An attention-based deep learning method for right ventricular quantification using 2D echocardiography: feasibility and accuracy

Polydoros N. Kampaktsis MD PhD[1*], Tuan A. Bohoran BSc[2,*], Mark Lebehn MD[1], Laura McLaughlin MD[3], Jay Leb MD[4], Zhonghua Liu PhD[5], Serafeim Moustakidis PhD[6], Athanasios Siouras PhD [6], Anvesha Singh MD [7], Rebecca T. Hahn MD[1], Gerry P. McCann MD[7], and Archontis Giannakidis PhD[2]

1 Division of Cardiology, Department of Medicine, Columbia University Irving Medical Center, New York City, NY, USA.
2 School of Science and Technology, Nottingham Trent University, Nottingham, UK.
3 Department of Medicine, Columbia University Irving Medical Center, New York City, NY, USA.
4 Department of Radiology, Columbia University Irving Medical Center, New York City, NY, USA.
5 Department of Biostatistics, Columbia University Irving Medical Center, New York City, NY, USA
6 AiDEAS, Tallinn, Estonia.
7 Department of Cardiovascular Sciences, University of Leicester and the NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Leicester, UK.

* Authors contributed equally

Corresponding author:
Polydoros N. Kampaktsis
177 Fort Washington Avenue, MHB2
10032 NY, NY
Columbia University Irving Medical Center
pkampaktsis@yahoo.com

Abstract

Aim: To test the feasibility and accuracy of a new attention-based deep learning (DL) method for right ventricular (RV) quantification using 2D echocardiography (2DE) with cardiac magnetic resonance imaging (CMR) as reference.

Methods and results: We retrospectively analyzed images from 50 adult patients (median age 51, interquartile range 32-62 42% women) who had undergone CMR within 1 month of 2DE. RV planimetry of the myocardial border was performed in end-diastole (ED) and end-systole (ES) for 8 standardized 2DE RV views with calculation of areas. The DL model comprised a Feature Tokenizer module and a stack of Transformer layers. Age, gender and calculated areas were used as inputs, and the output was RV volume in ED/ES. The dataset was randomly split into training, validation and testing subsets (35, 5 and 10 patients respectively).

Mean RVEDV, RVESV and RV ejection fraction (EF) were 163±70ml, 82±42ml and 51±8% respectively without differences among the subsets. The proposed method achieved good prediction of RV volumes (R^2=0.953, absolute percentage error [APE]=9.75±6.23%) and RVEF (APE=7.24±4.55%). Per CMR, there was 1 patient with RV dilatation and 3 with RV dysfunction in the testing dataset. The DL model detected RV dilatation in 1/1 case and RV dysfunction in 4/3 cases.

Conclusions: An attention-based DL method for 2DE RV quantification showed feasibility and promising accuracy. The method requires validation in larger cohorts with wider range of RV size and function. Further research will focus on the reduction of the number of required 2DE to make the method clinically applicable.

Keywords

Echocardiography
Machine learning
Deep learning
Right ventricle
Quantification

Introduction

Two-dimensional echocardiography (2DE) is the first line imaging technique for right ventricular (RV) evaluation as it widely available, portable and cost-effective (1). However, assessment of RV on 2DE is largely qualitative in clinical practice and suffers from inadequate accuracy and reproducibility (2, 3). Even when multiple 2DE quantitative indices are used, RV volumes cannot be directly and accurately calculated (3). Three-dimensional echocardiography (3DE) allows for quantitative RV assessment via calculation of RV volumes without geometric assumptions, but is commonly limited by poor image quality and requires special equipment and training (4). Cardiac magnetic resonance imaging (CMR) represents the gold standard for quantitative RV assessment (5, 6). However, CMR scanners are neither widely available nor portable; in 2019, only 5 CMR studies were performed for every 1,000 echocardiograms in the Medicare population (7). To bridge the gap between RV evaluation by 2DE and CMR, we sought to develop a novel, non geometric-, deep learning (DL)-based method for accurate and readily available 2DE RV quantification. In the current study, we tested the feasibility and accuracy of the proposed DL method.

Methods

Cohort

We retrospectively identified 50 adult patients who underwent 2DE and CMR within 30 days as part of routine cardiac care between 6/2020 – 9/2021 at Columbia University Irving Medical Center (CUIMC). In all cases, both CMR and 2DE were performed as part of an initial diagnostic evaluation and neither imaging modality was used for follow-up assessment. We excluded patients with prior cardiac surgery of the RV, tricuspid or pulmonic valve repair or replacement, complex congenital heart disease, more than small pericardial effusion, cardiac tamponade, atrial fibrillation at the time of either study, severe RV dysfunction or inadequate imaging quality. The study was approved by the CUIMC Institutional Review Board and informed consent was waived.

2DE image acquisition and pre-processing

Echocardiographic images were obtained by experienced sonographers with standard of care Philips Epic (7C and CVx) and i33 ultrasound systems. Scanning was performed according to the comprehensive protocol of the CUIMC Echocardiography Laboratory and the American Society of Echocardiography recommendations with a minimum of 22 anatomic views per study (8). The endocardial-myocardial RV interface was traced in end-systole (ES) and end-diastole (ED) by a single cardiologist [PNK] with expertise in cardiovascular ultrasound in the following 8 standardized views using a commercially available Syngo Dynamics workstation (Siemens): parasternal long axis (PLAX), RV-Inflow, parasternal short axis at the level of the aortic valve (PSAX-AV), base (PSAX-base), mid (PSAX-mid) and apex of the left ventricle (PSAX-distal), standard four-chamber (Four-C) and subcostal four chamber (Sub-C). If the standard four-chamber view was not available or had poor image quality, a focused apical RV view was used. An area was automatically calculated for each tracing by the workstation. The cardiologist was blinded to the CMR results.

Additional RV tracing measurements were performed for a subset of the patients (training subset, n=10, please see implementation subsection below) by [PNK] and independently by [ML] in order to evaluate for intra- and interobserver reliability respectively.

2DE detection of RV dilatation and dysfunction was qualitative, according to the judgement of experienced cardiologists. However, quantitative indices of RV dilatation (RV basal diameter in Four-C view) and dysfunction (tricuspid annular plane systolic excursion [TAPSE] and peak systolic RV tissue Doppler [RVS']) were calculated add hoc. We used a left ventricular ejection fraction of 50% as cutoff for left ventricular dysfunction.

Cardiac MRI image acquisition

Cardiac MRI studies were performed with breath holding and electrocardiogram gating using a Signa 1.5 Tesla MRI scanner (General Electric, Milwaukee, WI) with either sixteen or eight-channel phased array. Short-axis cine images were acquired using a steady-state free precession pulse sequence

with the following typical parameters: TR 3.0 ms, TE 1.0 ms, flip angle of 60°, 16 views per segment, field of view 35 x 35 mm, acquisition matrix 256 x 256, slice thickness 8 mm with no gap, and receiver bandwidth 125 kHz.

Cardiac MRI image analysis

CMR analysis was performed using cvi42 v5.11 (Circle Cardiovascular Imaging, Calgary) by one of four experienced cardiac radiologists with between 3 and 17 years of experience independently reading studies. Cine loops were used to select images at end-diastole and end-systole. End-diastole and end-systole were defined independently for both the right and left ventricles as the phases with the largest and smallest volumes, respectively. Endocardial segmentation was performed by manual tracing of each end-diastolic and end-systolic short-axis view and used to calculate right and left ventricular volumes. Four chamber cine images were used as a reference to help define the atrioventricular valves and apical planes. Right ventricular ejection fraction (RVEF) was calculated from end-diastolic and end-systolic volumes (RVEDV − RVESV)/RVEDV). By convention, trabeculations and papillary muscles were considered part of the ventricular blood pool in both systole and diastole. We applied gender specific CMR cutoffs for RV dilatation and dysfunction as proposed by Petersen et al (9).

Deep learning architecture

Areas derived from 2DE RV planimetry along with their corresponding cardiac phase (ED or ES), patient age and gender were used by a DL algorithm, namely a Feature Tokenizer (FT) and Transformer model, in order to predict CMR RV volumes (Figure 1). Age and gender were selected as input variables since they are available on every 2DE study and also determine normal RV cutoffs for volumes and RVEF. RVEF was then calculated from RV volumes using the formula (RVEDV-RVESV) / RVEDV. In summary, the FT-Transformer is a DL algorithm for tabular data where all categorical and numerical inputs (features) are tokenized and then forwarded to cascaded Transformer layers (10). A separate input variable specified the cardiac phase of the calculated areas (ED or ES).

In more detail, the Feature Tokenizer module turns inputs x into embeddings $T \in R^{h \times w}$. The embedding for a feature x_i is obtained by:

$$T_i = B_i + f_i(x_i) \in R^w \qquad f_i : Xi \rightarrow R^w$$

where B_i is the i^{th} feature bias, $f_{i\,(num)}$ is the element-wise multiplication of $x_{i\,(num)}$ with vector $W_{i(num)} \in R^w$ (numerical features) and $f_{i\,(cat)}$ is a lookup table with $W_{i\,(cat)} \in R^{Si \times w}$ for $x_{i\,(cat)}$ (categorical features). The module can be described overall as:

$$T_{i\,(num)} = B_{i\,(num)} + x_{i\,(num)} \cdot W_{i\,(num)} \quad \in R^w$$

$$T_{i\,(cat)} = B_{i\,(cat)} + e_i^T W_{i(cat)} \quad \in R^w$$

$$T = stack\,[T_{1\,(num)}, \ldots T_{k\,(num)}, T_{1\,(cat)}, \ldots T_{k\,(cat)}] \quad \in R^{h \times w}$$

where e_i^T is a one-hot vector for the corresponding categorical feature.

In the Transformer module, the embedding of [CLS] token (classification token) is appended to *T,* and *L* transformer layers $F_1, \ldots F_L$ are applied (11):

$$T_0 = \text{stack}\,[\,[CLS], T\,] \qquad T_i = F_i(T_{i-1})$$

The final representation of the [CLS] token is used for both RVEDV and RVESV prediction:

$$Y_{pred} = \text{Linear}(\,\text{ReLU}(\text{LayerNorm}\,(T_L{}^{[CLS]}))\,)$$

Normalization (12), removal of first Normalization, Feed Forward and Multi-Head Self-Attention (MHSA) (13) were implemented in each Transformer layer as described by Gorishniy et al (10).

Implementation

The cohort of 50 patients was randomly split into training (n=35), validation (n=5) and testing (n=10) datasets with equal number of ED and ES volumes corresponding to the same patients. The mean squared error loss function was used for training. We trained for 500 epochs with batch size=1. The hyperparameters were chosen using the random search method.

Statistical analysis & DL method evaluation

Baseline characteristics were compared between groups using the Kruskal-Wallis ANOVA test for continuous variables and Pearson's χ^2 test for categorical variables. Intraclass correlation coefficients (ICC) were calculated to evaluate the intra- and inter-observer reliability of measured RV tracings for the testing dataset using a 2-way fixed and a 2-way random effects model respectively. In both cases, an absolute agreement criterion was applied. ICC values of less than 0.5, 0.5-0.75, 0.75-0.9 and above 0.9 were used to indicate poor, moderate, good and excellent reliability, respectively. All tests were two-sided and $p<0.05$ was considered as statistically significant. STATA 17.0 BE (StataCorp LLC, Texas USA) was used for statistical analyses. Python programming language was used for implementation of the DL algorithms.

Predicted RV volumes and calculated RVEF were compared to CMR-derived RV volumes and RVEF (ground truth). State of the art machine learning algorithms for tabular data, namely XGBoost, CatBoost and Tab-Transformer (14), were also used to predict RV volumes based on the same 2DE planimetered areas. We compared predicted RV volumes against CMR-calculated volumes by using the R^2 coefficient, the absolute percentage error (APE) and performed Bland-Altman analysis. All state of the art methods were implemented according to the recommended values by Gorishniy et al (10). Predicted RV volumes and calculated RVEF were plotted against the CMR reference values (ground truth).

Results

Baseline characteristics

Table 1 summarizes the baseline clinical, echocardiographic and CMR characteristics of the study cohort. Patients had median age 51, interquartile range 32-62 and 42% were women. Twenty patients (40%) had a clinical history of heart failure and 6 (12%) had prior left sided cardiac surgery. Seven patients (14%) had hypertrophic cardiomyopathy and 6 (12%) had simple congenital heart disease (atrial/ventricular septal defect, anomalous coronary artery, bicuspid/unicuspid aortic valve). There were no significant differences in clinical characteristics between the training (n=40) and testing (n=10) subsets (Table 1).

By 2DE, 28% of patients had left ventricular systolic dysfunction. RV dilatation and dysfunction were identified in 20% and 16% of patients respectively, using 2DE. No patient had severe RV dilatation or dysfunction. Mean RV basal diameter was 3.8±0.8 cm. TAPSE and RVS' were 18.8±5.6 mm and 11.8±2.5 cm/s, respectively. The testing subset did not include any patient with RV dilatation by 2DE, however this difference was not statistically significant (p=0.317).

The mean time interval between 2DE and CMR was 9.9±5.6 days (median 6 days, interquartile range 2-20 days). By CMR, mean RVED and RVES volume was 163±70 ml and 86±45 ml, respectively. Mean RVEF was 50±8 %. Six patients (12%) had RV dilatation and 9 (18%) had RV dysfunction. According to the same cutoffs, there was no difference in RV dilatation or dysfunction between the training and testing subsets (13% vs. 10% and 18% vs. 20% respectively).

Final hyperparameters

The final hyperparemeters of the DL model are shown in Supplementary Table 1 and are compared to the respective values from the initial study that introduced its design.

Accuracy

The proposed DL method for RV volume prediction achieved good accuracy with R^2=0.953 (Figure 2A) and absolute percentage error [APE]=9.75±6.23%. It also outperformed state of the art machine learning algorithms for tabular data, namely XGBoost (R^2=0.684, APE=19.14±17.16%), CatBoost (R^2=0.749, APE=20.10±17.91%) and Tab-Transformer (14) (R^2=0.810, APE=16.22±13.17%) (Table 2). Similar accuracy was achieved for RVEF with APE=7.24±4.55% (Figure 2B) (Table 3). Bland-Altman analysis, presented as mean bias ± 95% limits of agreement, also revealed good agreement between CMR and the proposed method for RVED (1.27±23.35ml) and RVES (-2.61±19.63ml) volumes, and RVEF (-1.97±7.04%) (Figure 3).

Using CMR-derived RV volumes and RVEF, there was 1 patient with RV dilatation and 3 with RV dysfunction in the testing dataset (n=10). Of note, qualitative 2DE analysis did not show any patient with RV dilatation in the testing dataset and 2 patients with RV dysfunction were detected of whom none had RV dysfunction by CMR. Therefore, there was no correlation at all between qualitative 2DE analysis and CMR (0% accuracy for both dilatation and dysfunction). The FT-Transformer correctly classified the 1 patient with RV dilatation and did not detect any other patients with RV dilatation (100% diagnostic accuracy). The FT-Transformer correctly identified the 3 patients with RV dysfunction by CMR (100% sensitivity) and additionally detected a 4th patient that did not have RV dysfunction by CMR (75% specificity) yielding a diagnostic accuracy of 90%.

Intra- and interobserver variability

The results of intra- and interobserver reliability analyses of measured RV tracings from 8 different 2DE views in ED and ES are shown in Supplementary Tables 2 and 3, respectively. Measurements from the subcostal views showed poor intra- and interobserver reliability in both ES and ED. Parasternal short axis views of the RV at the mid and distal level showed poor to moderate intraobserver reliability in both ES and ED. On the contrary, only the parasternal mid axis view showed moderate interobserver reliability in ED. Measurements in all other views and cardiac phases showed in general good to excellent intra- and interobserver reliability.

Discussion

Machine learning applications, specifically DL, are increasing exponentially in cardiovascular medicine (15), and particularly in cardiovascular imaging (16). To our knowledge, we report the first non-geometric-based, DL method for volumetric and functional 2DE quantification of the RV. Specifically, we used a tabular FT-Transformer for the prediction of RV volumes -and therefore, RVEF- from eight planimetered 2DE views. Our results suggest an accuracy that is close to that of CMR and better compared to other state of the art DL algorithms. Importantly, the proposed method showed significantly improved diagnostic accuracy for the identification of RV dilatation and dysfunction compared to qualitative 2DE.

RV dysfunction was previously thought of as a bystander of other cardiopulmonary diseases, such as pulmonary hypertension, valvular or congenital heart disease (17-21). Today, its importance is well established and it is considered a strong and independent predictor of mortality (20, 22, 23). Furthermore, RV dysfunction is common; recent studies suggest that it is found in approximately 1 out of 5 people who undergo cardiac evaluation in the US (24, 25). Accurate RV size and function evaluation via non-invasive imaging is therefore critical for the timely diagnosis and treatment guidance of RV dysfunction.

CMR is the gold standard for RV evaluation, as it allows for accurate, reproducible and quantitative assessment of the complex RV shape without geometric assumptions (5, 6, 26). Despite its advantages though, CMR is not widely available; only 582 physicians provided CMR services in 45 states for the Medicare B population in 2017 (27) and annual CMR volume represents only 0.5% of the respective volume of echocardiograms in the US (7). Major barriers to expansion of CMR use are expensive infrastructure, specialized personnel and the high cost of CMR studies themselves.

On the other hand, 2DE is widely available, with over 23,000 echocardiograms performed in 2019 per 100,000 Medicare beneficiaries (7). The major limitations of 2DE RV evaluation in daily practice are low accuracy, poor reproducibility and inability to calculate RV volumes and RVEF (1, 2, 28). In a quality control study, 15 readers performed 2DE RV evaluation for 12 patients and results were compared to CMR. RV dilatation and low RVEF were missed in 4 out of 10, and 3 out of 10 times, respectively. When multiple 2DE quantitative indices were used (e.g. TAPSE, RVS'), accuracy improved but was not close to that of CMR. In addition, grading of RV dilatation and dysfunction remained problematic; mild or moderate RV dysfunction were correctly identified only 57% of time; and mildly or moderately dilated RV size only 47% of time (3). This lack of accuracy in 2DE RV evaluation can lead to under-diagnosis of RV dilatation/dysfunction, delay in treatment and potentially adverse clinical outcomes. Additionally, it constitutes the follow up of patients with RV dysfunction who receive specific therapy unreliable (29). 3DE can provide quantitative RV assessment without geometric assumptions, but has failed to bridge the gap between echocardiography and CMR in daily clinical practice, as it requires special equipment, advanced training and is significantly limited by image quality (4). In response to this unmet need for better RV evaluation, the National Heart, Lung and Blood Institute in collaboration with the National Institute of Biomedical Imaging and Bioengineering have recommended the development of new imaging methods of the RV (30).

A number of previous 2DE quantitative methods that used geometric assumptions to calculate RV volumes resulted in poor accuracy and reproducibility compared to CMR, most likely because of the geometrically complex RV shape (31, 32). A knowledge-based 2DE method to quantify RV volumes has also been proposed with reported good accuracy and reproducibility. The method interpolates between manual tracings from seven 2DE views by referencing against an online database of CMR-derived RV shapes (33). However, this method is tedious and time-consuming, requiring tracings from at least 6 different views and was not adopted for daily clinical practice. Recently, a DL was used for the first time to predict RVEF from 2DE (34). Although this study represents a very important step, it does not perform RV volume quantification and its overall accuracy was similar to 2DE.

The results of our pilot study suggest an agreement between the proposed DL method and CMR that is similar to the agreement between 3DE and CMR for RV evaluation (reported as -2.3±27.4ml and 5.2±19.0ml for RVED and RVES volumes per Bland-Altman analysis) (35). Therefore, the proposed DL method could break new ground in non-invasive RV evaluation, by offering widely available and accurate RV quantification based only on 2DE. Furthermore, it would potentially function as a gatekeeper for CMR and allow for more appropriate clinical follow up without the need for serial CMR imaging.

Limitations

We recognize 2 main limitations in the current pilot study. First, the cohort used was small and with only a small percentage of patients with dilated and dysfunctional RVs. Secondly, clinically established methods for quantification of left ventricular function such as the modified Simpson's rule and global longitudinal strain require tracings from two and three 2DE views respectively (36). Tracing of the endocardial-myocardial interface from 8 different views that is used in our method is time-consuming, can result in poor reliability as we noted from subcostal and mid-distal parasternal short-

axis views. Lastly, our results are based on CMR and 2DE that were performed within 30 days (median 6 days). Multimodality imaging is ideally performed on the same day or as close as possible to minimize physiological changes. However, experts in the field have previously used an interval up to 30 days in several studies combining 2DE and CMR (37-39).

Future directions

The results provide a strong foundation to expand the work in a larger dataset with further training and testing in much larger sample that will allow more robust performance statistics and an external validation cohort. Additionally we aim to implement DL models or other statistical methods for reduction of the number of required 2DE views to a maximum of 3 by analyzing which views carry the most importance. Finally, we aim to test the reproducibility of the method and investigate whether it can be combined with automated or semi-automated segmentation to reduce the manual burden from the physician.

Conclusions

In summary, the current study suggests that quantitative RV evaluation via a non-geometric, DL method using traced 2DE RV areas is feasible and increases diagnostic accuracy compared to qualitative 2DE for the detection of RV dilatation and dysfunction. Further research will aim at validating the method in larger cohorts with greater variation in RV size and function, as well as reducing the number of required 2DE views.

Funding:
Dr. Ted Goldstein funded this work with a generous private donation

Patent application:
Dr. Polydoros Kampaktsis and Dr. Archontis Giannakidis have filed a provisional patent for the method described in this manuscript along with the associated dataset.

Data sharing:
Authors elect to not share data

Disclosures:
The authors have nothing pertinent to disclose

Figures

Graphical abstract
2DE views of the RV are planimetered manually in ED and ES. Calculated areas are entered in the DL model along with age and gender. The DL model is trained to approximate CMR-derived RV volume in ED and ES. RVEF is then simply calculated from the RV volumes.

2DE = 2D echocardiography, 3D = 3 dimensional, CMR = cardiac magnetic resonance imaging, RV = right ventricle, ED = end diastole, ES = end systole, DL = deep learning

Figure 1. Attention-based deep learning architecture used for prediction of CMR RV volumes by traced 2DE RV areas. Features are turned into embeddings before a stack of Transformer layers is applied.

2DE = 2D echocardiography, CMR = cardiac magnetic resonance imaging, RV = right ventricle

Figure 2. Predicted **A)** RVED and RVES volumes and **B)** RVEF vs. ground truth (CMR) using the proposed attention-based deep learning architecture

RVED = right ventricular end diastolic, RVES = right ventricular end systolic, CMR = cardiac magnetic resonance imaging

Figure 3. Bland-Altman analysis for predicted RVED (A) and RVES (B) volume, as well as RVEF (C) vs. CMR

RV = right ventricle, RVED = right ventricular end diastolic, RVES = right ventricular end systolic RVEF = RV ejection fraction, CMR = cardiac magnetic resonance imaging

Tables

Table 1. Baseline clinical and imaging characteristics (presented as mean±SD or frequency (%))

Clinical Data	All (n=50)	Training (n=40)	Testing (n=10)	p-value
Clinical				
Female gender	21 (42)	16 (40)	5 (50)	0.567
Age	47±18	48±18	46±17	0.698
Coronary artery disease	11 (22)	9 (23)	2 (20)	0.864
Diabetes	7 (14)	6 (15)	1 (10)	0.684
Paroxysmal atrial fibrillation	6 (12)	5 (13)	1 (10)	0.828
Hypertrophic cardiomyopathy	7 (14)	5 (13)	2 (20)	0.541
Simple congenital heart disease (ASD,VSD, Bicuspid/unicuspid AV, ACA)	6 (12)	5 (13)	1 (10)	0.828
Heart failure	19 (40)	14 (38)	5 (50)	0.449
Prior cardiac surgery	6 (12)	5 (13)	1 (10)	0.828
Echocardiography				
LV systolic dysfunction	14 (28)	11 (28)	3 (30)	0.750
Mild	3 (6)	2 (5)	1 (10)	
Moderate	8 (16)	6 (15)	2 (20)	
Severe	3 (6)	3 (8)	0 (0)	
RV basal diameter [cm] (4-chamber)	3.8±0.8	3.9±0.7	3.8±0.9	0.698
RV dilatation	10 (20)	10 (26)	0 (0)	0.317
Mild	6 (12)	6 (15)	0 (0)	
Moderate	4 (8)	4 (10)	0 (0)	
RV dysfunction	8 (16)	6 (15)	2 (20)	0.843
Mild	7 (14)	5 (1)	2 (20)	
Moderate	1 (2)	1 (2)	0 (0)	
TAPSE [mm]	18.8±5.6	18.6±5.8	19.4±5.3	0.808
RVS' [cm/s]	11.8±2.7	11.8±2.6	11.9±1.8	0.634
FAC [%]	46±10	46±10	44±13	0.357
TR moderate or more	5 (10)	5 (13)	0 (0)	0.239
CMR				
Time between TTE & CMR [days]	9.9±5.6	10±10	8±8	0.488
RVEDV [ml]	163±70	164±76	160±39	0.628
RVESV [ml]	86±45	83±45	79±28	0.913
RVEF [%]	50±8	50±8	52±8	0.913
RV dilatation #	6 (12)	5 (13)	1 (10)	0.828
RV dysfunction #	9 (18)	7 (18)	2 (20)	0.854

* not available for 15 patients

^ all patients were in normal sinus rhythm at the time of TTE and CMR

gender-specific cutoffs were applied

Table 2. Quantitative comparison of the predicted RV volumes between the proposed and state of the art machine learning methods.

Method	R^2 Score	APE (%) mean (±SD)	*p*-value
Proposed	0.953	9.75 (±6.23)	—
TabTransformer	0.810	16.22 (±13.17)	0.13
CatBoost	0.749	20.10 (±17.91)	<0.001
XGBoost	0.684	19.14 (±17.16)	0.053

APE = absolute percentage error, SD = standard deviation. *p*-values were obtained from the Wilcoxon signed-rank test (α = 0.05).

Table 3. Quantitative comparison of the calculated RVEF between the proposed and state of the art machine learning methods.

Method	APE in RVEF (%) mean (±SD)	*p*-value
Proposed	7.24 (±4.55)	-
TabTransformer	11.42 (± 8.22)	0.160
CatBoost	19.15 (± 11.03)	0.006
XGBoost	24.21 (± 20.90)	0.038

APE = absolute percentage error, SD = standard deviation. *p*-values were obtained from the Wilcoxon signed-rank test (α = 0.05).

Data Availability Statement

The data underlying this article cannot be shared publicly in respect to the privacy of individuals that participated in the study. The data could be shared on reasonable request to the corresponding author

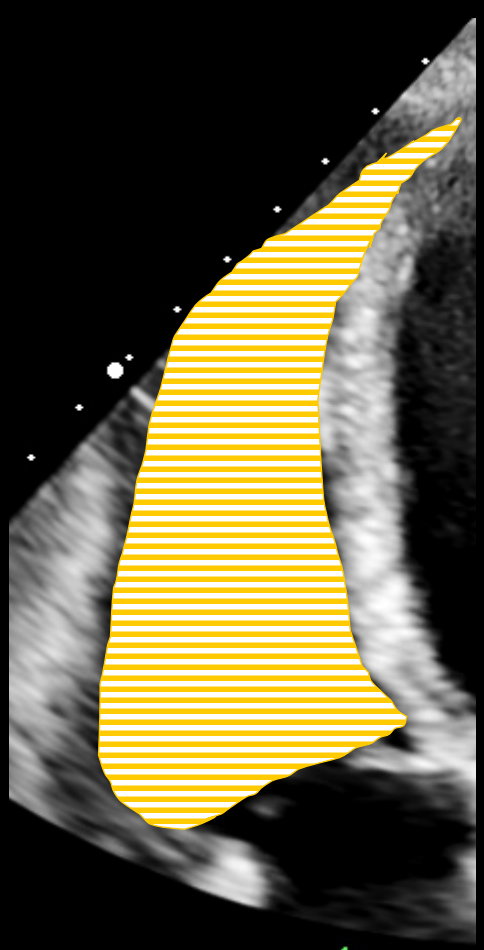
References

1. Rudski LG, Lai WW, Afilalo J, Hua L, Handschumacher MD, Chandrasekaran K, Solomon SD, Louie EK, Schiller NB. 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(7):685-713; quiz 86-8.
2. Lai WW, Gauvreau K, Rivera ES, Saleeb S, Powell AJ, Geva T. Accuracy of guideline recommendations for two-dimensional quantification of the right ventricle by echocardiography. Int J Cardiovasc Imaging. 2008;24(7):691-8.
3. Ling LF, Obuchowski NA, Rodriguez L, Popovic Z, Kwon D, Marwick TH. Accuracy and interobserver concordance of echocardiographic assessment of right ventricular size and systolic function: a quality control exercise. J Am Soc Echocardiogr. 2012;25(7):709-13.
4. Fernández-Golfín C, Zamorano JL. Three-Dimensional Echocardiography and Right Ventricular Function: The Beauty and the Beast? Circ Cardiovasc Imaging. 2017;10(2).
5. Geva T. Is MRI the preferred method for evaluating right ventricular size and function in patients with congenital heart disease?: MRI is the preferred method for evaluating right ventricular size and function in patients with congenital heart disease. Circ Cardiovasc Imaging. 2014;7(1):190-7.
6. Mooij CF, de Wit CJ, Graham DA, Powell AJ, Geva T. Reproducibility of MRI measurements of right ventricular size and function in patients with normal and dilated ventricles. J Magn Reson Imaging. 2008;28(1):67-73.
7. Reeves RA, Halpern EJ, Rao VM. Cardiac Imaging Trends from 2010 to 2019 in the Medicare Population. Radiol Cardiothorac Imaging. 2021;3(5):e210156.
8. Mitchell C, Rahko PS, Blauwet LA, Canaday B, Finstuen JA, Foster MC, Horton K, Ogunyankin KO, Palma RA, Velazquez EJ. Guidelines for Performing a Comprehensive Transthoracic Echocardiographic Examination in Adults: Recommendations from the American Society of Echocardiography. J Am Soc Echocardiogr. 2019;32(1):1-64.
9. Petersen SE, Aung N, Sanghvi MM, Zemrak F, Fung K, Paiva JM, Francis JM, Khanji MY, Lukaschuk E, Lee AM, Carapella V, Kim YJ, Leeson P, Piechnik SK, Neubauer S. Reference ranges for cardiac structure and function using cardiovascular magnetic resonance (CMR) in Caucasians from the UK Biobank population cohort. J Cardiovasc Magn Reson. 2017;19(1):18.
10. Gorishniy Y RI, Khrulkov V, Babenko A., editor Revisiting deep learningmodels for tabular data. Neural Information Processing Systems 34; 2021:18932-18943
11. Devlin J CM, Lee K, Toutanova C. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding2018.
12. Wang Q LB, Xiao T, Zhu J, Li C, Wong DF, Chao LS, editor Learning deep transformer models for machine translation. ACL; 2019.
13. Vaswani A SN, Parmar J, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I. , editor Attention is all you need. NIPS; 2017.
14. Huang X KA, Cvitkovic M, Karnin Z. TabTransformer: Tabular Data Modeling Using Contextual Embeddings2020.
15. Kampaktsis PN, Emfietzoglou M, Al Shehhi A, Fasoula NA, Bakogiannis C, Mouselimis D, Tsarouchas A, Vassilikos VP, Kallmayer M, Eckstein HH, Hadjileontiadis L, Karlas A. Artificial intelligence in atherosclerotic disease: Applications and trends. Front Cardiovasc Med. 2022;9:949454.
16. Litjens G, Ciompi F, Wolterink JM, de Vos BD, Leiner T, Teuwen J, Išgum I. State-of-the-Art Deep Learning in Cardiovascular Image Analysis. JACC Cardiovasc Imaging. 2019;12(8 Pt 1):1549-65.
17. Fontan F, Baudet E. Surgical repair of tricuspid atresia. Thorax. 1971;26(3):240.

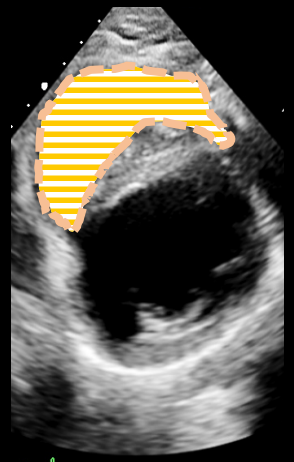
18. Starr I, Jeffers WA, Meade RH. The absence of conspicuous increments of venous pressure after severe damage to the right ventricle of the dog, with a discussion of the relation between clinical congestive failure and heart disease. American Heart Journal. 1943;26(3):291-301.
19. Kagan A. Dynamic responses of the right ventricle following extensive damage by cauterization. Circulation. 1952;5(6):816-23.
20. Konstam MA, Kiernan MS, Bernstein D, Bozkurt B, Jacob M, Kapur NK, Kociol RD, Lewis EF, Mehra MR, Pagani FD, Raval AN, Ward C. Evaluation and Management of Right-Sided Heart Failure: A Scientific Statement From the American Heart Association. Circulation. 2018;137(20):e578-e622.
21. Stout KK, Daniels CJ, Aboulhosn JA, Bozkurt B, Broberg CS, Colman JM, Crumb SR, Dearani JA, Fuller S, Gurvitz M, Khairy P, Landzberg MJ, Saidi A, Valente AM, Van Hare GF. 2018 AHA/ACC Guideline for the Management of Adults With Congenital Heart Disease: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation. 2019;139(14):e698-e800.
22. Kammerlander AA, Marzluf BA, Graf A, Bachmann A, Kocher A, Bonderman D, Mascherbauer J. Right ventricular dysfunction, but not tricuspid regurgitation, is associated with outcome late after left heart valve procedure. J Am Coll Cardiol. 2014;64(24):2633-42.
23. Haddad F, Doyle R, Murphy DJ, Hunt SA. Right ventricular function in cardiovascular disease, part II: pathophysiology, clinical importance, and management of right ventricular failure. Circulation. 2008;117(13):1717-31.
24. Padang R, Chandrashekar N, Indrabhinduwat M, Scott CG, Luis SA, Chandrasekaran K, Michelena HI, Nkomo VT, Pislaru SV, Pellikka PA, Kane GC. Aetiology and outcomes of severe right ventricular dysfunction. Eur Heart J. 2020;41(12):1273-82.
25. Nochioka K, Querejeta Roca G, Claggett B, Biering-Sørensen T, Matsushita K, Hung CL, Solomon SD, Kitzman D, Shah AM. Right Ventricular Function, Right Ventricular-Pulmonary Artery Coupling, and Heart Failure Risk in 4 US Communities: The Atherosclerosis Risk in Communities (ARIC) Study. JAMA Cardiol. 2018;3(10):939-48.
26. Khalique OK, Cavalcante JL, Shah D, Guta AC, Zhan Y, Piazza N, Muraru D. Multimodality Imaging of the Tricuspid Valve and Right Heart Anatomy. JACC Cardiovasc Imaging. 2019;12(3):516-31.
27. Goldfarb JW, Weber J. Trends in Cardiovascular MRI and CT in the U.S. Medicare Population from 2012 to 2017. Radiol Cardiothorac Imaging. 2021;3(1):e200112.
28. Damy T, Viallet C, Lairez O, Deswarte G, Paulino A, Maison P, Vermes E, Gueret P, Adnot S, Dubois-Randé JL, Hittinger L. Comparison of four right ventricular systolic echocardiographic parameters to predict adverse outcomes in chronic heart failure. Eur J Heart Fail. 2009;11(9):818-24.
29. Vanderpool RR, Hunter KS, Insel M, Garcia JGN, Bedrick EJ, Tedford RJ, Rischard FP. The Right Ventricular-Pulmonary Arterial Coupling and Diastolic Function Response to Therapy in Pulmonary Arterial Hypertension. Chest. 2022;161(4):1048-59.
30. Voelkel NF, Quaife RA, Leinwand LA, Barst RJ, McGoon MD, Meldrum DR, Dupuis J, Long CS, Rubin LJ, Smart FW, Suzuki YJ, Gladwin M, Denholm EM, Gail DB. Right ventricular function and failure: report of a National Heart, Lung, and Blood Institute working group on cellular and molecular mechanisms of right heart failure. Circulation. 2006;114(17):1883-91.
31. Jenkins C, Chan J, Bricknell K, Strudwick M, Marwick TH. Reproducibility of right ventricular volumes and ejection fraction using real-time three-dimensional echocardiography: comparison with cardiac MRI. Chest. 2007;131(6):1844-51.
32. Kim J, Srinivasan A, Seoane T, Di Franco A, Peskin CS, McQueen DM, Paul TK, Feher A, Geevarghese A, Rozenstrauch M, Devereux RB, Weinsaft JW. Echocardiographic Linear Dimensions for Assessment of Right Ventricular Chamber Volume as Demonstrated by Cardiac Magnetic Resonance. J Am Soc Echocardiogr. 2016;29(9):861-70.
33. Knight DS, Schwaiger JP, Krupickova S, Davar J, Muthurangu V, Coghlan JG. 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(8):989-98.

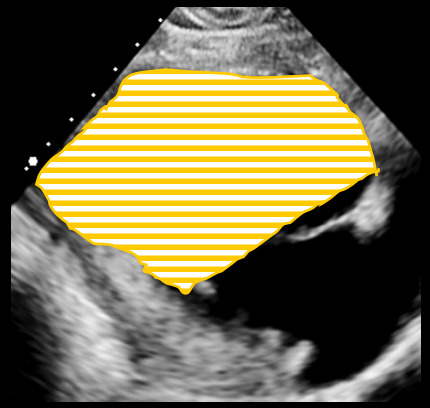
34. Tokodi M, Magyar B, Soós A, Takeuchi M, Tolvaj M, Lakatos BK, Kitano T, Nabeshima Y, Fábián A, Szigeti MB, Horváth A, Merkely B, Kovács A. Deep Learning-Based Prediction of Right Ventricular Ejection Fraction Using 2D Echocardiograms. JACC Cardiovasc Imaging. 2023;16(8):1005-18.
35. Knight DS, Grasso AE, Quail MA, Muthurangu V, Taylor AM, Toumpanakis C, Caplin ME, Coghlan JG, Davar J. Accuracy and reproducibility of right ventricular quantification in patients with pressure and volume overload using single-beat three-dimensional echocardiography. J Am Soc Echocardiogr. 2015;28(3):363-74.
36. Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, Flachskampf FA, Foster E, Goldstein SA, Kuznetsova T, Lancellotti P, Muraru D, Picard MH, Rietzschel ER, Rudski L, Spencer KT, Tsang W, Voigt JU. 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):1-39.e14.
37. Kochav J, Chen J, Nambiar L, Mitlak HW, Kushman A, Sultana R, Horn E, RoyChoudhury A, Devereux RB, Weinsaft JW, Kim J. Novel Echocardiographic Algorithm for Right Ventricular Mass Quantification: Cardiovascular Magnetic Resonance and Clinical Prognosis Validation. J Am Soc Echocardiogr. 2021;34(8):839-50.e1.
38. Rong LQ, Palumbo MC, Rahouma M, Lopes AJ, Devereux RB, Kim J, Girardi LN, Gaudino M, Weinsaft JW. Descending aortic strain quantification by intra-operative transesophageal echocardiography: Multimodality validation via cardiovascular magnetic resonance. Echocardiography. 2020;37(11):1820-7.
39. Kim J, Rodriguez-Diego S, Srinivasan A, Brown RM, Pollie MP, Di Franco A, Goldburg SR, Siden JY, Ratcliffe MB, Levine RA, Devereux RB, Weinsaft JW. Echocardiography-quantified myocardial strain-a marker of global and regional infarct size that stratifies likelihood of left ventricular thrombus. Echocardiography. 2017;34(11):1623-32.

2DE planimetry (ED/ES)

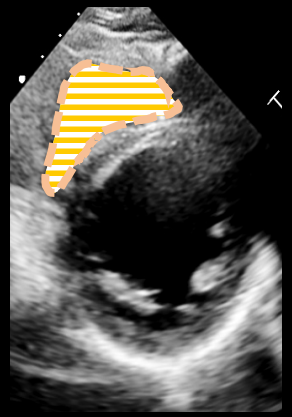


4-chamber

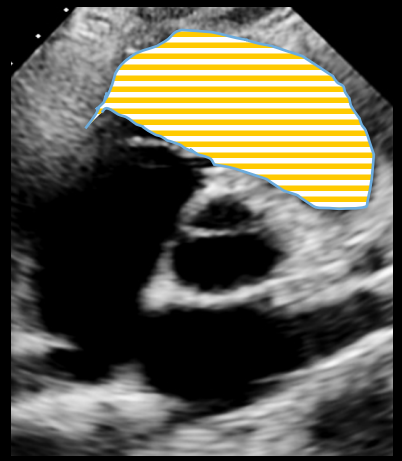




RV inflow



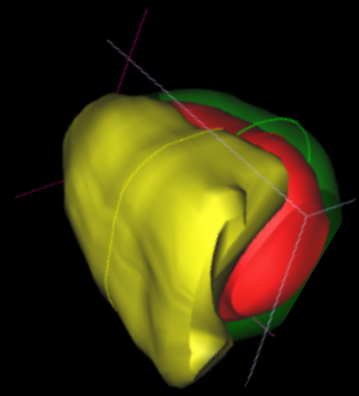
Short axis



RVOT

Training with CMR-derived RV 3D dataset

(yellow contour)



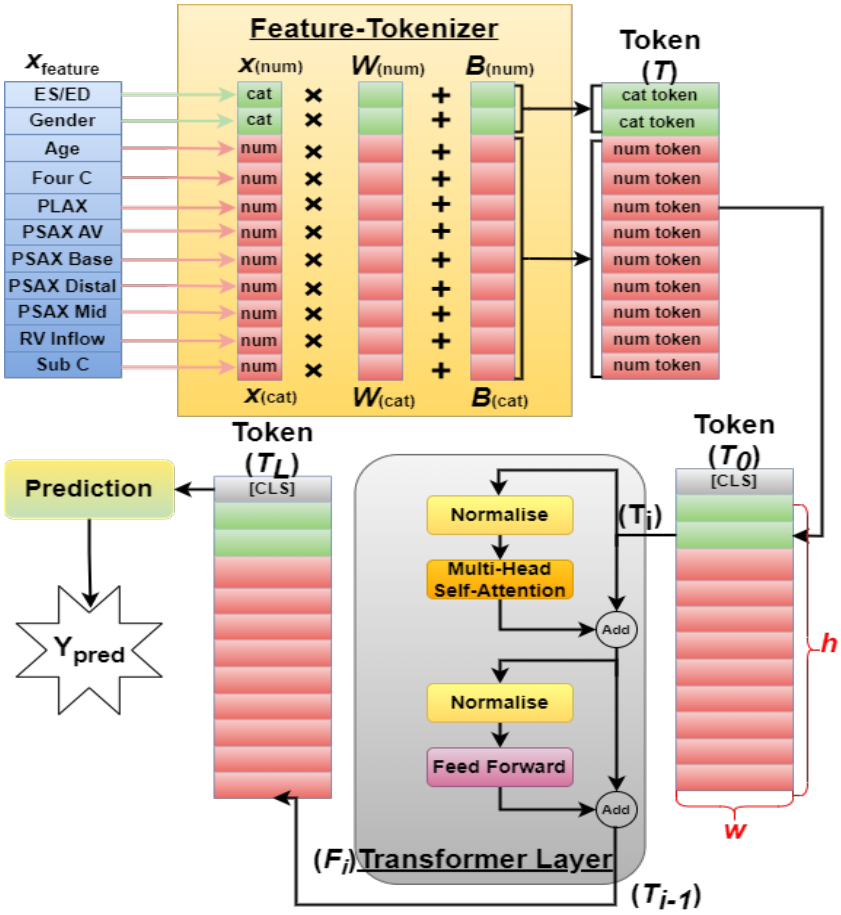
8 RV areas (ED+ES)

DL model

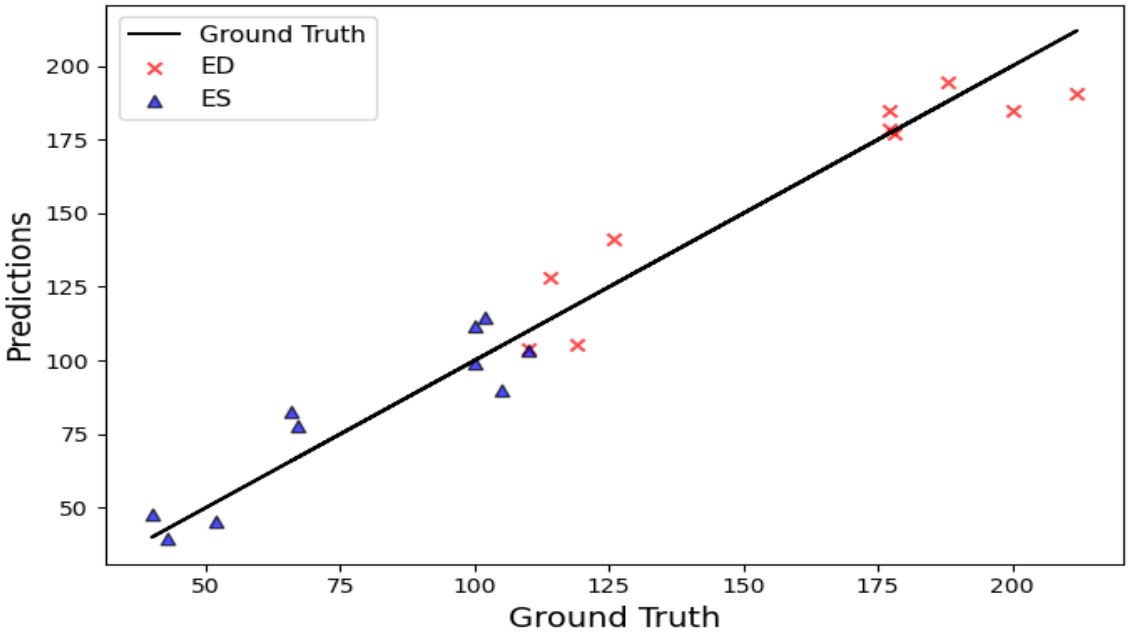
RV volume (ED/ES)

Age

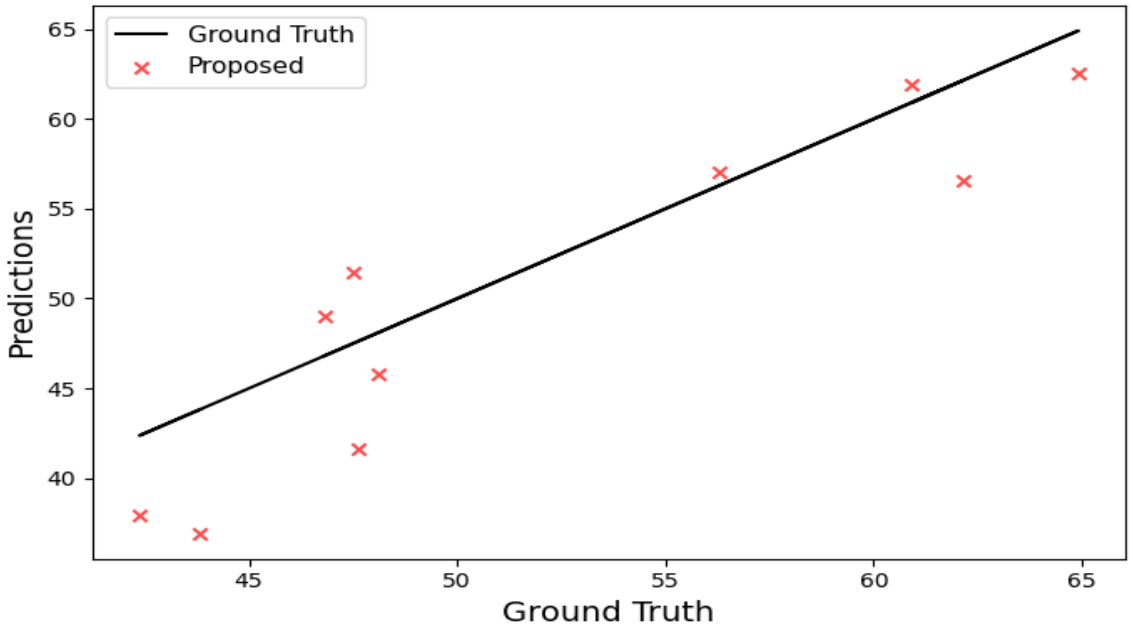
Gender


Feature-Tokenizer
$x_{feature}$
ES/ED
Gender
Age
Four C
PLAX
PSAX AV
PSAX Base
PSAX Distal
PSAX Mid
RV Inflow
Sub C
$x_{(num)}$
$W_{(num)}$
$B_{(num)}$
$x_{(cat)}$
$W_{(cat)}$
$B_{(cat)}$
cat
num
Token (T)
cat token
num token
Token (T_L)
Token (T_0)
[CLS]
Prediction
Y_{pred}
Normalise
Multi-Head Self-Attention
Add
Normalise
Feed Forward
Add
(T_i)
(F_i) Transformer Layer
(T_{i-1})
h
w

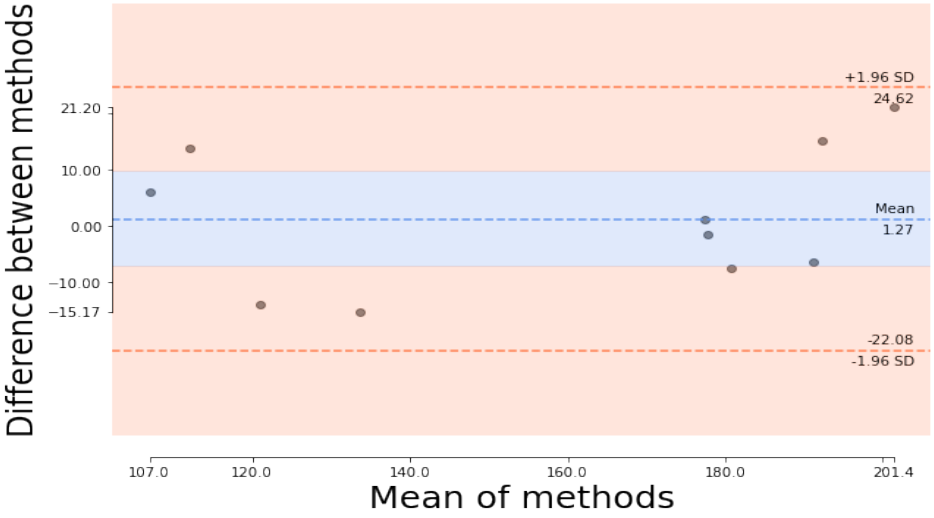
A. Predicted RV volumes vs. ground truth



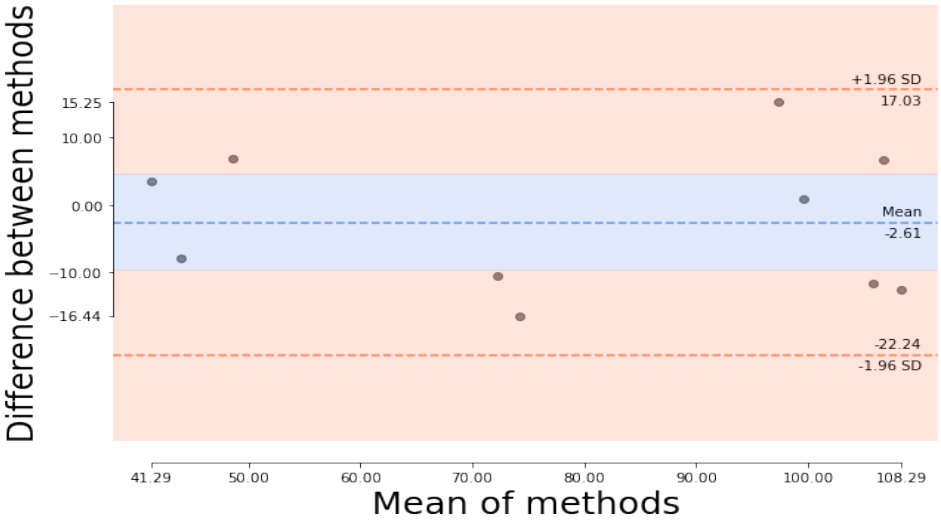
B. Predicted RVEF vs. ground truth



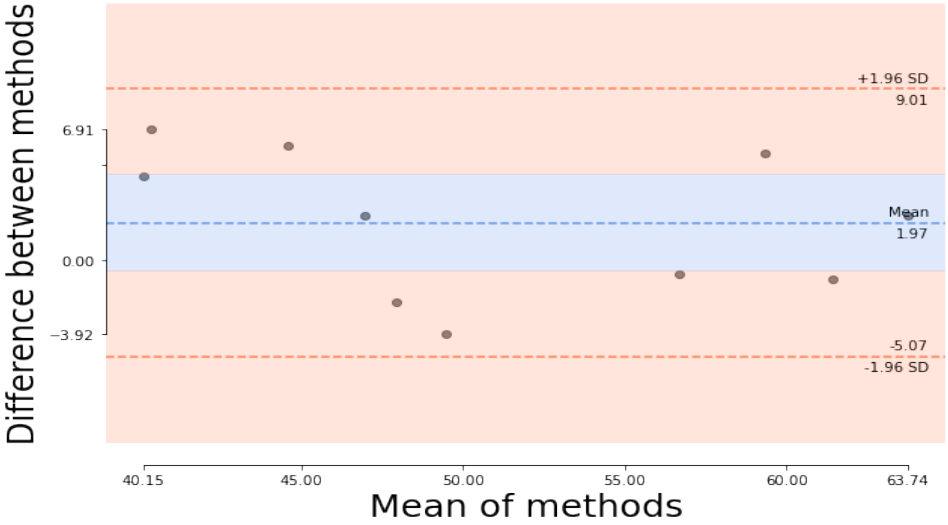
A. End-diastolic RV volume (cm^3)



B. End-systolic RV volume (cm^3)



C. RVEF (%)



Supplementary Table 1. Hyperparameters used in the final DL model for prediction of RV volumes by traced 2DE RV areas

Parameter	Used in our method	Recommended by Gorishniy et al. [12]
#Layers	3	6
Feature embedding size	16	512
Residual Dropout	0.3	0.2
Attention Dropout	0.3	0.5
FFN Dropout	0.3	0.5
FFN Factor	4/3	4/3
Learning Rate	0.01	LogUniform[3e-5,3e-4]
Weight Decay	0	LogUniform[3e-6,3e-3]
Optimizer	Adamax	AdamW

2DE = 2D echocardiography

Supplementary Table 3. Intraclass correlation coefficients for interobserver reliability of measured RV tracings. Based on 2-way random effects model and absolute agreement

A. End Diastole

RV tracing	Type	Intraclass correlation	95% Confidence Interval		F test with true value 0	
			Lower Bound	Upper Bound	Value	Sig
PLAX	Single	0.975	0.909	0.993	77.57	<0.001
	Average	0.987	0.952	0.997		
RVinflow	Single	0.937	0.782	0.983	32.45	<0.001
	Average	0.967	0.878	0.991		
PSAX_AV	Single	0.936	0.781	0.983	31.11	<0.001
	Average	0.967	0.877	0.991		
PSAX_base	Single	0.862	0.473	0.965	18.32	<0.001
	Average	0.926	0.642	0.982		
PSAX_mid	Single	0.648	0.040	0.901	4.33	0.020
	Average	0.786	0.077	0.947		
PSAX_distal	Single	0.944	0.781	0.986	42.15	<0.001
	Average	0.971	0.877	0.993		
Four_C	Single	0.930	0.723	0.982	34.66	<0.001
	Average	0.964	0.839	0.991		
SubC	Single	0.282	-0.352	0.752	1.80	0.197
	Average	0.440	-1.08	0.858		

B. End systole

RV tracing	Type	Intraclass correlation	95% Confidence Interval		F test with true value 0	
			Lower Bound	Upper Bound	Value	Sig
PLAX	Single	0.956	0.836	0.988	40.85	<0.001
	Average	0.977	0.910	0.994		
RVinflow	Single	0.931	0.750	0.982	25.78	<0.001
	Average	0.964	0.857	0.991		
PSAX_AV	Single	0.926	0.732	0.981	30.37	<0.001
	Average	0.961	0.845	0.990		
PSAX_base	Single	0.916	0.285	0.983	51.47	<0.001
	Average	0.956	0.443	0.991		
PSAX_mid	Single	0.836	0.470	0.956	10.30	0.001
	Average	0.911	0.639	0.977		
PSAX_distal	Single	0.974	0.906	0.993	82.21	<0.001
	Average	0.987	0.951	0.996		
Four_C	Single	0.913	0.709	0.977	22.46	<0.001
	Average	0.954	0.829	0.988		
SubC	Single	0.368	-0.334	0.797	2.09	0.144
	Average	0.538	-1.004	0.887		

Supplementary Table 2. Intraclass correlation coefficients for intraobserver reliability of measured RV tracings. Based on 2-way fixed effects model and absolute agreement.

A. End diastole

RV tracing	Type	Intraclass correlation	95% Confidence Interval		F test with true value 0	
			Lower Bound	Upper Bound	Value	Sig
PLAX	Single	0.891	0.452	0.975	27.59	<0.001
	Average	0.798	0.247	0.949		
RVinflow	Single	0.679	0.021	0.915	8.68	0.002
	Average	0.809	0.043	0.955		
PSAX_AV	Single	0.851	-0.032	0.972	43.99	<0.001
	Average	0.920	-0.065	0.985		
PSAX_base	Single	0.760	-0.069	0.952	32.40	<0.001
	Average	0.863	-0.148	0.975		
PSAX_mid	Single	0.455	-0.139	0.824	2.79	0.071
	Average	0.625	-0.323	0.903		
PSAX_distal	Single	0.570	0.021	0.868	4.27	0.021
	Average	0.726	0.040	0.929		
Four_C	Single	0.720	-0.076	0.937	16.59	<0.001
	Average	0.837	-0.164	0.967		
SubC	Single	0.362	-0.378	0.798	2.03	0.154
	Average	0.532	-1.218	0.888		

B. End systole

RV tracing	Type	Intraclass correlation	95% Confidence Interval		F test with true value 0	
			Lower Bound	Upper Bound	Value	Sig
PLAX	Single	0.882	0.230	0.975	32.68	<0.001
	Average	0.937	0.374	0.987		
RVinflow	Single	0.831	0.257	0.960	17.94	<0.001
	Average	0.907	0.409	0.979		
PSAX_AV	Single	0.860	-0.015	0.973	44.76	<0.001
	Average	0.925	-0.031	0.986		
PSAX_base	Single	0.658	-0.088	0.922	17.32	<0.001
	Average	0.794	-0.195	0.960		
PSAX_mid	Single	0.561	-0.048	0.870	5.45	0.009
	Average	0.719	-0.102	0.937		
PSAX_distal	Single	0.407	-0.162	0.800	2.58	0.087
	Average	0.578	-0.388	0.889		
Four_C	Single	0.682	-0.064	0.922	11.52	0.001
	Average	0.811	-0.138	0.959		
SubC	Single	0.454	-0.262	0.833	2.50	0.094
	Average	0.624	-0.712	0.909		