Published in final edited form as:

JACC Cardiovasc Imaging. 2021 November ; 14(11): 2105–2106. doi:10.1016/j.jcmg.2021.06.007.

Can machine learning help simplify the measurement of diastolic function in echocardiography?

Rima Arnaout, MD^{1,*}

¹Department of Medicine, Division of Cardiology, Bakar Computational Health Sciences Institute, University of California, San Francisco, San Francisco, CA

With every heartbeat, the heart has *two* jobs: it has to squeeze, and it has to relax. This is how I explain heart failure with preserved ejection fraction (HFpEF) to newly diagnosed patients. “That makes sense—I never really thought about that second part,” one patient told me in reply. He is not alone. Diastolic dysfunction underlying HFpEF is often under-recognized as a cause of heart failure, despite causing significant morbidity(1).

HFpEF can be multifactorial in etiology. It is a clinical diagnosis of heart failure in a patient with a left ventricular ejection fraction (LVEF) of 50 percent or above(2)22. Echocardiography can play an important role in establishing a HFpEF diagnosis by providing the requisite characterization of LV systolic function; it can give additional evidence for HFpEF by assessing for diastolic dysfunction. Updated in 2016, the joint American Society of Echocardiography/European Association of Cardiovascular Imaging guidelines for measuring diastolic dysfunction on echocardiogram are based on measurements of blood flow and tissue movement patterns that provide indirect information about the diastolic pressures in the left atrium and ventricle, as well as structural information that suggests long-standing pressure elevations(3). While the 2016 guidelines aimed to simplify previous guidance to increase evaluation of diastolic function in everyday practice, there are still over a dozen parameters that can still be used to generate an overall grading of diastolic function in a given echocardiogram(3). The cardiovascular imaging community can benefit from still further simplification of diastolic function assessment.

In the current issue of *JACC: Cardiovascular Imaging*, Chiou et. al. use now-established methods in machine learning on echocardiogram b-mode imaging(4–8) to design an automated method for binary classification of diastolic function. Specifically, they test whether their method can predict HFpEF vs. ‘asymptomatic.’ In this task, they achieve an accuracy, sensitivity, and specificity of 0.91, 0.96, and 0.85, respectively on an internal test set. On test data from an external medical center, accuracy, sensitivity, and specificity were 0.85, 0.79, and 0.89, respectively.

To do this, the authors made use of the fact that much of diastolic function is evaluated from the magnitude and timing of blood flow between left atrium and left ventricle as a surrogate for pressure differences between them. The authors cleverly hypothesized that if

*corresponding author: rima.arnaout@ucsf.edu.

Disclosures: None relevant

blood flow can be used in this way, then the size of the blood-filled chambers can also be used to evaluate diastolic function.

They first trained a neural network algorithm(7) to find the left atrium and the left ventricle from apical 4-chamber b-mode imaging. With these shapes, they could then calculate chamber area, length, width, and volume, for each frame in a given cine clip, and plot these measurements frame-by-frame across the cardiac cycle represented in the clip. This created a handful of channels of time-series data, which they then fed into a second simple algorithm that classified these time series as 'HFpEF' or 'not HFpEF,' achieving the performance mentioned above.

Even if it doesn't overturn clinical paradigms, machine learning may help practitioners and patients by providing modest improvements to the clinical workflow, based on existing clinical knowledge. Here, the authors have shown the potential for just this, re-imagining diastolic functional assessment on echo as a task a computer can accomplish with established segmentation and time-series classification methods, overall presenting the work in accordance with emerging guidance for scientific rigor in medical machine learning(9–11). Notably, this method leverages standard b-mode cine acquisitions of the apical 4-chamber view, and this is an asset. Also, where the presence of atrial fibrillation, common in HFpEF patients, often confounds traditional diastolic function assessment, an approach based on LA and LV chamber sizes over time may prove to be more tolerant to arrhythmia.

However, several questions and areas for future study remain. While the authors tested the model on a total of over 600 patients including external data, which is commendable, this is still a small fraction of the tens of millions of echocardiograms performed worldwide. The fact that the model performed worse on the external dataset indicates that it may not yet generalize for real-world use. Additionally, the patient characteristics of the control group in both internal and external test datasets (Tables 2–3) appear to have statistically significantly different distributions of many characteristics that co-exist with HFpEF, such as hypertension, diabetes, renal dysfunction, coronary artery disease, and atrial fibrillation. One might imagine that a simple classifier based on just these clinical characteristics alone may perform similarly to the authors' model in predicting HFpEF from this test set. In the real world, echocardiographers must distinguish diastolic dysfunction in patients who may have several HFpEF comorbidities, and clinicians who already have this clinical information are looking for value-add from an echocardiogram. For these reasons, a larger, more clinically challenging, and more racially and ethnically diverse test dataset is needed going forward.

It is yet to be determined where and how this very interesting proof of concept may provide value for patients and clinicians, and how that value will be measured(12). Will additional information from this model re-classify patients with suspected HFpEF more clearly into HFpEF vs non-HFpEF groups? Will this work be further developed to provide a grading of diastolic dysfunction rather than a binary output? Will it help to resolve cases where diastolic function by echocardiogram appears indeterminate? Will the model provide clear time savings to the clinical workflow? Time will tell, and this journal's readers will undoubtedly await new developments with excitement.

References

1. Yancy Clyde W, Jessup Mariell, Biykem Bozkurt, et al. 2013 ACCF/AHA Guideline for the Management of Heart Failure. *Journal of the American College of Cardiology* 2013;62:e147–e239. [PubMed: 23747642]
2. Biykem Bozkurt, Hershberger Ray E., Javed Butler, et al. 2021 ACC/AHA Key Data Elements and Definitions for Heart Failure. *Journal of the American College of Cardiology* 2021;77:2053–2150. [PubMed: 33250265]
3. Nagueh SF, Smiseth OA, Appleton CP, et al. Recommendations for the Evaluation of Left Ventricular Diastolic Function by Echocardiography: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *J Am Soc Echocardiogr* 2016;29:277–314. [PubMed: 27037982]
4. Arnaout R, Curran L, Chinn E, Zhao Y, Moon-Grady A. Deep-learning models improve on community-level diagnosis for common congenital heart disease lesions. *ArXiv e-prints* 2018.
5. Zhang J, Gajjala S, Agrawal P, et al. Fully Automated Echocardiogram Interpretation in Clinical Practice. *Circulation* 2018;138:1623–1635. [PubMed: 30354459]
6. Ouyang D, He B, Ghorbani A, et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature* 2020;580:252–256. [PubMed: 32269341]
7. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv:1505.04597 [cs]* 2015. Available at: <http://arxiv.org/abs/1505.04597>. Accessed May 28, 2018.
8. Arnaout R, Curran L, Zhao Y, Levine JC, Chinn E, Moon-Grady AJ. An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease. *Nat Med* 2021;27:882–891. [PubMed: 33990806]
9. Norgeot B, Quer G, Beaulieu-Jones BK, et al. Minimum information about clinical artificial intelligence modeling: the MI-CLAIM checklist. *Nat Med* 2020;26:1320–1324. [PubMed: 32908275]
10. Sengupta PP, Shrestha S, Berthon B, et al. Proposed Requirements for Cardiovascular Imaging-Related Machine Learning Evaluation (PRIME): A Checklist: Reviewed by the American College of Cardiology Healthcare Innovation Council. *JACC Cardiovasc Imaging* 2020;13:2017–2035. [PubMed: 32912474]
11. Kakarmath S, Esteva A, Arnaout R, et al. Best practices for authors of healthcare-related artificial intelligence manuscripts. *NPJ Digit Med* 2020;3:134. [PubMed: 33083569]
12. Quer G, Arnaout R, Henne M, Arnaout R. Machine Learning and the Future of Cardiovascular Care: JACC State-of-the-Art Review. *Journal of the American College of Cardiology* 2021;77:300–313. [PubMed: 33478654]