JACC: CARDIOVASCULAR IMAGING VOL. ■, NO. ■, 2021

PUBLISHED BY ELSEVIER

EDITORIAL COMMENT

Phenogrouping Diastolic Dysfunction by Artificial Intelligence

Learning From What We Teach the Machines*

Mani A. Vannan, MBBS,[a] Edgar Argulian, MD,[b] Patrizio Lancellotti, MD, PhD[c,d]

"Predicting the future isn't magic, it's artificial intelligence."

—Dave Waters (1)

Echocardiography is the cornerstone in the assessment of left ventricular diastolic dysfunction (LVDD) (2). This evaluation is based on an algorithmic approach to assess myocardial relaxation, myocardial stiffness, and left ventricular filling pressure (LVFP). LVDD frequently, but not always, accompanies heart failure with preserved ejection fraction (HFpEF) (3). About 50% of those with the clinical diagnosis of heart failure have HFpEF that carries a significant mortality, varying from 10% to 30% (higher in epidemiological studies than in clinical trials) (4). In addition, the degree of DD determined via echocardiography is predictive of clinical outcomes in HFpEF (2). Therefore, a reliable evaluation of DD by echocardiography is of paramount importance in the diagnosis and management of HFpEF.

Over the years, there have been many algorithmic pathways to assess DD with the use of echocardiography. The latest iteration of the systematic approach to evaluate for DD with the use of echocardiography was described in the 2016 American Society of Echocardiography (ASE) guidelines (5). This has been demonstrated to be more accurate and reliable compared with the previous approach. However, there are limitations to the ASE guidelines, with up to one-third of the individuals falling under the indeterminate classification, especially when applied to the general population (6). Furthermore, DD may include patients with abnormal relaxation but normal LVFP. Conversely, young patients with risk factors such as hypertension or diabetes mellitus may be graded as having worse DD than is truly the case (7). All of these, at least in part, attest to the fact that a stepwise approach does not account for the more complex interactions among the parameters. Finally, DD is only one component of HFpEF, which is a multisystem disorder, and it has been argued that assessment of DD might not even be necessary for the diagnosis of HFpEF (8).

In this issue of *iJACC*, Pandey et al (8) have proposed phenotyping DD by machine-learning (ML) methods using echocardiographic parameters and clinical features (9). They applied an unsupervised ML approach to retrospective echocardiographic data from patients with HF and identified low- and high-risk phenogroups, followed by developing a deep learning solution. This deep neural network (DeepNN) classifier was then used to predict the high- and low-risk phenogroups in a derivation cohort. Subsequently they applied the DeepNN classifier to predict elevated LVFP, adverse clinical outcomes, elevated cardiac biomarkers, exercise capacity, and response to therapy in validation cohorts. The predictive performance of the DeepNN classifier for elevated LVFP was tested in 84 patients who had invasive measurements in close temporal proximity to echocardiography. The DeepNN classifier performed significantly better than the currently recommended echocardiographic criteria for elevated LVFP. In a separate population of 219 patients with HF, the DeepNN classifier was able to identify the

*Editorials published in *JACC: Cardiovascular Imaging* reflect the views of the authors and do not necessarily represent the views of *iJACC* or the American College of Cardiology.

From the [a]Marcus Heart Valve Center, Piedmont Heart Institute, Atlanta, Georgia, USA; [b]Mount Sinai Morningside Heart, Mount Sinai Heart, New York, New York, USA; [c]Centre Hospitalier Universitaire Sart Tilman, GIGA Cardiovascular Sciences, University of Liège Hospital, Liege, Belgium; and [d]Gruppo Villa Maria Care and Research, Maria Cecilia Hospital, Cotignola, and Anthea Hospital, Bari, Italy.

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 1936-878X/$36.00 https://doi.org/10.1016/j.jcmg.2021.05.018

high- and low-risk phenogroups, with a higher rate of the cumulative composite end point of death and hospitalization in the high-risk group. This was true for the entire cohort and when applied to those with HFpEF (EF >50%). Finally, the DeepNN classifier was applied to 3 drug trial populations. In the TOPCAT population, the ML classifier categorized ~80% as high-risk and the cumulative incidence of primary endpoint of all-cause mortality and hospitalization (34%) was twice that of the low-risk phenogroup (17%). Furthermore, ~40% were classified as indeterminate DD by the ASE guidelines, and 80% of these were categorized in the high-risk phenogroup and had a higher incidence of adverse clinical outcome. It was also remarkable that nearly 75% of those classified as grade 1 DD by the ASE guidelines were identified as high-risk with significantly worse clinical outcomes. When the DeepNN classifier was applied to the combined RELAX and NEAT-HFpEF population, ~75% of the patients were identified as high-risk, which showed higher levels of circulating cardiac biomarkers of myocardial stress and injury, lower exercise capacity (VO_{2peak}) and worse Minnesota Living With Heart Failure quality of life score, independently from other variables.

The past few years has seen a plethora of publications reporting ML algorithms for almost the entire gamut of cardiovascular disease. This trend is set to expand and it is only a matter of time before ML approaches become a part of routine clinical practice. This is a much needed development, as the human brain can process only a finite number of elements to accurately diagnose, assign prognosis, or plan appropriate therapy for a given disease. HFpEF is a good example of the challenge that a clinician encounters for the diagnosis and management of disease. HFpEF is a multiorgan disorder that includes cardiac and vascular structural remodeling, endothelial and ventriculovascular functional perturbations, and metabolic and inflammatory abnormalities with differential dominance of circulating biomarkers in men and in women. ML algorithms should serve us well in clinical decision making, given that conventional multivariable logistic regression statistics do not account for the nonlinear complex interaction of the factors that ultimately determine the disease state, the outcome, and the effectiveness of a therapeutic plan for a HFpEF patient.

Pandey et al (8) have provided a good first step in that direction, but many issues remain to be addressed. The inclusion of parameters other than those that identify DD by echocardiography is important. For example, there is almost no information on vascular parameters, which are a key determinant of the pathophysiology and clinical course of HFpEF. There are numerous other reports of characterization of HFpEF phenotypes based on ML methods (10-12). There are similarities and differences in the phenotypes in these publications, perhaps caused by the inclusion of different variables in developing the ML algorithms, or due to unreliable and missing data. Although imputation methods may address the latter issue, the threshold at which imputation could account for the missing data is unclear. Also, ML phenotyping should not only be predictive of outcomes, but must aid in predicting utility or futility of therapy. In this regard, Pandey et al (8) have shown that spironolactone therapy in the high-risk TOPCAT phenogroup identified by the DeepNN classifier reduced the incidence of the primary adverse end point, an effect not seen in the low-risk phenogroup.

Although ML methods offer an ability to process data in multidimensional planes, their accuracy and reliability remain to be tested by prospective randomized application to larger and multi-institutional populations. ML tools must also offer transparency for the clinician to be able to audit and validate the models in the context of current clinical practice. This quote from Pedro Domingos is worth remembering: “The greatest benefit of machine learning may ultimately be not what the machines learn but what we learn by teaching them.” (The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World).

FUNDING SUPPORT AND AUTHOR DISCLOSURES

All authors have reported that they have no relationships relevant to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr Mani A. Vannan, Marcus Heart Valve Center, Piedmont Heart Institute, Atlanta, Georgia 30305, USA. E-mail: mvannan2560@gmail.com. Twitter: @ManiVannanheart.

REFERENCES

1. Best Artificial Intelligence Quotes. Accessed April 26, 2021. https://www.supplychaintoday.com/best-artificial-intelligence-quotes/

2. Nagueh SF. Left ventricular diastolic function: understanding pathophysiology, diagnosis, and prognosis with echocardiography. *J Am Coll Cardiol Img*. 2020;13:228–44.

3. Shah AM, Cikes M, Prasad N, et al., PARAGON-HF Investigators. Echocardiographic features of patients with heart failure and preserved left ventricular ejection fraction. *J Am Coll Cardiol*. 2019;74:2858–73.

4. Chan MM, Lam CS. How do patients with heart failure with preserved ejection fraction die? *Eur J Heart Fail*. 2013;15:604–13.

5. 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.

6. Almeida JG, Fontes-Carvalho R, Sampaio F, et al. Impact of the 2016 ASE/EACVI recommendations on the prevalence of diastolic dysfunction in the general population. *Eur Heart J Cardiovasc Imaging*. 2018;19:380–6.

7. Oh JK, Miranda WR, Bird JG, Kane GC, Nagueh SF. The 2016 diastolic function guideline: is it already time to revisit or revise them? *J Am Coll Cardiol Img*. 2020;13:327–35.

8. Zile MR, Gaasch WH, Carroll JD, et al. Heart failure with a normal ejection fraction: is measurement of diastolic function necessary to make the diagnosis of diastolic heart failure? *Circulation*. 2001;104:779–82.

9. Pandey A, Kagiyma N, Yanamala N, et al. Deep-learning models for the echocardiographic assessment of diastolic dysfunction. *J Am Coll Cardiol Img*. 2021;14:XXX-XX.

10. Shah SJ, Katz DH, Selvaraj S, et al. Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation*. 2015;131:269–79.

11. Lewis GA, Schelbert EB, Williams SG, et al. Biological phenotypes of heart failure with preserved ejection fraction. *J Am Coll Cardiol*. 2017;70:2186–200.

12. Schrub F, Oger E, Bidaut A, et al. Heart failure with preserved ejection fraction: a clustering approach to a heterogenous syndrome. *Arch Cardiovasc Dis*. 2020;113:381–90.

KEY WORDS artificial intelligence machine-learning, diastolic dysfunction, HFpEF