Deep learning to assess right ventricular ejection fraction from two-dimensional echocardiograms in precapillary pulmonary hypertension

Running title: RVEF estimation from 2D echocardiogram using deep learning

Michito Murayama, RMS, JRDCS, PhD[a,b], Hiroyuki Sugimori, PhD[c,d,e], Takaaki Yoshimura, PhD[d,e,f,g]*, Sanae Kaga, RMS, JRDCS, PhD[a,b], Hideki Shima, MD[i], Satonori Tsuneta, MD, PhD[j], Aoi Mukai[h], Yui Nagai, BS[h], Shinobu Yokoyama, RMS[b], Hisao Nishino, RMS[b], Junichi Nakamura, MD[i], Takahiro Sato, MD, PhD[i,k], Ichizo Tsujino, MD, PhD[i,k]

[a]Department of Medical Laboratory Science, Faculty of Health Sciences, Hokkaido University, N12, W5, Kita-ku, Sapporo 060-0812, Japan

[b]Diagnostic Center for Sonography, Hokkaido University Hospital, N14, W5, Kita-ku, Sapporo 060-8648, Japan

[c]Department of Biomedical Science and Engineering, Faculty of Health Sciences, Hokkaido University, Sapporo 060-0812, Japan.

[d]Clinical AI Human Resources Development Program, Faculty of Medicine, Hokkaido University, Sapporo 060-8648, Japan.

[e]Global Center for Biomedical Science and Engineering, Faculty of Medicine, Hokkaido University, Sapporo 060-8648, Japan

[f]Department of Health Sciences and Technology, Faculty of Health Sciences, Hokkaido University, Sapporo 060-0812, Japan.

[g]Department of Medical Physics, Hokkaido University Hospital, Sapporo 060-8648, Japan.

[h]Graduate School of Health Sciences, Hokkaido University, N12, W5, Kita-ku, Sapporo 060-0812, Japan

[i]Department of Respiratory Medicine, Faculty of Medicine, Hokkaido University, N14, W5, Kita-ku, Sapporo 060-8648, Japan

[j]Department of Radiology, Graduate School of Dental Medicine, Hokkaido University, N13 W7, Kita-Ku, Sapporo, Hokkaido 060-8586, Japan

[k]Division of Respiratory and Cardiovascular Innovative Research, Faculty of Medicine, Hokkaido University,

N14, W5, Kita-ku, Sapporo 060-8648, Japan

***Corresponding author:** Takaaki Yoshimura, PhD

Department of Health Sciences and Technology, Faculty of Health Sciences, N12, W5, Kita-ku, Sapporo 060-0812, Japan; Tel: +81-11-706-2851; Fax: None; Email: takaaki.ysm@med.hokudai.ac.jp

Data availability statement: The datasets generated and analyzed in the current study are available from the corresponding author upon reasonable request.

[Funding Statement]

This study was partly supported by the Charitable Trust Laboratory Medicine Research Foundation of Japan and Grants-in-Aid for the Regional R&D Proposal-Based Program from the Northern Advancement Center for Science & Technology of Hokkaido, Japan, to M. Murayama. The funders played no role in the study design, data collection and analysis, decision to publish, or manuscript preparation.

Conflict of interest disclosure: The authors declare that this research was conducted in the absence of commercial or financial relationships that could be construed as potential conflicts of interest.

Acknowledgements: We would like to thank Editage (www.editage.com) for the English language editing.

Deep learning to assess right ventricular ejection fraction from two-dimensional echocardiograms in precapillary pulmonary hypertension

Abstract

Purpose: Precapillary pulmonary hypertension (PH) is characterized by a sustained increase in right ventricular (RV) afterload, impairing systolic function. Two-dimensional (2D) echocardiography is the most performed cardiac imaging tool to assess RV systolic function; however, an accurate evaluation requires expertise. We aimed to develop a fully automated deep learning (DL)-based tool to estimate the RV ejection fraction (RVEF) from 2D echocardiographic videos of apical four-chamber views in patients with precapillary PH.

Methods: We identified 85 patients with suspected precapillary PH who underwent cardiac magnetic resonance imaging (MRI) and echocardiography. The data was divided into training (80%) and testing (20%) datasets, and a regression model was constructed using 3D-ResNet50. Accuracy was assessed using five-fold cross validation.

Results: The DL model predicted the cardiac MRI-derived RVEF with a mean absolute error of 7.67%. The DL model identified severe RV systolic dysfunction (defined as cardiac MRI-derived RVEF < 37%) with an area under the curve (AUC) of 0.84, which was comparable to the AUC of RV fractional area change (FAC) and tricuspid annular plane systolic excursion (TAPSE) measured by experienced sonographers (0.87 and 0.72, respectively). To detect mild RV systolic dysfunction (defined as RVEF ≤ 45%), the AUC from the DL-predicted RVEF also demonstrated a high discriminatory power of 0.87, comparable to that of FAC (0.90), and significantly higher than that of TAPSE (0.67).

Conclusion: The fully automated DL-based tool using 2D echocardiography could accurately estimate RVEF and exhibited a diagnostic performance for RV systolic dysfunction comparable to that of human readers.

Keywords: pulmonary hypertension, right ventricular ejection fraction, echocardiography, artificial intelligence, deep learning

Non-standard abbreviations and acronyms

AUC	Area under the curve
CNN	Convolutional neural network
DL	Deep learning
EDV	End-diastolic volume
ESV	End-systolic volume
FAC	Fractional area change
MRI	Magnetic resonance imaging
PA	Pulmonary arterial
PH	Pulmonary hypertension
ROC	Receiver operating characteristic
RV	Right ventricular
RVEF	RV ejection fraction
SD	Standard deviation
TAPSE	Tricuspid annular plane systolic excursion

INTRODUCTION

Precapillary pulmonary hypertension (PH) is hemodynamically defined as elevated mean pulmonary arterial (PA) pressure and pulmonary vascular resistance.[1] In PH, a sustained increase in the right ventricular (RV) afterload impairs RV systolic function. The prognostic significance of RV dysfunction is well recognized in patients with precapillary PH.[1] Therefore, screening for RV dysfunction is an integral part of precapillary PH care.

RV ejection fraction (RVEF) has prognostic significance in precapillary PH,[2,3] and cardiac magnetic resonance imaging (MRI) is the gold standard for accurate and precise measurement of RVEF; however, the routine use of cardiac MRI is limited by its cost, availability, and patient contraindications.[4] Thus, in daily clinical practice, two-dimensional (2D) transthoracic echocardiography is the most frequently performed cardiac imaging tool to assess RV function.[5,6] However, accurate evaluation of RV systolic function and diagnosis of reduced RV contraction by 2D echocardiography requires expertise and experience owing to the complex anatomy of the right ventricle.

Deep learning (DL) enables automated systems to meet, or even surpass, the performance of clinical experts across a range of cardiac image analysis tasks in cardiovascular medicine, from diagnosis to prognostication.[7] Numerous DL algorithms have been developed to automate the assessment of left ventricular ejection fraction using conventional echocardiographic acquisitions.[8-10] In contrast to the left ventricle, to the best of our knowledge, only one study has proposed a fully automated DL-based tool that estimates RVEF from 2D echocardiographic videos,[11] wherein it was highlighted that the DL model could predict RVEF from 2D echocardiography in healthy volunteers and patients with cardiovascular diseases; however, three-dimensional (3D) echocardiography was used rather than cardiac MRI.[11]

We hypothesized that developing a fully automated DL model that can predict cardiac MRI-driven RVEF in precapillary PH would bridge the gap between 2D echocardiography and cardiac MRI in RVEF assessment, resulting in improved precapillary PH care in daily practice. We, therefore, aimed to implement a DL-based tool to estimate the cardiac MRI-derived RVEF from 2D echocardiographic videos of patients with precapillary PH, in whom RVEF is the key determinant of symptoms and prognosis. Additionally, we benchmarked the performance of our system against the readings of experienced cardiac sonographers.

PATIENTS AND METHODS

Participants and overview of the study protocol

This study focused on building a fully automated DL model that analyzes 2D apical four-chamber view echocardiographic videos to predict cardiac MRI-derived RVEF as a continuous variable. We retrospectively identified 93 consecutive hospitalized patients with suspected or confirmed precapillary PH who underwent right heart catheterization, cardiac MRI, and echocardiography within 1 week under stable clinical conditions at Hokkaido University Hospital between May 2020 and June 2022. A flowchart illustrating the patients, video selection process, and dataset creation is shown in **Figure 1**. Among 93 enrolled consecutive patients, we excluded those with atrial fibrillation (n = 2), frequent premature ventricular or atrial contractions (n = 2), and tachycardia with a heart rate ≥ 100 bpm (n = 1). From 186 standard and RV-focused apical four-chamber view echocardiographic videos of 93 examinations of 73 patients fulfilling the inclusion criteria, we excluded eight examinations because of poor echocardiographic images (significant translational motion, gain changes, depth changes, or sector position changes). The final dataset contained 170 echocardiographic videos from 85 examinations of 69 patients.

The dataset was used to train and test the DL model. We evaluated the ability of the DL model to differentiate patients with RV systolic dysfunction. We also compared the performance of the DL model with that of experienced cardiac sonographers. This study was approved by the Hokkaido University Hospital Research Ethics Committee (023-0050) and was conducted in accordance with the 1964 Declaration of Helsinki and its subsequent amendments. As this was a retrospective study, the need for informed consent was waived by the ethics committee. Patients could opt out of participation, and they were informed of this right on our institutional website.

Right heart catheterization

Right heart catheterization was performed using a Swan-Ganz catheter. From the pressure records, we measured the mean right atrial, PA, and PA wedge pressure.[12] Stroke volume and cardiac output were measured using the thermodilution method, and pulmonary vascular resistance was calculated as follows: (mean PA pressure – mean PA wedge pressure)/cardiac output.

Cardiac MRI

Cardiac MRI studies were performed using a 1.5-Tesla MRI scanner (Achieva or Achieva dStream, Philips Healthcare, Best, Netherlands) or a 3.0-Tesla MRI scanner (Achieva Tx, Philips Medical Systems, Best, Netherlands) with electrocardiogram gating. Image acquisition and analysis were performed using a previously described protocol with high intra- and interobserver reproducibility, especially in patients with RV dilation.[13] Briefly, cine MRI with axial planes covering the entire heart was performed using a steady-state free-precession pulse sequence according to the Society for Cardiovascular Magnetic Resonance Task Force paper.[14] The details of these parameters are listed in **Table S1**. Cine MRI images were analyzed using commercially available analysis software (Extended MR WorkSpace or IntelliSpace Portal; Philips Healthcare, Best, The Netherlands). Using cine MRI, a radiologist (ST) with over 8 years of experience in cardiac image analysis manually traced the endocardial contours of the RV wall from the most apical slice to the uppermost slice, and the RV end-diastolic volume (EDV) and end-systolic volume (ESV) were calculated. The RVEF was calculated as follows: (EDV – ESV)/EDV × 100.

Based on recent guidelines from the European Society of Cardiology/European Respiratory Society,[1] MRI-derived RVEF was divided into three groups: >54% (estimated 1-year mortality <5%), 37–54% (estimated 1-year morality was 5–20%), and <37% (estimated 1-year mortality >20%). Owing to its prognostic significance, severe RV systolic dysfunction was defined as an MRI-derived RVEF <37% (primary analysis).[1] Because an RVEF of 45% is also meaningful and is the standard cutoff to diagnose RV systolic dysfunction,[15] a sensitivity analysis was performed using the definition of mild RV systolic dysfunction as an MRI-derived RVEF ≤45%.[16]

Echocardiography

All patients underwent transthoracic echocardiographic examination by registered medical sonographers in cardiology with more than 5 years of experience (MM, SK, SY, and HN) using commercially available ultrasound systems (Artida/Aplio i900 with a 3.0 MHz/i6SX1 probe [Canon Medical Systems, Otawara, Japan]; Vivid E9 ultrasound system with an M5S probe [GE Healthcare, Little Chalfont, UK]; iE33 ultrasound system with an S4 probe [Philips Medical Systems, Eindhoven, The Netherlands]; ACUSON SC2000 prime with 4V1c probe [SIEMENS Healthcare, Erlangen, Germany]; or Prosound F75 ultrasound system with a UST-52127 probe [Hitachi Ltd., Tokyo, Japan]) (**Table S2**). Cardiac chamber morphology and function were assessed according to

published guidelines.[5,17] Left ventricular ejection fraction was measured using the biplane disk summation method. Basal RV end-diastolic dimension was measured using an apical four-chamber view. The inferior vena cava dimension was measured from a subcostal longitudinal view. Tricuspid annular plane systolic excursion (TAPSE) and RV fractional area changes (FAC) were assessed as RV systolic function parameters.[18] All echocardiographic analyses were blinded to the cardiac MRI results.

DL model training and evaluation

In this study, we constructed a regression model for the estimation of RVEF from 2D echocardiographic images, necessitating the utilization of a regression convolutional neural network (CNN) based on the architecture of 3D-ResNet50. This publicly available 3D-CNN model, originally designed for classification tasks, was adapted for regression purposes. Specifically, the ultimate layer of ResNet50 capable of accommodating 3D inputs was substituted with a regression layer.[19] Both 2D echocardiographic videos and MRI-derived RVEF values were trained as pairs in the training phase. Moreover, DL-estimated RVEF in the test phase was output using this regression model from 2D echocardiographic videos that were not used for training. Finally, we evaluate the accuracy of the DL-estimated RVEF compared with MRI-derived RVEF as the gold standard (**Figure 2**). The details of data augmentation or model parameters in this regression model are shown below.

We obtained 18 continuous images from both standard and RV-focused apical four-chamber views and saved them in the Neuroimaging Informatics Technology Initiative file format. Data augmentation was applied to modify the duplicated image sequences individually to avoid the negative consequences of such duplications. Augmentation techniques improve network generalization by forcing DL models to recognize more general patterns and remain invariant to applied transformations.[11,20] The echocardiographic images were rotated by 1° from -5° to 5° and were also scaled from 0.7X to 1.3X in increments of 0.05X, yielding a total of 143 expansions.

We assessed the 3D-CNN using five-fold cross validation (training: test = 8:2) (**Figure 2**). During training, the CNN condition was a stochastic gradient descent with a momentum optimizer, and the initial learning rate, maximum number of epochs, and mini-batch size were set to 0.0001, 10, and 512, respectively. The training and analyses were performed using software developed in-house in MATLAB (2022a; The Mathworks, Natick, MA). Detailed information on the computer specifications used in this study has been described previously.[19] Five

subsets were created and assessed by sorting the MRI-derived RVEF values in descending order to avoid RVEF bias within the subset. Regression models were developed for the RVEF and the test. To evaluate the accuracy of the DL-estimated RVEF for each case, we compared the DL-RVEF with the cardiac MRI-derived RVEF values as the gold standard (**Figure 2**).

Statistical analysis

Statistical analyses were performed using IBM SPSS ver. 27 for Windows (IBM Co., Armonk, NY) and JMP Pro 17.0.0 (SAS Institute Inc., Cary, NC). All numerical data are presented as means ± standard deviation (SD) and categorical variables were expressed as a number (percentage). The Pearson's correlation coefficient was used. A Bland–Altman analysis was used to determine the bias and 95% limits of agreement between the DL-predicted RVEF and MRI-derived RVEF. A mean absolute error calculation was performed to assess the accuracy of the DL-estimated RVEF.[19] The diagnostic performance of the DL model was evaluated using receiver operating characteristic (ROC) analysis and pairwise comparisons of the area under the ROC curve (AUC). ROC curves were computed for two different thresholds of abnormal RV systolic function: MRI-derived RVEF <37% (primary analysis),[1] and MRI-derived RVEF ≤45% (sensitivity analysis).[11,16] Two-sided significance levels of 0.05 were used for all analyses.

RESULTS

Clinical characteristics

Table 1 summarizes the patient characteristics. Of 85 participants, 73 (86%) had PH. The remaining 12 participants (14%) did not meet the PH criteria, with mean PA pressure ≤ 20 mmHg. Most patients had pulmonary arterial hypertension, followed by chronic thromboembolic PH. Approximately half of the patients had substantial PH symptoms (World Health Organization functional classes III or IV). The RV cavity was enlarged (RV basal dimension > 42 mm) in 37 patients (44%). The MRI-derived RVEF values ranged from 10% to 69%, with a median of 38%. Severe RV systolic dysfunction (defined as MRI-derived RVEF < 37%) was identified in 41 participants (48%), and mild RV systolic dysfunction (defined as MRI-derived RVEF ≤ 45%) was present in 61 (72%) participants.

Estimation of gold standard RVEF by the DL model

Table 2 shows the correlation coefficients and mean absolute errors for MRI-derived RVEF prediction within each dataset. The DL model yielded a mean absolute error of 7.67% (r = 0.65). **Figure 3** shows the correlation between the MRI-derived RVEF and the DL-predicted RVEF across all participants. DL-predicted RVEF significantly correlated with MRI-derived RVEF (r = 0.63, p < 0.001) (**Figure 3A**). The Bland–Altman analysis showed no fixed bias (mean: 0.12%, 95% confidence interval: -1.95 to 2.19%; limits of agreement: -18.69 to 18.93%), but a proportional error was observed between MRI-derived RVEF and DL-predicted RVEF (p < 0.001) (**Figure 3B**).

Detection of RV systolic dysfunction by the DL model

The correspondence between DL-predicted RVEF and MRI-derived RVEF classification based on the guidelines was substantial, with a concordance rate of 73% and a κ value of 0.51, respectively (**Figure 4**). To detect severe RV systolic dysfunction, the AUC produced by DL-predicted RVEF showed high discriminatory power (0.84; 95% CI: 0.76–0.93), similar to AUC by FAC and TAPSE measured by expert sonographers (AUC = 0.84 for DL-predicted RVEF vs 0.87 for FAC, p = 0.468; AUC = 0.84 for DL-predicted RVEF vs 0.72 for TAPSE, p = 0.074) (**Figure 5A**). To detect mild RV systolic dysfunction, the AUC from DL-predicted RVEF also demonstrated high discriminatory power (0.87; 95% CI: 0.79–0.94), comparable to FAC measured by expert sonographers (AUC = 0.87 for DL-predicted RVEF vs 0.90 for FAC, p = 0.472), and significantly higher than AUC by TAPSE measured by expert sonographers (AUC = 0.87 for DL-predicted RVEF vs 0.67 for TAPSE; p = 0.007) (**Figure 5B**).

DISCUSSION

We propose a fully automated DL-based tool to analyze 2D RV standard and RV-focused apical four-chamber views and accurately estimate RVEF as a continuous variable in patients with precapillary PH. In addition, our system enables us to identify RV systolic dysfunction with a performance similar to that of expert cardiac sonographers. This is the first study to develop a fully automated DL model to assess RVEF from 2D echocardiograms in precapillary PH by comparing it to cardiac MRI-derived RVEF as the gold standard. Precapillary PH is characterized by alterations in the pulmonary vasculature, leading to increased pulmonary

vascular resistance and, ultimately, RV dysfunction. Chronically increased afterload on the RV leads to RV hypertrophy, dilation, and systolic dysfunction.[1,21] Therefore, RV systolic function is associated with the outcomes in patients with precapillary PH.[2,22] Although RVEF is an ejection phase parameter of the pre-and afterload-dependent RV pump function, it is the most widely used index of RV systolic function with promising prognostic roles.[2,3] Cardiac MRI is the reference imaging technique used to assess RVEF; however, factors such as availability, cost, portability, time consumption, and contraindications hinder its routine use in all patients with precapillary PH. Thus, 2D echocardiography is the first-line tool used to assess RV systolic function, and a variety of indices have been established, such as FAC, TAPSE, tissue Doppler-derived systolic tricuspid annular motion velocity, 2D speckle tracking myocardial strain, and 3D echocardiography-derived RVEF.[5,17] However, some parameters are not widely used in real-world situations.[23] A multinational online survey involving 1,150 participants from 109 countries revealed that visual assessment of RV systolic function (eyeballing) (72%) and TAPSE (69%) were the only parameters widely used worldwide in daily clinical routines.[23] Although visual assessment using the RV-focused apical four-chamber view was the simplest and most widely used method for the echocardiographic classification of RV systolic function, large-scale international data demonstrated its poor specificity for the detection of RV systolic dysfunction based on cardiac MRI-derived RVEF: sensitivity of 96% and specificity of 43% in beginners and sensitivity of 97% and specificity of 56% in experts.[24] Our DL model for RVEF estimation would be helpful in bridging the gap between the fact that eyeballing is the most used method, with poor specificity. TAPSE is simple and less dependent on optimal image quality and prolonged image analysis,[5] resulting in its widespread use around the world.[22] In addition, its feasibility has been validated by comparison with cardiac MRI-derived RVEF in patients with precapillary PH.[25,26] However, in precapillary PH, TAPSE has an important limitation. Ventricular apical longitudinal rotation, which occurs in PH, enhances TAPSE, resulting in the overestimation of RV systolic function.[27,28] The present study showed that our DL model for RVEF estimation was more accurate than TAPSE in detecting RV systolic dysfunction.

In contrast to the left ventricle, there are few studies on DL-based models that estimate the cardiac MRI-derived RVEF from 2D echocardiographic videos.[11,29] Tokodi et al. first proposed a fully automated DL-based prediction model of RVEF with the use of 2D echocardiography with large data sets.[11] They suggested that their method achieved good accuracy; the mean absolute error to predict RVEF was 4.57% in the internal data set and 5.54% in the external validation set.[11] Their study represents the first crucial step, indicating that DL improves 2D

echocardiographic RV assessment.[11] However, our study differs from previous studies in several ways. First, the gold standard of the RVEF in their study was 3D echocardiography and not cardiac MRI.[11] Second, their DL model was developed mainly from healthy volunteers, athletes, and patients with left-sided cardiovascular diseases rather than patients with precapillary PH, in whom RV dysfunction is the main pathology.[11] Kampaktsis et al. conducted a retrospective study of 50 patients with various cardiovascular diseases without severe RV dilation or systolic dysfunction (training:validation:testing = 35:5:10) in which the patients underwent cardiac MRI and echocardiography within a 1-month interval.[29] They developed an attention-based DL method to assess RV volumes and RVEF using 2D echocardiography with cardiac MRI as a reference,[29] and their method achieved good feasibility and promising accuracy in estimating RVEF (absolute percentage error = 7.24%). However, their method required RV tracing of the endocardial-myocardial RV interface in end-diastole and end-systole for eight standardized RV views: the parasternal long axis, RV inflow, parasternal short axis at three levels, standard four-chamber, and subcostal four-chamber views. Their method enabled accurate quantitative RV evaluation without a geometric assumption; however, it was not fully automated and was time-consuming owing to the large number of input variables. Previous studies by Tokodi et al.[11] and Kampaktsis et al.[29] are similar to the present study in that they adapted DL to estimate RVEF by 2D echocardiography; however, our study is the first to propose a DL-based tool capable of accurately estimating RVEF from 2D echocardiographic videos against cardiac MRI-derived RVEF as the gold standard in patients with precapillary PH.

The present study showed that the discriminatory power of our fully automated DL-based tool was comparable to that of experienced sonographers reading FAC and greater than that of experienced sonographers reading TAPSE in distinguishing patients with RV systolic dysfunction from those with normal RV systolic function. Unlike TAPSE, FAC reflects both the longitudinal and transverse components of RV contractions.[5] Consistent with our results, its clinical value in precapillary PH has been confirmed[25,30,31]; however, FAC measurements showed significant test-retest variability,[32] due to the difficulty in correctly tracing the RV endocardium in both systole and diastole. Because our DL model does not require tracing, it has the potential to shorten examination times and improve diagnostic accuracy, especially for beginner sonographers. In the present study, standard echocardiography machines from five ultrasound manufacturers were used, and their frequency of use was even, except for that of SIEMENS Healthcare, making the DL model highly versatile. In addition, our tool achieved remarkable performance with a mean absolute error of 7.67%, which is similar to the results of previous studies

on RVEF prediction using 2D echocardiograms.[11,29] These data suggest that our DL-RVEF prediction model is sufficiently accurate for routine clinical use. Although cardiac MRI is the gold standard for quantitative RV assessment, 2D echocardiography is the first-line imaging tool because it is widely available, portable, and cost-effective.[33,34] Therefore, our DL model bridges the gap between 2D echocardiography and cardiac MRI in RVEF assessment. Our system has the potential to reduce the time required for RV systolic function assessment, improve the diagnostic accuracy of RV systolic dysfunction, accelerate echocardiographic reporting, and reduce the workload of echocardiographers.

This study had some limitations. First, the retrospective design and inclusion of only hospitalized patients with suspected or confirmed precapillary PH with cardiac MRI datasets might represent a significant source of selection bias. Thus, whether our DL model could be applied to other cardiovascular diseases remains unknown. Second, cardiac MRI and echocardiography were not performed simultaneously, with a median time difference of 2 days (range: 0–7 days). Although we excluded patients with unstable hemodynamics and/or loading conditions between examinations, the possibility of hemodynamic and RV functional changes could not be completely excluded. Third, neither RV strain/3D echocardiographic analyses nor a head-to-head comparison of RV strain/3D echocardiographic parameters and the DL model to detect RV systolic dysfunction could be performed in this study. Although such new technologies have not been widely used in daily practice,[23] perhaps because of the time-consuming measurement procedures, recent innovations in technology allow for easy use in routine practice. Therefore, we need to compare the measurement time and diagnostic accuracy between the new technologies and our DL model. Fourth, we excluded examinations with poor echocardiographic images (significant translational motion, gain changes, depth changes, or sector position changes) when testing the DL algorithm, which may limit its real-world applicability. Fifth, the DL-predicted RVEF correlated well with the cardiac MRI-derived RVEF; however, there was a proportional error between them. A likely reason for this proportional error may be the small sample size, especially for patients with high RVEF: the present study included only four cases with cardiac MRI-derived RVEF > 54%. To build a more accurate DL model, larger cohorts with a wider range of RVEF, including high RVEF, should be used.

CONCLUSION

Our fully automated DL-based tool using 2D echocardiography accurately estimated cardiac MRI-derived RVEF

and showed a diagnostic performance similar to that of experienced cardiac sonographers. Further prospective studies are needed to explore whether our DL-based system significantly benefits the clinical workflow.

REFERENCES

1. Humbert M, Kovacs G, Hoeper MM, et al. 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. Eur Heart J. 2022;43:3618-3731.
2. Baggen VJ, Leiner T, Post MC, et al. Cardiac magnetic resonance findings predicting mortality in patients with pulmonary arterial hypertension: a systematic review and meta-analysis. Eur Radiol. 2016;26:3771-3780.
3. Lewis RA, Johns CS, Cogliano M, et al. Identification of cardiac magnetic resonance imaging thresholds for risk stratification in pulmonary arterial hypertension. Am J Resp Crit Care Med. 2020;201:458-468.
4. Hameed A, Condliffe R, Swift AJ, et al. Assessment of right ventricular function-a state of the art. Curr Heart Fail Rep. 2023;20:194-207.
5. Rudski LG, Lai WW, Afilalo J, et al. Guidelines for the echocardiographic assessment of the right heart in adults: a report from the American Society of Echocardiography endorsed by the European Association of Echocardiography, a registered branch of the European Society of Cardiology, and the Canadian Society of Echocardiography. J Am Soc Echocardiogr. 2010;23:685-713; quiz 86-8.
6. 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. 2015;16:233-270.
7. Kusunose K. Revolution of echocardiographic reporting: the new era of artificial intelligence and natural language processing. J Echocardiogr. 2023;21:99-104.
8. Kusunose K, Haga A, Yamaguchi N, et al. Deep learning for assessment of left ventricular ejection fraction from echocardiographic images. J Am Soc Echocardiogr. 2020;33:632-635.e1.
9. Liu X, Fan Y, Li S, et al. Deep learning-based automated left ventricular ejection fraction assessment using 2-D echocardiography. Am J Physiol Heart Circ Physiol. 2021;321:H390-H399.
10. He B, Kwan AC, Cho JH, et al. Blinded, randomized trial of sonographer versus AI cardiac function assessment. Nature. 2023;616:520-524.
11. Tokodi M, Magyar B, Soos A, et al. Deep learning-based prediction of right ventricular ejection

fraction using 2D echocardiograms. JACC Cardiovasc Imaging. 2023;16:1005-1018.

12. Shima H, Nakaya T, Tsujino I, et al. Accuracy of Swan–Ganz catheterization-based assessment of right ventricular function: Validation study using high-fidelity micromanometry-derived values as reference. Pulm Circ. 2022;12:e12078.

13. Clarke CJ, Gurka MJ, Norton PT, et al. Assessment of the accuracy and reproducibility of RV volume measurements by CMR in congenital heart disease. JACC Cardiovasc Imaging. 2012;5:28-37.

14. Schulz-Menger J, Bluemke DA, Bremerich J, et al. Standardized image interpretation and post-processing in cardiovascular magnetic resonance - 2020 update: Society for Cardiovascular Magnetic Resonance (SCMR): Board of Trustees Task Force on Standardized Post-Processing. J Cardiovasc Magn Reson. 2020;22:19.

15. Muraru D, Badano LP, Nagata Y, et al. Development and prognostic validation of partition values to grade right ventricular dysfunction severity using 3D echocardiography. Eur Heart J Cardiovasc Imaging. 2020;21:10-21.

16. Petersen SE, Khanji MY, Plein S, et al. European Association of Cardiovascular Imaging expert consensus paper: a comprehensive review of cardiovascular magnetic resonance normal values of cardiac chamber size and aortic root in adults and recommendations for grading severity. Eur Heart J Cardiovasc Imaging. 2019;20:1321-1331.

17. 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.e14.

18. Murayama M, Kaga S, Okada K, et al. Clinical utility of superior vena cava flow velocity waveform measured from the subcostal window for estimating right atrial pressure. J Am Soc Echocardiogr. 2022;35:727-737.

19. Inomata S, Yoshimura T, Tang M, et al. Estimation of left and right ventricular ejection fractions from cine-MRI using 3D-CNN. Sensors (Basel). 2023;23:6580.

20. Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning. J Big Data. 2019;6:1-48.

21. Konstam MA, Kiernan MS, Bernstein D, et al. Evaluation and management of right-sided heart

failure: a scientific statement from the American Heart Association. Circulation. 2018;137:e578-e622.

22. Ghio S, Klersy C, Magrini G, et al. Prognostic relevance of the echocardiographic assessment of right ventricular function in patients with idiopathic pulmonary arterial hypertension. Int J Cardiol. 2010;140:272-278.

23. Schneider M, Aschauer S, Mascherbauer J, et al. Echocardiographic assessment of right ventricular function: current clinical practice. Int J Cardiovasc Imaging. 2019;35:49-56.

24. Schneider M, Ran H, Aschauer S, et al. Visual assessment of right ventricular function by echocardiography: how good are we? Int J Cardiovasc Imaging. 2019;35:2001-2008.

25. Sato T, Tsujino I, Ohira H, et al. Validation study on the accuracy of echocardiographic measurements of right ventricular systolic function in pulmonary hypertension. J Am Soc Echocardiogr. 2012;25:280-286.

26. Sato T, Tsujino I, Oyama-Manabe N, et al. Simple prediction of right ventricular ejection fraction using tricuspid annular plane systolic excursion in pulmonary hypertension. Int J Cardiovasc Imaging. 2013;29:1799-1805.

27. Motoji Y, Tanaka H, Fukuda Y, et al. Association of apical longitudinal rotation with right ventricular performance in patients with pulmonary hypertension: insights into overestimation of tricuspid annular plane systolic excursion. Echocardiography. 2016;33:207-215.

28. Collier P, Xu B, Kusunose K, et al. Impact of abnormal longitudinal rotation on the assessment of right ventricular systolic function in patients with severe pulmonary hypertension. J Thorac Dis. 2018;10:4694-4704.

29. Kampaktsis PN, Bohoran TA, Lebehn M, et al. An attention-based deep learning method for right ventricular quantification using 2D echocardiography: feasibility and accuracy. Echocardiography. 2024;41:e15719.

30. Austin C, Alassas K, Burger C, et al. Echocardiographic assessment of estimated right atrial pressure and size predicts mortality in pulmonary arterial hypertension. Chest. 2015;147:198-208.

31. Hoette S, Creuzé N, Günther S, et al. RV fractional area change and TAPSE as predictors of severe right ventricular dysfunction in pulmonary hypertension: a CMR study. Lung. 2018;196:157-164.

32. Knight DS, Schwaiger JP, Krupickova S, et al. Accuracy and test-retest reproducibility of two-dimensional knowledge-based volumetric reconstruction of the right ventricle in pulmonary hypertension. J Am Soc Echocardiogr. 2015;28:989-998.

429 33. Reeves RA, Halpern EJ, Rao VM. Cardiac imaging trends from 2010 to 2019 in the Medicare
430 population. Radiol Cardiothorac Imaging. 2021;3:e210156.
431 34. Goldfarb JW, Weber J. Trends in cardiovascular MRI and CT in the U.S. Medicare Population from
432 2012 to 2017. Radiol Cardiothorac Imaging. 2021;3:e200112.

FIGURE LEGENDS

Figure 1. Flowchart illustrating patients and video selection process and data set creation

PH, pulmonary hypertension; MRI, magnetic resonance imaging; RV, right ventricle.

Figure 2. Illustration of the CNN for regression, and the data set creation and analysis method for the RVEF prediction

2D, two-dimensional; CNN, convolutional neural network; MRI, magnetic resonance imaging; RVEF, right ventricular ejection fraction.

Figure 3. Relationships between gold standard and DL-predicted RVEF values

(A) Correlation analyses. (**B)** Bland–Altman analyses. DL, deep learning; RVEF, right ventricular ejection fraction; MRI, magnetic resonance imaging; SD, standard deviation.

Figure 4. Agreement plots depicting the concordance of DL-predicted RVEF and MRI-derived RVEF classification based on the guidelines

RVEF, right ventricular ejection fraction; DL, deep learning; MRI, magnetic resonance imaging.

Figure 5. Receiver operating characteristic curves illustrating the performance of DL-predicted RVEF values and conventional echocardiographic parameters measured by sonographers in identifying RV systolic dysfunction

(A) Area under the receiver operating characteristic (AUC) curves to detect MRI-derived RVEF <37% (primary analysis). (**B)** AUC curves to detect MRI-derived RVEF ≤45% (sensitivity analysis). Sonographer FAC, right ventricular fractional area change measured by registered medical sonographers; Sonographer TAPSE, tricuspid annular plane systolic excursion measured by registered medical sonographers; DL, deep learning; RVEF, right ventricular ejection fraction.

Graphical abstract

Using a dataset of patients with precapillary pulmonary hypertension with two-dimensional (2D)

echocardiographic videos of apical four-chamber views and cardiac magnetic resonance imaging (MRI)-derived right ventricular (RV) ejection fraction (RVEF), we developed a deep learning-based model to predict RVEF from 2D echocardiographic videos alone. We also compared the diagnostic performance of the deep learning-based model for RV dysfunction with that of experienced cardiac sonographers.

465 **Table 1.** Clinical, cardiac MRI, and echocardiographic characteristics of the study cohort

Variable	Mean ± SD	Range
Baseline characteristics		
Age (years)	59 ± 17	23–84
Gender, female, n (%)	59 (69)	
Heart rate (beats/min)	72 ± 11	50–104
Body surface area (m^2)	1.55 ± 0.18	1.18–1.93
Systolic blood pressure (mmHg)	111 ± 16	82–152
World Health Organization functional class III/IV	37/3	
Pulmonary hypertension subtype		
1	44	
3	11	
4	18	
Cardiac Catheterization		
Mean right atrial pressure (mmHg)	4 ± 2	-2–11
Mean pulmonary atrial pressure (mmHg)	32 ± 12	12–68
Mean pulmonary atrial wedge pressure (mmHg)	7 ± 2	2–12
Cardiac index (L/min/m^2)	2.7 ± 0.6	1.6–4.7
Pulmonary vascular resistance (Wood units)	6.5 ± 0.7	0.7–23.1
Cardiac magnetic resonance imaging		

Right ventricular end-diastolic volume (mL)	164 ± 73	79–570
Right ventricular end-systolic volume (mL)	108 ± 67	26–448
Right ventricular ejection fraction (%)	37 ± 12	10–69
Echocardiography		
Left ventricular end-diastolic dimension (mm)	43 ± 6	24–55
Left ventricular ejection fraction (%)	66 ± 8	35–84
Right ventricular basal dimension (mm)	42 ± 8	27–74
Tricuspid annular plane systolic excursion (mm)	19 ± 5	9–31
Right ventricular fractional area change (%)	29 ± 10	7–57
Inferior vena cava dimension (mm)	13 ± 4	6–29

466 Data are mean ± SD or number of patients.

467 **Table 2.** Results of gold standard right ventricular ejection fraction prediction on five-fold cross validation

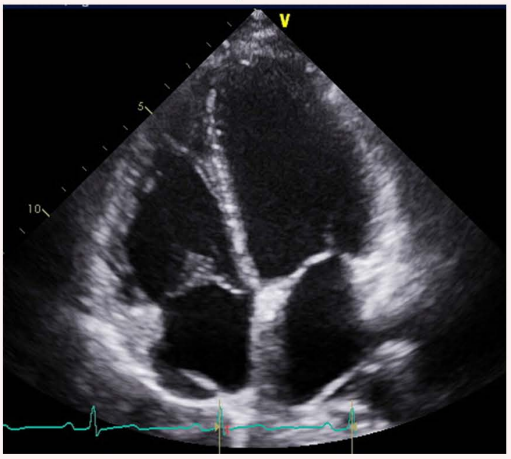
	Correlation Coefficient R (95% confidence interval)	p	Mean absolute error
Dataset A	0.773 (0.464 to 0.914)	<0.001	7.435%
Dataset B	0.648 (0.243 to 0.860)	0.005	7.248%
Dataset C	0.610 (0.183 to 0.843)	0.009	7.709%
Dataset D	0.690 (0.313 to 0.879)	0.002	7.367%
Dataset E	0.539 (0.079 to 0.810)	0.026	8.600%
Mean	0.650		7.670%

468

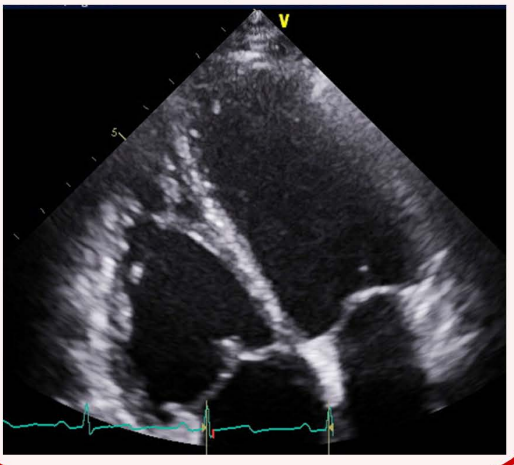
469

2D echo videos

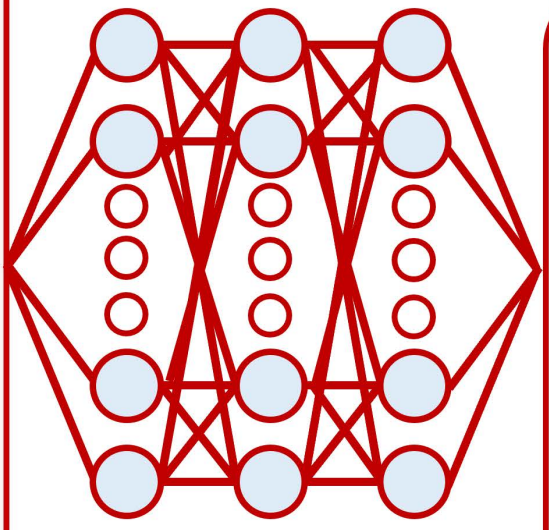
standard RV apical 4-chamber



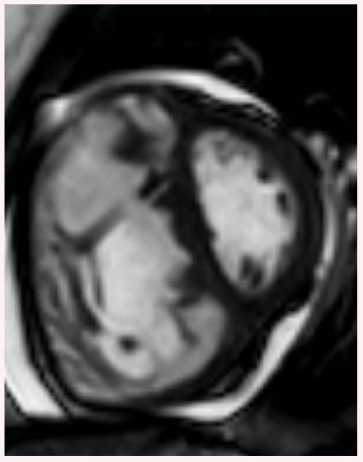
RV-focused apical 4-chamber



Regression CNN based on 3D-ResNet50

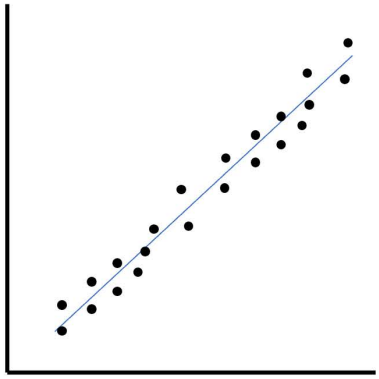


MRI

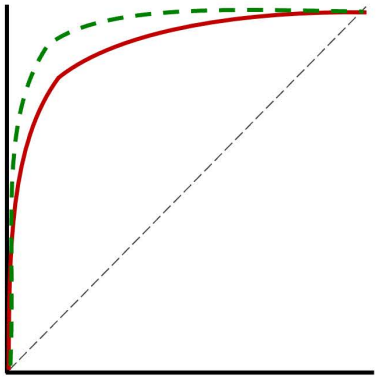


gold standard RVEF

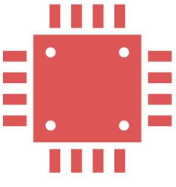
Prediction of RVEF

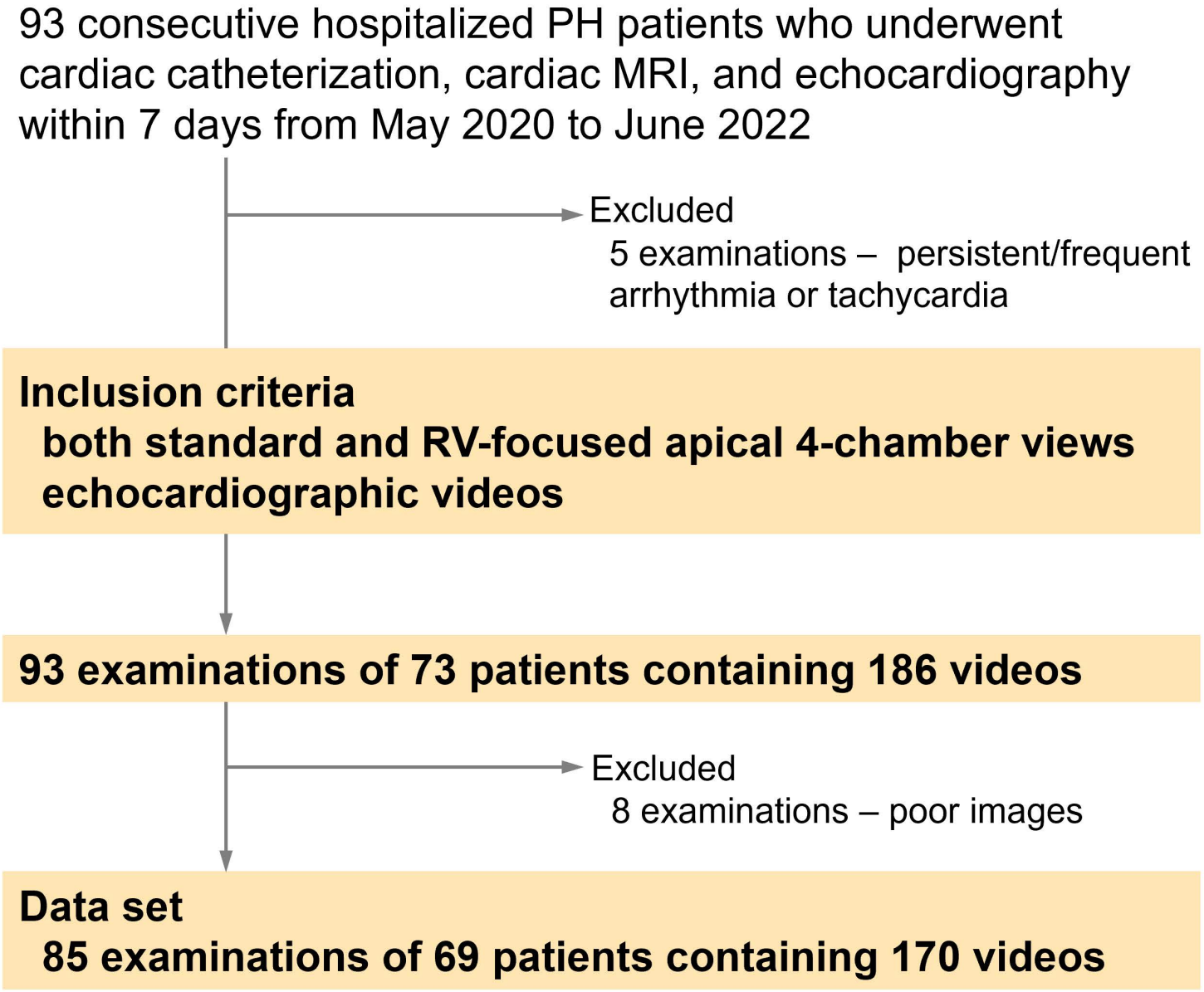


Detection of RV dysfunction

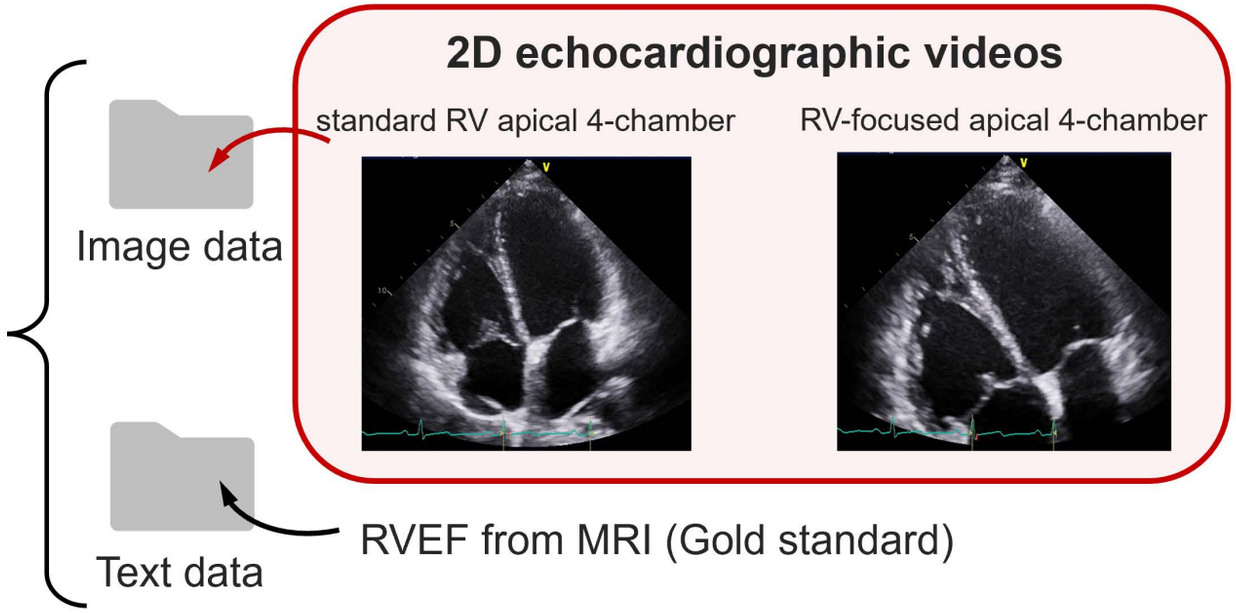


versus




93 consecutive hospitalized PH patients who underwent cardiac catheterization, cardiac MRI, and echocardiography within 7 days from May 2020 to June 2022
Excluded
5 examinations – persistent/frequent arrhythmia or tachycardia
Inclusion criteria
both standard and RV-focused apical 4-chamber views echocardiographic videos
93 examinations of 73 patients containing 186 videos
Excluded
8 examinations – poor images
Data set
85 examinations of 69 patients containing 170 videos

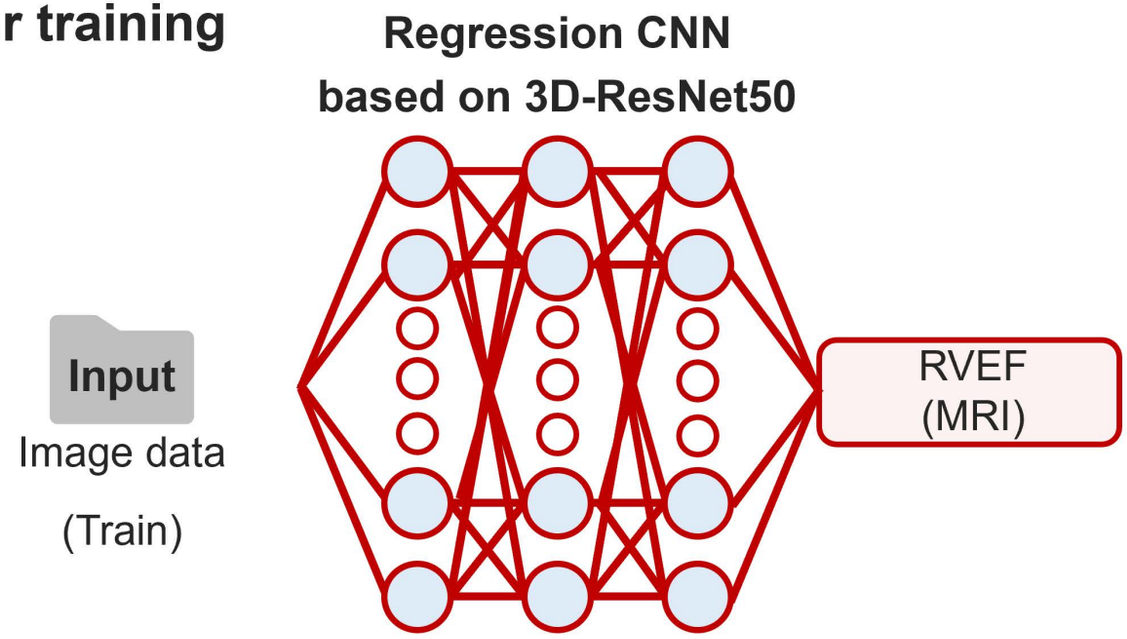
Dataset for each patient



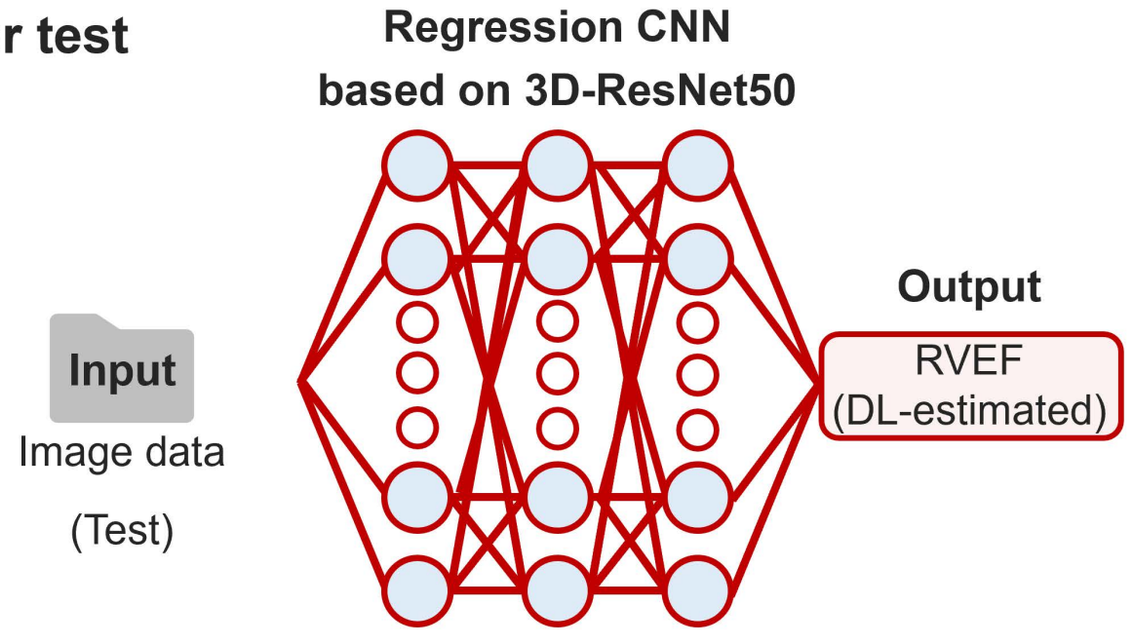
5-fold cross-validation

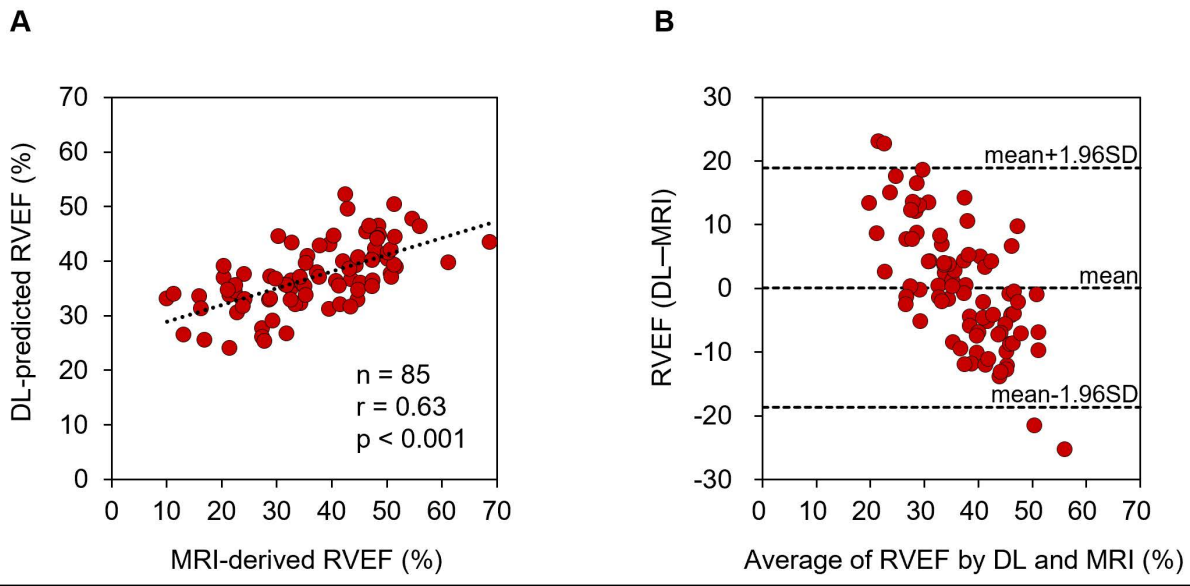
Dataset	Examination ID				
	1 – 17	18 – 34	35 – 51	52 – 68	69 – 85
Fold A	Test	Train	Train	Train	Train
Fold B	Train	Test	Train	Train	Train
Fold C	Train	Train	Test	Train	Train
Fold D	Train	Train	Train	Test	Train
Fold E	Train	Train	Train	Train	Test

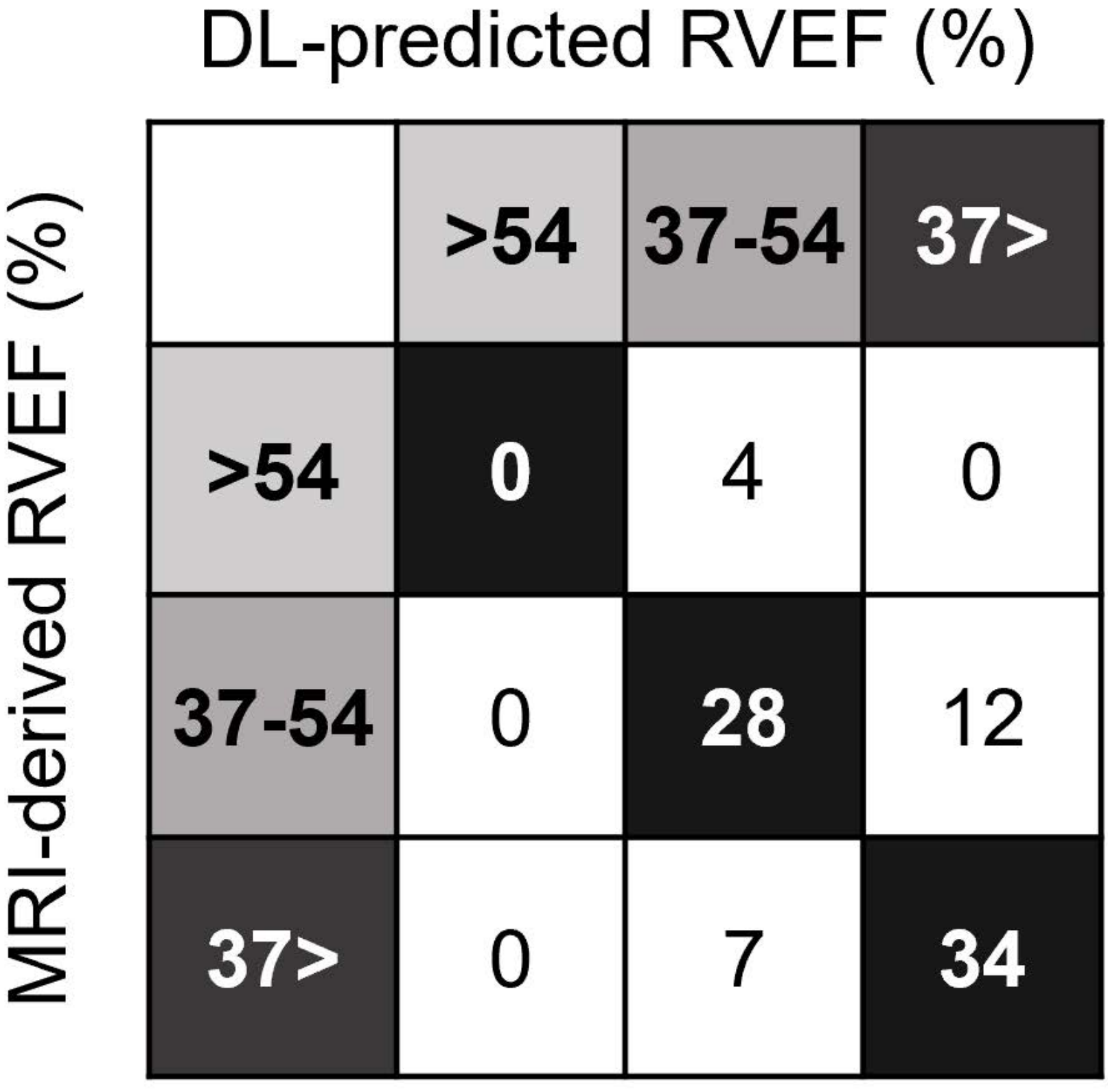
For training

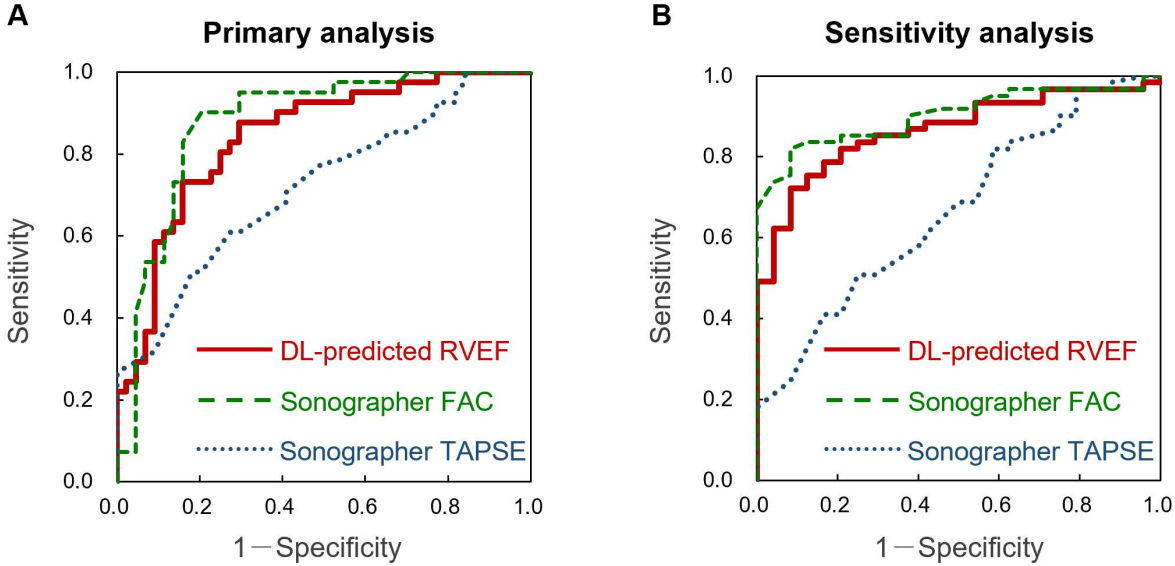


For test




A
B
DL-predicted RVEF (%)
MRI-derived RVEF (%)
n = 85
r = 0.63
p < 0.001
RVEF (DL−MRI)
Average of RVEF by DL and MRI (%)
mean+1.96SD
mean
mean-1.96SD


DL-predicted RVEF (%)
MRI-derived RVEF (%)
>54
37-54
37>
>54
0
4
0
37-54
0
28
12
37>
0
7
34


A
Primary analysis
Sensitivity
1 — Specificity
0.0
0.2
0.4
0.6
0.8
1.0
DL-predicted RVEF
Sonographer FAC
Sonographer TAPSE
B
Sensitivity analysis
Sensitivity
1 — Specificity
DL-predicted RVEF
Sonographer FAC
Sonographer TAPSE