The Journal of Emergency Medicine, Vol. ■, No. ■, pp. 1–8, 2023

0736-4679/$ - see front matter

ELSEVIER

https://doi.org/10.1016/j.jemermed.2023.02.005

Ultrasound in Emergency Medicine

AI-ENABLED ASSESSMENT OF CARDIAC FUNCTION AND VIDEO QUALITY IN EMERGENCY DEPARTMENT POINT-OF-CARE ECHOCARDIOGRAMS

Bryan He,[*,1] Dev Dash, MD,[†,1] Youyou Duanmu, MD,[†] Ting Xu Tan,[†] David Ouyang, MD,[‡,2] and James Zou, PHD[§,2]

*Department of Computer Science, Stanford University, Stanford, California, †Department of Emergency Medicine, Stanford University, Stanford, California, ‡Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, California, and §Department of Biomedical Data Science, Stanford University, Stanford, California

Reprint Address: Bryan He, Department of Computer Science, Stanford University, Stanford, CA

□ Abstract—Background: The adoption of point-of-care ultrasound (POCUS) has greatly improved the ability to rapidly evaluate unstable emergency department (ED) patients at the bedside. One major use of POCUS is to obtain echocardiograms to assess cardiac function. Objectives: We developed EchoNet-POCUS, a novel deep learning system, to aid emergency physicians (EPs) in interpreting POCUS echocardiograms and to reduce operator-to-operator variability. Methods: We collected a new dataset of POCUS echocardiogram videos obtained in the ED by EPs and annotated the cardiac function and quality of each video. Using this dataset, we train EchoNet-POCUS to evaluate both cardiac function and video quality in POCUS echocardiograms. Results: EchoNet-POCUS achieves an area under the receiver operating characteristic curve (AUROC) of 0.92 (0.89–0.94) for predicting whether cardiac function is abnormal and an AUROC of 0.81 (0.78–0.85) for predicting video quality. Conclusions: EchoNet-POCUS can be applied to bedside echocardiogram videos in real time using commodity hardware, as we demonstrate in a prospective pilot study. © 2023 Elsevier Inc. All rights reserved.

□ Keywords—point-of-care ultrasound; cardiac function; machine learning

[1]These authors contributed equally to this work.
[2]Co-senior authors.

INTRODUCTION

Bedside point-of-care ultrasound (POCUS) can greatly improve emergency medicine through its cost-effectiveness, prompt availability to augment diagnostic workups, and the ability to evaluate unstable patients at the bedside (1–3). Diagnostically, POCUS can be used in trauma for rapid assessment of hemorrhage, evaluation of gallbladder and kidney disease, cardiac function through echocardiographic techniques, evaluation of pregnancy and pelvic pathology, deep venous thrombosis, among other emerging modalities (4,5). Procedurally, it is critical for the placement of central lines, nerve blocks for regional anesthesia, and needle guidance for paracentesis and thoracentesis (6).

In recent years, formal ultrasound training has been incorporated into the curriculum of many medical schools and residency programs, including emergency medicine and internal medicine programs (7,8). Additionally, ultrasound fellowship training programs have been developed to meet this increasing need for bedside ultrasound, as well as to develop centers of excellence for training and research. At the same time, the cost of ultrasound hardware has decreased such that many emergency departments (EDs) and intensive care units have POCUS machines. The utilization of ultrasound can streamline pa-

RECEIVED: 20 July 2022; FINAL SUBMISSION RECEIVED: 11 January 2023;
ACCEPTED: 17 February 2023

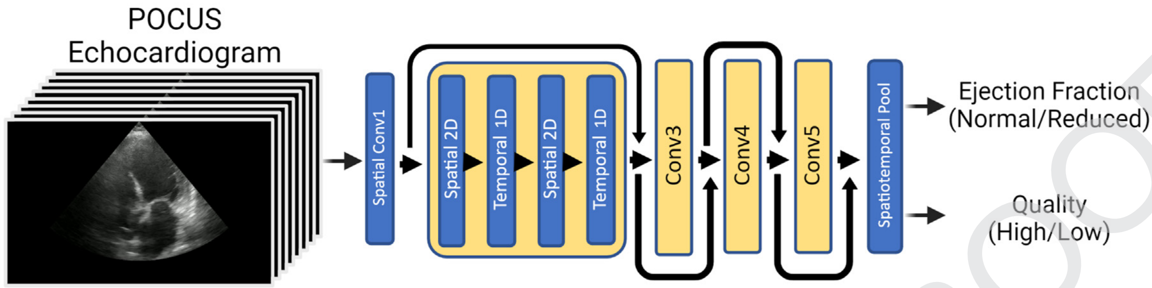

Figure 1. Overview of EchoNet-POCUS. Our system takes a raw apical four-chamber view echocardiogram video as input to a neural network and predicts the ejection fraction and video quality as output. POCUS = point-of-care ultrasound.

tient care with its ability to provide diagnostic data in real time at bedside and help avoid sending unstable patients to the radiology department or echocardiography lab for testing.

One major use of POCUS is echocardiography to evaluate cardiac function, including the assessment of cardiac contractility for the diagnosis of life-threatening pathology such as acute heart failure (9). Cardiovascular emergencies are urgent and require prompt intervention, and high-quality real-time bedside echocardiographic information can aid treatment in EDs and private offices (10). Emergency physicians (EPs) can estimate cardiac contractility using POCUS, but the bedside evaluation of ejection fraction has significant variation and is operator dependent, so streamlining and standardizing this assessment would be useful in both the clinical and the education settings (11). Additionally, other works seeking to aid in the interpretation of echocardiograms have focused on formal echocardiogram studies, which are expected to have less variation than POCUS echocardiograms (12–15).

To overcome these challenges, we present EchoNet-POCUS, a deep learning model to provide clinicians with automated interpretation of cardiac function and video quality from POCUS echocardiogram videos (Figure 1). We show that the predictions from EchoNet-POCUS can help identify patients with low ejection fractions and videos of poor quality to streamline the workflow of EPs. We additionally demonstrate that EchoNet-POCUS can be applied at bedside with commodity hardware in a prospective pilot study.

METHODS

OVERVIEW OF ECHONET-POCUS

EchoNet-POCUS is a deep learning-based model that takes raw POCUS echocardiogram videos as input and predicts whether the ejection fraction is normal and whether the video quality is sufficient to accurately interpret. EchoNet-POCUS uses a convolutional neural network with spatiotemporal convolutions as its backbone (16).

We collected 927 POCUS echocardiograms in the apical four-chamber view from 573 patients who underwent treatment at our ED. Each video was visually assessed by a consensus of three EPs for normal or low ejection fraction, which is a key metric for cardiac function. The videos were also assessed for quality and were marked as sufficient if the left ventricle was visible and the video could be accurately assessed.

DATA COLLECTION

For training and evaluating EchoNet-POCUS, we collected all POCUS echocardiograms in the apical four-chamber view recorded in our ED between January and November 2020, including 927 videos from 574 patients. The videos were recorded on machines from three different manufacturers (496 on Mindray [Mindray North America, Rahway, NJ]; 339 on GE Venue [GE HealthCare, Chicago, IL], and 92 on Philips [Philips N.V., Amsterdam, The Netherlands]). This research was approved by our Institutional Review Board.

Two physicians separately labeled each video with its cardiac function (normal/reduced) and its quality (sufficient/insufficient) (Table 1). The physicians visually assessed the cardiac function of each video based on its left ventricular ejection fraction, the ratio of change in the end-systolic and end-diastolic volumes, which is a key metric of function (17,18). A video was labeled as insufficient quality if the physicians were unable to confidently evaluate the cardiac function based on the video. Videos with disagreements between the original two annotations were adjudicated by a third physician, who had undergone fellowship training in POCUS.

MODEL TRAINING AND EVALUATION

Model design and training was done in Python using the PyTorch deep learning library (19). EchoNet-POCUS

Table 1. Summary Statistics of Patients and Videos for the Retrospective Dataset*

Statistic	Total
Patients	573
Videos	927
Cardiac function	
Normal	775
Reduced	152
Quality	
Sufficient	768
Insufficient	159

* Multiple videos can be collected for a single patient.

uses a R2+1D model using pretrained weights from Ouyang et al. (14,16).

For predicting both ejection fraction and video quality, the model was trained to minimize the cross-entropy loss between the prediction and ground-truth annotation using a stochastic gradient descent optimizer with an initial learning rate of 0.0001, momentum of 0.9, and batch size of 16 for 45 epochs. The learning rate was decayed by a factor of 0.1 every 15 epochs. Video clips of 32 frames generated by sampling every other frame (sampling period of 2) were used as the model input.

To evaluate the model, we used 10-fold cross validation, where 80% of the patients were used for training, 10% were used for validation, and 10% were used for testing. This process was repeated 10 times for different samples of patients. For each fold, the weights from the epoch with the lowest validation loss were selected for final testing.

PROSPECTIVE FEASIBILITY PILOT IN THE ED

Patients for the bedside usage of ED-POCUS were scanned between September and November 2021. Patients whose condition was stable and who were not actively undergoing a clinical procedure were considered and videos were passively collected, some, in part, for POCUS training. All patients were scanned using a GE Venue ultrasound machine.

RESULTS

OVERVIEW OF PERFORMANCE

We used EchoNet-POCUS to predict the ejection fraction and quality of the videos in our dataset, and we evaluated its performance using 10-fold cross-validation. For predicting ejection fraction on videos labeled as sufficient quality by the physicians, EchoNet-POCUS achieved an area under the receiver operating characteristic curve (AUROC) of 0.92 (0.89–0.94), and for predicting video quality, EchoNet-POCUS achieved an AUROC of 0.81 (0.78–0.85) (Figure 2). Additionally, we evaluate the calibration of the predictions made by EchoNet-POCUS by binning the videos by the probability predicted by EchoNet-POCUS. We find that the predictions are reasonably calibrated—that is, the probability estimated by EchoNet-POCUS is close to the real probability—for both cardiac function and quality. This means that EchoNet-POCUS accurately estimates its uncertainty. EchoNet-POCUS was able to predict ejection fraction across all three ultrasound machines in our dataset, with test AUROCs of 0.93 (0.88–0.94), 0.88 (0.82–0.93), and 0.85 (0.69–0.97) on Mindray, GE Venue, and Philips machines, respectively.

USAGE FOR TRIAGE AND CONSENSUS

One potential use for EchoNet-POCUS is to reduce the labor needed for physicians to identify patients with reduced ejection fractions. EchoNet-POCUS can be used to identify patients that are likely to have a reduced ejection fraction, which can be confirmed by a physician. In this use case, EchoNet-POCUS can identify 90% of patients with reduced ejection fractions, while requiring a physician to read only 35% of videos (Table 2). In challenging and more ambiguous videos where the initial two readers disagreed, EchoNet-POCUS was able to serve as an additional reader, agreeing with the consensus-providing EP in 80% of videos. EchoNet-POCUS is calibrated to be the most consistent with a consensus read, suggesting that the model can serve as an additional reader or diagnostic/teaching aid.

STRATIFICATION BY VIDEO QUALITY

We use EchoNet-POCUS to group videos by quality and find that EchoNet-POCUS can predict the ejection fraction of videos in the top quartile of quality with an AUROC of 0.94 (0.91–0.97) (Figure 3). In contrast, the EchoNet-POCUS can predict the ejection fraction of videos in the bottom quartile of quality with an AUROC of 0.64 (0.61–0.67). This suggests that the video quality can be used as feedback to prompt physicians to retake videos with poor quality, resulting in videos that can be evaluated more easily.

COMPARISON OF TRAINING METHODS

We additionally compared against several other training methods in Table 3 (14). The weights from EchoNet-POCUS were trained starting with the weights from and

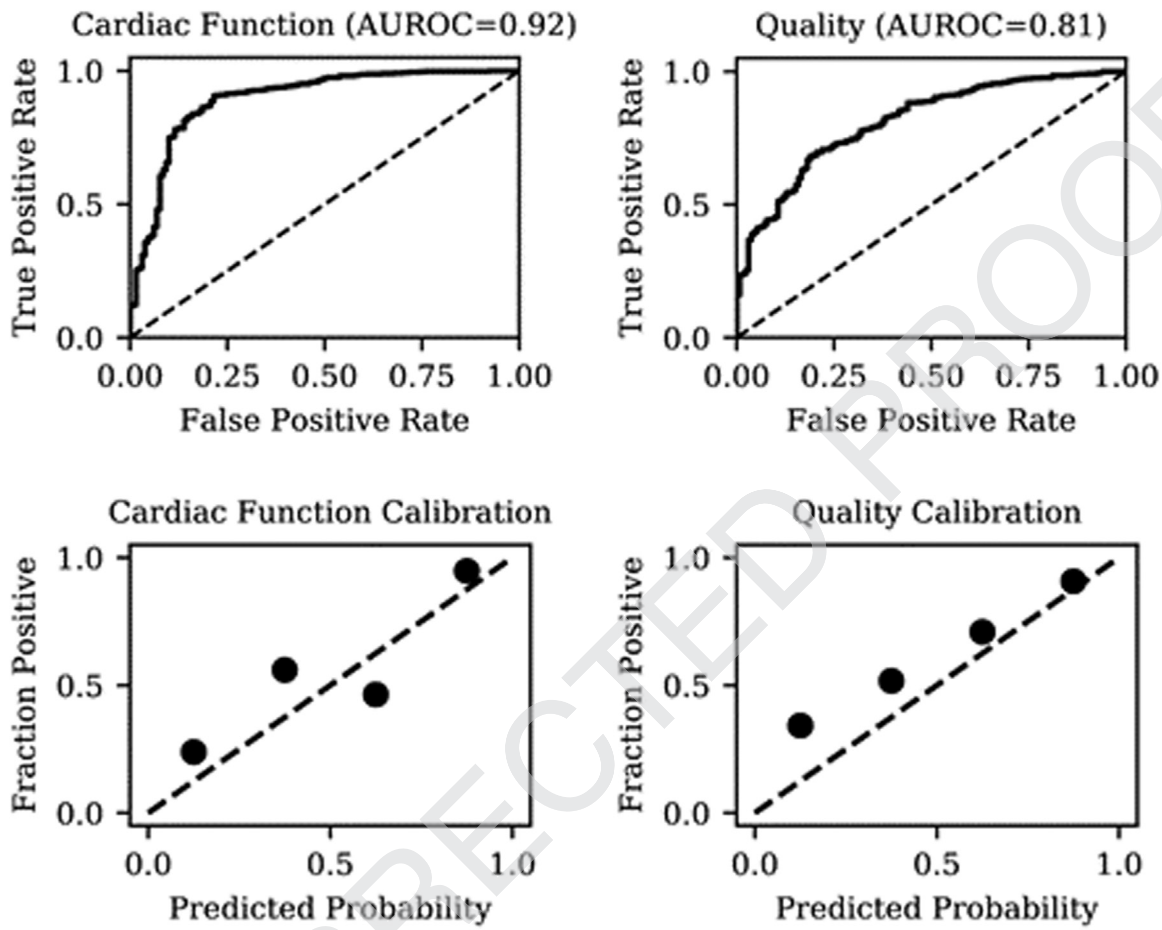

Figure 2. **Receiver operating characteristic (top) and calibration analysis (bottom) for predicting cardiac function (left) and video quality (right). AUROC = area under the receiver operating characteristic curve.**

Table 2. Confusion Matrix for Predicting Ejection Fractions on Videos With Sufficient Quality*

		Human Labeled EF	
		Normal	Low
Predicted EF	Normal	490	13
	Low	138	127

* Threshold selected for a sensitivity of 90%.EF = ejection fraction.

fine-tuning all weights, which resulted in the best performance. Using the weights from Ouyang et al. directly without retraining resulted in significantly poorer performance on predicting cardiac function and no ability to predict video quality (14). Fine-tuning only the final layer did not result in significant improvements. Finally, training EchoNet-POCUS from scratch resulted in almost the same performance as starting from pretrained weights, suggesting that the dataset is sufficiently large to capture most variation present in apical four-chamber views.

DEPLOYMENT AT BEDSIDE

To make EchoNet-POCUS accessible to clinicians, we designed a system to provide deep learning predictions at bedside (Figure 4). Ultrasound machines commonly provide video output through various video output ports. We use a high-definition multimedia interface (HDMI) connection and capture card to provide this video to a Raspberry Pi, which preserves privacy by cropping all protected health information and extracts clips of interest. Q7

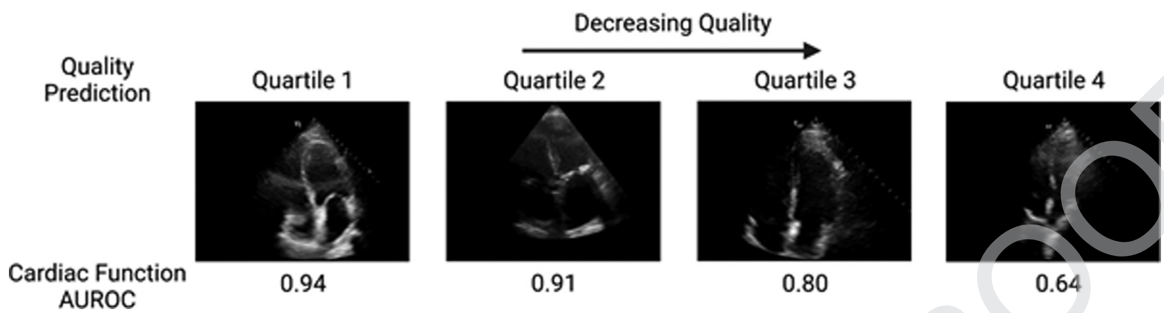

Figure 3. Performance for predicting cardiac function stratified by EchoNet-POCUS's predicted video quality. Example videos are shown for different quality levels. The performance is significantly higher for high quality videos (area under the receiver operating characteristic curve [AUROC] 0.94 for the top-quality quartile) compared with lower-quality videos (AUROC 0.64 for the bottom quartile). POCUS = point-of-care ultrasound.

Table 3. Performance of Different Methods of Training

Training Method	AUC	
	EF	Quality
No retraining	0.81 (0.77–0.85)	–
Fine-tune (last)	0.80 (0.76–0.84)	0.72 (0.67–0.76)
Fine-tune (all)	0.91 (0.89–0.94)	0.82 (0.78–0.87)
Trained without cardiology weights	0.89 (0.86–0.93)	0.81 (0.76–0.84)

No retraining: weights from Ouyang et al. (2020) used without modification; Fine-tune (last): final layer of weights retrained; Fine-tune (all): all weights retrained; Trained without cardiology weights: model trained without using weights from Ouyang et al. (2020) (14).
AUC = area under the curve; EF = ejection fraction. Q8

The predictions from EchoNet-POCUS are then provided back to the user.

To demonstrate the feasibility of deploying EchoNet-POCUS, we prospectively applied the algorithm in real time to analyze videos from 47 patients collected at ED bedside. The patients were scanned by an EP, and EchoNet-POCUS produced nearly instantaneous assessment of the echocardiograms. We generated consensus expert labels using the same procedure as the training videos to measure the prospective performance of the model. On this additional set of prospective patients, EchoNet-POCUS achieves an AUROC of 0.96 (0.90–1.00) for predicting ejection fractions and an AUROC of 0.89 (0.81–0.95) for prediction of video quality.

DISCUSSION

Although point-of-care echocardiography has become widely used in the ED and other clinical settings, there remains a need for the standardization of cardiac contractility assessment. We trained EchoNet-POCUS to aid in both interpreting POCUS echocardiograms and assessing video quality. Given the portability of the model, this feedback can occur in real time, as opposed to other models not designed to operate at the bedside.

EchoNet-POCUS can interpret ED echocardiogram videos with a high degree of accuracy and can be run on commodity hardware in real time, aiding in ultrasound education and significantly decreasing the barrier to entry into ED POCUS. The videos identified as high quality by EchoNet-POCUS have greater model cardiac function prediction accuracy, indicating that the quality predicted by the model can be given as direct feedback to physicians in real time to aid with improving video quality and image acquisition at the bedside. Given real-time feedback regarding video quality, we envision user interface adaptations on the commodity hardware that promote sequential collection of bedside echocardiogram videos to obtain high model ejection fraction prediction accuracy.

Due to the low cost of additional parts needed to run EchoNet-POCUS in real time, these devices could be made commonly available. Additionally, the portability of the commodity hardware along with ease of customizability of the user interface lends the ability of the same underlying technology to be deployed in a variety of

Please cite this article as: B. He et al., AI-Enabled Assessment of Cardiac Function and Video Quality in Emergency Department Point-of-Care Echocardiograms, Journal of Emergency Medicine, https://doi.org/10.1016/j.jemermed.2023.02.005

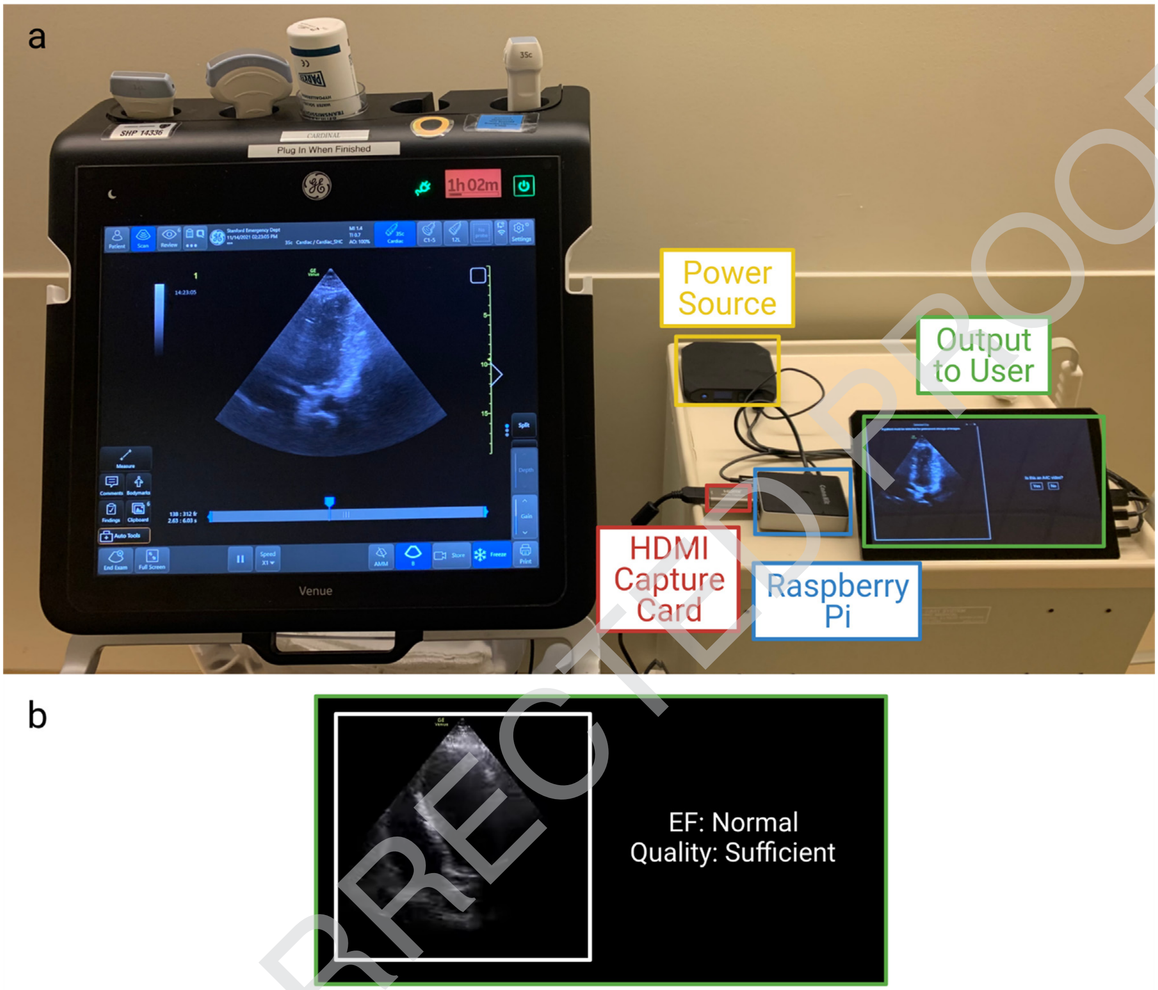

Figure 4. (a) Hardware used to run EchoNet-POCUS in real time at bedside. An HDMI capture card and Raspberry Pi are used to extract the echocardiogram video from the ultrasound machine, run EchoNet-POCUS, and present predictions to the EP. (b) Example of the output provided to the emergency physician. EF = ejection fraction; POCUS = point-of-care ultrasound.

clinical environments. Potential uses of EchoNet-POCUS include primary educational purposes at large academic centers and providing additional patient level contextual information at community hospitals with lower POCUS proficiency rates. This could result in increased utilization of bedside POCUS wherever these machines are deployed.

LIMITATIONS

Although we are limited in the number of patients in our prospective study, the pilot demonstrates the feasibility of using EchoNet-POCUS in real time at the bedside. However, validating the model in a large prospective study across multiple sites is an important direction of future work to ensure that EchoNet-POCUS can generalize across hospitals.

CONCLUSIONS

EPs are increasingly using POCUS to evaluate unstable patients, but considerable operator-to-operator variability remains. We curated a dataset of POCUS echocardiograms to develop EchoNet-POCUS. We find that EchoNet-POCUS can improve the consistency of evaluating cardiac function and can screen for video with poor quality. EchoNet-POCUS can be readily run in real time using low-cost hardware.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation CAREER 1942926, National Institutes of Health P30AG059307, UF1 AG057707, and grants from the

Silicon Valley Foundation and the Chan Zuckerberg Initiative.

Q9

REFERENCES

1. Noble V, Nelson B. Manual of emergency and critical care ultrasound. Cambridge, UK: Cambridge University Press; 2011.
2. Van Schaik GWW, Van Schaik KD, Murphy MC. Point-of-care ultrasonography (POCUS) in a community emergency department: an analysis of decision making and cost savings associated with POCUS. J Ultrasound Med 2019;38:2133–40.
3. Whitson MR, Mayo PH. Ultrasonography in the emergency department. Crit Care 2016;20:227.
4. Sorensen B, Hunskaar S. Point-of-care ultrasound in primary care: a systematic review of generalist performed point-of-care ultrasound in unselected populations. Ultrasound J 2019;11:31.
5. Zieleskiewicz L, Muller L, Lakhal K, et al. Point-of-care ultrasound in intensive care units: assessment of 1073 procedures in a multicentric, prospective, observational study. Intensive Care Med 2015;41:1638–47.
6. Phelan M, Hagerty D. The oblique view: an alternative approach for ultrasound-guided central line placement. J Emerg Med 2009;37:403–8.
7. Ahern M, Mallin MP, Weitzel S, Madsen T, Hunt P. Variability in ultrasound education among emergency medicine residencies. West J Emerg Med 2010;11:314–18.
8. Hayward M, Chan T, Healey A. Dedicated time for deliberate practice: one emergency medicine program's approach to point-of-care ultrasound (PoCUS) training. CJEM 2015;17:558–61.
9. Gaspari R, Weekes A, Adhikari S, et al. Emergency department point-of-care ultrasound in out-of-hospital and in-ED cardiac arrest. Resuscitation 2016;109:33–9.
10. Kirkpatrick JN, Grimm R, Johri AM, et al. Recommendations for echocardiography laboratories participating in cardiac point of care cardiac ultrasound (POCUS) and critical care echocardiography training: report from the American Society of Echocardiography. J Am Soc Echocardiogr 2020;33:409–22 e4.
11. Albaroudi B, Haddad M, Albaroudi O, Abdel-Rahman ME, Jarman R, Harris T. Assessing left ventricular systolic function by emergency physician using point of care echocardiography compared to expert: systematic review and meta-analysis. Eur J Emerg Med 2022;29:18–32.
12. Hughes JW, Yuan N, He B, et al. Deep learning evaluation of biomarkers from echocardiogram videos. EBioMedicine 2021;73.
13. Madani A, Arnaout R, Mofrad M, Arnaout R. Fast and accurate view classification of echocardiograms using deep learning. NPJ Dig Med 2018;1:6 Available atAccessed March 8, 2023. doi:10.1038/s41746-017-0013-1. Q10
14. Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of cardiac function. Nature 2020;580:252–6.
15. Zhang J, Gajjala S, Agrawal P, et al. Fully automated echocardiogram interpretation in clinical practice. Circulation 2018;138:1623–35.
16. Tran D, Wang H, Torresani L, Ray J, LeCun Y, Paluri M. A closer look at spatiotemporal convolutions for action recognition. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, June 2018. Q11
17. Huang H, Nijjar PS, Misialek JR, et al. Accuracy of left ventricular ejection fraction by contemporary multiple gated acquisition scanning in patients with cancer: comparison with cardiovascular magnetic resonance. J Cardiovasc Magn Reson 2017;19:34.
18. Loehr LR, Rosamond WD, Chang PP, Folsom AR, Chambless LE. Heart failure incidence and survival (from the Atherosclerosis Risk in Communities study). Am J Cardiol 2008;101:1016–22.
19. Paszke A, Gross S, Massa F, et al. Advances in neural information processing systems, 32 Wallach H, Larochelle H, Beygelzimer A, d'Alché-Buc F, Fox, Garnett R, eds. Curran Associates, Inc.; 2019 Available at https://proceedings.neurips.cc/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf Accessed Xxxx XX, 202X. Q12

ARTICLE SUMMARY

1. Why is this topic important?

Point-of-care ultrasound (POCUS) is becoming increasingly available to assess cardiac function at bedside. However, there exists considerable operator-to-operator variability among readers.

2. What does this study attempt to show?

Our study shows that EchoNet-POCUS can aid emergency physicians in interpreting POCUS echocardiograms. The low cost of hardware required will allow the model to be easily accessible.

3. What are the key findings?

We find that EchoNet-POCUS can accurately evaluate both cardiac function and video quality. Using EchoNet-POCUS to filter low-quality videos further increases its accuracy.

4. How is patient care impacted?

Patients receiving POCUS echocardiograms can receive faster and more accurate diagnoses when emergency physicians are aided by EchoNet-POCUS.

Please cite this article as: B. He et al., AI-Enabled Assessment of Cardiac Function and Video Quality in Emergency Department Point-of-Care Echocardiograms, Journal of Emergency Medicine, https://doi.org/10.1016/j.jemermed.2023.02.005