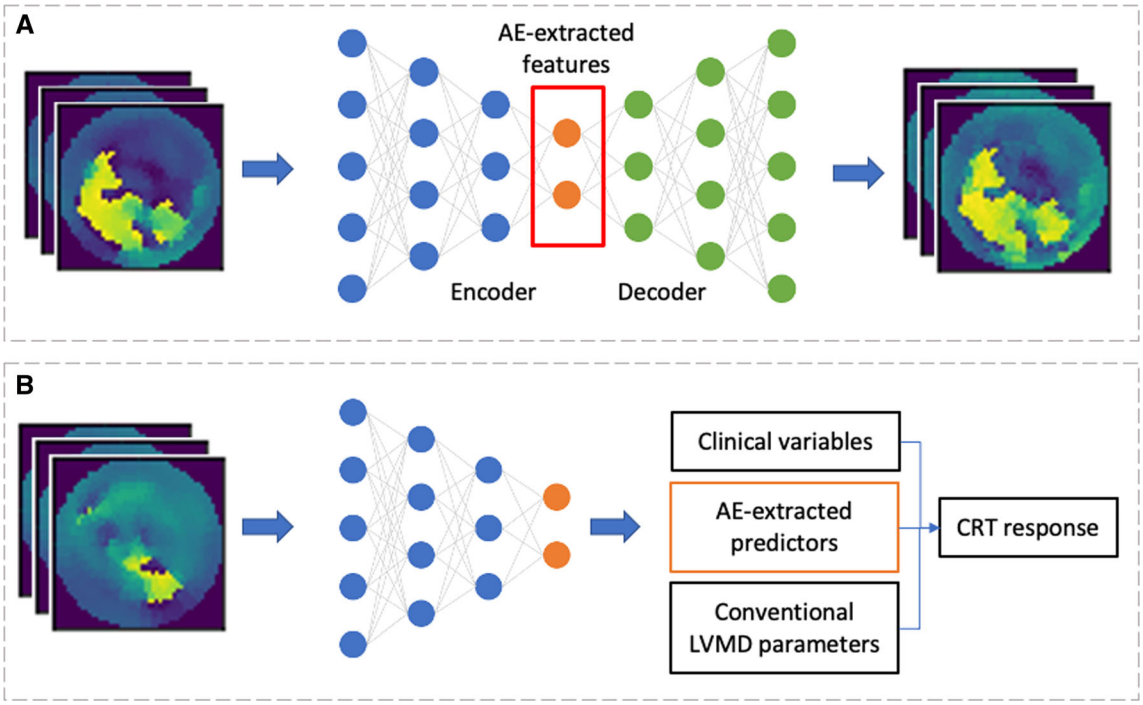

Figure 1. Training an autoencoder (AE) model from left-ventricular mechanical dyssynchrony (LVMD) (**A**) and building a prediction model (**B**) for CRT response. In (**A**), the multi-layer AE model is applied to the systolic phase polarmaps to extract compressed features and reduce the dimensions. In (**B**), clinical variables, AE-extracted predictors, and conventional LVMD parameters (phase standard deviation and bandwidth) are used to build the prediction model for CRT response.

considered statistically significant. PCC was also used to determine whether there were significant correlations between the selected AE-extracted predictors and clinical parameters, i.e., LVMD parameters (PSD and PBW), as well as to rank the importance of all variables in the prediction of the CRT response. The Akaike information criterion (AIC) and likelihood ratio (LR) test were used to determine whether the prediction model with AE-extracted predictors is statistically significantly favored over the model with significant clinical parameters. The predictive performances of all significant variables were evaluated by the area under the curve (AUC) of the receiver operating characteristic (ROC) of binary logistic regression. Statistical analysis was performed by IBM SPSS Statistics software version 26 (SPSS Inc, Chicago, Illinois) and Python Statsmodels package.[33]

External validation

In order to confirm the generalizability of the method proposed in this paper, the data from the multicenter VISION-CRT trial was used for external validation. The complete study design and preliminary results of VISION-CRT were previously published.[14,34,35] Because of the data error (n = 22), death of patients before follow-up (n = 11), and extremely small heart (ESV = 8ML) as outliner (n = 1), 165 of 199 patients were selected as the external validation dataset for this study, which was processed in the same way as in the study by He et al[35] However, because of the mismatch of patient IDs in the clinical records and SPECT MPI (n = 17), an external validation dataset of 148 cases (responders: n = 66, 44.6%) containing the complete records and SPECT MPIs was finally obtained. Due to the lack of pre- and post-CRT echocardiographic data, the CRT response in the external validation dataset was defined as a $> 5\%$ increase in LVEF measured by gated SPECT MPI at 6-month follow-up compared to pre-CRT SPECT MPI. This definition of CRT response was calculated in the same way as the definition in the training dataset, except for the modality used to measure LVEF (training dataset: echo, external validation dataset: gated SPECT MPI). Clinical variables (e.g., gender, age, NYHA), ECG parameters (QRS duration, LBBB), and LV measurements from gated SPECT MPI at baseline and 6-month follow-up (e.g., LVEF, ESV) were included in this study.

Interpretability of AE-extracted feature

To determine the contribution of the attributes in the original systolic phase polarmap image to the value of the AE-extracted LVMD predictors, heatmaps of weights were used to map directly from the input LVMD polarmap to each AE feature extracted from the hidden layers.

We inputted an image into our AE network and visualized the network diagram by ''unrolling'' the pixel into a single column of neurons, as shown in Fig. 1. The AE model ''forces'' the network to learn the features of the image itself. The weights of each neuron in the input layer are connected to each AE feature through multiple hidden layers. We can explain these weights as forming templates of the output features. If an image is largely sympathetic to a filter, the image will generate high activation in a particular neuron in the input layer. Thus, the neurons in the hidden layer reflect the presence of those features in the original image. In the output layer, a single neuron, corresponding to the different AE features, is a weighted combination of those previous hidden activations. Therefore, we visualized the weights in the input layer corresponding to the neurons in the hidden layer that were selected as AE-extracted features.

RESULTS

A total of 157 patients underwent CRT, but complete clinical assessment data, baseline SPECT MPI, and 6-month follow-up data were obtained in 130 patients. The baseline characteristics are shown in Table 1. The average age was 62.3 ± 12.1 years, 91 (70.0%) patients were male, and 67 (51.9%) patients were classified as NYHA functional class III. After a 6-month follow-up, 89 of the 130 patients (68.5%) were considered as CRT responders, and the rest were considered as non-responders to CRT. Significant differences of EDV (286.0 ± 102.5 vs 349.0 ± 156.8, $P = .008$) and ESV (228.3 ± 97.3 vs 282.0 ± 147.6, $P = .016$) were noted between responders and non-responders. However, there were no significant differences between the two groups in baseline PSD and PBW.

All 157 pre-CRT LVMD polarmaps were used to train the AE model for feature extraction. After the training, 32 features of LVMD were extracted from the hidden layers of the autoencoder model as the AE-extracted LVMD predictors.

Twenty-seven of the 157 subjects did not have a follow-up on their CRT response, so the AE-extracted predictors from phase polarmaps of the remaining 130 patients were analyzed. Only 4 of the 32 AE-extracted

Table 1. Baseline characteristics of the enrolled patients

Variables	All (n = 130)	Response (n = 89, 68.5%)	Non-response (n = 41, 31.5%)	*P* value
Age	62.3 ± 12.1	63.4 ± 11.0	59.9 ± 13.8	.128
CKD	9 (6.9%)	6 (6.7%)	3 (7.3%)	.801
DM	28 (21.5%)	18 (20.2%)	10 (24.4%)	.759
QRS duration	151.5 ± 21.2	153.7 ± 19.7	146.7 ± 23.4	.084
LVEDV	305.9 ± 126.4	286.0 ± 102.5	349.0 ± 158.6	.008
LVESV	245.2 ± 118.2	228.3 ± 97.3	282.0 ± 147.6	.016
LVEF	21.5 ± 7.8	21.7 ± 7.5	21.1 ± 8.3	.68
Gender	91 (70.0%)	62 (69.7%)	29 (70.7%)	.934
NYHA				.104
II	43 (33.1%)	33 (37.1%)	10 (24.4%)	
III	67 (51.5%)	46 (51.7%)	21 (51.2%)	
IV	20 (15.4%)	10 (11.2%)	10 (24.4%)	
SRS	18.4 ± 9.8	17.4 ± 8.8	20.5 ± 11.2	.09
PBW	203.5 ± 73.9	205.4 ± 73.7	199.4 ± 74.2	.666
PSD	60.1 ± 18.5	60.2 ± 17.7	59.8 ± 19.9	.913

Data are expressed as mean ± standard deviation or number (percentage)
CKD, chronic kidney disease; *DM*, diabetes mellitus; *LVEDV*, left ventricular end-diastolic volume; *LVESV*, left ventricular end-systolic volume; *LVEF*, left ventricular ejection fraction; *NYHA*, New York Heart Association; *SRS*, summed rest score; *PBW*, phase bandwidth; *PSD*, phase standard deviation

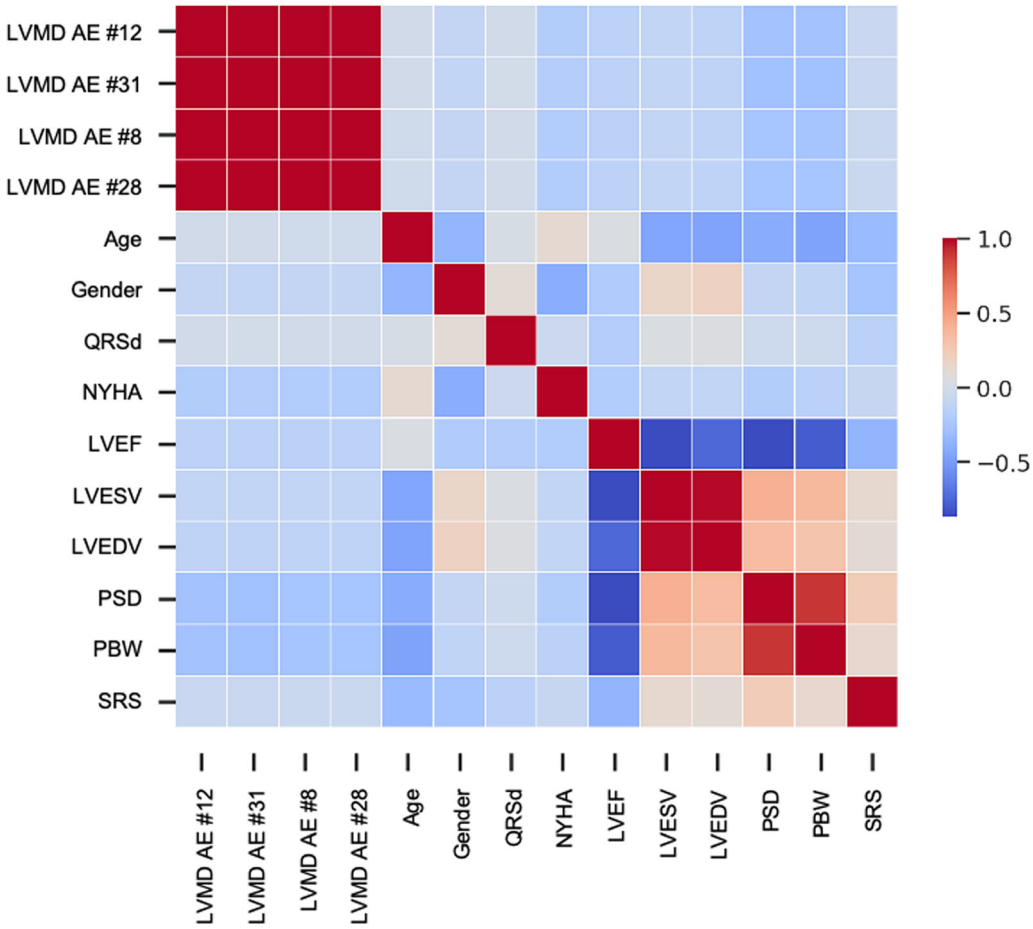

Figure 2. Pearson's correlations between clinical variables, phase standard deviation (PSD), phase bandwidth (PBW), and AE-extracted left ventricular mechanical dyssynchrony (LVMD) parameters. Only the AE-extracted LVMD parameters which were significant in the univariate analysis, are displayed. There are strong correlations between these AE-extracted LVMD parameters (all Pearson correlation coefficient [PCC] > 0.99, P values $< .05$), so PCC between the significant AE-extracted LVMD parameters and the CRT response is further applied to select only one AE-extracted LVMD predictor (LVMD AE # 31), which has the highest correlation (PCC = 0.20) with the CRT response. This AE-extracted LVMD predictor is used in the subsequent statistical analysis.

features showed statistical significance in the univariate analysis. As shown in Fig. 2, these four AE-extracted LVMD predictors had a strong correlation with each other (all PCC > 0.99, P values $< .05$), but they were not correlated with any clinical variables and conventional LVMD parameters (PSD and PBW). Only one AE-extracted LVMD predictor, which has the highest Pearson correlation with the CRT response (PCC 0.20), was selected for the subsequent statistical analysis to characterize the polarmaps of LVMD.

In the univariate analysis, LVEDV (OR 1, 95% CI 0.99-1.00, $P = .013$), LVESV (OR 1, 95% CI 0.99-1.00, $P = .023$), NYHA (OR 0.56, 95% CI 0.32-0.99, $P = .045$), and the AE-extracted LVMD predictor (OR 2.00, 95% CI 1.08-3.67, $P = .026$) had significantly predictive values to CRT response, as shown in Table 2. Since there was a strong correlation between LVEDV and LVESV (PCC = 0.98, P values = .032), only the LVESV, which had a stronger correlation with CRT response than LVEDV (PCC = 0.51 vs PCC = 0.47), was included in the multivariate analysis. In the multivariate analysis, LVESV (OR 1.00, 95% CI 1.0-1.0, $P = .013$) and AE-extracted LVMD predictor (OR 1.11, 95% CI 1.02-1.23, $P = .021$) had significant predictive values for CRT response, as shown in Table 2.

In Fig. 3, the fitting performance of the predictive model with the AE-extracted LVMD predictor (AIC 162.04) was better than the model with PBW (AIC 164.68). The AE-extracted LVMD predictor brought incremental value over clinical variables (LR 5.52,

Table 2. Univariate and multivariate logistic regression analysis

	Univariate analysis			Multivariate analysis (AE #30)		
Variables	**OR**	**95% CI**	***P* value**	**OR**	**95% CI**	***P* value**
Age	1.02	0.99-1.06	.131			
CKD	0.92	0.22-3.86	.904			
DM	0.79	0.33-1.9	.592			
LVEDV	1	0.99-1.0	.013			
LVESV	1	0.99-1.0	.023	1.00	1.00-1.00	.013
LVEF	1.01	0.96-1.06	.678			
Gender	0.95	0.42-2.14	.902			
NYHA	0.56	0.32-0.99	.045	0.89	0.80-1.00	.052
PBW	1	1.0-1.01	.664			
PSD	1	0.98-1.02	.912			
QRSd	1.02	1.0-1.04	.086			
SRS	0.97	0.93-1.01	.091			
LVMD AE #12	2.00	1.08-3.67	.026	1.11	1.02-1.23	.021
LVMD AE #31	1.77	1.07-2.94	.026			
LVMD AE #8	1.38	1.04-1.84	.028			
LVMD AE #28	1.31	1.03-1.68	.028			

CKD, chronic kidney disease; *DM*, diabetes mellitus; *LVEDV*, left ventricular end-diastolic volume; *LVESV*, left ventricular end-systolic volume; *LVEF*, left ventricular ejection fraction; *NYHA*, New York Heart Association; *SRS*, summed rest score; *PBW*, phase bandwidth; *PSD*, phase standard deviation; *AE*, autoencoder

P = .019) and clinical variables with PBW (LR 7.33, P = .007). In addition, the AUCs of the clinical variables (95% CI 0.57-0.78, AUC 0.66), clinical variables combined with PBW (95% CI 0.60-0.80, AUC 0.69), clinical variables combined with the AE-extracted LVMD predictor (95% CI 0.61-0.81, AUC 0.72), and clinical variables combined with both PBW and the AE-extracted LVMD predictor (95% CI 0.64-0.83, AUC 0.74) increased sequentially, as shown in Fig. 4.

The well-trained model was also used in the external validation group to test the performance of the AE-extracted LVMD predictor. The baseline characteristic and statistical results have been previously published.[35] It is worth noting that there is a big difference in the distribution of patient features between the external validation dataset and the training set. Examples include QRS duration (training set: response [n = 89, 68.5%] vs non-response [n = 41, 31.5%]: 153.7 ± 19.7 vs 146.7 ± 23.4; external validation set: response [n = 66, 44.6%] vs non-response [n = 82, 55.4%]: 161.7 ± 22.4 vs 160.0 ± 27.4), PSD (training set: 60.2 ± 17.7 vs $59.8 \pm .9$; external validation set: 46.8 ± 21.0 vs 50.6 ± 19.6), and PBW (training set: 205.4 ± 73.7 vs 199.4 ± 74.2; external validation set: 142.4 ± 76.0 vs 159.9 ± 73.0). In the univariate analysis, MI (OR 0.2, 95% CI 0.07-0.55, P = .002), CAD (OR 0.3, 95% CI 0.14-0.65, P = .002), EDV (OR 1, 95% CI 0.99-1.00, P = .013), ACEI or ARB (OR 3.44, 95% CI 1.3-9.13, P = .013), scar score (OR 0.96, 95% CI 0.94-0.99, P = .016) showed significant predictive values to CRT response. The AE-extracted predictor showed significant predictive value at the significance level of 0.1 (95% CI 0.95-2.12, P = .092), and it alone achieved an AUC of 0.57 (95% CI 0.48-0.66).

Figure 5 illustrates four patient examples. According to the conventional LVMD parameters (PSD $> 21^\circ$ and PBW $> 112^\circ$;[6] PSD$> 43^\circ$ and PBW $> 128^\circ$;[7] PSD $> 43^\circ$ and PBW $> 135^\circ$[10]), A and C might be responders to CRT, and B and D might be non-responders to CRT. However, post-CRT follow-ups showed that A and B were CRT responders, and C and D were CRT non-responders; the AE-extracted feature successfully predicted the CRT response for these patients.

DISCUSSION

In this study, we proposed and externally validated the use of AE techniques to extract new parameters from baseline gated SPECT MPI and discover predictors for CRT patient selection. The AE-extracted LVMD predictor was significantly different from and superior to

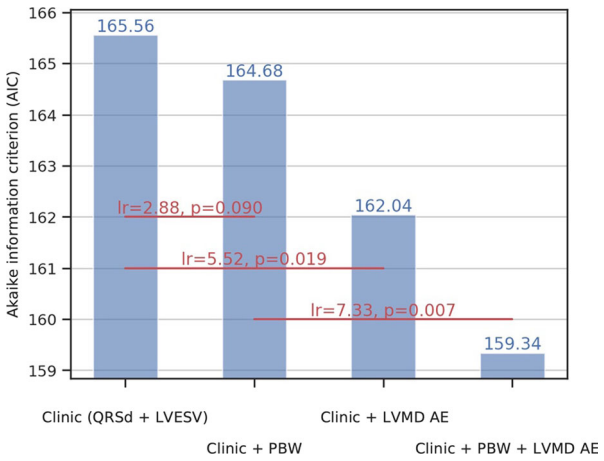

Figure 3. Fitting performance and incremental values of the AE-extracted LVMD predictor in the prediction of CRT response. Akaike information criterion (AIC) reflected the fitting performance of the model. The larger the value, the better the fitting performance of the model. The likelihood ratio test compared the goodness of fit of two nested models (two models were connected by a red line) and reflected the incremental predictive value of the newly added variables. AE-extracted LVMD had incremental predictive value over both the clinic parameters (LR = 5.52, P = .019) and the combination of clinical variables and PBW (LR = 7.33, P = .007). *QRSd*, QRS duration; *LVESV*, left ventricular end-systolic volume; *PBW*, phase bandwidth; *LVMD*, left ventricular mechanical dyssynchrony; *AE*, autoencoder; *LR*, likelihood ratio. Clinic parameters include QRSd and LVESV.

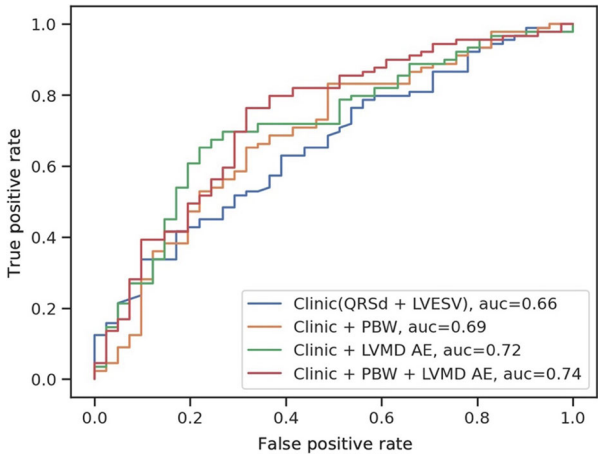

Figure 4. Receiver-operating characteristic curves to predict the CRT response. *QRSd*, QRS duration; *LVESV*, left ventricular end-systolic volume; *PBW*, phase bandwidth; *LVMD*, left ventricular mechanical dyssynchrony; *AE*, autoencoder.

conventional LVMD parameters. Moreover, it had an incremental value over conventional LVMD parameters and clinical variables.

Limitations of conventional LVMD parameters

Early retrospective studies suggested a significant association between baseline LVMD and CRT response.[6–8,10,36] Three studies proposed three different PBW thresholds (112° vs 128° vs 135°) for the prediction of CRT response, and they all claimed to have good prediction results (sensitivity, 72% vs 86% vs 70%, specificity, 70% vs 80% vs 70%); it is worth noting that these thresholds were all defined from small sample populations (n = 30, 32, 42, respectively).[6,7,10] PSD had similar results in these three articles. However, the data from GUIDE-CRT clinical trial showed that PSD and PBW were not associated with volumetric CRT response, referred to as a reduction of $\geq 15\%$ in LVESV.[16,37] Another multicenter trial (VISION-CRT) showed that baseline PSD and PBW were not associated with the CRT response, which was defined by any improvement in one or more of the following: decrease of ≥ 1 in NYHA class, increase of $\geq 5\%$ in LVEF, decrease of $\geq 15\%$ in left ventricular end-systolic volume (LVESV), or decrease of ≥ 5 points in Minnesota Living With Heart Failure Questionnaire.[14] Gendre et al[15] demonstrated that baseline LVMD parameters, including PSD and PBW, could not predict the response to CRT defined by a reduction of LVESV $\geq 15\%$ or improve peak $VO_2 \geq 10\%$, even though the study was a single-center study with a small number of patients (n = 42). Zhang et al[16] also found similar results in a multicenter study with 79 CRT patients, in which a reduction of LVESV $\geq 15\%$ was used to define the volumetric response to CRT. Existing global parameters (PSD and PBW) characterizing LVMD are easily affected by the outliers of phase measurement.[11–13] PSD may be deceptive for characterizing the widely distributed and multi-modal distributions in phase histograms; PBW includes almost the entire distribution range of the histogram (95%).[13,38] The statistical analysis relies heavily on pre-assumed relationships between factors, and there are problems related to identifying appropriate data, processing interconnected rather than independent factors, and even violating assumptions.[39] A method that comprehensively considers the relationship between variables and extracts potential, new, and more predictive features for CRT is needed.

Advantages and interpretability of AE model for feature learning

Compared with conventional statistical methods, data-driven feature learning through deep learning has higher performance. The advantage of deep learning for feature learning is a layered architecture similar to the

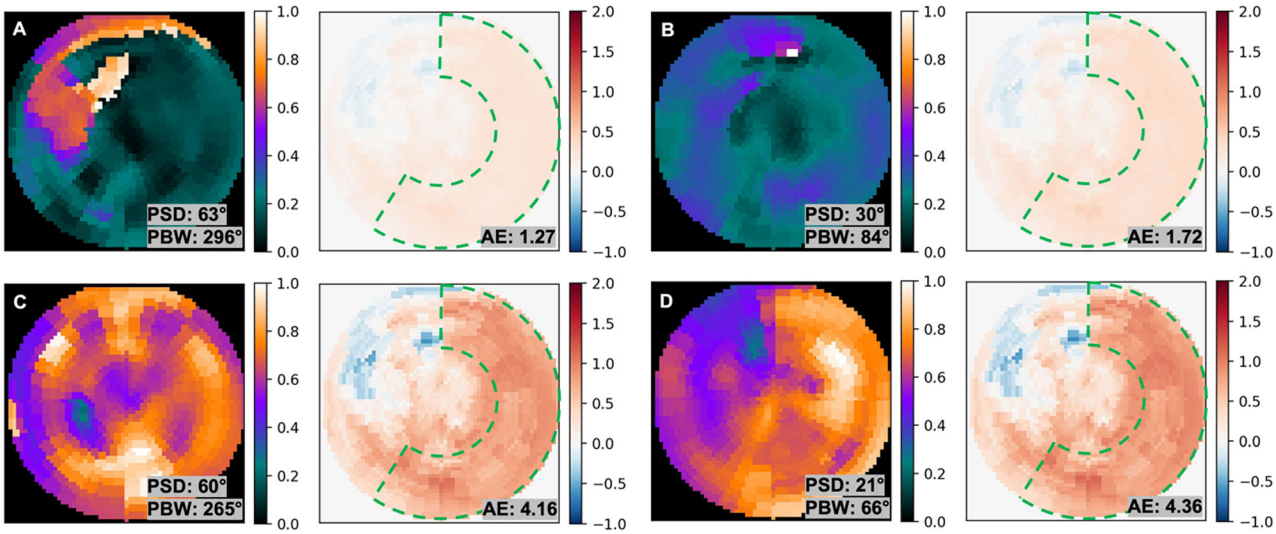

Figure 5. Illustrations of PSD and PBW vs AE-extracted LVMD predictor for 4 patients (Patients A and B: CRT non-responders; Patient C and D: CRT responders). The left graph for each patient is the systolic phase polarmap and the right graph is the weight heatmap. For the weight heatmap, the higher saturation of the color indicates the higher absolute value of the weights in the deep neural networks. Red and blue colors indicate positive and negative values, respectively. The green dashed box indicates the half-moon shaped region, including part of the anterior wall, the complete lateral and inferior walls, excluding the septum and the apex. *AE*, autoencoder; *LVMD*, left ventricular mechanical dyssynchrony; *PSD*, Phase histogram standard deviation; *PBW*, Phase histogram bandwidth; *CRT*, cardiac resynchronization therapy.

human brain.[40] Through deep learning, simple features are extracted from the raw data, and then more complex features are learned through multiple layers. Finally, a considerable number of robust features are generated through multiple iterations of learning. Specifically, feature learning is classified into two categories, supervised learning and unsupervised learning.

In supervised learning, labeled data are forwarded from the input to the output for prediction. Backpropagation is used to optimize the parameters of the deep learning model by minimizing the cost function between the target value and the predicted value. Betancur et al[22] proposed a deep convolutional neural network for predicting the probability of obstructive coronary artery disease in the left anterior descending artery, left circumflex artery, and right coronary artery that had better per-patient sensitivity (from 79.8% to 82.3%, $P < .05$) and per-vessel sensitivity (from 64.4% to 69.8%, $P < .01$) than total perfusion deficit. Such end-to-end supervised learning neural network avoids the tricky problem of determining which components are needed to perform a learning task and how those components interact, but also brings a lack of interpretability and requires significant amount of data. In unsupervised learning, unlabeled data are used to learn new features and find new patterns on their own, which can yield undiscovered information that might require human intervention to understand the hidden patterns and correlate them with the domain knowledge. Cikes et al[41] proposed an unsupervised approach to analyze the phenotype of heterogeneous HF by integrating clinical variables and eight echocardiographic descriptors (traces) extracted from full cardiac cycle echocardiographic images. They emphasized that their unsupervised method was not trained based on a priori knowledge, and the interpretability of their model was based on the distribution of existing prior variables in different phenogroups.

Interpretable machine learning approaches based on complex mathematical formulas are rendered black boxes by the complexity and scale of their structures.[30,42] Some state-of-the-art models, including deep learning and ensemble models, limit the clinical actionability of model predictions due to the ''black box'' structure, which further undermines their usefulness to clinicians.[43] Heatmaps of the weights used in this paper is a straightforwardly interpretable method that shows the relationship from the input LVMD polarmap, mapped to each AE-extracted predictor from the hidden layer. The brightness signifies the weights associated with single filters' activations (a specific template) from the hidden layer.

Typically, LV lead of the CRT is implanted in the lateral wall, or a region adjacent to the lateral wall, which can encompass the anterior or inferior walls of the left ventricle (depending on coronary sinus venous anatomy). Thus, the ideal site for LV lead placement could be in a half-moon shaped region of the polar map, encompassing anterior, lateral, and inferior walls, while excluding the septum. Studies have indicated that LV lead placement in the apex is associated with increased risk of HF and death[44–46], and thus that is also excluded from this half-moon of the polar map. Therefore, ideal AE-extracted predictors of positive CRT response would align on the half-moon of polar map.

As shown in Fig. 5, this weight heatmap is ''responsible'' for classifying the half-moon shaped phase polarmap (roughly two thirds of a circle without the apex of the LV), and its goal is to output a high value for half-moon shaped polarmap and a low value for non-half-moon shaped polarmap. Suppose that we receive an input phase polarmap like Fig. 5D, we can anticipate that the neurons responsible for classifying half-moon shaped polarmap should have high values, because their weights are such that high weights tend to align with pixels tending to be high in half-moon shaped polarmap. For other non-half-moon shaped phase polarmap, such as Fig. 5A, most of the pixels would not line up with a half-moon shaped polar map, so less overlap would negate high-valued pixels in those images by low weights. The better the input image matches this half-moon shaped template, the higher the AE score obtained. It can be observed that this half-moon shaped polarmap template has higher weights in the bright anterior, lateral, and inferior myocardial walls, and lower weights in the dark apical, septal, and anteroseptal walls, which is consistent to clinical experience. Accordingly, when we design the CRT predictor based on LVMD, we should put more focus on the anterior, lateral, and inferior myocardial walls.

In this paper, we proposed and externally validated the use of an unsupervised learning algorithm, autoencoder, for feature learning from SPECT MPI to predict CRT response. Our new AE-extracted predictor in this study is significantly different from conventional LVMD parameters. This data-driven approach of unsupervised learning requires only unlabeled training images, is able to learn non-linear relationships and avoids the loss of important information in feature extraction. The model has been tested in an external validation dataset, and the results seen here are reproducible and have significant predictive value for CRT response. The performance of AE-extracted predictors can be improved with increase of the patient sample size. More importantly, unlike the global variables PSD and PBW, AE-extracted feature assigns higher weights to anterior, lateral, and inferior myocardial walls of interest, which are consistent with the recommended pacing sites of the LV lead in clinical practice.

LIMITATIONS

Several limitations of this study should be acknowledged. First, this study enrolled a relatively small number of patients from multiple medical centers with the inherent limitation of such study design. The performance and feasibility of the data-driven system may be affected by the quality of data. Second, the lack of post-CRT SPECT MPI in the training set and the lack of pre- and post-CRT echocardiography in the external validation dataset led us to use the same threshold ($a > 5\%$ increase LVEF) but by different imaging modalities (training: echo vs external validation: gated SPECT MPI) to define the CRT response. Although many studies demonstrated a good correlation between LVEF measured by gated SPECT MPI and echocardiography,[47–50] the different definitions of CRT response between the two trials may introduce bias in the prediction performance. Moreover, some clinical parameters in the training dataset were not available in the external validation dataset (e.g., CKD), and the distributions of the patient characteristics (e.g., QRS morphology, QRS duration, PSD, and PBW) were substantially different between the training and external validation dataset. Therefore, the performances of the AE-extracted predictor were different between the two datasets.

NEW KNOWLEDGE GAINED

This paper proposed a method using deep learning to discover new LVMD predictors from gated SPECT MPI. The new predictor is different from existing LVMD parameters and has significant predictive value for CRT response.

CONCLUSION

AE techniques have significant value in the discovery of new clinical parameters. The new AE-extracted LVMD predictor extracted from the baseline gated SPECT MPI has the potential to improve the prediction of CRT response.

Acknowledgments

This research was supported by a grant from the American Heart Association (Project Number: 17AIREA33700016, PI: Weihua Zhou), a new faculty grant from Michigan Technological University Institute of

Computing and Cybersystems (PI: Weihua Zhou), a Michigan Technological Research Excellence Fund Research Seed grant (PI: Weihua Zhou), and a grant from the National Nature Science Foundation of China (Project Number: 82070521, PI: Jiangang Zou). The authors would like to thank the International Atomic Energy Agency (IAEA) for providing access to the data of the multicenter trial: ''Value of intraventricular synchronism assessment by gated-SPECT myocardial perfusion imaging in the management of heart failure patients submitted to cardiac resynchronization therapy'' (IAEA VISION-CRT), Coordinated Research Protocol E1.30.34.

Author contributions

ZH: Conception design, and analysis and interpretation of data; drafting of the manuscript and revising it critically for important intellectual content; and final approval of the manuscript submitted. XZ: Active involvement in collecting data and performing experiments with subsequent participation in data analysis, drafting the manuscript, and revising it critically for important intellectual content. CZ: Conception design, and analysis and interpretation of data. XL: Conception design, and analysis and interpretation of data. SM: Analysis and interpretation of data. ZQ: Active involvement in collecting data and performing experiments with subsequent participation in data analysis. YW: Active involvement in collecting data and performing experiments with subsequent participation in data analysis. XH: Active involvement in collecting data and performing experiments with subsequent participation in data analysis. JZ: Active involvement in collecting data and performing experiments with subsequent participation in data analysis; and final approval of the manuscript submitted. WZ: Conception design, and analysis and interpretation of data; drafting of the manuscript and revising it critically for important intellectual content; and final approval of the manuscript submitted.

Disclosures

All authors declare that there are no conflict of interest.

References

1. Tracy CM, Epstein AE, Darbar D, DiMarco JP, Dunbar SB, Estes NAM. 2012 ACCF/AHA/HRS focused update incorporated into the ACCF/AHA/HRS 2008 guidelines for device-based therapy of cardiac rhythm abnormalities: A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. J Am Coll Cardiol 2013;61:e6-75.
2. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J 2016;37:2129-2200m.
3. Zhou W, Garcia EV. Nuclear image-guided approaches for cardiac resynchronization therapy (CRT). Curr Cardiol Rep 2016;18:7.
4. He Z, Garcia EV, Zhou W. Nuclear imaging guiding cardiac resynchronization therapy. In: Nuclear cardiology: Basic and advanced concepts in clinical practice. Berlin: Springer; 2021.
5. Chung ES, Leon AR, Tavazzi L, Sun JP, Nihoyannopoulos P, Merlino J, et al. Results of the predictors of response to CRT (PROSPECT) trial. Circulation 2008;117:2608-16.
6. Azizian N, Rastgou F, Ghaedian T, Golabchi A, Bahadorian B, Khanlarzadeh V, et al. LV dyssynchrony assessed with phase analysis on gated myocardial perfusion spect can predict response to CRT in patients with end-stage heart failure. Res Cardiovasc Med 2014;3:6.
7. Mukherjee A, Patel CD, Naik N, Sharma G, Roy A. Quantitative assessment of cardiac mechanical dyssynchrony and prediction of response to cardiac resynchronization therapy in patients with nonischaemic dilated cardiomyopathy using gated myocardial perfusion SPECT. Nucl Med Commun 2015. https://doi.org/10.1097/MNM.0000000000000282.
8. Friehling M, Chen J, Saba S, Bazaz R, Schwartzman D, Adelstein EC, et al. A prospective pilot study to evaluate the relationship between acute change in left ventricular synchrony after cardiac resynchronization therapy and patient outcome using a single-injection gated SPECT protocol. Circ Cardiovasc Imaging 2011;4:532-9.
9. Tsai SC, Chang YC, Chiang KF, Lin WY, Huang JL, Hung GU, et al. LV dyssynchrony is helpful in predicting ventricular arrhythmia in ischemic cardiomyopathy after cardiac resynchronization therapy a preliminary study. Med U S 2016;95:e2840.
10. Henneman MM, Chen J, Dibbets-Schneider P, Stokkel MP, Bleeker GB, Ypenburg C, et al. Can LV dyssynchrony as assessed with phase analysis on gated myocardial perfusion SPECT preferably predict response to CRT? J Nucl Med 2007;48:1104-11.
11. O'Connell JW, Schreck C, Moles M, Badwar N, DeMarco T, Olgin J, et al. A unique method by which to quantitate synchrony with equilibrium radionuclide angiography. J Nucl Cardiol 2005;12:441-50.
12. Van Kriekinge SD, Nishina H, Ohba M, Berman DS, Germano G. Automatic global and regional phase analysis from gated myocardial perfusion SPECT imaging: Application to the characterization of ventricular contraction in patients with left bundle branch block. J Nucl Med 2008;49:1790-7.
13. Nakajima K, Okuda K, Matsuo S, Kiso K, Kinuya S, Garcia EV. Comparison of phase dyssynchrony analysis using gated myocardial perfusion imaging with four software programs: Based on the Japanese Society of Nuclear Medicine working group normal database. J Nucl Cardiol 2017;24:611-21.
14. Peix A, Karthikeyan G, Massardo T, Kalaivani M, Patel C, Pabon LM, et al. Value of intraventricular dyssynchrony assessment by gated-SPECT myocardial perfusion imaging in the management of heart failure patients undergoing cardiac resynchronization therapy (VISION-CRT). J Nucl Cardiol 2019. https://doi.org/10.1007/s12350-018-01589-5.
15. Gendre R, Lairez O, Mondoly P, Duparc A, Carrié D, Galinier M, et al. Research of predictive factors for cardiac resynchronization therapy: a prospective study comparing data from phase-analysis of gated myocardial perfusion single-photon computed tomography and echocardiography: Trying to anticipate response to CRT. Ann Nucl Med 2017;31:218-26.
16. Zhang X, Qian Z, Tang H, Hua W, Su Y, Xu G, et al. A new method to recommend left ventricular lead positions for improved CRT volumetric response and long-term prognosis. J Nucl Cardiol 2019;28:372-84.
17. Tajik AJ. Machine learning for echocardiographic imaging: Embarking on another incredible journey. J Am Coll Cardiol 2016;68:2296-8.

18. Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives. IEEE Trans Pattern Anal Mach Intell 2013;35:1798-828.
19. Betancur J, Rubeaux M, Fuchs TA, Otaki Y, Arnson Y, Slipczuk L, et al. Automatic valve plane localization in myocardial perfusion SPECT/CT by machine learning: Anatomic and clinical validation. J Nucl Med 2017;58:961-7.
20. Lecun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436-44.
21. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature 2017;542:115-8.
22. Betancur J, Commandeur F, Motlagh M, Sharir T, Einstein AJ, Bokhari S, et al. Deep learning for prediction of obstructive disease from fast myocardial perfusion SPECT. A Multicenter Study. JACC Cardiovasc Imaging 2018;11:1654-63.
23. Xu Y, Mo T, Feng Q, Zhong P, Lai M, Chang EI. Deep learning of feature representation with multiple instance learning for medical image analysis. State Key Laboratory of Software Development Environment , Key Laboratory of Biomechanics and Mechanobiology of Ministry of Education , Beihang University M. *IEEE Int Conf Acoust Speech Signal Process* 2014; 1645–49.
24. Bleeker GB, Bax JJ, Fung JW-H, van der Wall EE, Zhang Q, Schalij MJ, et al. Clinical versus echocardiographic parameters to assess response to cardiac resynchronization therapy. Am J Cardiol 2006;97:260-3.
25. Bax JJ, Marwick TH, Molhoek SG, Bleeker GB, van Erven L, Boersma E, et al. Left ventricular dyssynchrony predicts benefit of cardiac resynchronization therapy in patients with end-stage heart failure before pacemaker implantation. Am J Cardiol 2003;92:1238-40.
26. Henzlova MJ, Duvall WL, Einstein AJ, Travin MI, Verberne HJ. ASNC imaging guidelines for SPECT nuclear cardiology procedures: Stress, protocols, and tracers. J Nucl Cardiol 2016;23:606-39.
27. Chen J, Kalogeropoulos AP, Verdes L, Butler J, Garcia EV. Left-ventricular systolic and diastolic dyssynchrony as assessed by multi-harmonic phase analysis of gated SPECT myocardial perfusion imaging in patients with end-stage renal disease and normal LVEF. J Nucl Cardiol 2011;18:299-308.
28. Chen J, Garcia EV, Folks RD, Cooke CD, Faber TL, Tauxe EL, et al. Onset of left ventricular mechanical contraction as determined by phase analysis of ECG-gated myocardial perfusion SPECT imaging: Development of a diagnostic tool for assessment of cardiac mechanical dyssynchrony. J Nucl Cardiol 2005;12:687-95.
29. Goodfellow I, Bengio Y, Courville A. Autoencoders. In: Deep Learning. Cambridge: MIT Press; 2016.
30. Goodfellow I, Yoshua B, Courville A. Deep learning. Cambridge: MIT Press; 2016.
31. Zhang Y, Chen W, Yeo CK, Lau CT, Lee BS. Detecting rumors on Online Social Networks using multi-layer autoencoder. In: 2017 IEEE Technology & Engineering Management Conference (TEMSCON). Santa Clara: IEEE; 2017. p. 437-41.
32. Paszke A, Gross S, Chintala S, Chanan G, Yang E, DeVito Z, et al. Automatic differentiation in PyTorch; 2017
33. Seabold S, Perktold J. Statsmodels: Econometric and statistical modeling with python. *Proc 9th Python Sci Conf* 2010; p. 92–96.
34. Jimenez-Heffernan A, Butt S, Mesquita CT, Massardo T, Peix A, Kumar A, et al. Technical aspects of gated SPECT MPI assessment of left ventricular dyssynchrony used in the VISION-CRT study. J Nucl Cardiol 2020. https://doi.org/10.1007/s12350-020-02122-3.
35. He Z, de Amorim Fernandes F, do Nascimento EA, Garcia EV, Mesquita CT, Zhou W. Incremental value of left ventricular shape parameters measured by gated SPECT MPI in predicting the super-response to CRT. J Nucl Cardiol 2021. https://doi.org/10.1007/s12350-020-02469-7.
36. Nakamura K, Takami M, Shimabukuro M, Maesato A, Chinen I, Ishigaki S, et al. Effective prediction of response to cardiac resynchronization therapy using a novel program of gated myocardial perfusion single photon emission computed tomography. Europace 2011;13:1731-7.
37. Tao N, Qiu Y, Tang H, Qian Z, Wu H, Zhu R, et al. Assessment of left ventricular contraction patterns using gated SPECT MPI to predict cardiac resynchronization therapy response. J Nucl Cardiol 2018;25:2029-38.
38. Zhou W, Hung G-U. Left-ventricular mechanical dyssynchrony in the prognosis of dilated cardiomyopathy: Which parameter is more useful? J Nucl Cardiol 2018;25:1688-91.
39. Spicker P. The real dependent variable problem: The limitations of quantitative analysis in comparative policy studies. Soc Policy Adm 2018;52:216-28.
40. Chen M, Shi X, Zhang Y, Wu D, Guizani M. Deep features learning for medical image analysis with convolutional autoencoder neural network. IEEE Trans Big Data 2017;7790:1-1.
41. Cikes M, Sanchez-Martinez S, Claggett B, Duchateau N, Piella G, Butakoff C, et al. Machine learning-based phenogrouping in heart failure to identify responders to cardiac resynchronization therapy. Eur J Heart Fail 2019;21:74-85.
42. Petch J, Di S, Nelson W. Opening the black box: The promise and limitations of explainable machine learning in cardiology. Can J Cardiol 2022;38:204-13.
43. Guidi JL, Clark K, Upton MT, Faust H, Umscheid CA, Lane-Fall MB, et al. Clinician perception of the effectiveness of an automated early warning and response system for sepsis in an Academic Medical Center. Ann Am Thorac Soc 2015;12:1514-9.
44. Becker M, Altiok E, Ocklenburg C, Krings R, Adams D, Lysansky M, et al. Analysis of LV lead position in cardiac resynchronization therapy using different imaging modalities. JACC Cardiovasc Imaging 2010;3:472-81.
45. Zhou W, Hou X, Piccinelli M, Tang X, Tang L, Cao K, et al. 3D fusion of LV venous anatomy on fluoroscopy venograms with epicardial surface on SPECT myocardial perfusion images for guiding CRT LV lead placement. JACC Cardiovasc Imaging 2014;7:1239-48.
46. Singh JP, Klein HU, Huang DT, Reek S, Kuniss M, Quesada A, et al. Left ventricular lead position and clinical outcome in the multicenter automatic defibrillator implantation trial-cardiac resynchronization therapy (MADIT-CRT) trial. Circulation 2011;123:1159-66.
47. Mizunobu M, Sakai J, Sasao H, Murai H, Fujiwara H. Assessment of left ventricular systolic and diastolic function using ECG-gated technetium-99m tetrofosmin myocardial perfusion SPECT: Comparison with ultrasound echocardiography. Int Heart J 2013;54:212-5.
48. Garg N, Dresser T, Aggarwal K, Gupta V, Mittal MK, Alpert MA. Comparison of left ventricular ejection fraction values obtained using invasive contrast left ventriculography, two-dimensional echocardiography, and gated single-photon emission computed tomography. SAGE Open Med 2016;4:205031211665594.
49. Gimelli A, Landi P, Marraccini P, Sicari R, Frumento P, L'Abbate A, et al. Left ventricular ejection fraction measurements: accuracy and prognostic implications in a large population of patients with known or suspected ischemic heart disease. Int J Cardiovasc Imaging 2008;24:793-801.

50. Shojaeifard M, Ghaedian T, Yaghoobi N, Malek H, Firoozabadi H, Bitarafan-Rajabi A, et al. Comparison of gated SPECT myocardial perfusion imaging with echocardiography for the measurement of left ventricular volumes and ejection fraction in patients with severe heart failure. Res Cardiovasc Med 2015. https://doi.org/10.5812/cardiovascmed.29005.