JOURNAL OF THE AMERICAN COLLEGE OF CARDIOLOGY VOL. 82, NO. 20, 2023

PUBLISHED BY ELSEVIER

EDITORIAL COMMENT

A New Hope for Deep Learning–Based Echocardiogram Interpretation

The DROIDs You Were Looking For*

Márton Tokodi, MD, PhD, Attila Kovács, MD, PhD

In recent years, the confluence of technology and medicine has yielded remarkable advancements that were once relegated to the realm of science fiction. From genomics to telemedicine, the health care landscape is continuously evolving, with artificial intelligence (AI) emerging as a transformative force. With this fast-paced development, are we approaching a world like that portrayed in *Star Wars*, where only droid medics deliver care? Although this is a very unlikely scenario, AI-powered machines will undoubtedly play an increasing role in health care, not as a replacement but rather as a valuable supplement to health care professionals. Considering the high volume of studies ordered, echocardiography would significantly benefit from AI-powered solutions. Such tools have the potential to optimize scheduling, aid image acquisition, automate the assessment of routine echocardiographic parameters, and improve diagnostics and prognostication.[1-7] With the first blinded, randomized trial proving that initial assessment of left ventricular ejection fraction (LVEF) by AI is noninferior to assessment by sonographers, the interest has recently surged in automated echocardiogram interpretation tools, especially in deep learning (DL) models for predicting routine echocardiographic parameters.[8]

In this issue of the *Journal of the American College of Cardiology*, Lau et al[9] proposed 2 DL-based echocardiogram interpretation models, DROID-LA and DROID-LV, to automate the assessment of standard measurements of left atrial (LA) and left ventricular (LV) structure and function. The DROID models were

SEE PAGE 1936

trained, tested, and internally validated on 64,028 echocardiographic studies of 27,135 patients included in the Enterprise Warehouse of Cardiology (EWOC) sample and externally validated on 9,248 studies from the Community Care Cohort Project (C3PO) cohort and more than 10,000 studies from the EchoNet-Dynamic and EchoNet-LVH data sets.[4,5,10] The investigators designed fully automated end-to-end models that, after performing view classification, video type identification, and image quality and axis assessment for all echocardiographic videos in the given study, predict LVEF, LV end-diastolic and end-systolic dimensions, interventricular septal and posterior wall thicknesses, and LA anteroposterior dimension at the study level. In the C3PO external validation set, DROID-LA and DROID-LV accurately predicted LA and LV linear measures (R^2 values ranging from 0.59 to 0.82) and LVEF (mean absolute error: 4.23 percentage points, $R^2 = 0.74$), whereas their performance was slightly lower but still acceptably accurate in the EchoNet-Dynamic and EchoNet-LVH data sets. Moreover, the DROID-derived LA and LV measurements were found to be independently associated with incident cardiovascular outcomes, including heart failure, atrial fibrillation, myocardial infarction, and all-cause death, in a subset of the C3PO cohort.

Several DL models have recently been proposed to predict LV measurements from echocardiographic videos.[4,5,11-21] Many of these models perform image

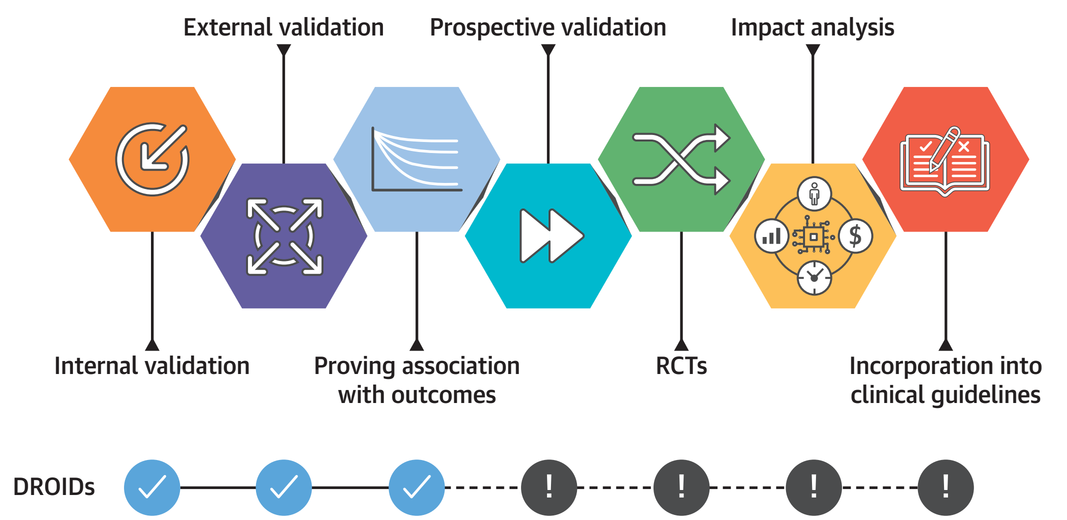
*Editorials published in the *Journal of the American College of Cardiology* reflect the views of the authors and do not necessarily represent the views of the *Journal of the American College of Cardiology* or the American College of Cardiology.

From the Heart and Vascular Center, Semmelweis University, Budapest, Hungary.

The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the Author Center.

ISSN 0735-1097/$36.00 https://doi.org/10.1016/j.jacc.2023.09.799

FIGURE 1 Evolution of Studies Developing DL Models for Automated Echocardiogram Interpretation



The first published deep learning (DL) models for the automated measurement of echocardiographic parameters were developed in single-center databases and were validated only internally. With the increasing recognition of its importance and the growing number of public data sets, external validation has been included in more and more recently published studies. Nevertheless, only a few have proved the association of the predicted echocardiographic parameters with outcomes so far. Although the DROID models have passed these first 3 stages, further validation will be required through prospective studies and randomized controlled trials. Moreover, impact analyses should also be performed to investigate whether applying these DL models improves patient outcomes and cost efficiency. Only after successfully completing these stages and obtaining the required approvals from the corresponding agencies (eg, U.S. Food and Drug Administration) could DL models be considered for incorporation into clinical guidelines. RCT = randomized controlled trial.

segmentation on echocardiographic video frames and then calculate the target echocardiographic parameters from these segments.[5,11-15] Nevertheless, there is a growing number of solutions that do not rely on such an intermediate segmentation step.[16-21] Although segmentation-based algorithms produce clinically interpretable intermediate results by delineating the regions of interest and provide an opportunity for quality control as erroneous segmentations can be easily identified and corrected, segmentation-free approaches have several advantages over them, including but not limited to their ability to learn global features and latent information within the entire image rather than relying on specific boundaries or regions. By recognizing these strengths, the investigators opted for a segmentation-free design.

But what makes the DROIDs stand out from the other previously published DL models predicting LV echocardiographic parameters? First, they were trained on a well-curated data set—the EWOC, a retrospective multi-institutional ambulatory cardiology electronic health record sample, which is 1 or even 2 magnitudes larger than the data sets used to train the previously published models. Second, the DROIDs have a unique 2-component architecture intentionally designed to mimic the clinical workflow of echocardiogram interpretation: they encode important information from all available echocardiographic videos via 3-dimensional convolutional neural networks (ie, MoViNet-A2) and then they use attention heads to translate the encoded information from the most relevant videos within the same study into study-level predictions. Thus, the DROID models are very flexible as they can simultaneously process up to 40 parasternal long-axis, apical 4-chamber, and apical 2-chamber view videos from the same echocardiographic study, and they can estimate left heart measurements even if only 1 of 3 views is available. Moreover, fusing this innovative architecture with the extensive EWOC data set yielded a robust model that even surpassed the EchoNet-Dynamic model's performance in predicting LVEF within the C3PO external validation set. Third, as a cherry on top, the investigators also showed that the DL-derived LV measurements are associated with cardiovascular outcomes, which has been only done in a handful of studies focusing on the DL-based assessment of echocardiographic parameters.[22]

There are 2 other important aspects of the study that should be pointed out: the investigators made the source code along with the model weights publicly available and placed emphasis on the explainability of the model by determining which echocardiographic videos contribute the largest attention weights, and also by overlaying saliency maps on top of the echocardiographic videos to highlight anatomical structures with the greatest impact on the predictions.[23] The former enables other groups of researchers to validate the DROID models independently, whereas the latter provides insights into the prediction process and increases transparency. Both are critical to earning the trust of physicians and patients and enabling the widespread use and clinical adoption of DL models.

This elegant study represents a major step forward in the evolution of clinical investigations targeting the development and evaluation of DL models for the automated interpretation of echocardiograms (**Figure 1**). Considering the remarkable performance of the DROID models during external validation, one of the logical next steps would be conducting prospective randomized trials and impact analyses to investigate whether applying these models improves patient outcomes and cost efficiency. Another direction worth exploring would be training the proposed model architecture to predict additional echocardiographic measurements, especially right heart parameters (eg, right ventricular diameters or fractional area change), which has been targeted by only a very limited number of studies so far.[22,24]

In conclusion, although the vision of health care dominated by droid medics remains firmly in the realm of science fiction, the reality of AI's role is evolving rapidly. Echocardiography, as shown by the success of the DROID models, is poised to reap substantial benefits from these advancements. As we move forward, we should embrace these AI-powered tools as invaluable partners, enhancing our ability to provide exceptional care to our patients. The force of innovation is with us, and together, we can shape a future where artificial and human intelligence work synergistically to transform cardiology and improve the lives of countless patients.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

Project no. RRF-2.3.1-21-2022-00004 (MILAB) has been implemented with the support provided by the European Union. Dr Kovács has received grant support from the National Research, Development, and Innovation Office (NKFIH) of Hungary (FK 142573); and has been supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. Dr Tokodi has received consulting fees from CardioSight Inc outside the submitted work. Dr Kovács has received personal fees from Argus Cognitive Inc and CardioSight Inc outside the submitted work.

ADDRESS FOR CORRESPONDENCE: Dr Márton Tokodi, Heart and Vascular Center, Semmelweis University, 68 Városmajor Street, Budapest 1122, Hungary. E-mail: tokmarton@gmail.com. @tokmarton, @kovatti87.

REFERENCES

1. Sánchez-Puente A, Dorado-Díaz PI, Sampedro-Gómez J, et al. Machine-learning to optimize the echocardiographic follow-up of aortic stenosis. *J Am Coll Cardiol Img*. 2023;16(6):733–744.

2. Kovács A, Tokodi M. Refining echocardiographic surveillance of aortic stenosis using machine learning. *J Am Coll Cardiol Img*. 2023;16(6):745–748.

3. Narang A, Bae R, Hong H, et al. Utility of a deep-learning algorithm to guide novices to acquire echocardiograms for limited diagnostic use. *JAMA Cardiol*. 2021;6:624–632.

4. Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature*. 2020;580:252–256.

5. Duffy G, Cheng PP, Yuan N, et al. High-throughput precision phenotyping of left ventricular hypertrophy with cardiovascular deep learning. *JAMA Cardiol*. 2022;7:386–395.

6. Shad R, Quach N, Fong R, et al. Predicting post-operative right ventricular failure using video-based deep learning. *Nat Comm*. 2021;12:5192.

7. Valsaraj A, Kalmady SV, Sharma V, et al. Development and validation of echocardiography-based machine-learning models to predict mortality. *eBioMedicine*. 2023;90.

8. He B, Kwan AC, Cho JH, et al. Blinded, randomized trial of sonographer versus AI cardiac function assessment. *Nature*. 2023;616:520–524.

9. Lau ES, Di Achille P, Kopparapu K, et al. Deep learning-enabled assessment of left heart structure and function predicts cardiovascular outcomes. *J Am Coll Cardiol*. 2023;82(18):1936–1948.

10. Khurshid S, Reeder C, Harrington LX, et al. Cohort design and natural language processing to reduce bias in electronic health records research. *NPJ Digit Med*. 2022;5:47.

11. Leclerc S, Smistad E, Pedrosa J, et al. Deep learning for segmentation using an open large-scale dataset in 2D echocardiography. *IEEE Trans Med Imaging*. 2019;38:2198–2210.

12. Liu X, Fan Y, Li S, et al. Deep learning-based automated left ventricular ejection fraction assessment using 2-D echocardiography. *Am J Physiol Heart Circ Physiol*. 2021;321:H390–H399.

13. Smistad E, Østvik A, Salte IM, et al. Real-time automatic ejection fraction and foreshortening detection using deep learning. *IEEE Trans Ultrason Ferroelectr Freq Control*. 2020;67:2595–2604.

14. Zhang J, Gajjala S, Agrawal P, et al. Fully automated echocardiogram interpretation in clinical practice. *Circulation*. 2018;138:1623–1635.

15. Salte IM, Østvik A, Smistad E, et al. Artificial intelligence for automatic measurement of left ventricular strain in echocardiography. *J Am Coll Cardiol Img*. 2021;14:1918–1928.

16. Kazemi Esfeh MM, Luong C, Behnami D, Tsang T, Abolmaesumi P. A Deep Bayesian Video Analysis Framework: Towards a More Robust Estimation of Ejection Fraction. In: Martel AL, Abolmaesumi P, Stoyanov D, et al., eds. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*. Cham: Springer International Publishing; 2020:582–590.

17. Behnami D, Liao Z, Girgis H, et al. Dual-View Joint Estimation of Left Ventricular Ejection Fraction with Uncertainty Modelling in Echocardiograms. In: Shen D, Liu T, Peters TM, et al., eds. *Medical Image Computing and Computer Assisted Intervention — MICCAI 2019*. Cham: Springer International Publishing; 2019:696–704.

18. Ghorbani A, Ouyang D, Abid A, et al. Deep learning interpretation of echocardiograms. *NPJ Digit Med*. 2020;3:10.

19. Reynaud H, Vlontzos A, Hou B, Beqiri A, Leeson P, Kainz B. Ultrasound Video Transformers for Cardiac Ejection Fraction Estimation. In: de Bruijne M, Cattin PC, Cotin S, et al., eds. *Medical Image Computing and Computer Assisted Intervention — MICCAI 2021*. Cham: Springer International Publishing; 2021:495–505.

20. Kusunose K, Haga A, Yamaguchi N, et al. Deep learning for assessment of left ventricular ejection fraction from echocardiographic images. *J Am Soc Echo Cardiogr*. 2020;33:632–635.e1.

21. Silva JF, Silva JM, Guerra A, Matos S, Costa C. Ejection Fraction Classification in Transthoracic Echocardiography Using a Deep Learning Approach. 2018 IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), June 18-21. Sweden: Karlstad; 2018:123–128.

22. Tokodi M, Magyar B, Soós A, et al. Deep learning-based prediction of right ventricular ejection fraction using 2D echocardiograms. *J Am Coll Cardiol Img*. 2023;16:1005–1018.

23. broadinstitute/m14h. https://github.com/broadinstitute/ml4h/tree/master/model_zoo/DROID

24. Beecy AN, Bratt A, Yum B, et al. Development of novel machine learning model for right ventricular quantification on echocardiography — a multimodality validation study. *Echocardiography*. 2020;37:688–697.

KEY WORDS deep learning, echocardiography, left ventricle, outcomes