Diagnostic and prognostic Evaluation of an Echocardiography-based Artificial Intelligence algorithm for detecting HFpEF: A Case-Control Analysis

Katarina Yaros[1], Matthew Segar[2], Vinayak Subramanian[1], Alvin Chandra[1], Thomas Koshy[1], Ross Upton[3], Ashley Akerman[3], Ambarish Pandey[1]

[1]*UT Southwestern Medical Center, Dallas, TX*

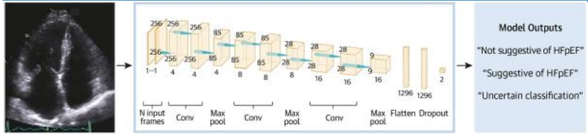
[2]*Texas Heart Institute, Houston, TX*

[3]*Ultromics Ltd, Oxford, United Kingdom*

Introduction

- Heart failure (HF) with preserved Ejection Fraction (HFpEF) is common among older adults an is associated with a high burden of morbidity and mortality.
- Establishing HFpEF diagnosis remains challenging and often requires left ventricular filling pressure assessment by invasive or non-invasive approaches.
- Echocardiography-based artificial intelligence (AI) HFpEF model that uses a 3-dimensional convolutional neural network (3D CNN) to detect HFpEF using a single 4-chamber clip from a resting echocardiogram has been cleared by the FDA.
- External validation of this algorithm against clinically adjudicated and confirmed HFpEF cases is limited.
- **Purpose:** To evaluate the diagnostic and prognostic performance of the echocardiography-based AI HFpEF model in a cohort of HFpEF patients and matched controls.

Methods

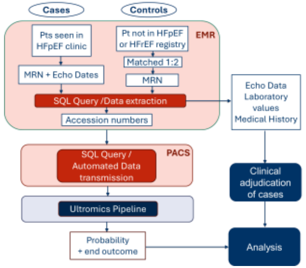


Adapted from Akerman AP, et al. JACC Adv. 2023;2(6):100452

Schematic design of AI model employed for HFpEF detection. 3D CNN model uses a single apical 4-chamber video clip.

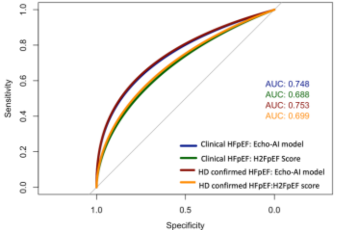
- **Cases:** Clinically adjudicated HFpEF
  - clinical history, signs and symptoms of HF
  - LV ejection fraction >45%
  - evidence of elevated filling pressures by resting (PCWP > 15 mm Hg) or exercise invasive hemodynamics (PCWP > 25 mm Hg) or echocardiogram (E/e' >14)
- **Controls**: Age, sex, and BMI-matched participants without HF and a normal echo within EHR.

Schematic of workflow. Clinical and echo data were extracted from EHR and automated transmission of echo data to Ultromics was used.



Echo-based AI HFpEF model can reliably identify HFpEF patients

Results



AUROC for predicting HFpEF outcome using the Echo AI model and H2FpEF score. In the matched cohorts of patients with clinically adjudicated HFpEF and matched control individuals (N = 122 each), the AI algorithm-based probability of HFpEF demonstrated good performance in identifying clinically adjudicated and hemodynamically confirmed HFpEF (AUROC: 0.75 for each) that was greater than H2FpEF score (0.69 and 0.70).

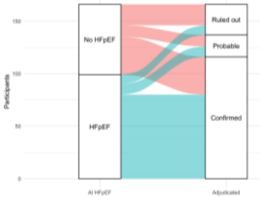
- Based on Youden's index, the AI algorithm-based probability threshold of >0.75 was identified as the optimal cutoff for detecting HFpEF by the AI algorithm, with high sensitivity (0.85) and accuracy (0.74) and, adequate specificity (0.66).

Statistical Analysis:

- The performance of the AI HFpEF model was evaluated using receiver operator curves.
- Among patients with clinically adjudicated HFpEF, the association of the AI-HFpEF phenotype with elevated resting/exercise PCWP and peak exercise oxygen uptake (VO2peak) was assessed using multivariable logistic and linear regression models adjusting for age, sex, race, ethnicity, and comorbidities (diabetes, hypertension, BMI, kidney disease, atrial fibrillation).

	Controls			Cases		
	Not Detected	Detected	P	Not Detected	Detected	P
n	100	22		67	99	
age (mean (SD))	70.02 (11.86)	70.74 (12.85)	0.800	66.84 (13.03)	73.87 (10.48)	<0.001
gender = MALE (%)	32 (32.0)	3 (13.6)	0.143	28 (41.8)	20 (20.2)	0.005
bmi (mean (SD))	34.86 (8.98)	36.50 (9.10)	0.452	38.07 (9.93)	34.48 (9.42)	0.020
afib = 1 (%)	15 (15.0)	5 (22.7)	0.570	23 (34.3)	56 (56.6)	0.008
dm = 1 (%)	32 (32.0)	7 (31.8)	1.000	37 (55.2)	57 (57.6)	0.888
ckd = 1 (%)	14 (14.0)	3 (13.6)	1.000	15 (22.4)	46 (46.5)	0.003
htn = 1 (%)	76 (76.0)	16 (72.7)	0.961	61 (91.0)	92 (92.9)	0.882
obesity = 1 (%)	82 (82.0)	16 (72.7)	0.487	59 (88.1)	80 (80.8)	0.304
cont_ultromics (median [IQR])	0.20 [0.11, 0.41]	0.90 [0.85, 0.94]	<0.001	0.17 [0.09, 0.35]	0.93 [0.90, 0.97]	<0.001
cat_ultromics = detectedHFpEF (%)	0 (0.0)	22 (100.0)	<0.001	0 (0.0)	99 (100.0)	<0.001

- Among 166 patients referred for evaluation of HFpEF, 82% had clinically adjudicated HFpEF, and 69.8% had elevated LV filling pressure at rest or exercise.



Alluvial Plot Demonstrating Reclassification of Patients Using Ultromics HFpEF AI Model and adjudicated outcomes. Patients were clinically adjudicated into No HFpEF, Probable HFpEF, and Confirmed HFpEF based on symptoms, imaging and laboratory findings.

Association between AI-HFpEF phenotype and invasive hemodynamics. In the HFpEF referral cohort, a higher probability of HFpEF based on the AI-algorithm was associated with lower VO2 peak (β [95% CI] per 5% higher probability: -0.11 [-0.21 to -0.01, P-value: 0.03] and greater odds of elevated PCWP (Odds ratio [95% CI] per 5% higher probability: 1.07 [1.01 – 1.15, P-value: 0.04] at rest or exercise after adjusting for age, sex, race, ethnicity, and other comorbidities.

	Model	
	Est (95% CI)	P
Categorical Ultromics		
Probable HFpEF	7.73 (3.72, 16.06)	<0.001
Confirmed HFpEF	9.11 (4.18, 19.84)	<0.001
Rest E/e	5.46 (3.93, 6.99)	<0.001
Continuous Ultromics (per 5% increase)		
Probable HFpEF	1.14 (1.08, 1.2)	<0.001
Confirmed HFpEF	1.14 (1.09, 1.21)	<0.001
Rest E/e	0.34 (0.24, 0.45)	<0.001

Discussion

Echo-based AI algorithm has potential to be clinically valuable tool in HFpEF diagnostic process given its discrimination performance against established clinical HF2PEF score. Its simple input makes it very appealing relative to the complex scoring systems.