BRIEF RESEARCH COMMUNICATION

Automated Echocardiographic Detection of Heart Failure With Preserved Ejection Fraction Using Artificial Intelligence Is Associated With Cardiac Mortality and Heart Failure Hospitalization

Heart failure with preserved ejection fraction (HFpEF) accounts for approximately 50% of diagnoses of heart failure (HF) and frequently leads to hospitalization. Clinical algorithms developed for diagnosis have been applied to stratify risk for HF hospitalization or death.[1-3] Deep learning has been applied to the automated interpretation of echocardiograms, but limited information exists regarding the potential of the learning models for predicting clinical outcomes.[4] An artificial intelligence (AI) model was recently developed to identify patients with HFpEF using a single apical four-chamber video clip from a standard transthoracic echocardiographic examination.[5] A convolutional neural network was applied to the video clip. The model comprised a series of three-dimensional convolutional layers designed to operate on two-dimensional videos over two in-plane spatial dimensions within the image frames and across the time dimension. The present study was conducted to assess the association between the model output and other HF biomarkers, risk for HF hospitalization and cardiac mortality, and to compare its performance with two clinical scores: H2FPEF (heavy, hypertensive, atrial fibrillation, pulmonary hypertension, elder, and filling pressure)[6] and HFA-PEFF (Heart Failure Association pretest assessment, echocardiography and natriuretic peptide score, functional testing, and final etiology).[7]

This retrospective, multisite study was approved by our institutional review board.

The model was developed to classify patients with HFpEF vs individuals without HFpEF (control subjects). Patients with HFpEF were defined according to guidelines and included a diagnosis by the treating physician[1] within 1 year of an echocardiographic examination demonstrating elevated left ventricular filling pressure. Control subjects were patients undergoing clinically indicated echocardiography who lacked these features (Supplemental Appendix). All patients had left ventricular ejection fractions ≥50%. The present analysis used the second version of the AI model (Supplemental Appendix). In the previously described independent test population consisting of 646 patients with HFpEF and 638 control subjects, the updated model produced 95 uncertain outputs (7.4%); in the remaining 607 patients and 582 control subjects, sensitivity was 89.8% (95% CI, 87.5%-92.5%), specificity was 86.3% (95% CI, 83.6%-89.7%), negative predictive value was 89.0% (95% CI, 87.0%-91.7%), and positive predictive value was 87.2% (95% CI, 84.7%-89.8%).

Incident HF hospitalization was obtained from electronic health record chart review using standardized definitions, using the first event after the echocardiographic examination. Mortality was obtained from the National Death Index, and causes of cardiac deaths were manually reviewed. End points were adjudicated by investigators blinded to AI analysis results. Cardiac mortality and HF hospitalization were plotted accounting for death as a competing risk. The method of

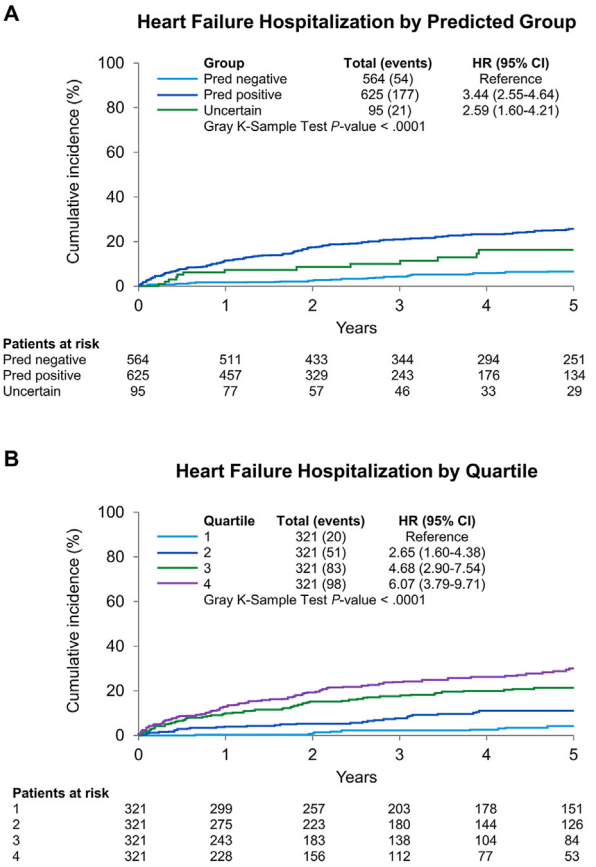

Figure 1 Risk for HF hospitalization was higher in patients with positive and uncertain AI model output *(top)*. Those with output scores that placed them in higher quartiles had incrementally greater risk for HF hospitalization *(bottom)*, accounting for competing risk for death.

Fine and Gray was used to estimate the hazard ratios (HRs) adjusted for differences in age and sex.

Among 1,284 patients followed for a median of 3.4 years (interquartile range, 1.7-6.5 years), there were 252 HF hospitalizations and 540 deaths. Figure 1 demonstrates the risk for HF hospitalization on the basis of HF categorical AI output (top) and quartiles of continuous probability output (bottom). After adjustment for age and sex, positive AI output was associated with a higher risk for HF hospitalization than negative output (HR, 3.76; 95% CI, 2.71-5.21; $P < .001$) and likewise for uncertain output (HR, 2.79; 95% CI, 1.60-4.62; $P < .001$). Cardiac deaths ($n = 135$) were attributable to HF in 63 patients (47%), to coronary artery disease in 55 (41%), to valve disease in five (4%), to arrhythmia in five (4%), and to other causes in seven (5%). Again adjusting for age and sex, cardiac mortality was higher in patients with positive output (HR, 5.55; 95% CI, 3.28-9.37; $P < .001$); patients with an uncertain output tended to have a higher mortality (HR, 2.22; 95% CI, 0.94-5.24; $P = .07$). Patients with higher continuous probability outputs demonstrated incrementally higher risk for cardiac mortality (fourth quartile vs first quartile: HR, 11.65; 95% CI, 4.65-29.20; $P < .0001$). Figure 2 demonstrates the risk for HF hospitalization on the basis of clinical H2FPEF score[6] (top). The clinical score differentiated high and low risk for HF hospitalization, but 776 of 1,284 patients (60%) were indeterminate. Application of the AI model to the nondiagnostic H2FPEF outputs (bottom) allowed

This study was supported by grants from the American Society of Echocardiography Foundation and Ultromics. Dr. Pellikka is supported as the Betty Knight Scripps Professor of Cardiovascular Disease Clinical Research.

Federico M. Asch, MD, served as guest editor for this report.

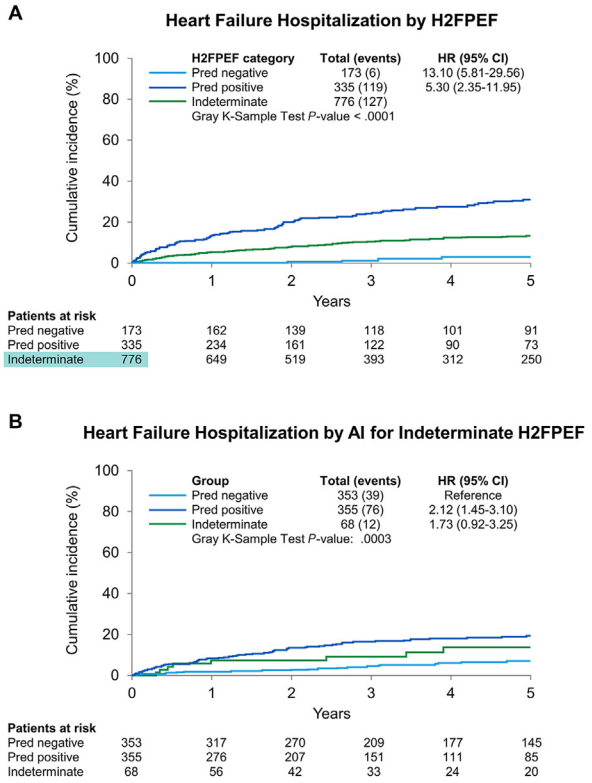

Figure 2 Risk for HF hospitalization was higher in patients with positive H2FPEF outputs *(top)*, but many (n = 776) had nondiagnostic outputs. Application of the AI model was able to reclassify 708 (91%) of the nondiagnostic H2FPEF outputs *(bottom)*.

the classification of all but 68 of the 776 patients (8.8%). The AI model demonstrated a similar relationship between output and risk for HF hospitalization in patients with and those without diagnostic H2FPEF output. Findings were similar when patients were stratified according to HFA-PEFF score. HFA-PEFF score, brain natriuretic peptide, and N-terminal pro–brain natriuretic peptide also differed according to the AI model prediction (Table). Few patients underwent exercise testing; differences in exercise capacity were not significantly different.

In this study we assessed the ability of a novel, HFpEF AI model using a single echocardiographic video clip to identify patients at increased risk for HF hospitalization and cardiac mortality. In summary, (1) positive model output was associated with higher risks for HF hospitalization and cardiac mortality, (2) patients with uncertain outputs demonstrated intermediate risks for these end points, (3) HF hospitalization and cardiac mortality risk were incrementally associated with higher model probability output scores, and (4) the AI model reclassified HF hospitalization risk in nondiagnostic clinical scores, including 91% for H2FPEF outputs and 92% for HFA-PEFF. This is the first AI echocardiographic model to produce outputs discriminating a specific disease (HFpEF) that are incrementally associated with risk for HF hospitalization and cardiac mortality. Prospective studies are required to confirm these retrospective results, to externally validate the AI model's outputs in other echocardiographic laboratories, and to understand the implications for patient management. Studies using a broad representation of HFpEF phenotypes should be undertaken to understand the generalizability of this model in a naturally heterogeneous clinical syndrome.

REVIEW STATEMENT

Given her role as *JASE* Editor-in-Chief, Patricia A. Pellikka, MD, had no involvement in the peer review of this article and has no access to information regarding its peer review. Full responsibility for the editorial process for this article was delegated to Federico M. Asch, MD.

CONFLICTS OF INTEREST

Drs. Akerman, Porumb, Hawkes, Woodward, and Upton are employed by Ultromics.

SUPPLEMENTARY DATA

Supplementary data related to this article can be found at https://doi.org/10.1016/j.echo.2024.06.016.

Claire Cassianni, MSc, Mayo Clinic Alix School of Medicine, Rochester, Minnesota
Geoffrey D. Huntley, MD, Department of Cardiovascular Medicine, Mayo Clinic, Rochester, Minnesota
Matteo Castrichini, MD, Department of Medicine, Mayo Clinic, Rochester, Minnesota
Ashley P. Akerman, PhD, Mihaela Porumb, PhD, Ultromics, Oxford, United Kingdom

Table Additional testing within 1 year of the qualifying echocardiographic study

	AI model prediction			
	Negative (*n* = 564)	Positive (*n* = 625)	Uncertain (*n* = 95)	*P*
H2FpEFF category, *n* (%)				<.0001*
Prediction negative	161 (28.5)	6 (1.0)	6 (6.3)	
Prediction positive	50 (8.9)	264 (42.2)	21 (22.1)	
Nondiagnostic	353 (62.6)	355 (56.8)	68 (71.6)	
HFA-PEFF category, *n* (%)				<.0001*
Prediction negative	292 (51.8)	27 (4.3)	26 (27.4)	
Prediction positive	20 (3.5)	207 (33.1)	11 (11.6)	
Nondiagnostic	252 (44.7)	391 (62.6)	58 (61.1)	
BNP, pg/mL				<.0001†
Median (IQR)	63.0 (22-148)	352 (200-636)	115 (27-211)	
n	49	114	10	
NT-proBNP, pg/mL				<.0001†
Median (IQR)	210 (74-572)	1,941 (697-5,866)	675 (206-3,306)	
n	87	284	26	
Exercise test workload, METs				
Mean ± SD	9.1 ± 2.8	7.4 ± 3.2	11.1 ± 2.2	.13‡
n	49	8	3	
Exercise test FAC, %				
Mean ± SD	100.8 ± 28.9	86.0 ± 28.0	128.0 ± 25.5	.11‡
n	43	7	3	

BNP, Brain natriuretic peptide; *FAC*, Functional aerobic capacity; *IQR*, interquartile range; *METs*, metabolic equivalents; *NT-proBNP*, N-terminal pro–brain natriuretic peptide.
*Chi-Square *P* value.
†Kruskal-Wallis *P* value.
‡Analysis of variance *P* value.

Christopher G. Scott, MS, Department of Quantitative Health Sciences, Mayo Clinic, Rochester, Minnesota
Halley N. Davison, MBA, Department of Cardiovascular Medicine, Mayo Clinic, Rochester, Minnesota
William Hawkes, PhD, Gary Woodward, PhD, Ultromics, Oxford, United Kingdom
Barry Borlaug, MD, Department of Cardiovascular Medicine, Mayo Clinic, Rochester, Minnesota
Ross Upton, PhD, Ultromics, Oxford, United Kingdom
Patricia A. Pellikka, MD, Department of Cardiovascular Medicine, Mayo Clinic, Rochester, Minnesota

REFERENCES

1. Heidenreich PA, Bozkurt B, Aguilar D, et al. 2022 AHA/ACC/HFSA guideline for the Management of heart failure: executive summary: a report of the American College of Cardiology/American Heart Association Joint Committee on clinical practice guidelines. J Am Coll Cardiol 2022;79:1757-80.
2. Selvaraj S, Myhre PL, Vaduganathan M, et al. Application of diagnostic algorithms for heart failure with preserved ejection fraction to the community. JACC Heart Fail 2020;8:640-53.
3. Verbrugge FH, Reddy YNV, Sorimachi H, et al. Diagnostic scores predict morbidity and mortality in patients hospitalized for heart failure with preserved ejection fraction. Eur J Heart Fail 2021;23:954-63.
4. 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;82:1936-48.
5. Akerman AP, Porumb M, Scott CG, et al. Automated echocardiographic detection of heart failure with preserved ejection fraction using artificial intelligence. JACC: Advances 2023;2:100452.
6. Reddy YNV, Carter RE, Obokata M, et al. A simple, evidence-based approach to help guide diagnosis of heart failure with preserved ejection fraction. Circulation 2018;138:861-70.
7. Pieske B, Tschope C, de Boer RA, et al. How to diagnose heart failure with preserved ejection fraction: the HFA-PEFF diagnostic algorithm: a consensus recommendation from the Heart Failure Association (HFA) of the European Society of Cardiology (ESC). Eur Heart J 2019;40:3297-317.

https://doi.org/10.1016/j.echo.2024.06.016