Study Design

DIGITAL HEALTH

DIGITAL HEALTH
Volume 10: 1–12

DOI: 10.1177/20552076241260557
journals.sagepub.com/home/dhj

Sage

Assessment of left ventricular ejection fraction in artificial intelligence based on left ventricular opacification

Ye Zhu[1,2,3,†], Zisang Zhang[1,2,3,†], Junqiang Ma[4,5,†], Yiwei Zhang[1,2,3], Shuangshuang Zhu[1,2,3], Manwei Liu[1,2,3], Ziming Zhang[1,2,3], Chun Wu[1,2,3], Chunyan Xu[1,2,3], Anjun Wu[1,2,3], Chenchen Sun[1,2,3], Xin Yang[6], Yonghuai Wang[7,8], Chunyan Ma[7,8], Jun Cheng[4,5], Dong Ni[4,5], Jing Wang[1,2,3], Mingxing Xie[1,2,3], Wufeng Xue[4,5] and Li Zhang[1,2,3]

Abstract

Background: Left ventricular opacification (LVO) improves the accuracy of left ventricular ejection fraction (LVEF) by enhancing the visualization of the endocardium. Manual delineation of the endocardium by sonographers has observer variability. Artificial intelligence (AI) has the potential to improve the reproducibility of LVO to assess LVEF.

Objectives: The aim was to develop an AI model and evaluate the feasibility and reproducibility of LVO in the assessment of LVEF.

Methods: This retrospective study included 1305 echocardiography of 797 patients who had LVO at the Department of Ultrasound Medicine, Union Hospital, Huazhong University of Science and Technology from 2013 to 2021. The AI model was developed by 5-fold cross validation. The validation datasets included 50 patients prospectively collected in our center and 42 patients retrospectively collected in the external institution. To evaluate the differences between LV function determined by AI and sonographers, the median absolute error (MAE), spearman correlation coefficient, and intraclass correlation coefficient (ICC) were calculated.

Results: In LVO, the MAE of LVEF between AI and manual measurements was 2.6% in the development cohort, 2.5% in the internal validation cohort, and 2.7% in the external validation cohort. Compared with two-dimensional echocardiography (2DE), the left ventricular (LV) volumes and LVEF of LVO measured by AI correlated significantly with manual measurements. AI model provided excellent reliability for the LV parameters of LVO (ICC > 0.95).

Conclusions: AI-assisted LVO enables more accurate identification of the LV endocardium and reduces observer variability, providing a more reliable way for assessing LV function.

[1]Department of Ultrasound, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[2]Clinical Research Center for Medical Imaging in Hubei Province, Wuhan, China
[3]Hubei Province Key Laboratory of Molecular Imaging, Wuhan, China
[4]School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, China
[5]National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, Shenzhen, China
[6]Electronics and Information Engineering Department, Huazhong University of Science and Technology, Wuhan, China
[7]Department of Cardiovascular Ultrasound, The First Hospital of China Medical University, Shenyang, Liaoning, China
[8]Clinical Medical Research Center of Imaging in Liaoning Province, Shenyang, Liaoning, China

[†]Contributed equally.

Corresponding author:
Mingxing Xie, Department of Ultrasound, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1277 Jiefang Avenue, Wuhan 430022, China.
Email: xiemx@hust.edu.cn;

Wufeng Xue, School of Biomedical Engineering, Health Science Center, Shenzhen University, 3688 Nanhai Avenue, Shenzhen 518060, China,
Email: xuewf@szu.edu.cn;

Li Zhang, Department of Ultrasound, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, 1277 Jiefang Avenue, Wuhan 430022, China.
E-mail: zli429@hust.edu.cn.

Keywords

Echocardiography, left ventricular opacification, left ventricular ejection fraction, artificial intelligence

Submission date: 10 December 2023; Acceptance date: 23 May 2024

Introduction

Accurate assessment of cardiac function is crucial for diagnosing cardiovascular diseases[1] and decision-making regarding the clinical management of patients with heart failure,[1–4] myocardial infarction,[5] heart valve disease,[6,7] cancer, and so on.[8,9] Left ventricular ejection fraction (LVEF) is the most commonly used echocardiographic parameter to assess left ventricular (LV) function in clinical routine.[10] The current guidelines for echocardiography highlight the necessity of precise measurement of LVEF, as its cutoff points are essential indications for therapeutic interventions.[11] However, traditional methods to assess LVEF highly depend on individual experience and have considerable observer variability. Recent advances have demonstrated the ability of artificial intelligence (AI) to accurately identify endocardial borders and measure LV volumes and function with minimal manual assistance, with proven accuracy and reliability similar to expert-provided LVEF reference values.[12–14]

The interface between blood and tissue, combined with the presence of endocardial trabeculae, significantly influenced the evaluation of LVEF by sonographers.[15] Contrast echocardiography, especially left ventricular opacification (LVO), has been proven to improve 70%∼90% of poor echocardiographic images, which helps sonographers quantify LV volumes and LVEF more accurately.[16,17] However, whether AI can automatically analyze LVO echocardiograms, reduce observer variability in measuring LVEF, and improve measurement reproducibility remains unclear.

Accordingly, we presented a temporally-consistent AI algorithm to estimate LVEF automatically based on the LVO dataset. The dataset comprised two-dimensional echocardiography (2DE) and LVO videos obtained from 1100 patients. The feasibility and reproducibility of automated quantification of LV volumes and LVEF in LVO were evaluated, and results obtained by the AI-assisted approach were compared with those of experienced sonographers.

Methods

The study was approved by the Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China. The ethics committee waived the need for informed consent because the images were obtained during routine clinical practice.

Datasets

The study consisted of three cohorts. The development cohort 1 retrospectively included 1000 patients who had undergone LVO in our center from 2013 to 2021. We excluded 203 studies with incomplete LV structures or non-standard views. Poor image quality was not an exclusion standard. Finally, cohort 1 included 634 apical four-chamber (A4C) videos and 295 apical two-chamber (A2C) videos for 2DE and 671 A4C videos and 319 A2C videos for LVO from 797 patients. The echocardiograms were mainly acquired with EPIQ 7C, Philips iE33, and IE Elite. The prospective internal validation cohort 2 comprised 50 patients from our center who underwent both 2DE and LVO examinations, including A4C and A2C videos. LVO was performed utilizing SonoVue® (Bracco, Italy) as the ultrasonic contrast agent. The independent external validation cohort 3 included 50 patients with 2DE and LVO A4C videos from the Department of Ultrasound Medicine, First Hospital of China Medical University. Due to incomplete 2DE and LVO views, 8 patients in cohort 3 were excluded.

Data preprocessing and annotation

All datasets were anonymized and sensitive information around the images was removed. This study employed an independently developed convolutional neural network to classify 2DE and LVO A4C videos[18] automatically. Due to the differences in image sizes captured by different equipment, all videos were resized to 256×256 pixels by linear interpolation, and the intensity was normalized to [–1, 1]. Using the end of diastole (ED) as the start frame and the end of systole (ES) as the end frame, 10 frames are sampled from each video at equal intervals to form a sequence.

Two sonographers independently completed data annotation. Data quality control was conducted by a senior sonographer. Cohort 1 and Cohort 3 were randomly assigned to

two sonographers. A representative cardiac cycle was chosen from each video, and the LV endocardium was manually traced at ED and ES. According to the recommendations of the American Society of Echocardiography,[19] the papillary muscles and trabeculations were contained within the LV cavity. Two sonographers independently annotated cohort 2 to assess the test-retest variability, in which the LV endocardium was manually traced at ED and ES. For intra-observer variability, the same observer re-analyzed the data within one month, blinded to previous measurements. For inter-observer variability, two observers independently labeled the videos in cohort 2. Annotation results were randomly sampled at a 10% proportion and checked by the senior sonographer.

Temporal-consistent deep learning model development and training

We presented the Co-Learning of segmentation and tracking on appearance and shape level (CLAS) algorithm based on 3D U-Net to perform bidirectional motion tracking and video segmentation tasks.[20] CLAS used 3D U-Net as a shared feature extractor. The model adopted ED and ES frames to extract rich semantic information from intermediate frames through bidirectional motion tracking. This improved the accuracy of segmentation results and enhanced the temporal consistency of sequence segmentation results. Based on the segmentation of the LV, the monoplane Simpson's method was used to calculate the LV end-diastolic volume (LVEDV) and LV end-systolic volume (LVESV) and assess LVEF, as shown in Figure 1. During training, 5-fold cross-validation was adopted. We split the dataset at the patient level to ensure the reliability of the results. The training hyperparameters and learning curve of CLAS were detailed in Appendix Page 1 and Figure S1.

Statistical analysis

The Shapiro-Wilk test was used to test normality. Continuous variables were expressed as median and interquartile ranges (IQR). Categorical data were presented as a percentage. The accuracy of LV segmentation was evaluated using the Dice similarity coefficient (DSC), mean surface distance (MAD), and Hausdorff distance (HD) (calculation formulas shown in Appendix Page 1-2). The median absolute error (MAE) was used to compare the differences between AI and manual measurements of LVEDV, LVESV, and LVEF. The relative bias was determined by dividing the absolute difference between AI and manual measurements by the manual measurements. The reliability of LV volumes and LVEF in 2DE or LVO was evaluated using linear regression, Spearman's correlation coefficient, Bland Altman analysis, and intraclass correlation coefficients. The bias (mean difference), limits of agreement (LOA), and standard deviation of bias (SD of Bias) were calculated. The LOA and SD of Bias were used to estimate the precision or random error of the measurements around the bias. For all statistical tests, P values <0.05 were considered statistically significant. Statistical analyses were performed with R (version 4.2.2) and Python (version 3.7).

Results

Study population

The study included 566 men and 231 women aged 41 to 62 (median age 53) in cohort 1. The dataset included different diseases, such as hypertrophic cardiomyopathy (HCM), ischemic cardiomyopathy (ICM), dilated cardiomyopathy (DCM), cardiac space-occupying lesions (CSOL), and others. The detailed baseline characteristics were summarized in Appendix Table S1. 50 patients (29 male, 21 female) aged 41–59 were in cohort 2 (median age 53), as shown in Table 1. And cohort 3 comprised 42 patients (34 male and 16 female) aged 38–65 (median age 56.5).

LV segmentation results of the development and internal validation cohort

The AI model utilized the myocardial motion information captured by the optical flow algorithm to reconstruct intermediate frames, assist with ED and ES segmentation, and improve the temporal consistency of the segmentation results of echocardiographic sequences, as shown in Figure 2. For segmentation, LV endocardium labels included 1342 frame-level segmentations of LVO and 1268 frame-level segmentations of 2DE performed by sonographers. These sparse labels were used to train CLAS to generate frame-level segmentation of the entire sequence.

LV segmentation with LVO was better than that with 2DE. In the development cohort, the DSC for ED segmentation in LVO was 0.96, and the DSC for ES was 0.93. In 2DE, the DSC for ED segmentation was 0.94, and the DSC for ES was 0.89. The detailed results are shown in Appendix Table S2. In the internal validation cohort, the segmentation results for LVO (ED: DSC, 0.96, MAD, 1.08 mm, HD, 4.11 mm; ES: DSC, 0.91, MAD, 1.46 mm, HD, 5.29 mm) were better than those obtained by 2DE (ED: DSC, 0.94, MAD, 1.61 mm, HD, 5.23 mm; ES: DSC, 0.88, MAD, 2.1 mm, HD, 6.66 mm), as shown in Appendix Table S3.

The variability analysis of LV volumes and LVEF

To investigate the feasibility of AI in quantifying cardiac function, we compared the measurements obtained by CLAS with those manually measured by sonographers using MAE, correlation coefficient, and agreement analysis.

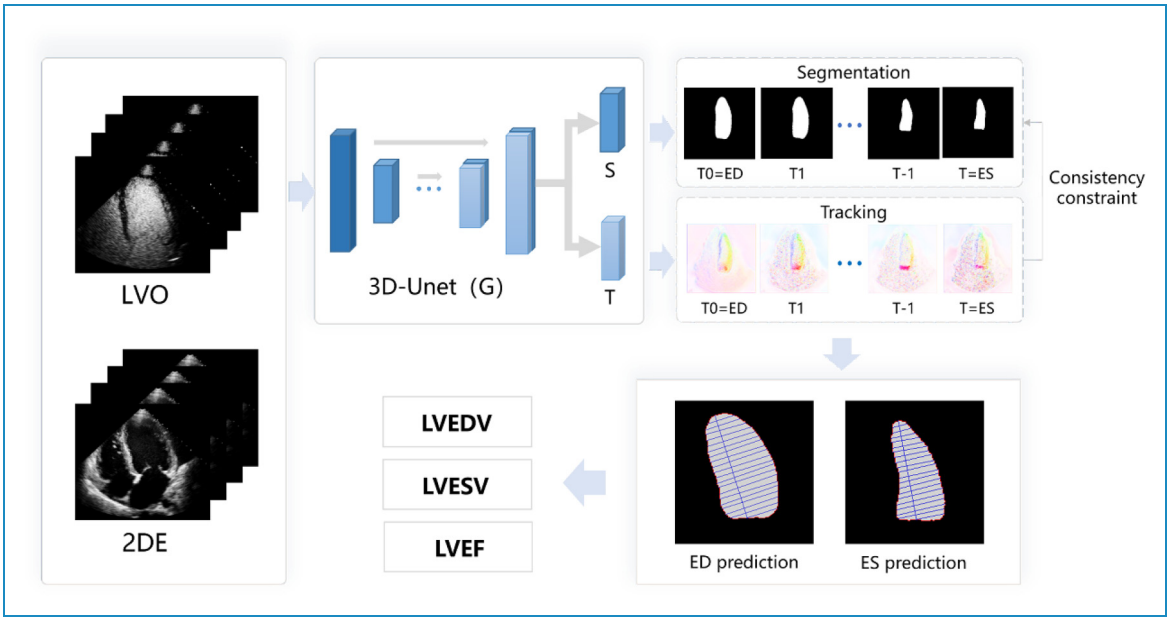

Figure 1. The co-learning of segmentation and tracking on the appearance and shape level algorithm was employed to analyze the sequences of 2DE and LVO echocardiograms and identify the endocardial contour of LV by integrating the temporal and spatial information. The LV volumes and LVEF were quantified based on the single-plane Simpson method. 3D-UNet (G) as backbone for feature extraction, a segmentation head S that outputs the prediction of LV, and a tracking head T that outputs the optical flow fields.

Table 1. Patients' baseline characteristics.

Variable	Cohort 1 (N = 797)	Cohort 2 (N = 50)
Demographics		
Age (years)	53.0 (42, 62.0)	53.00 (41.3, 59.0)
Sex(male)	566 (71.0%)	29 (58.0%)
Diagnosis		
HCM	296 (37.1%)	36 (72.0%)
DCM	57 (7.2%)	4 (8.0%)
ICM	189 (23.7%)	2 (4.0%)
CSOL	54 (6.8%)	1 (2.0%)
Other	201 (25.2%)	7 (14.0%)

N, number of studies; HCM, hypertrophic cardiomyopathy; DCM, dilated cardiomyopathy; ICM, ischemic cardiomyopathy; CSOL, cardiac space-occupying lesions. Other, cohort 1 mainly included 24 cases of valvular disease, 33 cases of left ventricular hypertrophy, 56 cases of left ventricular noncompaction, and 57 cases without obvious abnormality in cardiac morphology and structure and 31 cases without clear ultrasound diagnosis. And cohort 2 included 4 cases of left ventricular hypertrophy, 2 cases of hypertension, and 1 case without obvious abnormality in cardiac morphology and structure.

The results showed that AI-driven LVO measurements had strong correlation and good agreement with sonographer measurements in the development and internal validation cohorts.

Development cohort. The median LV volumes and LVEF determined by AI and sonographers were presented in Appendix Table S4. The differences between AI and sonographer measurements for LV volumes (LVEDV: MAE, 5.6 ml; LVESV: MAE, 3.4 ml) were smaller than those obtained by 2DE (LVEDV: MAE, 11.3 ml; LVESV: MAE, 7.5 ml). The MAE of LVEF measurements obtained by sonographers and AI was 2.6% for LVO and 5.2% for 2DE ($P < 0.05$), as shown in Table 2.

In 2DE, the correlations between AI and manual measurements of LVEDV, LVESV, and LVEF were high (LVEDV: r = 0.91, LVESV: r = 0.892, LVEF: r = 0.823; $P < 0.05$ for all). In LVO, the correlations between AI and manual measurements for LV volumes and LVEF were more excellent (LVEDV: r = 0.967, LVESV: r = 0.971, LVEF: r = 0.961; $P < 0.05$ for all). The detailed results are shown in Figure 3 and Table 2.

LV volumes were slightly underestimated by automated measurements of both 2DE and LVO compared to manual reference values, with tolerable biases (2DE: −2.2 ml for LVEDV, −1.9 ml for LVESV; LVO: −1.8 ml for

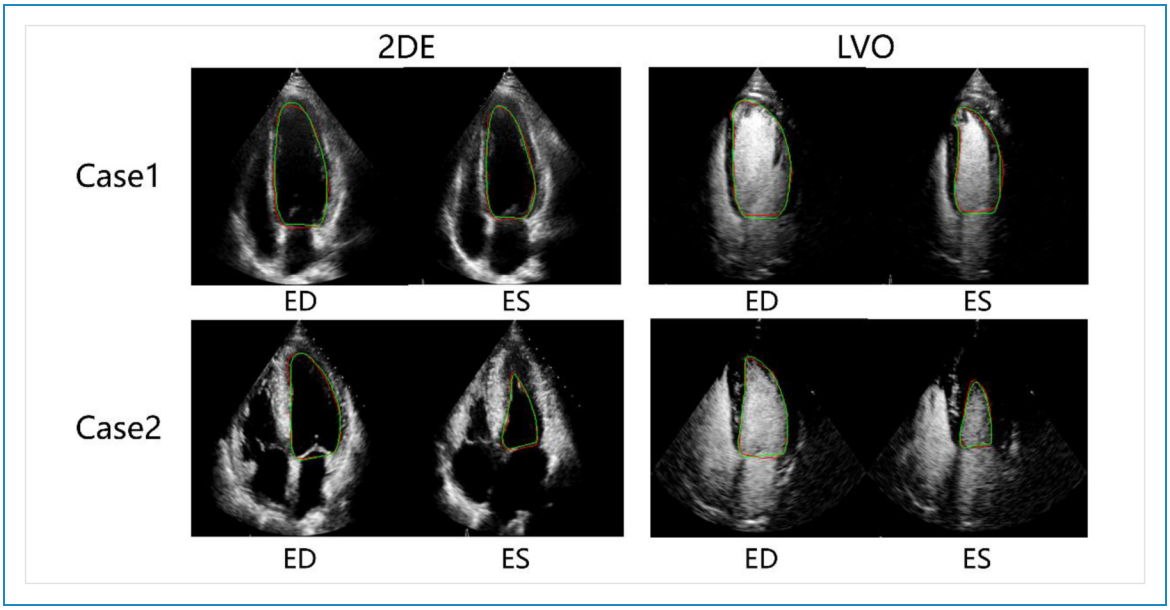

Figure 2. Two examples of left ventricle endocardium segmentation. The top row displays the segmentation for case 1, while the bottom row shows the segmentation for case 2. The red line represents the AI prediction results, and the green line represents the manual segmentation results.

LVEDV, −1.8 ml for LVESV). LVO outperformed 2DE, with a reduction in the 95% LOA for LVEDV, LVESV, and LVEF (from 72.4% to 42.0%, 53.0% to 26.2%, and 32.4% to 17.1%, respectively). The absolute value of relative bias for LV volumes was 0.3% to 2.9%. The SD of the Bias of LVO was smaller compared to 2DE, as shown in Figure 4 and Appendix Table S5. The ICCs between AI and manual measurements of cardiac function for both 2DE and LVO were above 0.86.

Internal validation cohort. In 2DE, the MAE of LVEDV, LVESV, and LVEF was 7.6 ml, 5.5 ml, and 6.8%, while the MAE for LVO were 5.5 ml, 2.7 ml, and 2.5%, respectively. The AI and manual measurements of LVEDV and LVESV in 2DE showed a high correlation (LVEDV: r = 0.86, LVESV: r = 0.85; $P < 0.05$ for all), while LVEF also showed a strong correlation (r = 0.69, $P < 0.05$) between the two methods. The results are presented in Appendix Table S6 and Figure 5. The correlation coefficients between AI and sonographers for LVO cardiac function parameters (LVEDV: r = 0.93, LVESV: r = 0.95, LVEF: r = 0.91; $P < 0.05$ for all) were higher than the values obtained by 2DE.

The difference between AI and manual measurements of LVEDV, LVESV, and LVEF for LVO was smaller than that for 2DE (LVEDV: Bias = -2.1 ml, SD = 9.9%; LVESV: Bias = -1.0 ml, SD = 6.1%; LVEF: Bias = -0.3%, SD = 4.4%), as shown in Table 3.

Furthermore, we used the biplane Simpson's method to measure the LVEDV, LVESV, and LVEF. The results showed that in LVO, the MAE between AI and manual measurements of LVEDV and LVEF were lower than those of 2DE ($P < 0.05$), and the correlation was better ($P < 0.05$) (Table S16). Meanwhile, when measuring cardiac function using the biplane Simpson's method, the agreement was better than the monoplane Simpson's method (Figure S2).

Cardiac function measurements in the subgroups

LVEDV, LVESV, and LVEF were measured by AI and sonographers in subgroups with different cardiac diseases, as shown in Appendix Table S7. In 2DE, the MAE of LVEDV, LVESV, and LVEF were 11.7 ml, 8.7 ml, and 5.2%, respectively. In LVO, the MAE of LVEDV, LVESV, and LVEF were 6.3 ml, 4.1 ml, and 2.5%, respectively. In each subgroup, the MAE of LVO cardiac function parameters between AI and sonographers' measurements were less than 2DE. The results are presented in Figure 6 and Appendix Table S8.

In 2DE, the correlation between AI and manual LVEF measurements was lowest in the HCM group at 0.64. The correlation of LVEF between AI and manual measurements was 0.87 in LVO. In subgroups of LVO, AI and manual measurements of LVEF had a higher correlation ($P < 0.05$) compared with 2DE, as shown in Appendix Table S9.

Furthermore, we divided 50 patients from cohort 2 into three groups based on image quality: good, fair, and poor, to evaluate the differences of LVEF between AI and

Table 2. Difference between AI and manual for LVEDV, LVESV, and LVEF in cohort 1.

	2DE			LVO		
Parameter	MAE	IQR	r	MAE	IQR	r
LVEDV	11.3	(5.3,19.8)	0.910	5.6*	(2.8,10.8)	0.967*
LVESV	7.5	(3.8,13.9)	0.892	3.4*	(1.4,6.6)	0.971*
LVEF	5.2	(2.5,8.7)	0.823	2.6*	(1.3,4.6)	0.961*

LVEDV, left ventricular end diastolic volume; LVESV, left ventricular end systolic volume; LVEF, left ventricular ejection fraction; 2DE, 2D echocardiography; LVO, left ventricular opacification; R, Spearman correlation coefficient; MAE, median absolute error; IQR, interquartile range. N = 508
*P < 0.05, 2DE versus LVO.

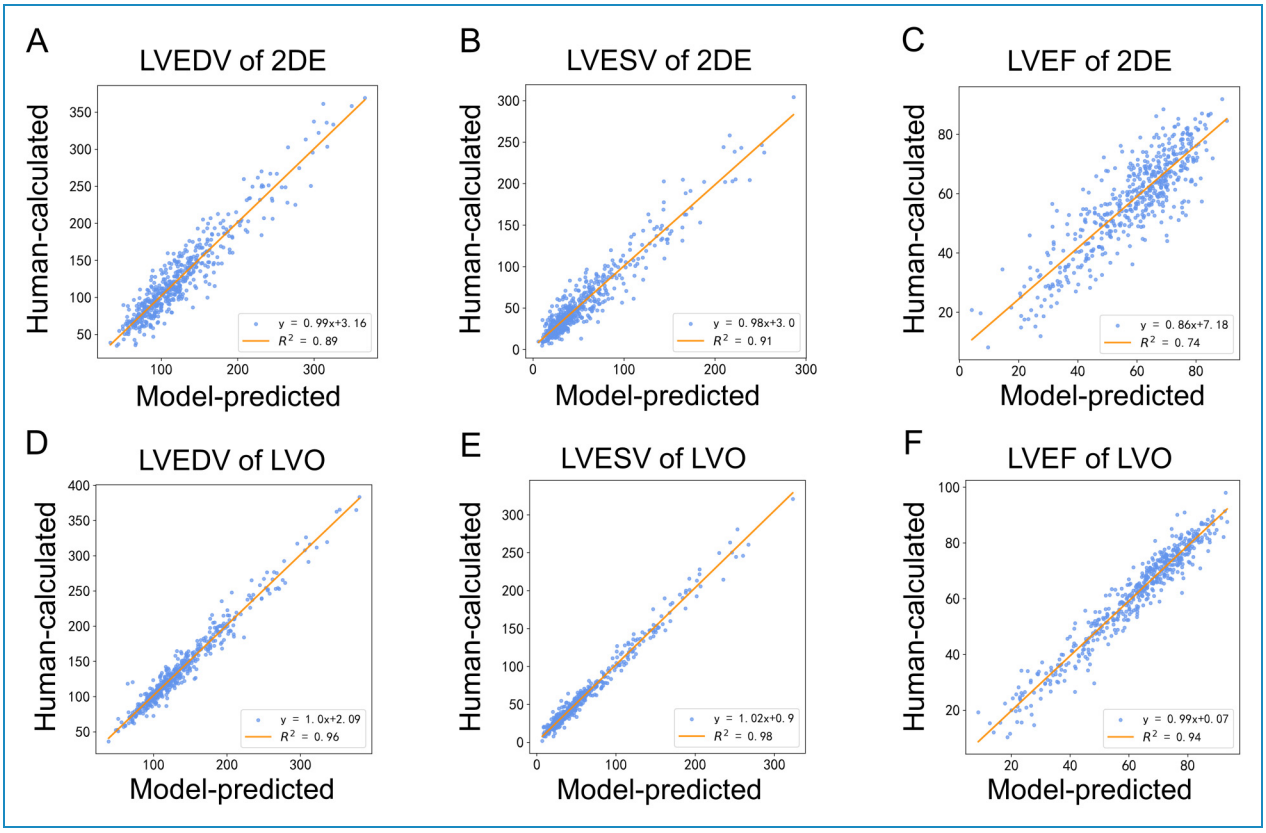

Figure 3. Linear regression results between AI prediction and manual measurements in cohort 1. A, LVEDV of 2DE; B, LVESV of 2DE; C, LVEF of 2DE; D, LVEDV of LVO; E, LVESV of LVO; F, LVEF of LVO.

manual measurements in different subgroups of image quality (Appendix Page 3-4).

Inter-observer and intra-observer variability analysis

Compared with 2DE (Intra-observer: MAE = 4.2%, r = 0.89; Inter-observer: MAE = 6.0%, r = 0.80), there was a significant improvement in the observer variability of LVEF in LVO (Intra-observer: MAE = 1.9%, r = 0.97; Inter-observer: MAE = 3.3%, r = 0.92, $P < 0.05$), as shown in Table 4.

In LVO, LVEF measurements showed lower intra-observer variability and higher reproducibility (ICC = 0.99). Compared to 2DE, the inter-observer Bias (−1.2%) and LOA (21.4%) for LVEF in LVO were lower. Whereas in 2DE, the values were −2.1% and

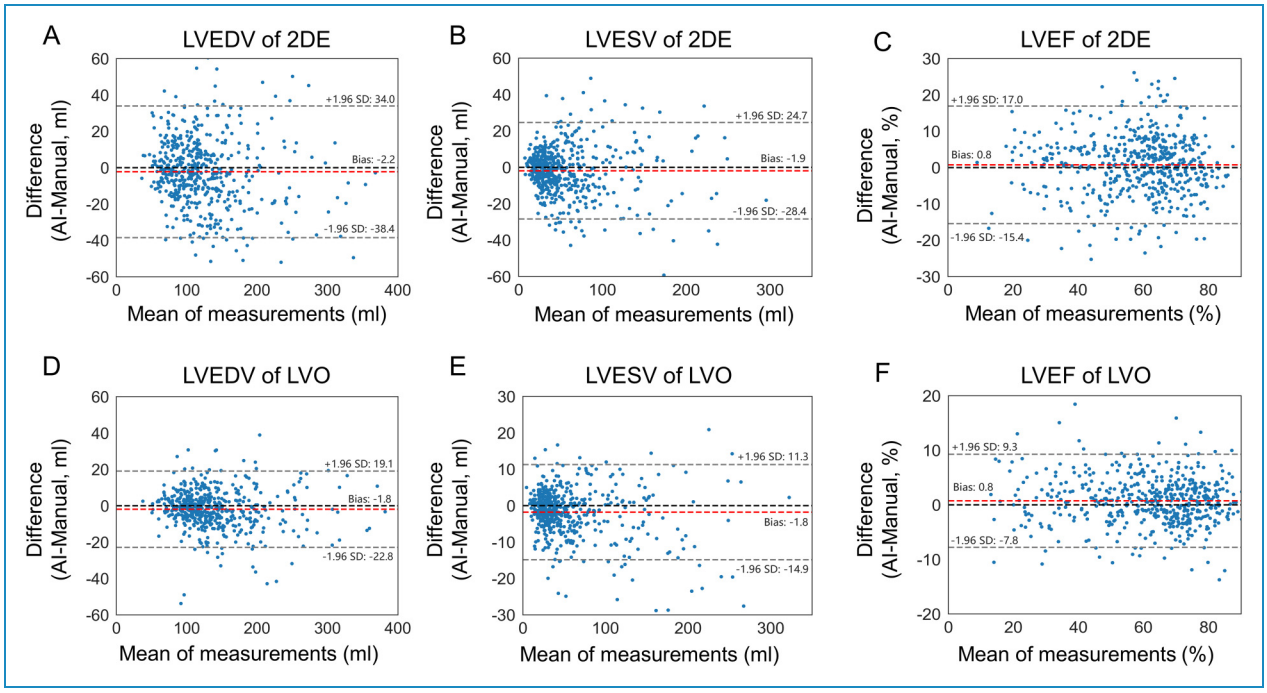

Figure 4. Bland-Altman plot. Bland-Altman plots comparing fully automated AI to manual measurements in cohort 1. A, LVEDV of 2DE; B, LVESV of 2DE, C, LVEF of 2DE; D, LVEDV of LVO; E, LVESV of LVO; F, LVEF of LVO. The middle line denotes the mean difference. Dashed lines denote ± 1.96 standard deviations.

38.3%, respectively. The ICCs of LVEDV and LVESV for both inter-observer and intra-observer were above 0.86. Detailed results are shown in Appendix Table S10.

In terms of AI measurements, the model parameters were optimized and fixed during the training process, ensuring identical outputs when the same image was input multiple times, thus eliminating observer differences. This guaranteed the reproducibility of AI in the assessment of cardiac function.

Generalization of the AI model

We used an external validation dataset consisting of 84 echocardiographic videos from 42 patients to assess the generalizability of the AI model.

The segmentation performance was similar to that of Cohort 1. The mean DSC, mean MAD and mean HD of LV segmentation for 2DE were 0.92, 1.76 mm, and 5.52 mm, respectively, and for LVO of 0.95, 1.32 mm, and 4.99 mm, as shown in Appendix Table S11.The correlation between AI and manual measurements of LV volumes and LVEF for both 2DE and LVO was over 0.8. The MAE of LVEF was 4.8% in 2DE and 2.7% in LVO, as shown in Appendix Table S12.

The agreement analysis indicated that LV volumes and LVEF measurements in LVO showed less variability (LVEDV: Bias = -1.7 ml, SD = 5.9 ml; LVESV: Bias = 0.7 ml, SD =5.1 ml; LVEF: Bias = -1.3%, SD = 4.2%) compared to 2DE (LVEDV: Bias = 5.0 ml, SD of Bias = 12.3 ml; LVESV: Bias = 2.7 ml, SD = 10.4 ml; LVEF: Bias = -1.0%, SD = 6.7%), as shown in Appendix Table S13.

Discussion

This study aimed to investigate the capability of AI to precisely integrate temporal and spatial information for assessing LVEF in LVO. The study found that (1) LV endocardial segmentation results were more accurate with LVO compared to 2DE. (2) Compared to 2DE, AI's predictions of LV volumes and LVEF in LVO had less variability and were within acceptable range. (3) Compared to 2DE, there was a significant improvement in observer variability of LVEF in LVO. AI measurements enhanced reproducibility by eliminating observer differences, ensuring consistent outputs for the same input image. (4) The results from internal and external validation datasets were consistent with those from the development cohort, suggesting that the AI model demonstrates a degree of generalization capability. Overall, AI can improve the reproducibility of quantitative cardiac function parameters in LVO.

Segmentation of the LV endocardium and assessment of cardiac function

Previous studies have utilized AI models to extract spatial features of ED and ES of 2DE, automatically recognize the LV endocardium, and assess LVEF. One study

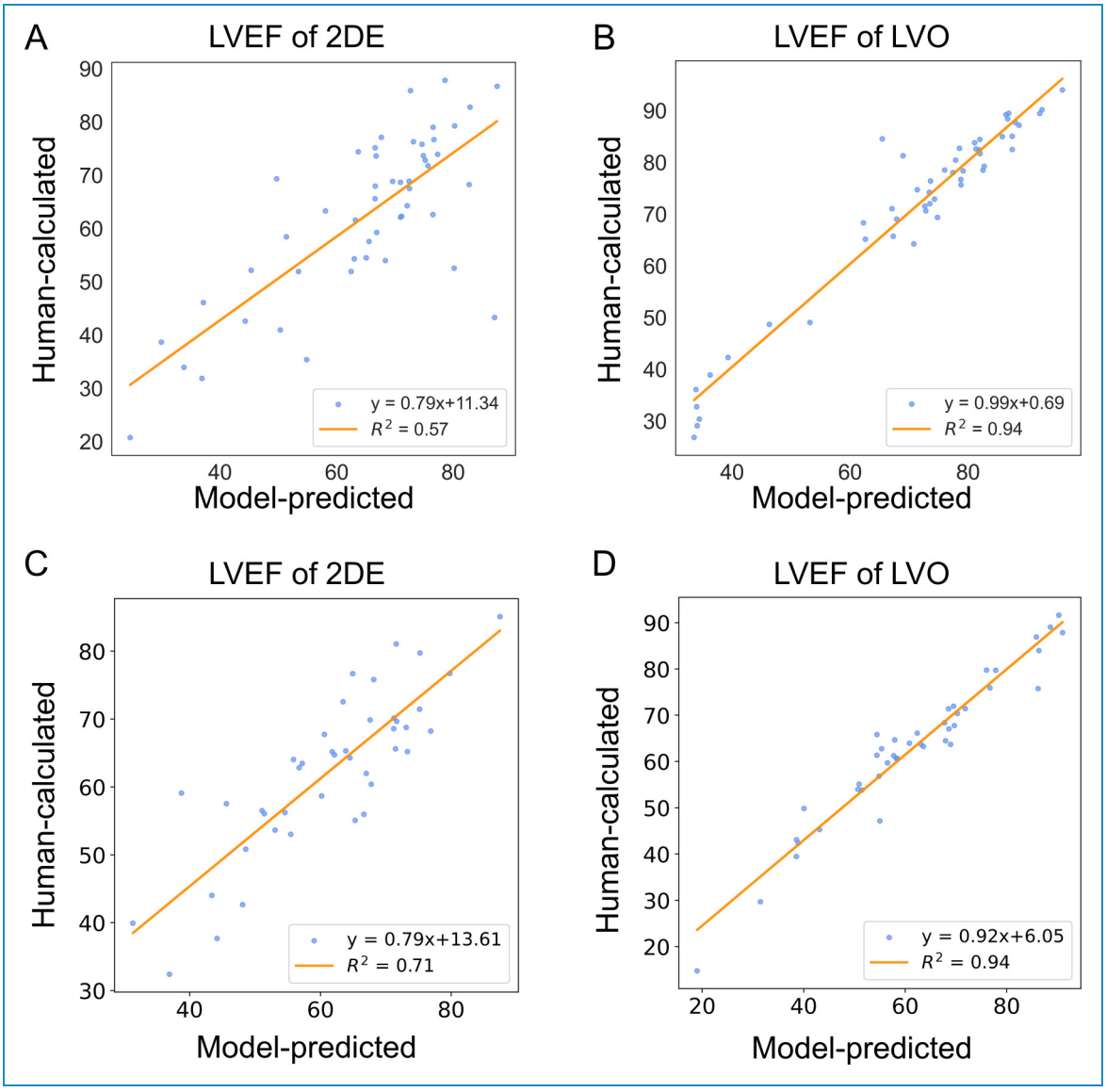

Figure 5. The linear regression results of association between CLAS and manual measurements of LVEF. A, LVEF of 2DE in cohort 2; B, LVEF of LVO in cohort 2; C, LVEF of 2DE in cohort 3; D, LVEF of LVO in cohort 3.

indicated that the mean correlation coefficient was 0.8 and the mean absolute error was 5.6%.[21] Due to the rich temporal information of echocardiography, Ouyang et al.[12] developed the EchoNet-Dynamic model based on 10,030 2DE sequences and sparse labels to realize dynamic tracking of LV endocardium. The model implicitly captured the temporal domain information of cardiac motion and predicted LVEF with a mean absolute error of 4.1%. Despite the high accuracy in segmenting LV and quantifying LVEF tasks, the segmentation still lacked temporal consistency. We employed the optical flow algorithm to capture the myocardial motion features, explicitly representing LV motion information. By combining spatial and temporal information, and using the sparse labels of ED and ES, we segmented the LV endocardium in the intermediate frames in a semi-supervised manner. This approach improved the temporal consistency of segmentation results, accurately segmenting key frames and enabling more precise quantification of cardiac function parameters.

Ultrasonic contrast agents can fill the intertrabecular space to improve the definition of the outermost layer of the LV endocardium.[16] Several studies have shown that LVO improves image quality, leading to more precise LV and LVEF measurements and a much better correlation and agreement of cardiac function measurements with cardiac magnetic resonance.[22,23] The measurements of LV volumes and LVEF were more consistent between and within observers when compared to 2DE. Manual endocardial tracing is time-

Table 3. Comparison of LVEDV, LVESV and LVEF measured by AI versus manual echocardiographic measurements in cohort 2.

	Parameter	Bias ± LOA	SD of Bias	Relative bias (%)	ICC
2DE	LVEDV	4.7 ± 55.0	14.0	7.5	0.89
	LVESV	−0.6 ± 32.8	8.4	4.8	0.94
	LVEF	2.6 ± 40.9	10.4	6.3	0.75
LVO	LVEDV	−2.1 ± 38.6	9.9	−1.0	0.95
	LVESV	−1.0 ± 24.0	6.1	1.0	0.98
	LVEF	−0.3 ± 17.2	4.4	0.3	0.97

LVEDV, left ventricular end diastolic volume; LVESV, left ventricular end systolic volume; LVEF, left ventricular ejection fraction; 2DE, 2D echocardiography; LVO, left ventricular opacification. Relative bias = (AI - Manual) / Manual. ICC, intraclass correlation coefficients.

consuming and labor-intensive regardless of 2DE or LVO. We applied an AI model to assess cardiac function by LVO, enabling automated analysis of LVO and reducing observer variability in traditional methods. Compared to 2DE, we observed a significant enhancement in the correlation and agreement of cardiac function parameters between AI and manual measurements by LVO.

Previous studies showed variability in LVEF assessment using single four-chamber echocardiography, ranging from 7.6–13.9%.[23,24] In cohort 2, LV volumes and LVEF measurements from LVO had a smaller range of variation between and within observers compared to 2DE.

In an external dataset, the MAE of LVEF between AI and manual measurements in LVO was significantly smaller than that of 2DE. The correlation between AI and manual measurements was very strong, and the results were similar to those of the development cohort. The AI model performed well in the external dataset, demonstrating a certain degree of generalization ability.

Subgroup analysis

In the development cohort, the HCM subgroup showed a significant reduction in MAE for LVO cardiac function measurements compared to 2DE. The explanations are: (1) LVO improves LV endocardial boundary delineation in cases where 2DE images of LV apex or lateral wall of HCM are unclear. (2) HCM patients have thicker papillary muscles and increased trabeculae, which can obscure LV endocardium. LVO facilitates endocardial boundary determination by incorporating trabeculae, LV, and strengthened papillary muscles. (3) LV chamber size during systole is significantly reduced in HCM, even close to occlusion, making it challenging to discern endocardial boundary. LVO improves the distinction between the cardiac cavity and the myocardium. In ICM and CSOL subgroups, cardiac structures can vary due to thrombus, infarction, or space occupancy, and AI with LVO reduces the variability of cardiac function measurements.

In the DCM subgroup, 2DE and LVO showed larger MAEs in EDV and ESV than in other subgroups. The main reason is that in patients with DCM, the cardiac cavity significantly enlarges, especially during ED, when the apical lateral wall and the mid anterolateral wall are often outside the sector area. This out-of-range situation results in the unclear display of these two ventricular wall segments, thereby increasing the difficulty of endocardial recognition. In this scenario, AI may tend to over-segment, resulting in contours larger than the actual size. In addition, since the AI model focuses on capturing continuous information between frames, when the ED segmentation contour is oversized, this bias extends to ES, resulting in oversized ES contours as well. Therefore, the MAEs of both EDV and ESV are relatively significant. However, since LVEF is calculated by comparing the volumes at ED and ES, the deviations are partially offset to each other, resulting in no significant increase in the MAE of LVEF. It is noteworthy that when comparing LVO with 2DE in patients with DCM, the MAE of LVO was significantly lower than that of 2DE. This is attributed to the fact that the contrast agent fills the cardiac cavity, resulting in a clearer delineation of the endocardial contour.

AI quantified LVO cardiac function parameters with fewer outliers

Although the results of cardiac function parameters were convincing, outliers still existed. The combination of AI and LVO results in a more accurate quantification of cardiac function parameters with fewer outliers. In cohort 2 and cohort 3, the outliers for the difference between AI and manual measurements of quantitative LVO cardiac function were less than 2DE. The detailed results are shown in Appendix Table S14 and Table S15.

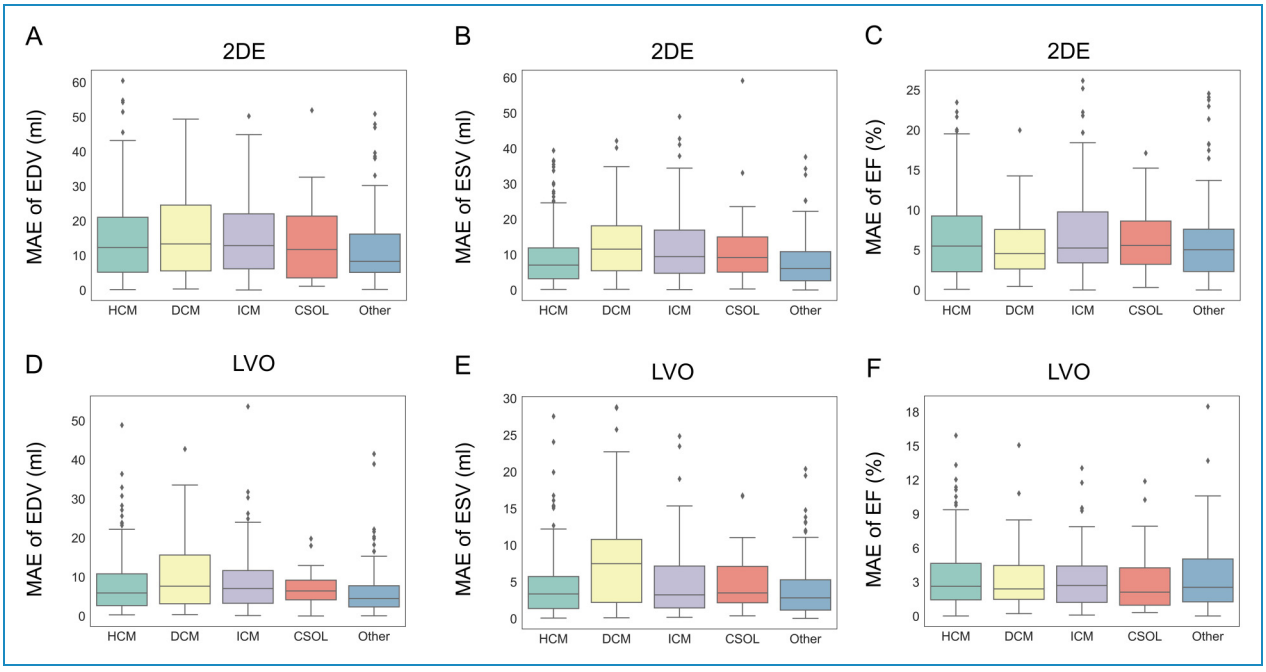

Figure 6. Subgroup analysis in cohort 1. A, B, C, the MAE of LVEDV, LVESV and LVEF in 2DE, respectively; D, E, F, the MAE of LVEDV, LVESV and LVEF in LVO. MAE, median absolute error. HCM, hypertrophic cardiomyopathy. DCM, dilated cardiomyopathy. ICM, ischemic cardiomyopathy. CSOL, cardiac space-occupying lesions. Other, mainly included valvular disease, left ventricular hypertrophy, and left ventricular noncompaction, without obvious abnormality in cardiac morphology and structure, and without clear ultrasound diagnosis.

Table 4. Inter-observer and intra-observer variability analysis.

	LVEDV			LVESV			LVEF		
	MAE	IQR	r	MAE	IQR	r	MAE	IQR	r
2DE									
Intra-observer	7.2*	(4.0,10.4)	0.933	3.1*	(1.4,6.1)	0.961	4.2*	(2.3,6.0)	0.886*
Inter-observer	10.3*	(5.0,16.0)	0.863	5.3*	(2.6,9.3)	0.869	6.0*	(3.8,10.6)	0.795*
LVO									
Intra-observer	4.2	(1.7,7.7)	0.968	1.9	(1.1,3.2)	0.98	1.9	(1.2,2.7)	0.974
Inter-observer	6.2	(3.6,9.1)	0.923	3.7	(2.3,6.2)	0.926	3.3	(1.9,5.2)	0.919

*$P < 0.05$, 2DE versus LVO.

In cohort 1's subgroup analysis, HCM showed the weakest correlation between AI and manual LVEF, with more AI-based cardiac function measurement outliers in 2DE. However, LVO significantly decreased measurement variability and improved endocardial boundary recognition. Injecting microbubble contrast agents into the blood circulation can help LVO generate strong backscattering signals, sharpen the visibility of the endocardial border and allow for more precise LV volumes and LVEF measurements.[17] AI combined with LVO considerably improved the correlation and agreement of LVEF, demonstrating its potential to accurately quantify cardiac function parameters for individuals with cardiac structure abnormalities while reducing measurement variability.

Limitation

Firstly, our center used SonoVue exclusively to assess LV function for LVO. Further research is needed to confirm the

utility of AI in LVO when using other contrast agents. Secondly, this study did not include cardiac magnetic resonance imaging data. Subsequent research will consider incorporating gold standard data to further validate the accuracy of the results. Thirdly, the development cohort comprised images obtained from multiple vendors' instruments with varying image quality and other characteristics, but the study demonstrated that AI had excellent generalizability and could effectively interpret images from various instruments. Finally, the validation dataset had a significantly smaller sample size than the development cohort. Consequently, we intend to acquire more extensive and comprehensive datasets in the near future to conduct a comprehensive and systematic quantitative analysis of cardiac function.

Conclusion

This study proposed an AI algorithm based on temporal consistency to improve the continuity between frames. The AI algorithm can automatically analyze contrast echocardiography, significantly reducing the need for manual intervention and improving the efficiency of LVEF measurements. Compared with 2DE, AI exhibits superior correlation and agreement in assessing cardiac function in LVO. The reliable measurements provided by AI can help reduce inter- and intra-observer variability and provide cardiologists with more accurate cardiac function assessments, thereby contributing to the optimization of treatment strategy.

Clinical perspectives

LVO improves the clarity of the endocardial boundary and contributes to the reproducibility of cardiac function parameters. Despite this, the LVEF measurements require the operator to delineate the LV endocardial boundary at the end-diastolic and end-systolic of LV. Previous studies have demonstrated that AI can efficiently segment the LV and assess LV volumes and functional parameters in non-contrast echocardiograms. Further research is needed to determine the accuracy of AI in assessing cardiac function in LVO.

This study initially demonstrated that it is possible to use AI to incorporate temporal information to assess LV volumes and LVEF. Meanwhile, the AI's prediction showed a strong correlation to the measurements taken by the sonographer, and its performance was consistent across the external dataset, demonstrating that the AI model is capable of generalizing. The AI model can automate the monitoring of the LV endocardial boundary and reduce the variability of observers, which has important clinical significance.

Future research will conduct more validation studies in diverse patient populations, such as those with noncompaction of ventricular myocardium, ventricular aneurysm, cardiac masses, and thrombosis, to evaluate the applicability and accuracy of AI-assisted LVO analysis in different scenarios. Meanwhile, the value of AI-assisted LVO on clinical decision-making will be thoroughly investigated. We will design experiments to compare the accuracy of AI-assisted LVO with doctors' independent assessment of cardiac function, and explore how AI can assist cardiologists in improving the precision of diagnosis and the effectiveness of treatment strategy.

Abbreviations

2DE	Two Dimensional Echocardiography
DSC	Dice Similarity Coefficients
HD	Hausdorff Distance
ICC	Intraclass Correlation Coefficients
LVEDV	Left Ventricular End-Diastolic Volume
LVEF	Left Ventricular Ejection Fraction
LVESV	Left Ventricular End-Systolic Volume
LVO	Left Ventricular Opacification
MAD	Mean Surface Distance
MAE	Median Absolute Error

Acknowledgements: Our research team would like to express our sincere gratitude to the sonographers at the Department of Ultrasound Medicine, Union Hospital Tongji Medical College Huazhong University of Science and Technology, for providing us with the essential data resources and support that enabled us to successfully complete this study. We would also like to acknowledge the sonographers at the Department of Ultrasound Medicine, First Hospital of China Medical University, which provided us with external dataset. The support from these entities and individuals played a crucial role in the completion of our research, and we extend our heartfelt appreciation to all who have contributed to this study.

Author contributions: Y.Z. (Ye Zhu), Z.S.Z. (Zisang Zhang), J.Q.M. (Junqiang Ma), D.N. (Dong Ni), M.X.X. (Mingxing Xie), W.F.X. (Wufeng Xue) and L.Z. (Li Zhang) conceived of the study; Y.W.Z. (Yiwei Zhang), Z.S.Z. (Zisang Zhang), and M.W.L. (Manwei Liu) labelled data; Y.Z. (Ye Zhu), Z.S.Z. (Zisang Zhang) and J.Q.M. (Junqiang Ma) created and ran the data processing pipeline; Y.Z. (Ye Zhu) and J.Q.M. (Junqiang Ma) designed and evaluated the deep learning models; Y.Z. (Ye Zhu), Z.S.Z. (Zisang Zhang), J.Q.M. (Junqiang Ma) and Y.W.Z. (Yiwei Zhang) wrote the manuscript; C.W. (Chun Wu), Z.M.Z. (Ziming Zhang), S.S.Z. (Shuangshaung Zhu), J.C. (Jun Cheng), X.Y. (Xin Yang), J.W. (Jing Wang), M.X.X. (MingXing Xie) and L.Z. (Li Zhang) provided language help and proofed the article. Y.H.W (Yonghuai Wang) and C.Y.M (Chunyan Ma) provided external validation datasets. C.W. (Chun Wu), C.Y.X (Chunyan Xu), A.J.W (Anjun Wu) and C.C.S (Chen chen Sun) provided internal validation datasets.

Conflict of interest: The authors of this manuscript declare that there is no conflict of interest.

Data availability: The data underlying this article cannot be shared publicly due to privacy concerns of individuals participating in the study. The anonymized imaging data will be shared on reasonable request to the corresponding author pending approval by the involved health care trusts.

Ethical approval: The ethics committee of the Union Hospital Tongji Medical College Huazhong University of Science and Technology approved this study.

Funding: This work has been supported by the National Natural Science Foundation of China [grant numbers 82151316, 82171964, 82230066, U21A20387, 82171961, and 81771851].

Guarantor: Li Zhang

Patient consent: The risk to the subjects in the study was no greater than the minimum risk, and all echocardiographic data from the study was anonymous to ensure that the relevant study could be conducted without violating patient privacy. The study successfully applied for blanket consent.

Supplemental material: Supplemental material for this article is available online.


ORCID iD: Ye Zhu https://orcid.org/0009-0005-2860-9402

References

1. Ziaeian B and Fonarow GC. Epidemiology and aetiology of heart failure. *Nat Rev Cardiol* 2016; 13: 368–378.
2. Lupón J and Bayés-Genís A. Left ventricular ejection fraction in heart failure: a clinician's perspective about a dynamic and imperfect parameter, though still convenient and a cornerstone for patient classification and management. *Eur J Heart Fail* 2018; 20: 433–435.
3. Mele D, Nardozza M and Ferrari R. Left ventricular ejection fraction and heart failure: an indissoluble marriage? *Eur J Heart Fail* 2018; 20: 427–430.
4. Alhakak AS, Teerlink JR, Lindenfeld J, et al. The significance of left ventricular ejection time in heart failure with reduced ejection fraction. *Eur J Heart Fail* 2021; 23: 541–551.
5. Dagres N and Hindricks G. Risk stratification after myocardial infarction: is left ventricular ejection fraction enough to prevent sudden cardiac death? *Eur Heart J* 2013; 34: 1964–1971.
6. Donal E, Magne J and Cosyns B. Left ventricular ejection fraction thresholds reappraisal: also for bicuspid valve disease? *J Am Coll Cardiol* 2022; 80: 1085–1087.
7. Hecht S, Butcher SC, Pio SM, et al. Impact of left ventricular ejection fraction on clinical outcomes in bicuspid aortic valve disease. *J Am Coll Cardiol* 2022; 80: 1071–1084.
8. Thavendiranathan P, Grant AD, Negishi T, et al. Reproducibility of echocardiographic techniques for sequential assessment of left ventricular ejection fraction and volumes: application to patients undergoing cancer chemotherapy. *J Am Coll Cardiol* 2013; 61: 77–84.
9. Singh JP, Solomon SD, Fradley MG, et al. Association of cardiac resynchronization therapy with change in left ventricular ejection fraction in patients with chemotherapy-induced cardiomyopathy. *JAMA* 2019; 322: 1799–1805.
10. Asch FM, Poilvert N, Abraham T, et al. Automated echocardiographic quantification of left ventricular ejection fraction without volume measurements using a machine learning algorithm mimicking a human expert. *Circ Cardiovasc Imaging* 2019; 12: e009303.
11. Lang RM, Badano LP, Mor-Avi V, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American society of echocardiography and the European association of, cardiovascular imaging. *Eur Heart J Cardiovasc Imaging* 2016; 17: 12.
12. Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature* 2020; 580: 252–256.
13. Zhang J, Gajjala S, Agrawal P, et al. Fully automated echocardiogram interpretation in clinical practice: feasibility and diagnostic accuracy. *Circulation* 2018; 138: 1623–1635.
14. Asch FM, Mor-Avi V, Rubenson D, et al. Deep learning-based automated echocardiographic quantification of left ventricular ejection fraction: a point-of-care solution. *Circ Cardiovasc Imaging* 2021; 14: e012293.
15. Mor-Avi V, Jenkins C, Kühl HP, et al. Real-time 3-dimensional echocardiographic quantification of left ventricular volumes: multicenter study for validation with magnetic resonance imaging and investigation of sources of error. *JACC Cardiovasc Imaging* 2008; 1: 413–423.
16. Senior R, Becher H, Monaghan M, et al. Clinical practice of contrast echocardiography: recommendation by the European association of cardiovascular imaging (EACVI) 2017. *Eur Heart J Cardiovasc Imaging* 2017; 18: 1205–1205af.
17. Porter TR, Mulvagh SL, Abdelmoneim SS, et al. Clinical applications of ultrasonic enhancing agents in echocardiography: 2018 American society of echocardiography guidelines update. *J Am Soc Echocardiogr* 2018; 31: 241–274.
18. Zhu Y, Ma J, Zhang Z, et al. Automatic view classification of contrast and non-contrast echocardiography. *Front Cardiovasc Med* 2022; 9: 989091.
19. Lang RM, Badano LP, Mor-Avi V, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American society of echocardiography and the European association of cardiovascular imaging. *J Am Soc Echocardiogr* 2015; 28: 1–39 14.
20. Wei H, Cao H, Cao Y, et al. Temporal-consistent Segmentation of Echocardiography with Co-learning from Appearance and Shape. In: Medical Image Computing and Computer Assisted Intervention, 2020.
21. Leclerc S, Smistad E, Pedrosa J, et al. Deep learning for segmentation using an open large-scale dataset in 2D echocardiography. *IEEE Trans Med Imaging* 2019; 38: 2198–2210.
22. Kurt M, Shaikh KA, Peterson L, et al. Impact of contrast echocardiography on evaluation of ventricular function and clinical management in a large prospective cohort. *J Am Coll Cardiol* 2009; 53: 802–810.
23. Malm S, Frigstad S, Sagberg E, et al. Accurate and reproducible measurement of left ventricular volume and ejection fraction by contrast echocardiography: a comparison with magnetic resonance imaging. *J Am Coll Cardiol* 2004; 44: 1030–1035.
24. Farsalinos KE, Daraban AM, Ünlü S, et al. Head-to-Head comparison of global longitudinal strain measurements among nine different vendors: the EACVI/ASE inter-vendor comparison study. *J Am Soc Echocardiogr* 2015; 28: 1171–1181, e1172.