Video-Based Deep Learning for Automated Assessment of Left Ventricular Ejection Fraction in Pediatric Patients

Charitha D. Reddy, MD, Leo Lopez, MD, David Ouyang, MD, James Y. Zou, PhD, and Bryan He, BS, *Palo Alto, Los Angeles, and Stanford, California*

Background: Significant interobserver and interstudy variability occurs for left ventricular (LV) functional indices despite standardization of measurement techniques. Artificial intelligence models trained on adult echocardiograms are not likely to be applicable to a pediatric population. We present EchoNet-Peds, a video-based deep learning algorithm, which matches human expert performance of LV segmentation and ejection fraction (EF).

Methods: A large pediatric data set of 4,467 echocardiograms was used to develop EchoNet-Peds. EchoNet-Peds was trained on 80% of the data for segmentation of the left ventricle and estimation of LVEF. The remaining 20% was used to fine-tune and validate the algorithm.

Results: In both apical 4-chamber and parasternal short-axis views, EchoNet-Peds segments the left ventricle with a Dice similarity coefficient of 0.89. EchoNet-Peds estimates EF with a mean absolute error of 3.66% and can routinely identify pediatric patients with systolic dysfunction (area under the curve of 0.95). EchoNet-Peds was trained on pediatric echocardiograms and performed significantly better to estimate EF ($P < .001$) than an adult model applied to the same data.

Conclusions: Accurate, rapid automation of EF assessment and recognition of systolic dysfunction in a pediatric population are feasible using EchoNet-Peds with the potential for far-reaching clinical impact. In addition, the first large pediatric data set of annotated echocardiograms is now publicly available for efforts to develop pediatric-specific artificial intelligence algorithms. (J Am Soc Echocardiogr 2023;36:482-9.)

Keywords: Pediatric cardiology, Echocardiography, Deep learning, Left ventricular function, Artificial intelligence

Evaluation of left ventricular (LV) size and systolic function is crucial in the diagnosis and management of children with congenital and acquired heart disease.[1] It is used for following disease evolution and monitoring the effects of interventions in the pediatric population.[1-3] Chemotherapy dosing for pediatric cancer patients,[4] pacemaker placement in patients with arrhythmias,[5] medication titration in the setting of heart failure,[6] ventricular remodeling after surgical repair,[7] and follow-up care for patients with genetic abnormalities[8] are a few examples that demonstrate the wide range of uses for ventricular functional assessment. Echocardiography is noninvasive, safe, and readily available, and it is the standard modality for serial assessment of cardiac function. Automated quantification of ventricular function may improve variability, increase efficiency, and democratize specialized care in areas where access is limited.[9-11]

Machine learning and artificial intelligence technologies have been successfully used in adults to improve the reliability and accuracy of LV functional assessment by echocardiography.[9,11-13] Large-scale studies utilizing machine learning algorithms in pediatric echocardiography are scarce, although there has been promising work in LV segmentation with other imaging modalities.[14,15] Machine learning is more challenging in children, partly because the rarity of pediatric heart disease precludes the collection of the large data sets needed to validate machine learning models. Children have increased variability in anatomic abnormality, heart rate, size, and ability to cooperate, all contributing to a wide range of spatial and temporal resolution as well as image quality.[16] It is unclear whether models trained on adult data are generalizable to pediatrics given this increased variability. Automated LV functional assessment is an important initial step in applying machine learning to pediatric echocardiography. It can serve as a foundation for models that focus on single ventricles,

From the Division of Pediatric Cardiology, Department of Pediatrics, Stanford University School of Medicine, Palo Alto, California (C.D.R., L.L.); Department of Cardiology, Cedars-Sinai Medical Center, Los Angeles, California (D.O.); Department of Biomedical Data Science (J.Y.Z.) and Department of Computer Science (B.H.), Stanford University, Stanford, California.

Drs. Ouyang, Zou, and He contributed equally to this work.

Conflicts of Interest: None.

Reprint requests: Charitha D. Reddy, MD, Division of Cardiology, Department of Pediatrics, Stanford School of Medicine, 750 Welch Road, Suite 305, Palo Alto, CA 94034 (E-mail: *reddyc@stanford.edu*).

0894-7317/$36.00

https://doi.org/10.1016/j.echo.2023.01.015

Abbreviations
A4C = Apical 4-chamber
BSA = Body surface area
EF = Ejection fraction
LV = Left ventricular, ventricle
LVEF = Left ventricular ejection fraction
MAE = Mean absolute error
PSAX = Parasternal short-axis
RMSE = Root mean squared error

cardiomyopathies, and heart transplants, and it has the potential to impact models that focus on disease progression and outcomes.[11]

A standard parameter of LV systolic function is ejection fraction (EF), but studies have revealed significant measurement variability in adults[17,18] and in children with acquired and congenital heart disease.[19-28] A human expert usually performs manual segmentation to trace LV diastolic and systolic areas in single or multiple planes in order to calculate LV volumes.[1,2,29] The segmentation process is time-consuming and a major source of measurement variability for EF.[28,30,31] Although the standard recommendation involves the average of EF measurements over 3 to 5 separate cardiac cycles, this is rarely done in clinical practice.[21,32] A few single-center studies have published normative values for children, but these are limited because of homogenous patient populations, small sample size, variable measurement performance methodologies, and other confounding factors, thereby decreasing the generalizability of the final results.[26,27,33,34] Additionally, pediatric echocardiography often prioritizes the area-length or bullet method to calculate LVEF,[21,28] whereas adult echocardiography has trained models using the method of disks to calculate LVEF.[13,35,36]

We present EchoNet-Peds, a deep learning model in pediatric echocardiography that automates LV segmentation and calculation of EF from echocardiographic images and video clips. This is performed by using 2 standard echocardiographic views, namely, the parasternal short-axis (PSAX) and apical 4-chamber (A4C) views. The 5/6 area-length method to calculate LV volumes and EF was used because it is more reliable than other methods for evaluating systolic function, particularly in pediatrics.[21,28]

METHODS

Study Population

The study population consisted of patients who underwent an echocardiographic evaluation at Lucile Packard Children's Hospital Stanford as part of routine clinical care from 2014 to 2021. Data collection was approved by the Stanford University Institutional Review Board.

The patients were divided into 2 subsets: patients with anatomically normal hearts with normal EF and patients with anatomically normal hearts with systolic dysfunction in the absence of congenital heart disease (e.g., dilated cardiomyopathy and chemotherapy-induced systolic dysfunction). The patients with an EF by 5/6 area-length >55% were classified as having normal systolic function, and those with EF <55% were classified as abnormal. We included patients with EFs ranging from 5% to 75% on the interpretation report. All echocardiograms were acquired by trained sonographers or physicians using Philips iE33 (Koninklijke Philips N.V.), Siemens Acuson SC2000 (Siemens Medical Solutions), or Philips Epiq 7 (Koninklijke Philips N.V.) ultrasound machines with videos subsequently stored and analyzed in a Siemens Syngo Dynamics (Siemens Medical Solutions) picture archiving and communication system. To better reflect true clinical practice, video clips were not excluded on the basis of variable image quality.

Echocardiographic Data and Labeling

The echocardiographic data included grayscale two-dimensional video clips from A4C and PSAX views for calculation of LVEF using the 5/6 area-length method.[29] Views were identified using a programmed query within each study based on which video frames had tracings of LV volumes by a sonographer or cardiologist. The final measurement included in the report was extracted from the digital imaging and communications in medicine file linked to the EF. We used the calculated EF on the finalized clinical report as the gold standard by which to train and benchmark the success of EchoNet-Peds. The echocardiographic videos were cropped to remove patient identifying information and the electrocardiographic tracing, resulting in square videos. The videos were then downsampled to 112×112 pixel videos for training and testing. Videos were randomly split into groups for training, testing, and validation.

Model Training

EchoNet-Peds model development included training and validation phases for segmentation and automated EF estimation. We used 10-fold cross-validation to train and evaluate EchoNet-Peds, where 80% of video clips were used for training, 10% for testing, and 10% for validation. There was no overlap of video clips within the training, testing, and validation sets. Details of the model development process are reported in the Supplemental Index. For both tasks, we initialized model weights using the EchoNet-Dynamic model,[13] which was trained on adult echocardiograms, and fine-tuned the weights using our pediatric data set. The validation set was used to select the training epoch with the lowest loss for deployment. We used a stochastic gradient descent optimizer to train EchoNet-Peds to segment the left ventricle (LV), minimizing the binary cross-entropy loss between the generated estimations and sonographer tracing of the LV. The systolic and diastolic frames selected by the clinicians during the clinical workflow were used as input, with segmentation throughout the cardiac cycle as output. We also used a stochastic gradient descent optimizer to train EchoNet-Peds to estimate EF, minimizing the mean squared error between the generated estimations and human-measured EF. Video clips of 32 frames using every other frame were used as input, with a percentage EF as output.

EchoNet-Peds

EchoNet-Peds is made up of 2 separate components (Figure 1): a segmentation task and an EF estimation task. Ground truth for video clip labels were defined by the sonographer's or cardiologist's measurements. The segmentation task replicates the human workflow of tracing LV areas in A4C and PSAX views. Each training video clip included frames with manual tracings at end systole and end diastole that were obtained during clinical workflow; these comprised the 17,600 "labeled images." Using a strategy previously described by Tran *et al.*,[37] EchoNet-Peds automatically generates additional areas on the intermediate frames of the video clip as well. For example, if the LV end-diastolic area human tracing existed on frame 34/244 and the LV end-systolic area human tracing was on frame 120/244, EchoNet-Peds generates new estimates for the area on each of the frames between 34 and 120 as the LV is in various stages of

HIGHLIGHTS

- EchoNet-Peds is the first pediatric trained AI model to assess ventricular function.
- EchoNet-Peds estimates EF with accuracy similar to human experts.
- EchoNet-Peds is rapid and reproducible among a range of pediatric ages and sizes.
- Training EchoNet-Peds with pediatric data improved performance over adult models.

contraction and expansion, utilizing the whole echocardiographic video clip rather than 2 isolated frames (Videos 1 and 2). This task allows for the model to identify ventricular contractions and for the clinician to visualize the task accuracy.

Second, EchoNet-Peds performs automatic EF estimation across frames to mimic the EF expected from a 5/6 area-length calculation, utilizing only the video clip as input. This task most closely resembles a workflow in which a human expert visually estimates EF. EchoNet-Peds is best trained to approximate 5/6 area-length EF when both A4C and PSAX video clips are present. However, EchoNet-Peds can also successfully measure 5/6 area-length EF when only 1 of those views is available.

Statistical Analysis

Patient statistics are expressed as mean ± SD. The performance for EF estimation is measured in mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination (R^2). The performance for segmenting the LV is measured in terms of the Dice similarity coefficient, a widely used statistical tool in machine learning–derived segmentation to measure the similarity between 2 data sets by comparing the proportion of true positive pixels to the total number of pixels in an image (e.g., Dice similarity coefficient of 1 suggests identical segmentation between 2 data sets).[38] Confidence intervals were computed using 10,000 bootstrapped samples and obtaining 95th percentile ranges for each generated estimation. Comparisons of training methods were performed using a 2-tailed binomial test.

EchoNet-Peds Performance

Because of the larger variation in size among pediatric patients compared to adult patients, the performance of EchoNet-Peds among the full range of patients was evaluated by stratifying the patients by sex, age, and body surface area (BSA). In addition, the performance of EchoNet-Peds was compared with the EchoNet-Dynamic[13] model to better understand the importance of training pediatric models on pediatric data versus extrapolating from adult data-derived models. Comparison models are included. Because the success of a model is also dependent on weights assigned to various factors determined during the development of the model, we also compared EchoNet-Peds, which uses weights similar to EchoNet-Dynamic, against a model using the baseline weights from the Kinetics-400 data set,[39] which does not include medical videos.

RESULTS

Study Population and Data Set

The study data set consisted of 4,467 echocardiograms from 1,958 patients (43% female) ranging from 0 to 18 years (10 ± 5.4) as noted in Table 1. The A4C video clips and PSAX video clips were extracted from the echocardiograms, resulting in 7,643 video clips and 17,600

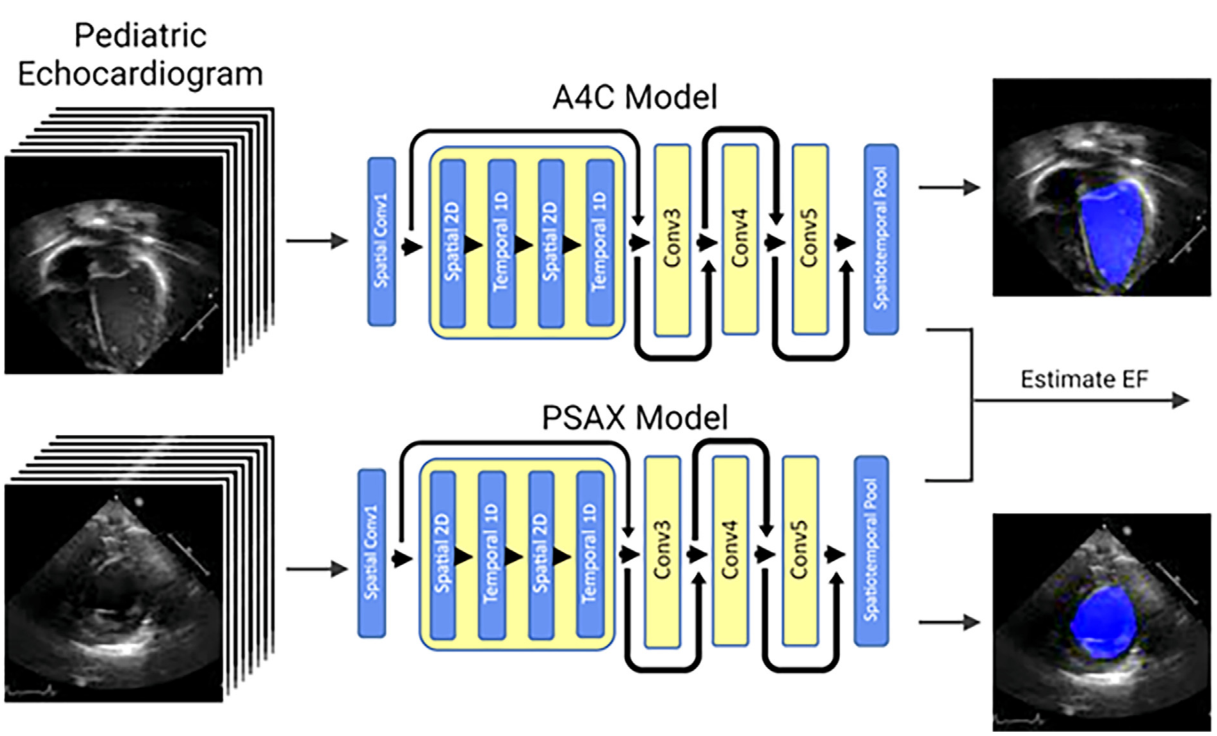

Figure 1 EchoNet-Peds uses standard A4C and PSAX echocardiogram video clips as input data. EchoNet-Peds generates segmentations *(blue)* of the LV through each frame of the echocardiogram video clip, bounded by the single-frame tracings done at end systole and end diastole. EchoNet-Peds then generates an EF by aggregating information across both A4C and PSAX views.

Table 1 Demographics of patients and visits from Lucile Packard Children's Hospital at Stanford

Statistic	
Patients	1,958
Gender, female	830 (42)
Age, years	10.2 ± 5.4
Total visits	4,467
Visits with A4C	3,176 (71)
Visits with PSAX	4,424 (99)
Visits with both	3,133 (70)
Normal EF (>55%)	3,854 (86)
EF (%)	61 ± 10

Data are presented as mean ±SD or n (%). Patients may have multiple visits included in the data set.

labeled images. Of the 7,643 video clips, 6,114 (80%) were used for training, 765 (10%) for testing, and 764 (10%) for validation (Figure 2). Eighty-six percent of studies had an EF ≥55%.

Evaluation of Model Performance

EchoNet-Peds estimated EF with an MAE of 3.66% (3.53%-3.82%), RMSE of 4.98% (4.75%-5.33%), and R^2 of 0.77 (0.74-0.80) compared to human experts when presented with input from both A4C and PSAX videos (Figure 3A). As expected, the model performs best when both views are used. However, the model also estimated an accurate EF compared with the 5/6 area-length EF when presented with only 1 view. When EchoNet-Peds is presented with only an A4C video clip, the MAE is 4.15% (4.02%-4.28%), RMSE is 5.70% (5.41%-6.00%), and R^2 is 0.70 (0.66-0.74). If presented with only a PSAX video clip, the MAE is 3.80% (3.69%-3.91%), RMSE is 5.14% (4.95%-5.33%), and R^2 is 0.74 (0.71-0.77).

A secondary analysis separately evaluated the model's ability to differentiate systolic dysfunction from normal function. The combined estimation achieves an area under the curve receiver-operating characteristic of 0.954 (0.942-0.965) for diagnosing the echocardiograms from the data set with reduced EF (<55%; Figure 3B) and increases further to 0.980 (0.970-0.988) for EFs of <50%. EchoNet-Peds also segmented the LV with Dice similarity coefficients of 0.891 (95% CI, 0.889-0.894) and 0.896 (95% CI, 0.894-0.898) for A4C and PSAX views, respectively.

EchoNet-Peds can estimate EF and segment a video clip in 0.062 seconds using one NVIDIA GeForce GTX 1080 Ti GPU (NVIDIA).

Consistent Performance Across Population

EchoNet-Peds performed well in estimating EF (Table 2) and segmenting the LV (Table 3) among all patient groups stratified by sex, age, and BSA. EchoNet-Peds's ability to accurately estimate EF was particularly high for patients <1 year old or with a low BSA.

Comparison of Training Methods

EchoNet-Peds significantly outperformed EchoNet-Dynamic[13] in estimating EF ($P <$ 1e-100). Comparisons were made using only the A4C view as EchoNet-Dynamic was not trained with PSAX images (Supplemental Table 1). Additionally, EchoNet-Peds, which is fine-tuned from the weights of EchoNet-Dynamic,[13] sees small improvements over training from Kinetics-400[39] weights ($P = .09$), suggesting that our pediatric data set is large enough for stand-alone training. Left ventricular segmentation (Supplemental Table 2), when compared to the Common Objects in Context (COCO) dataset,[40] shows similar trends, although the difference in performance is smaller.

DISCUSSION

EchoNet-Peds is the first artificial intelligence model trained entirely on a large pediatric data set to assess ventricular function. EchoNet-Peds performed well with the identification of patients with decreased systolic function, particularly for patients with an EF <50%. More clinically significant is the fact that estimations generated by EchoNet-Peds have the same or less variability compared with human expert interobserver variability,[20,28,30] which has been reported in large national data sets to be 4% to 14%.[20,21] Patient subgroups stratified by age and weight had an MAE < 5%, indicating precise estimation of EF, suggesting that EchoNet-Peds could be used to identify subtle changes in EF on serial examinations. Moreover, EchoNet-Peds was trained against a gold standard of the reported 5/6 area-length EF for each echocardiogram, which required both the A4C and PSAX views. However, once trained, EchoNet-Peds generated EF with

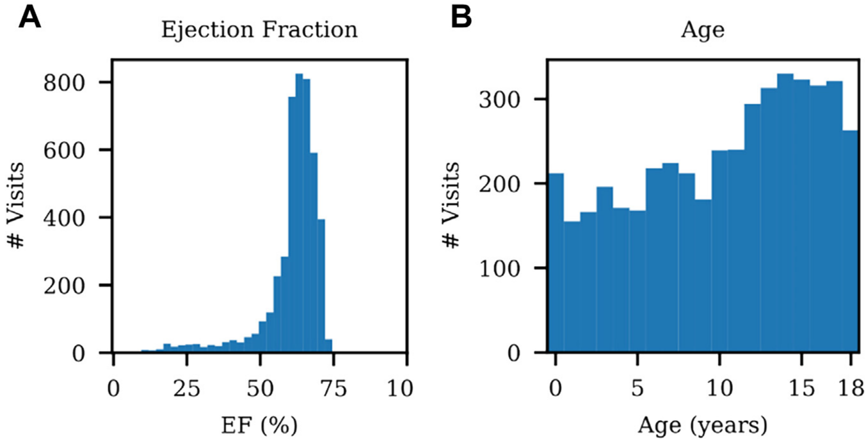

Figure 2 **(A)** Distribution of EFs measured at each visit. **(B)** Distribution of age at each visit.

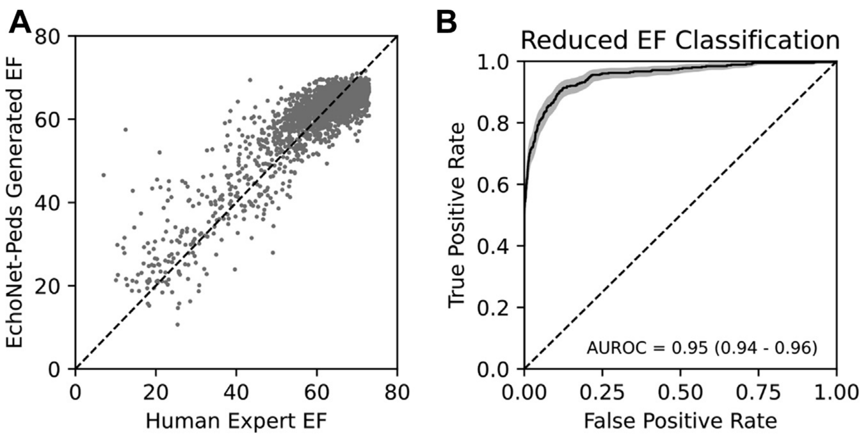

Figure 3 **(A)** The estimated EF vs the human-measured EF. **(B)** Receiver-operating characteristic curve for classifying reduced EF (<55%).

reasonable accuracy if only 1 view was available, allowing for its utility when echocardiographic windows are limited or 1 view is suboptimal.

The segmentation task does not directly provide the values required to calculate EF. Instead, EchoNet-Peds is conceptually closer to an "expert qualitative assessment," but with improved accuracy and decreased variability since it is trained against objective measures. The segmentation task remains important as it provides a visual representation of EchoNet-Peds and demonstrates how EF estimations are generated. The segmentation task generates the equivalent of a manual tracing for each frame between end systole and end diastole; for echocardiographic clips with multiple beats, it can average the EF values for each beat. Our echocardiographic clips had variable numbers of cardiac cycles, often depending on patient heart rate and clip length, but EchoNet-Peds performed consistently across all environments. The success of the segmentation is defined by the high Dice similarity coefficient, a measure of spatial overlap and reproducibility broadly used in artificial intelligence segmentation tasks across imaging modalities.[38,41] Similarly, we purposefully used a large data set that was representative of the clinical variability typically seen in a pediatric echocardiography laboratory, with different ultrasound machines and image acquisition settings and variable image quality. EchoNet-Peds successfully segmented the LV and generated accurate EF measurements despite this variability. In addition, EchoNet-Peds can complete both tasks in 0.06 seconds per video clip, which is more efficient than human manual assessment. EchoNet-Peds is accurate, reproducible, and rapid, which lays the foundation for an end-to-end clinical aid that could be inserted directly into the existing workflow.

To date, most machine learning tools used in pediatric echocardiography have been trained on adult echocardiographic data.[42] This is primarily related to the relative paucity of data available for pediatric acquired and congenital heart diseases needed to create sufficient training data sets. Additional limitations include the wide variety of sizes and ages in pediatric patients, patient movement during

Table 2 Stratified performance for generated EF by sex, age, and BSA

Group	No. of echocardiograms	EF, %	R^2	MAE, %
All echocardiograms	4,467	61 ± 10	0.77 (0.74-0.80)	3.66 (3.56-3.80)
Gender, male	2,557	62 ± 9	0.74 (0.69-0.78)	3.65 (3.50-3.89)
Gender, female	1,898	60 ± 11	0.79 (0.74-0.84)	3.67 (3.45-3.95)
Age:				
<1	211	54 ± 18	0.85 (0.79-0.90)	4.99 (4.37-5.97)
1-5	843	62 ± 10	0.79 (0.70-0.85)	3.42 (3.18-3.70)
6-11	1,291	62 ± 9	0.76 (0.69-0.83)	3.42 (3.21-3.61)
12-14	920	61 ± 10	0.72 (0.62-0.80)	3.79 (3.51-4.31)
>15	1,202	61 ± 9	0.71 (0.64-0.76)	3.75 (3.55-4.10)
BSA:				
Quartile 1	1,111	60 ± 12	0.84 (0.78-0.88)	3.70 (3.37-3.96)
Quartile 2	1,111	62 ± 9	0.76 (0.67-0.82)	3.46 (3.22-3.72)
Quartile 3	1,111	62 ± 9	0.72 (0.66-0.78)	3.68 (3.42-4.10)
Quartile 4	1,112	61 ± 9	0.68 (0.55-0.75)	3.77 (3.54-4.09)

Patients with missing demographic information are omitted from the stratified analysis.

Table 3 Stratified performance for segmenting the LV by sex, age, and BSA

	A4C		PSAX	
Group	No. of echocardiograms	Dice coefficient	No. of echocardiograms	Dice coefficient
All echocardiograms	3,176	0.891 (0.889-0.894)	4,424	0.896 (0.894-0.898)
Gender, male	1,813	0.891 (0.888-0.894)	2,533	0.895 (0.891-0.898)
Gender, female	1,351	0.892 (0.889-0.895)	1,880	0.896 (0.893-0.900)
Age:				
<1	169	0.860 (0.848-0.878)	206	0.861 (0.848-0.875)
1-5	631	0.894 (0.891-0.897)	834	0.893 (0.888-0.899)
6-11	890	0.899 (0.897-0.902)	1,283	0.902 (0.898-0.907)
12-14	639	0.892 (0.888-0.896)	910	0.898 (0.893-0.902)
>15	847	0.886 (0.881-0.890)	1,191	0.894 (0.890-0.899)
BSA:				
Quartile 1	842	0.887 (0.881-0.893)	1,096	0.888 (0.882-0.892)
Quartile 2	770	0.899 (0.895-0.901)	1,104	0.901 (0.897-0.905)
Quartile 3	752	0.894 (0.891-0.897)	1,103	0.903 (0.900-0.907)
Quartile 4	797	0.885 (0.880-0.890)	1,099	0.891 (0.885-0.895)

Patients with missing demographic information are omitted from the stratified analysis.

evaluation, and image artifact.[43] Recent focus has been placed on alternative methods, such as iterative segmentation, to augment limited pediatric data.[44] However, our study clearly illustrates the benefits of training a model on pediatric data; EchoNet-Peds performed significantly better among all age groups and sizes than a corresponding model trained on adult data. There was better EF estimation by R^2 in patients <1 year of age compared with the other ages. Segmentation performance also improved, although less obviously, suggesting the model is more robust in adapting to the differences of pediatric and adult populations. The best segmentation performance occurred in patients age 6 to 11 years old, potentially correlating to the ages during which improved patient cooperation and optimal acoustic windows coincide. Previous literature has also demonstrated that small changes in segmentation can result in exponential errors in EF estimation.[45,46] In a Pediatric Heart Network study by Frommelt *et al.*,[28] there was good reproducibility between measurements of LV volume, but significant variability in LV function calculations, suggesting that clinical practice may adjust measurements to correlate to qualitative assessment of function. Therefore, we hypothesize that the adult model performs less well overall on the pediatric data due to variability in patient size, weight, heart rates, and movement, despite relative accuracy with segmentation. It is important to note, however, that EchoNet-Peds did benefit from "fine-tuning" by utilizing weights derived from EchoNet-Dynamic. EchoNet-Peds is an important extension of the work done with EchoNet-Dynamic, but the use of a wholly pediatric data set and the introduction of a novel echocardiographic view (PSAX) illustrate the benefits of a pediatric-centric approach, rather than forcibly applying adult models to pediatric questions. The strategy we used in this study may serve as a model for adapting adult-trained algorithms more accurately to pediatric data while retaining the benefits of larger adult data sets.

We worked with a team at Stanford Medicine to make our deidentified data set of 4,467 pediatric echocardiograms publicly available to the research community at large. At the time of writing, this is the largest known labeled video data set of pediatric echocardiograms that include human expert tracings. We aim to make this data set a resource for clinicians and researchers interested in validating pediatric-specific algorithms, and we hope to encourage further growth in our field.

Limitations

In this study, we used pediatric echocardiographic data derived from A4C and PSAX views to estimate EF against a gold standard of the 5/6 area-length method. While large studies have shown that the 5/6 area-length method is more accurate in pediatric patients,[20,21] we recognize that many clinicians still use other methods to estimate EF. We also limited our algorithm to two-dimensional imaging only. Other studies have used three-dimensional echocardiography to develop machine learning algorithms for estimating EF because it correlates better with EF derived from cardiac magnetic resonance imaging than from two-dimensional echocardiography. Because three-dimensional echocardiography is still not routinely used in most clinical practice settings due to the need for special training, we opted to focus our model on a solution with the broadest possible impact. EchoNet-Peds utilizes supervised learning for the initial segmentation task with human-labeled images, so there is inherent introduction of human bias. Lastly, we validated EchoNet-Peds using data obtained at only 1 institution, mostly because of the challenges in accessing data from other hospitals. We tried to mitigate this limitation by showing that EchoNet-Peds works well across a broad range of study quality and ultrasound machines. Zuercher *et al.*[47] demonstrated a similar MAE on a single view (A4C) and a smaller subset of approximately 300 echocardiograms at a different institution, suggesting the generalizability of the algorithm to broader populations. However, widespread external validation with particular focus on adequate representation of various demographics and diagnoses is an important direction for future work.

Future Directions

While machine learning has been used widely in adult cardiology over the past few years, its use in pediatrics has lagged. Left ventricular functional assessment has been an important area of study, particularly in point-of-care ultrasounds performed by noncardiologists or in rural areas without access to experts.[48-50] EchoNet-Peds has an end-to-end workflow that can easily be implemented at the bedside to make accurate assessments of cardiac function and guide clinical decisions rapidly. Moreover, Dykes *et al.*[51] showed that parent-driven acquisition of LV A4C and PSAX video clips on handheld ultrasounds is feasible, and EchoNet-Peds has the potential to improve this clinical workflow. Since EchoNet-Peds performs very well in estimating EF even when only 1 view of the LV is obtained, this could increase utility, decrease parental burden, and provide more data. Since our data were derived from expert-obtained ultrasounds, further testing will need to be performed to assess generalizability in these limited settings. We additionally hope to expand to patients with congenital heart disease and explore validity in patients with varied LV or single ventricular geometry. Lastly, we hope the difference in performance between the adult and pediatric models convinces the pediatric community of the importance of creating large-scale data sets to better explain these differences and minimize bias. Our hope is that the release of the EchoNet-Peds data set and code for the algorithm will allow for broad exposure, implementation, and improvement to have the highest clinical benefit for patients.

CONCLUSION

EchoNet-Peds is an automated system for evaluating LV function in pediatric patients. EchoNet-Peds achieves good performance in estimating EF and segmenting the LV across all subgroups of pediatric patients. EchoNet-Peds accurately segments the LV and estimates LVEF when compared to a human expert, even when only a single view is available, which is particularly beneficial when appropriate image acquisition is limited by patient movement or inexperienced imagers. It outperforms models trained on adult echocardiograms, highlighting the benefit of training pediatric algorithms with dedicated pediatric data. EchoNet-Peds can be used to reduce the time and manual labor needed to measure LV function, potentially reduce errors from human variability, and improve access to care in areas without local experts.

CODE AND DATA SET AVAILABILITY

We have made source code and data set available at https://echonet.github.io/pediatric.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online at https://doi.org/10.1016/j.echo.2023.01.015.

REFERENCES

1. Lopez L, Colan SD, Frommelt PC, et al. Recommendations for quantification methods during the performance of a pediatric echocardiogram: a report from the pediatric measurements writing group of the American Society of echocardiography pediatric and congenital heart disease council. J Am Soc Echocardiogr 2010;23:465-95; quiz 576-467.
2. Lai WW, Geva T, Shirali GS, et al. Guidelines and standards for performance of a pediatric echocardiogram: a report from the task force of the pediatric council of the American Society of echocardiography. J Am Soc Echocardiogr 2006;19:1413-30.
3. Van der Hauwaert LG, Denef B, Dumoulin M. Long-term echocardiographic assessment of dilated cardiomyopathy in children. Am J Cardiol 1983;52:1066-71.
4. Leerink JM, van der Pal HJH, Kremer LCM, et al. Refining the 10-year prediction of left ventricular systolic dysfunction in long-term Survivors of childhood cancer. JACC CardioOncol 2021;3:62-72.
5. Shah MJ, Silka MJ, Silva JNA, et al. 2021 PACES expert consensus statement on the indications and management of cardiovascular implantable electronic devices in pediatric patients. Cardiol Young 2021;31:1738-69.
6. Watanabe K, Shih R. Update of pediatric heart failure. Pediatr Clin North Am 2020;67:889-901.
7. Buddhe S, Du W, Walters HL 3rd, et al. Predictors of left ventricular remodeling after aortic valve replacement in pediatric patients with isolated aortic regurgitation. Congenit Heart Dis 2013;8:167-73.
8. Djordjevic SA, Milic-Rasic V, Brankovic V, et al. Cardiac findings in pediatric patients with spinal muscular atrophy types 2 and 3. Muscle Nerve 2021;63:75-83.
9. Madani A, Arnaout R, Mofrad M, et al. Fast and accurate view classification of echocardiograms using deep learning. NPJ Digit Med 2018;1:6.
10. Madani A, Ong JR, Tibrewal A, et al. Deep echocardiography: data-efficient supervised and semi-supervised deep learning towards automated diagnosis of cardiac disease. NPJ Digit Med 2018;1:59.
11. Zhang J, Gajjala S, Agrawal P, et al. Fully automated echocardiogram interpretation in clinical practice. Circulation 2018;138:1623-35.
12. Ghorbani A, Ouyang D, Abid A, et al. Deep learning interpretation of echocardiograms. NPJ Digit Med 2020;3:10.
13. Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of cardiac function. Nature 2020;580:252-6.
14. Karimi-Bidhendi S, Arafati A, Cheng AL, et al. Fullyautomated deeplearning segmentation of pediatric cardiovascular magnetic resonance of patients with complex congenital heart diseases. J Cardiovasc Magn Reson 2020;22:80.
15. Yoshida ALY, Yoshimura N, kuramoto T, et al. Automated heart segmentation using U-Net in pediatric cardiac CT. Measurement: Sensors 2021;18:100-27.
16. Power A, Poonja S, Disler D, et al. Echocardiographic image quality deteriorates with age in children and young adults with duchenne muscular dystrophy. Front Cardiovasc Med 2017;4:82.
17. De Geer L, Oscarsson A, Engvall J. Variability in echocardiographic measurements of left ventricular function in septic shock patients. Cardiovasc Ultrasound 2015;13:19.
18. Wood PW, Choy JB, Nanda NC, et al. Left ventricular ejection fraction and volumes: it depends on the imaging method. Echocardiography 2014;31:87-100.
19. Margossian R, Chen S, Sleeper LA, et al. The reproducibility and absolute values of echocardiographic measurements of left ventricular size and function in children are algorithm dependent. J Am Soc Echocardiogr 2015;28:549-58.e1.
20. Selamet Tierney ES, Hollenbeck-Pringle D, Lee CK, et al. Reproducibility of left ventricular dimension versus area versus volume measurements in pediatric patients with dilated cardiomyopathy. Circ Cardiovasc Imaging 2017;10:e006007.
21. Colan SD, Shirali G, Margossian R, et al. The ventricular volume variability study of the Pediatric Heart Network: study design and impact of beat averaging and variable type on the reproducibility of echocardiographic measurements in children with chronic dilated cardiomyopathy. J Am Soc Echocardiogr 2012;25:842-54.e6.
22. Thompson AJ, O'Leary PW, Miller A, et al. Inter-observer and inter-vendor variability in strain measurements in patients with single right ventricular anatomy. Pediatr Cardiol 2021;42:1341-9.
23. Lee CK, Margossian R, Sleeper LA, et al. Variability of M-mode versus two-dimensional echocardiography measurements in children with dilated cardiomyopathy. Pediatr Cardiol 2014;35:658-67.
24. Norgard G, Johannessen KA. Variability of digitized left ventricular M-mode echocardiography: a study in healthy subjects and patients with repaired tetralogy of fallot. Clin Physiol 1993;13:373-83.

25. Lipshultz SE, Easley KA, Orav EJ, et al. Reliability of multicenter pediatric echocardiographic measurements of left ventricular structure and function: the prospective P(2)C(2) HIV study. Circulation 2001;104:310-6.
26. Lai WW. Echocardiography in pediatric and congenital heart disease from fetus to adult. Chichester, West Sussex, Hoboken, NJ: John Wiley and Sons Ltd; 2015.
27. Lopez L, Colan S, Stylianou M, et al. Relationship of echocardiographic Z scores adjusted for body surface area to age, sex, race, and ethnicity: the pediatric heart network normal echocardiogram database. Circ Cardiovasc Imaging 2017;10:e006979.
28. Frommelt PC, Minich LL, Trachtenberg FL, et al. Challenges with left ventricular functional parameters: the pediatric heart Network normal echocardiogram database. J Am Soc Echocardiogr 2019;32:1331-8.e1.
29. Nielsen JC, Lytrivi ID, Ko HH, et al. The accuracy of echocardiographic assessment of left ventricular size in children by the 5/6 area x length (bullet) method. Echocardiography 2010;27:691-5.
30. 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;44:1030-5.
31. Pellikka PA, She L, Holly TA, et al. Variability in ejection fraction measured by echocardiography, gated single-photon emission computed tomography, and cardiac magnetic resonance in patients with coronary artery disease and left ventricular dysfunction. JAMA Netw Open 2018;1:e181456.
32. 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.
33. Cantinotti M, Giordano R, Scalese M, et al. Nomograms for two-dimensional echocardiography derived valvular and arterial dimensions in Caucasian children. J Cardiol 2017;69:208-15.
34. Pettersen MD, Du W, Skeens ME, et al. Regression equations for calculation of z scores of cardiac structures in a large cohort of healthy infants, children, and adolescents: an echocardiographic study. J Am Soc Echocardiogr 2008;21:922-34.
35. Levy F, Dan Schouver E, Iacuzio L, et al. Performance of new automated transthoracic three-dimensional echocardiographic software for left ventricular volumes and function assessment in routine clinical practice: comparison with 3 Tesla cardiac magnetic resonance. Arch Cardiovasc Dis 2017;110:580-9.
36. Knackstedt C, Bekkers SC, Schummers G, et al. Fully automated versus standard tracking of left ventricular ejection fraction and longitudinal Strain: the FAST-EFs multicenter study. J Am Coll Cardiol 2015;66:1456-66.
37. Tran D, Bourdev L, Fergus R, et al. Learning Spatiotemporal Features with 3D Convolutional Networks. Paper presented at: 20015 IEEE International Conference on Computer Vision (ICCV); 2015. Santiago, Chile.
38. Yao AD, Cheng DL, Pan I, et al. Deep learning in neuroradiology: a systematic review of current algorithms and approaches for the new wave of imaging technology. Radiol Artif Intell 2020;2:e190026.
39. Kay W, Carreira J, Simonyan K, et al. The kinetics human action dataset. arXiv. https://doi.org/10.48550/arXiv.1705.06950. Accessed March 6, 2023.
40. Lin T-Y, Maire M, Belongie S, et al. Microsoft COCO: common objets in contextcomputer vision. ECCV. 2014. pp. 740-55. https://doi.org/10.48550/arXiv.1405.0312. Accessed March 6, 2023.
41. Ono S, Komatsu M, Sakai A, et al. Automated endocardial border detection and left ventricular functional assessment in echocardiography using deep learning. Biomedicines 2022;10:1082.
42. Tandon A, Mohan N, Jensen C, et al. Retraining convolutional neural networks for specialized cardiovascular imaging tasks: lessons from tetralogy of fallot. Pediatr Cardiol 2021;42:578-89.
43. Herz C, Pace DF, Nam HH, et al. Segmentation of tricuspid valve leaflets from transthoracic 3D echocardiograms of children with hypoplastic left heart Syndrome using deep learning. Front Cardiovasc Med 2021;8:735587.
44. Pace DF, Dalca AV, Brosch T, et al. Iterative segmentation from limited training data: applications to congenital heart disease. Deep Learn Med Image Anal Multimodal Learn Clin Decis Support (2018 2018;11045:334-42.
45. O'Dell WG. Accuracy of left ventricular cavity volume and ejection fraction for conventional estimation methods and 3D surface fitting. J Am Heart Assoc 2019;8:e009124.
46. Bloomgarden DC, Fayad ZA, Ferrari VA, et al. Global cardiac function using fast breath-hold MRI: validation of new acquisition and analysis techniques. Magn Reson Med 1997;37:683-92.
47. Zuercher M, Ufkes S, Erdman L, et al. Retraining an artificial intelligence algorithm to calculate left ventricular ejection fraction in pediatrics. J Cardiothorac Vasc Anesth 2022;36:3610-6.
48. De Marchi L, Meineri M. POCUS in perioperative medicine: a North American perspective. Crit Ultrasound J 2017;9:19.
49. Singh Y, Tissot C, Fraga MV, et al. International evidence-based guidelines on point of care ultrasound (POCUS) for critically ill neonates and children issued by the POCUS working group of the European Society of paediatric and neonatal intensive care (ESPNIC). Crit Care 2020;24:65.
50. Blanco P, Volpicelli G. Common pitfalls in point-of-care ultrasound: a practical guide for emergency and critical care physicians. Crit Ultrasound J 2016;8:15.
51. Dykes JC, Kipps AK, Chen A, et al. Parental acquisition of echocardiographic images in pediatric heart transplant patients using a handheld device: a pilot telehealth study. J Am Soc Echocardiogr 2019;32:404-11.