



## OPEN Comparison of the 2021 ESC guideline with the HFA-PEFF and H<sub>2</sub>FPEF scores in diagnosing heart failure with preserved ejection fraction

Minjae Yoon<sup>1,2,4</sup>, Jaewon Oh<sup>3,4</sup>, Chan Joo Lee<sup>3</sup> & Seok-Min Kang<sup>3</sup>✉

The HFA-PEFF and H<sub>2</sub>FPEF scores were recently proposed to help diagnose heart failure (HF) with preserved ejection fraction (HFpEF). We aimed to evaluate HFpEF prevalence according to the 2021 European Society of Cardiology (ESC) HF guideline and the HFA-PEFF and H<sub>2</sub>FPEF scores in patients with unexplained dyspnea and to compare the concordance between these three algorithms in diagnosing HFpEF. We analyzed 992 patients with unexplained dyspnea suspected of having HFpEF, who underwent echocardiography, cardiopulmonary exercise testing, and N-terminal pro-brain natriuretic peptide measurement at a single tertiary center. Patients were classified as having confirmed, suspected, or no HFpEF on high, intermediate, or low scores according to HFA-PEFF or H<sub>2</sub>FPEF. Additionally, patients were classified into three categories according to elevated natriuretic peptide levels and echocardiographic parameters following the 2021 ESC HF guideline. Among the 992 patients included, confirmed prevalence HFpEF ranged from 28.5% (2021 ESC) to 21.1% (HFA-PEFF) and 4.7% (H<sub>2</sub>FPEF). Significant differences in the prevalence of hypertension, atrial fibrillation, number of antihypertensive medications, natriuretic peptide levels, and echocardiographic parameters of diastolic dysfunction were observed among the three HFpEF groups, with the highest burden in the group defined by the H<sub>2</sub>FPEF score. Comparing the HFA-PEFF and H<sub>2</sub>FPEF scores, 40.2% of patients were classified into different likelihood categories for HFpEF depending on the score used. The overlap of patients diagnosed with confirmed HFpEF according to the three algorithms was limited to 3.5%. HFpEF prevalence in patients with unexplained dyspnea varied significantly depending on the algorithm applied.

**Keywords** Heart failure with preserved ejection fraction, HFA-PEFF, H<sub>2</sub>FPEF, 2021 ESC heart failure guideline, unexplained dyspnea

The prevalence of heart failure (HF) with preserved ejection fraction (HFpEF) is increasing<sup>1–4</sup>. HFpEF accounts for approximately 50% of all patients with HF, with > 4% of the general population aged ≥ 60 years identified as having HFpEF<sup>3,5,6</sup>. However, diagnosing HFpEF remains challenging. There is no consensus on how to define HFpEF, with wide variation in diagnostic criteria across various society guidelines and clinical trials<sup>5,7–10</sup>. This lack of uniformity in disease definition may stem partly from an incomplete understanding of the pathobiology and phenotypic heterogeneity of HFpEF<sup>11</sup>. Invasive or exercise hemodynamic assessments have emerged as the gold standard for identifying HFpEF. However, their high cost, procedural risk, and technical complexity make them impractical for routine evaluation and preclude their use in large clinical trials<sup>7,12,13</sup>.

The HFA-PEFF and H<sub>2</sub>FPEF scores are two recently proposed, score-based algorithms designed to aid in the diagnosis of HFpEF<sup>14,15</sup>. The H<sub>2</sub>FPEF score was derived from patients with unexplained dyspnea who were referred for invasive hemodynamic exercise testing and includes four clinical and two echocardiographic

<sup>1</sup>Division of Cardiology, Department of Internal Medicine, Seoul National University Bundang Hospital, Seongnam, Republic of Korea. <sup>2</sup>Department of Medicine, Yonsei University College of Medicine, Seoul, Republic of Korea. <sup>3</sup>Division of Cardiology, Department of Internal Medicine, Severance Hospital, Yonsei University College of Medicine, 50-1, Yonsei-Ro, Seodaemun-gu, Seoul 03722, Republic of Korea. <sup>4</sup>Minjae Yoon and Jaewon Oh contributed equally to this work. ✉email: smkang@yuhs.ac

items<sup>14</sup>. More recently, the Heart Failure Association of the European Society of Cardiology (ESC) published a separate diagnostic algorithm, the HFA-PEFF score, which contains three domains: functional, morphological, and biomarker<sup>15</sup>. Additionally, the 2021 ESC HF guideline diagnoses HFpEF by incorporating symptoms and signs of HF, echocardiographic parameters, and natriuretic peptides (NP); this approach does not represent new diagnostic score but rather a simplified approach<sup>8</sup>.

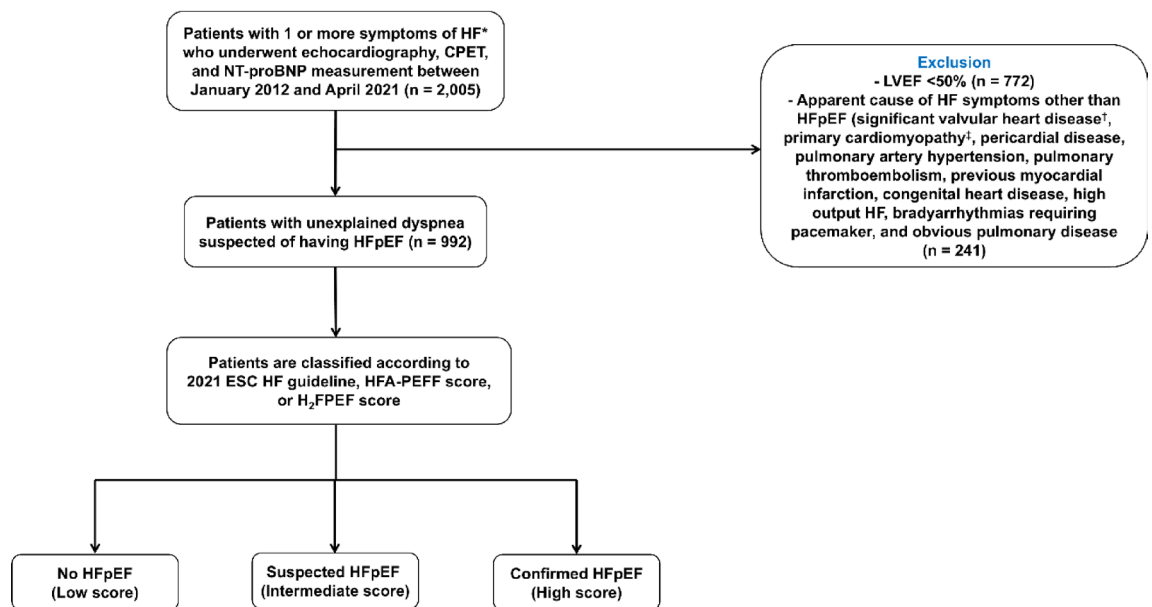
Several recent studies have evaluated the diagnostic performance of the HFA-PEFF and H<sub>2</sub>FPEF scores in previous HFpEF trials and registries, demonstrating the generalizability and validity of these scores<sup>16–22</sup>. However, some studies found limited sensitivities of these two scores for the diagnosis of HFpEF<sup>17,21</sup> and significant discrepancies in classifying patients with suspected HFpEF<sup>23</sup>. Additionally, discordance between the ESC HF guideline algorithm and HFA-PEFF and H<sub>2</sub>FPEF scores in diagnosing HFpEF has been reported<sup>24</sup>. Therefore, we aimed to evaluate HFpEF prevalence using the three different algorithms in patients with unexplained dyspnea and to compare their concordance in diagnosing HFpEF (Fig. 1).

## Results

### Baseline characteristics according to three HFpEF diagnostic algorithms

Among the 992 patients with unexplained dyspnea suspected of having HFpEF, the mean age was 60.9 years, and 63.7% were female (Table 1). The overall prevalence rates of hypertension, diabetes mellitus, and atrial fibrillation (AF) were 72.6%, 21.6%, and 6.6%, respectively. The mean body mass index was 24.6 kg/m<sup>2</sup> and mean systolic blood pressure was 122.6 mmHg. The median N-terminal pro-B-type natriuretic peptide (NT-proBNP) level was 71 pg/mL and mean number of antihypertensive medications was 1.3.

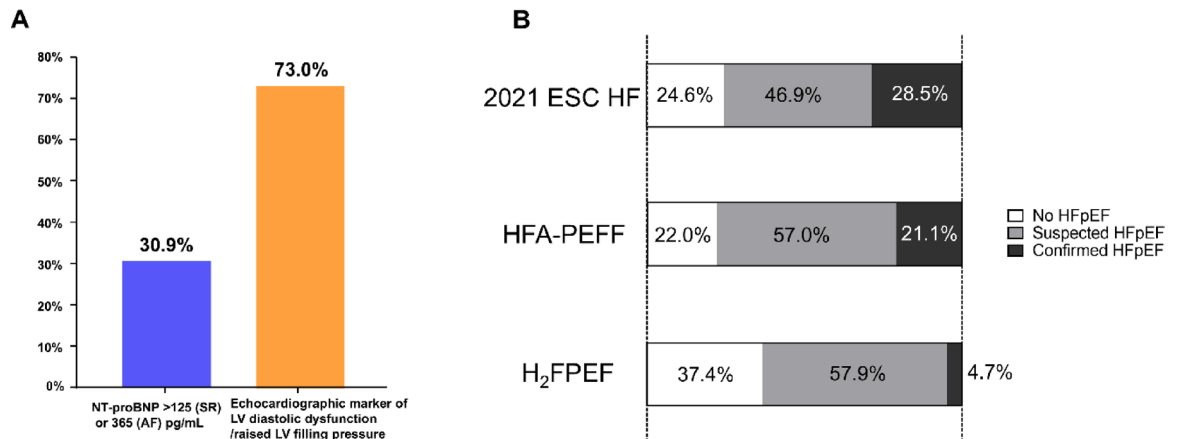
According to the 2021 ESC HF guideline, 724 (73.0%) patients had an echocardiographic marker of left ventricular (LV) diastolic dysfunction or raised LV filling pressure, and 307 (30.9%) had elevated NT-proBNP levels (Fig. 2A). Thus, 283 (28.5%), 465 (46.9%), and 244 (24.6%) patients were classified as having confirmed HFpEF, suspected HFpEF, and no HFpEF, respectively, according to the 2021 ESC HF guideline (Fig. 2B). According to the HFA-PEFF and H<sub>2</sub>FPEF scores, 209 (21.1%), 565 (57.0%), and 218 (22.0%) patients and 47 (4.7%), 574 (57.9%), and 371 (37.4%) patients were classified as having confirmed HFpEF, suspected HFpEF, and no HFpEF, respectively. The baseline characteristics of patients with confirmed HFpEF using the three different diagnostic algorithms are shown in Table 1. Hypertension and AF prevalence, number of antihypertensive medications, NT-proBNP, left atrial volume index (LAVI), e' velocity, E/e', peak tricuspid regurgitation (TR) velocity, and pulmonary artery systolic pressure (PASP) were significantly different among the three confirmed



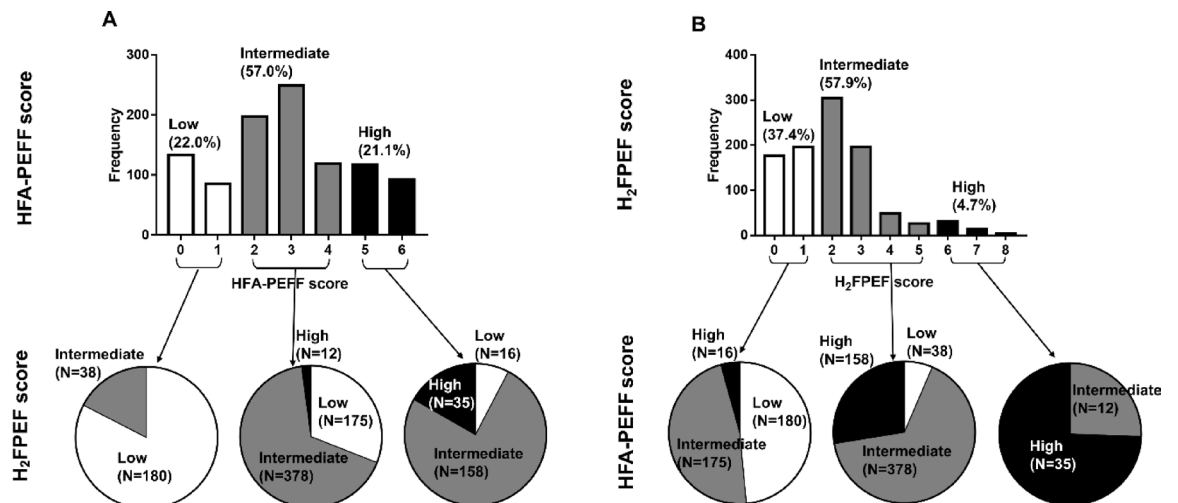
**Fig. 1.** Patients enrollment algorithm. According to the 2021 ESC HF guideline, patients who had symptoms of HF and LVEF  $\geq 50\%$ , along with both elevated NT-proBNP and echocardiographic markers of LV diastolic dysfunction or raised LV filling pressure, were classified as having confirmed HFpEF. Patients meeting one of these criteria were classified as having suspected HFpEF, while those meeting neither were classified as having no HFpEF. Based on the HFA-PEFF or H<sub>2</sub>FPEF scores, patients with low, intermediate, or high scores were classified as having no HFpEF, suspected HFpEF, or confirmed HFpEF, respectively. \*Symptoms of HF: dyspnea on exertion, paroxysmal nocturnal dyspnea, or orthopnea. †Significant valvular heart disease: greater than mild stenosis, greater than moderate regurgitation, or valve surgery. ‡Primary cardiomyopathy: hypertrophic cardiomyopathy, restrictive cardiomyopathy, or arrhythmogenic right ventricular cardiomyopathy. CPET cardiopulmonary exercise testing, ESC European Society of Cardiology, HF heart failure, HFpEF heart failure with preserved ejection fraction, LV left ventricular, LVEF left ventricular ejection fraction, NT-proBNP N-terminal pro-B type natriuretic peptide.

	All patients ( <i>n</i> = 992)	Confirmed HFpEF by 2021 ESC HF guideline ( <i>n</i> = 283)	Confirmed HFpEF by HFA-PEFF (score $\geq 5$ ) ( <i>n</i> = 209)	Confirmed HFpEF by H <sub>2</sub> FPEF (score $\geq 6$ ) ( <i>n</i> = 47)	<i>P</i> value across three groups
Age, years	60.9 $\pm$ 13.9	67.6 $\pm$ 10.7	67.5 $\pm$ 10.8	70.5 $\pm$ 8.4	0.182
Female	632 (63.7)	185 (65.4)	140 (67.0)	25 (53.2)	0.196
BMI, kg/m <sup>2</sup>	24.6 $\pm$ 3.8	24.5 $\pm$ 3.9	24.7 $\pm$ 4.1	25.4 $\pm$ 4.8	0.369
Systolic blood pressure, mmHg	122.6 $\pm$ 16.3	126.4 $\pm$ 17.7	127.7 $\pm$ 17.8	121.0 $\pm$ 16.8	0.065
Heart rate, beat/min	73.5 $\pm$ 12.9	71.4 $\pm$ 14.0	70.2 $\pm$ 13.5	71.4 $\pm$ 15.9	0.628
<b>Comorbidities</b>					
Hypertension	720 (72.6)	256 (90.5)	194 (92.8)	47 (100.0)	0.048
Diabetes mellitus	214 (21.6)	87 (30.7)	66 (31.6)	15 (31.9)	0.974
Dyslipidemia	495 (49.9)	156 (55.1)	120 (57.4)	24 (51.1)	0.706
Chronic kidney disease	101 (10.2)	61 (21.6)	49 (23.4)	7 (14.9)	0.436
Atrial fibrillation	65 (6.6)	43 (15.2)	37 (17.7)	44 (93.6)	<0.001
Pulmonary disease*	70 (7.1)	17 (6.0)	10 (4.8)	3 (6.4)	0.811
Stroke	79 (8.0)	36 (12.7)	27 (12.9)	7 (14.9)	0.919
<b>Medication</b>					
Number of antihypertensive medications	1.3 $\pm$ 1.3	2.0 $\pm$ 1.4	2.2 $\pm$ 1.4	2.8 $\pm$ 0.9	0.001
RAS inhibitor (ACEi/ARB)	398 (40.1)	171 (60.4)	135 (64.6)	40 (85.1)	0.005
Beta-blocker	259 (26.1)	123 (43.5)	100 (47.8)	31 (66.0)	0.016
Calcium channel blocker	274 (27.6)	112 (39.6)	88 (42.1)	22 (46.8)	0.610
Diuretics	344 (34.7)	161 (56.9)	123 (58.9)	39 (83.0)	0.003
Other antihypertensive medications	10 (1.0)	7 (2.5)	7 (3.3)	0 (0)	0.553
<b>Laboratory data</b>					
Serum creatinine level, mg/dL	0.77 (0.65–0.95)	0.81 (0.68–1.06)	0.81 (0.70–1.27)	0.82 (0.73–1.06)	0.927
eGFR, ml/min/1.73 m <sup>2</sup>	87.7 $\pm$ 20.7	76.9 $\pm$ 24.7	75.5 $\pm$ 26.1	76.9 $\pm$ 21.1	0.818
Hemoglobin level, g/dL	13.4 $\pm$ 1.8	12.8 $\pm$ 1.8	12.7 $\pm$ 1.9	13.7 $\pm$ 1.9	0.003
NT-proBNP, pg/mL	71 (33–155)	284 (172–640)	362 (246–785)	706 (401–1420)	<0.001
<b>Echocardiographic parameters</b>					
LVEF, %	66.8 $\pm$ 6.3	66.1 $\pm$ 7.4	66.2 $\pm$ 7.6	66.3 $\pm$ 8.1	0.986
LVEDD, mm	47.8 $\pm$ 4.6	48.7 $\pm$ 4.9	49.3 $\pm$ 4.9	48.7 $\pm$ 5.3	0.331
LVESD, mm	31.1 $\pm$ 4.2	31.8 $\pm$ 4.7	32.2 $\pm$ 4.9	31.8 $\pm$ 5.3	0.720
IVSd, mm	9.3 $\pm$ 1.4	9.8 $\pm$ 1.4	10.0 $\pm$ 1.5	9.8 $\pm$ 1.5	0.430
PWd, mm	9.2 $\pm$ 1.3	9.7 $\pm$ 1.3	9.8 $\pm$ 1.4	9.7 $\pm$ 1.3	0.559
Relative wall thickness	0.39 $\pm$ 0.06	0.40 $\pm$ 0.10	0.40 $\pm$ 0.07	0.40 $\pm$ 0.07	0.961
LVMI, g/m <sup>2</sup>	91.2 $\pm$ 22.1	103.2 $\pm$ 24.9	107.4 $\pm$ 25.9	99.9 $\pm$ 26.3	0.086
LAVI, mL/m <sup>2</sup>	30.1 $\pm$ 10.7	38.3 $\pm$ 13.4	41.7 $\pm$ 13.5	50.7 $\pm$ 16.7	<0.001
E-wave velocity, m/s	0.63 $\pm$ 0.17	0.69 $\pm$ 0.21	0.70 $\pm$ 0.22	0.86 $\pm$ 0.23	<0.001
Septal e' velocity, cm/s	6.5 $\pm$ 2.4	5.5 $\pm$ 1.7	5.3 $\pm$ 1.6	6.0 $\pm$ 1.7	0.014
E/e'	10.7 $\pm$ 4.1	13.3 $\pm$ 5.1	13.9 $\pm$ 5.2	15.2 $\pm$ 5.3	0.043
Peak TR velocity, m/s ( <i>n</i> = 913)	2.26 $\pm$ 0.35	2.42 $\pm$ 0.40	2.46 $\pm$ 0.44	2.61 $\pm$ 0.37	0.018
PASP, mmHg ( <i>n</i> = 913)	26.1 $\pm$ 7.1	29.6 $\pm$ 9.2	30.6 $\pm$ 9.9	34.2 $\pm$ 9.7	0.007

**Table 1.** Baseline characteristics of patients according to three HFpEF diagnostic algorithms. Values are expressed as number (%), mean  $\pm$  standard deviation, or median (interquartile range). *ACEi* angiotensin-converting enzyme inhibitor, *ARB* angiotensin receptor blocker, *BMI* body mass index, *e'* early diastolic mitral annular velocity, *E* peak velocity of early diastolic trans-mitral flow, *eGFR* estimated glomerular filtration rate, *ESC* European Society of Cardiology, *HF* heart failure, *HFpEF* heart failure with preserved ejection fraction, *IVSd* interventricular septal end-diastolic dimension, *LAVI* left atrial volume index, *LVEDD* left ventricle end-diastolic dimension, *LVEF* left ventricular ejection fraction, *LVESD* left ventricle end-systolic dimension, *LVMI* left ventricular mass index, *NT-proBNP* N-terminal pro-brain natriuretic peptide, *PASP* pulmonary artery systolic pressure, *PWd* posterior wall end-diastolic dimension, *RAS* renin-angiotensin system, *TR* tricuspid regurgitation. \*Pulmonary disease is defined as interstitial lung disease, chronic obstructive pulmonary disease or asthma that is not severe enough to explain current dyspnea.



**Fig. 2.** Prevalence of HFpEF according to the 2021 ESC HF guideline, HFA-PEFF score, and H<sub>2</sub>FPEF score in patients with unexplained dyspnea. **(A)** Prevalence of elevated NT-proBNP and echocardiographic marker of LV diastolic dysfunction/raised LV filling pressure according to 2021 ESC HF guideline, **(B)** prevalence of HFpEF according to three categories. *AF* atrial fibrillation, *ESC* European Society of Cardiology, *HF* heart failure, *HFpEF* heart failure with preserved ejection fraction, *LV* left ventricular, *SR* sinus rhythm.



**Fig. 3.** Reclassification of patients into likelihood categories (low, intermediate, and high) from the HFA-PEFF score to the H<sub>2</sub>FPEF score **(A)** and vice versa **(B)**. White, low likelihood (no HFpEF); light gray, intermediate likelihood (suspected HFpEF); dark gray, high likelihood of HFpEF (confirmed HFpEF), estimated by each score. *HFpEF* heart failure with preserved ejection fraction.

HFpEF groups according to each diagnostic algorithm, with the highest values observed in those classified as confirmed HFpEF according to the H<sub>2</sub>FPEF score.

### Reclassification of HFpEF according to 2021 ESC HF guideline, HFA-PEFF, and H<sub>2</sub>FPEF scores

The reclassification of patients into likelihood categories (low, intermediate, and high) from the HFA-PEFF score to the H<sub>2</sub>FPEF score, or vice versa, is presented in Fig. 3; Table 2. Among the 209 patients with confirmed HFpEF according to the HFA-PEFF score, only 35 (16.7%) were reclassified as having confirmed HFpEF according to the H<sub>2</sub>FPEF score, whereas 158 (75.6%) and 16 (7.7%) were reclassified as having suspected and no HFpEF, respectively. Conversely, among the 47 patients with confirmed HFpEF according to the H<sub>2</sub>FPEF score, 35 (74.5%) and 12 (25.5%) were recategorized as having confirmed and suspected HFpEF according to the H<sub>2</sub>FPEF score, respectively. Among the 565 patients with suspected HFpEF based on the HFA-PEFF score, 175 (31.0%) were reclassified as having no HFpEF based on the H<sub>2</sub>FPEF score. Among 371 patients with no HFpEF, 191 (51.5%) were categorized as having confirmed HFpEF or suspected HFpEF according to the H<sub>2</sub>FPEF score. In summary, 399 (40.2%) patients were classified into different likelihood categories, depending on the score used.

HFA-PEFF score	H <sub>2</sub> FPEF score			Total
	Confirmed HFpEF (high score)	Suspected HFpEF (intermediate score)	No HFpEF (low score)	
Confirmed HFpEF (high score)	35 (3.5%)	158 (15.9%)	16 (1.6%)	209 (21.0%)
Suspected HFpEF (intermediate score)	12 (1.2%)	378 (38.1%)	175 (17.6%)	565 (57.0%)
No HFpEF (low score)	0 (0%)	38 (3.8%)	180 (18.1%)	218 (22.0%)
Total	47 (4.7%)	574 (57.9%)	371 (37.4%)	992 (100%)
2021 ESC HF	HFA-PEFF score			Total
	Confirmed HFpEF (high score)	Suspected HFpEF (intermediate score)	No HFpEF (low score)	
Confirmed HFpEF	208 (21.0%)	75 (7.6%)	0 (0%)	283 (28.5%)
Suspected HFpEF	1 (0.1%)	404 (40.7%)	60 (6.0%)	465 (46.9%)
No HFpEF	0 (0%)	86 (8.7%)	158 (15.9%)	244 (24.6%)
Total	209 (21.1%)	565 (57.0%)	218 (22.0%)	992 (100%)
2021 ESC HF	H <sub>2</sub> FPEF score			Total
	Confirmed HFpEF (high score)	Suspected HFpEF (intermediate score)	No HFpEF (low score)	
Confirmed HFpEF	36 (3.6%)	217 (21.9%)	30 (3.0%)	283 (28.5%)
Suspected HFpEF	11 (1.1%)	315 (31.8%)	139 (14.0%)	465 (46.9%)
No HFpEF	0 (0%)	42 (4.2%)	203 (20.5%)	244 (24.6%)
Total	47 (4.7%)	574 (57.9%)	371 (37.4%)	992 (100%)

**Table 2.** Reclassification table of HFpEF prevalence using 2021 ESC HF guideline and the HFA-PEFF and H<sub>2</sub>FPEF scores. *ESC* European Society of Cardiology, *HF* heart failure, *HFpEF* heart failure with preserved ejection fraction.

Table 2 shows the reclassification and concordance of the HFpEF prevalence using the three diagnostic algorithms. The concordance between the HFA-PEFF and H<sub>2</sub>FPEF scores, when comparing the three HFpEF categories, was moderate (Cohen's weighted kappa = 0.47). The concordance between the 2021 ESC HF guideline and HFA-PEFF scores was good (Cohen's weighted kappa = 0.77), while the concordance between the 2021 ESC HF guideline and H<sub>2</sub>FPEF scores was moderate (Cohen's weighted kappa = 0.45).

The distribution of H<sub>2</sub>FPEF scores according to different body mass index (BMI) cut-off values is shown in Fig. S1. Among the total of 992 patients, 72 (7.3%), 234 (23.6%), and 420 (42.3%) had BMI of  $\geq 30$ ,  $\geq 27$ , or  $\geq 25$  kg/m<sup>2</sup>, respectively. Using a lower BMI cut-off, the prevalence of confirmed HFpEF and suspected HFpEF according to the H<sub>2</sub>FPEF score revealed a slightly increasing trend, while the prevalence of no HFpEF revealed a decreasing trend. The concordance between the H<sub>2</sub>FPEF and HFA-PEFF scores and the H<sub>2</sub>FPEF score and the 2021 ESC HF guideline at the BMI cut-off of 27–25 kg/m<sup>2</sup> was moderate (Cohen's weighted kappa = 0.49 and 0.47 for 27 kg/m<sup>2</sup> and 0.48 and 0.44 for 25 kg/m<sup>2</sup>, respectively), indicating that concordance did not significantly change with different BMI cut-offs.

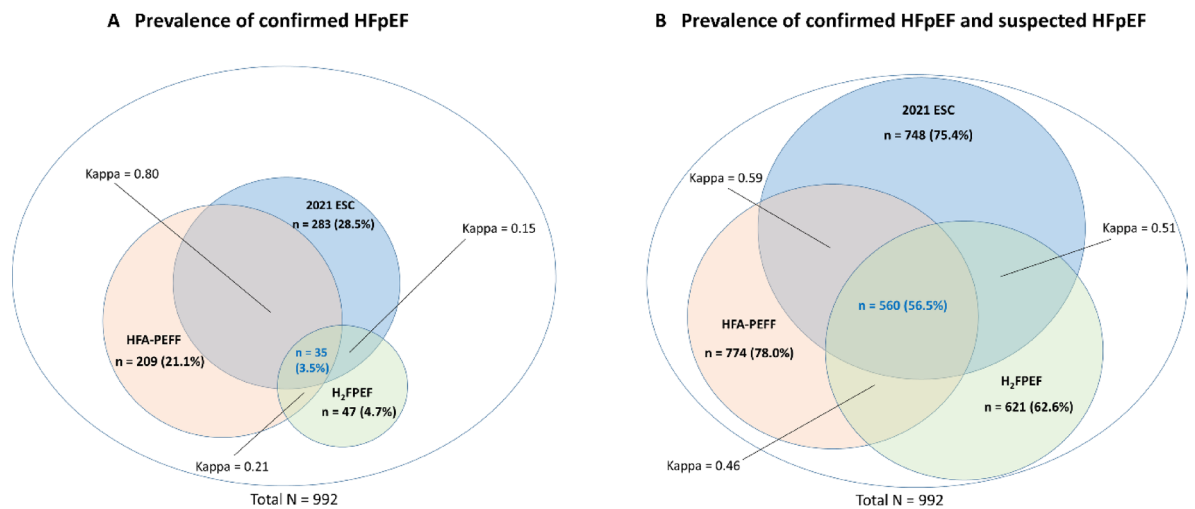
### Concordance of three HFpEF diagnostic algorithms

The prevalence and concordance of confirmed HFpEF in patients with unexplained dyspnea, according to the three HFpEF diagnostic algorithms are shown in Fig. 4A. Among the 992 patients with unexplained dyspnea, the prevalence of confirmed HFpEF verified: 28.5% ( $n = 283$ ) based on the 2021 ESC HF guideline, 21.1% ( $n = 209$ ) based on the HFA-PEFF, and 4.7% ( $n = 47$ ) according to the H<sub>2</sub>FPEF. The concordance of confirmed HFpEF was highest between the 2021 ESC HF guideline and HFA-PEFF score, reflected by Cohen's unweighted kappa coefficient of 0.80 (good) and a reclassification rate of 7.3%. In contrast, the concordance between 2021 ESC guideline and H<sub>2</sub>FPEF score, as well as between the HFA-PEFF score and H<sub>2</sub>FPEF score, was poor and fair, with Cohen's unweighted kappa coefficient of 0.15 and 0.21, respectively. The overlap of participants diagnosed with confirmed HFpEF according to the three algorithms was limited to 3.5%.

The prevalence of combined confirmed and suspected HFpEF was 78.0% ( $n = 774$ ) for HFA-PEFF, 75.4% ( $n = 749$ ) for the 2021 ESC HF guideline, and 62.6% ( $n = 621$ ) for the H<sub>2</sub>FPEF (Fig. 4B). The reclassification of HFpEF by incorporating the confirmed and suspected HFpEF categories into a single "possible HFpEF" category is shown in Table S1. The concordance between the 2021 ESC and HFA-PEFF scores, 2021 ESC and H<sub>2</sub>FPEF scores, and HFA-PEFF and H<sub>2</sub>FPEF scores was moderate with Cohen's unweighted kappa value of 0.59, 0.51, and 0.46, respectively.

### Discussion

This study applied and compared three current algorithms for diagnosing HFpEF in patients with unexplained dyspnea. We found that the prevalence of HFpEF significantly varies depending on the diagnostic algorithm applied, with the prevalence of confirmed HFpEF ranging from 28.5% (2021 ESC guideline) to 21.1% (HFA-PEFF) and 4.7% (H<sub>2</sub>FPEF). While 40.2% of the patients were classified into different likelihood categories when comparing the HFA-PEFF and H<sub>2</sub>FPEF scores, only 3.5% of patients had a consistent confirmed HFpEF according to the three algorithms. Second, each algorithm identified different HFpEF subgroups, characterized by differing comorbidities and echocardiographic parameters related to diastolic dysfunction.



**Fig. 4.** Prevalence and concordance of HFpEF according to the three HFpEF algorithms in patients with unexplained dyspnea. **(A)** Of the 992 patients with unexplained dyspnea, the prevalence of confirmed HFpEF ranged from 28.5% ( $n = 283$ , 2021 ESC HF guideline) to 21.1% ( $n = 209$ , HFA-PEFF) and 4.7% ( $n = 47$ , H<sub>2</sub>FPEF). **(B)** The prevalence of combined confirmed HFpEF and suspected HFpEF ranged from 78.0% ( $n = 774$ , HFA-PEFF) to 75.4% ( $n = 749$ , 2021 ESC HF guideline) and 62.6% ( $n = 621$ , H<sub>2</sub>FPEF). ESC European Society of Cardiology, HF heart failure, HFpEF heart failure with preserved ejection fraction.

Notably, recent studies demonstrated that sodium-glucose cotransporter 2 inhibitors reduce cardiovascular events in patients with HFpEF<sup>25,26</sup>. Therefore, early and accurate diagnosis of HFpEF may help reduce its medical and economic burden by allowing the timely initiation of proper medications. However, diagnosing HFpEF remains challenging. Some patients are overdiagnosed with HFpEF, as various conditions share similar symptoms, mainly in patients with many comorbidities<sup>11</sup>. However, in other cases, HFpEF may be missed<sup>27</sup>. Moreover, no clear consensus exists on how HFpEF should be defined, and invasive or exercise-based hemodynamic assessments are difficult to use as routine evaluations in real clinical practice<sup>7,11</sup>. Thus, some score-based algorithms, such as HFpEF, HFA-PEFF, and H<sub>2</sub>FPEF scores, have been proposed to aid in HFpEF diagnosis. These tools represent a shift from the classical binary diagnostic approach to one based on the likelihood estimation of HFpEF.

However, significant discrepancies have been reported in the classification of patients with suspected HFpEF when using the two novel scores and the 2021 ESC HF guidelines<sup>23,24</sup>. In our study, the prevalence and characteristics of confirmed HFpEF were significantly different across the three diagnostic algorithms, ranging from 4.7 to 28.5%. Although the overall proportion of confirmed HFpEF among patients with unexplained dyspnea was similar to a previous study, there were slight differences regarding which scores had the highest prevalence of HFpEF<sup>24</sup>. Additionally, when using either the HFA-PEFF or H<sub>2</sub>FPEF scores, 40.2% of patients were reclassified into different likelihood categories of HFpEF depending on the score used. Additionally, the overlap of confirmed HFpEF cases across the three different algorithms was minimal, at 3.5%. This highlights how different diagnostic scores can result in different patient groups being classified as having HFpEF, being excluded from HFpEF, or being referred for further testing. This variability limits the clinical applicability of these diagnostic algorithms and reflects the continuing diagnostic uncertainty surrounding HFpEF.

The composition and elements of each diagnostic algorithm differ significantly. The H<sub>2</sub>FPEF score includes four clinical comorbidities and two echocardiographic items, whereas the HFA-PEFF score and 2021 ESC guideline focus more on echocardiographic structural and functional parameters, along with NP levels. Thus, AF, hypertension, and the number of antihypertensive medications emerged as the main factors driving the discrepancy between the algorithms, as they were key items in the H<sub>2</sub>FPEF score. In contrast, AF tends to raise the thresholds for HFpEF in the HFA-PEFF score and 2021 ESC HF guideline. In addition, patients with confirmed HFpEF according to the H<sub>2</sub>FPEF score had the highest NP levels and more advanced echocardiographic markers of diastolic dysfunction, presumably because the H<sub>2</sub>FPEF score applied the most stringent criteria. Additionally, the HFA-PEFF score and 2021 ESC HF guideline revealed relatively high agreement and concordance in our analysis, possibly due to an overlap in the items containing echocardiographic parameters.

Currently, the two newer diagnostic scores for HFpEF are not widely adopted in real-world clinical practice due to the large number of variables they involve, and they are not the primary recommendation in the 2021 ESC HF guideline due to their complexity<sup>8</sup>. Additionally, both scores have the limitation of classifying a large proportion of patients as intermediate likelihood, requiring further diastolic stress testing. Aizpurua et al. reported that 36% of patients with known HFpEF fell into the intermediate likelihood category according to the HFA-PEFF score, implying that over 50% of patients with unexplained dyspnea would have to undergo further testing before HFpEF could be diagnosed or excluded<sup>18</sup>. To date, HFpEF diagnosis is challenging and fraught with obstacles and inconsistencies. Future studies will help improve these diagnostic algorithms to increase their accuracy and facilitate their use in real-world clinical practice.

Although patients with HFpEF in Western countries are predominantly obese or overweight, obesity is less common in Asian patients<sup>28,29</sup>. Particularly, the DELIVER study, which evaluated the effect of dapagliflozin in HF with mildly reduced or preserved ejection fraction, reported that Asian and non-Asian patients had mean BMIs of 25.1 and 31.0 kg/m<sup>2</sup>, respectively<sup>30</sup>. In our study, the mean BMI of patients with confirmed HFpEF according to the 2021 ESC HF guideline, HFA-PEFF or H<sub>2</sub>FPEF scores was approximately 25 kg/m<sup>2</sup>, which is consistent with previous data. Although the H<sub>2</sub>FPEF score uses a BMI cut-off of 30 kg/m<sup>2</sup><sup>14</sup>, the cohort from which this score was developed is primarily included patients from North America. Although previous studies have evaluated the diagnostic performance and validity of the HFA-PEFF and H<sub>2</sub>FPEF scores in Asian populations<sup>20–22</sup>, the H<sub>2</sub>FPEF score may not accurately reflect the characteristics of Asian HFpEF patients with relatively low BMI. Here, the use of different BMI cut-offs slightly changed HFpEF prevalence but did not significantly alter the concordance between HFpEF diagnostic algorithms. Further studies on HFpEF and BMI cut-off values in Asian populations are warranted.

### Limitation

This study had some limitations. First, as this study was a retrospective analysis of a single-center cohort composed almost entirely of Korean individuals, the generalizability of our findings is limited. Additionally, the patients in this study were limited to those who underwent echocardiography, cardiopulmonary exercise testing (CPET), and NP measurements, which limits the interpretation and significance of our findings. Nevertheless, by including patients who underwent CPET, we believe our analysis more accurately represented patients with suggestive HFpEF. Second, we lacked echocardiographic data, including LV global longitudinal strain (GLS), or lateral e', which are components of the HFA-PEFF score. Therefore, we could not use these variables in our calculations, possibly resulting in a lower HFpEF prevalence. However, in an HFA-PEFF validation study, the inclusion of global longitudinal strain as a criterion resulted in only a minimal reclassification of individuals<sup>18</sup>. Third, 79 patients (8.0%) had missing data for TR velocity and PASP, and these patients were assumed to have a normal value for TR velocity and PASP. This assumption may have resulted in an underestimation of HFpEF prevalence. Fourth, we were unable to present data on confirmatory tests for HFpEF, such as diastolic stress echocardiography or invasive catheterization, as these procedures were rarely performed at our institution. Consequently, we could not definitively confirm whether the patients truly had HFpEF. Instead, we estimated the likelihood using diagnostic scores and categorized the patients into three groups. Therefore, the diagnostic accuracy of the three algorithms in identifying true HFpEF could not be assessed, and some cases may have been misclassified. Fifth, in our analysis, patients were classified into confirmed HFpEF, suspected HFpEF, and no HFpEF, according to the 2021 ESC HF guideline. This classification does not strictly align with the 2021 ESC HF guideline, which states that an elevated NP is not mandatory for the diagnosis of HFpEF. However, for simplicity in analysis, patients with both an elevated NP and an echocardiographic marker of LV diastolic dysfunction or raised LV filling pressure were classified as having confirmed HFpEF in this study. Sixth, NP levels can be influenced by various comorbidities (e.g., anemia and impaired renal function), medications, and the timing of measurement in relation to treatment. As this study was not prospectively designed, NP could not be measured in a standardized manner. However, we included patients who had NT-proBNP, transthoracic echocardiography, and CPET data available within a 3-month interval. In addition, we were unable to comprehensively assess or categorize the nature and severity of patients' HF symptoms. Therefore, we could not analyze potential differences in HF symptom profiles across the various HFpEF classification criteria, which represents another limitation of the study.

Despite these limitations, this study has certain strengths, including the large number of participants suspected of having HFpEF with available data on echocardiography and NP levels. Therefore, our study provides important insights into the concordance of three diagnostic algorithms for HFpEF in patients with unexplained dyspnea.

### Conclusions

Our study provides new data on the application of the 2021 ESC HF guidelines, the HFA-PEFF score, and the H<sub>2</sub>FPEF score for diagnosing HFpEF. The prevalence of HFpEF in patients with unexplained dyspnea largely varied significantly depending on the algorithm applied. The differences in HFpEF definitions across the guideline and various scoring systems have led to considerable differences in the stratification of HFpEF.

### Methods

#### Study population

We retrospectively analyzed 2005 patients with one or more symptoms of HF who underwent echocardiography, CPET, and NT-proBNP measurement between January 2012 and April 2021 at a single tertiary center (Fig. 1). Symptoms of HF included dyspnea upon exertion, paroxysmal nocturnal dyspnea, or orthopnea. We excluded patients with a left ventricular ejection fraction (LVEF) < 50% or an apparent cause of HF symptoms other than HFpEF, such as significant valvular heart disease, primary cardiomyopathy, pericardial disease, pulmonary artery hypertension, pulmonary thromboembolism, previous myocardial infarction, congenital heart disease, high-output HF, bradyarrhythmias requiring a pacemaker, or evident pulmonary disease<sup>31</sup>. Significant valvular heart disease was defined as valvular stenosis greater than mild or valvular regurgitation greater than moderate. Finally, 992 participants with unexplained dyspnea suspected of having HFpEF were included in this study.

This study was approved by the Institutional Review Board of the Yonsei University Health System (4-2021-0605) and was conducted in accordance with the tenets of the Declaration of Helsinki. The requirement for informed consent was waived by the Institutional Review Board of the Yonsei University Health System because of the retrospective nature of the study, and personal identifying information was removed after cohort creation.

## 2021 ESC HF guideline and HFA-PEFF and H<sub>2</sub>FPEF scores

The HFA-PEFF score contains minor and major criteria across three domains: functional (septal or lateral  $e'$  velocity,  $E/e'$  ratio, TR velocity or PASP, and LV GLS, morphological (rhythm-specific LAVI, sex-specific measures of LV mass index (LVMI), LV wall thickness, and relative wall thickness) and biomarker (rhythm-specific NP) domains<sup>15</sup>. Within each domain, a major criterion scores 2 points and a minor criterion scores 1 point. Each domain can contribute a maximum of 2 points if any major criterion is met, or 1 point if no major criterion but any minor criterion is met. Due to a lack of echocardiographic data on lateral  $e'$  and GLS, these parameters were not included in the calculation of the HFA-PEFF score for the current study.

The H<sub>2</sub>FPEF score is the sum of four clinical variables—BMI > 30 kg/m<sup>2</sup> [2 points], treatment with ≥ 2 antihypertensive medicines [1 point], AF [3 points], and age > 60 years [1 point]—and two echocardiographic items:  $E/e'$  ratio > 9 [1 point] and PASP > 35 mmHg [1 point]<sup>14</sup>.

The HFA-PEFF and H<sub>2</sub>FPEF scores estimate the likelihood of HFpEF and categorize patients suspected of having HFpEF into low, intermediate, or high likelihood of HFpEF<sup>14,15,31</sup>. Low scores (HFA-PEFF or H<sub>2</sub>FPEF ≤ 1) are considered to exclude HFpEF. Intermediate scores (H<sub>2</sub>FPEF score of 2–5 or HFA-PEFF score of 2–4) indicate an intermediate likelihood of HFpEF, wherein case further diastolic stress testing is needed to confirm HFpEF. High scores (H<sub>2</sub>FPEF ≥ 6 or HFA-PEFF ≥ 5) are considered highly suggestive of HFpEF or diagnosis of HFpEF. Here, patients with high, intermediate, or low scores according to HFA-PEFF or H<sub>2</sub>FPEF scores were classified as having confirmed HFpEF, suspected HFpEF, or no HFpEF, respectively, for ease of analysis.

According to the 2021 ESC HF guideline<sup>8</sup>, the diagnosis of HFpEF requires: (1) symptoms and signs of HF, (2) an LVEF ≥ 50%, and (3) objective evidence of cardiac structural or functional abnormalities consistent with the presence of LV diastolic dysfunction or raised LV filling pressures, including raised NPs. For simplicity in analysis, patients with symptoms of HF and LVEF ≥ 50%, along with both elevated NT-proBNP levels and echocardiographic markers of diastolic dysfunction, were classified as having confirmed HFpEF in this study. The NT-proBNP level and echocardiographic parameter thresholds specified in the guideline were applied<sup>8</sup>. Patients meeting only one of these criteria (NT-proBNP levels or echocardiographic parameters) were classified as having suspected HFpEF, while those meeting neither were classified as not having HFpEF.

## Echocardiographic parameters

All echocardiographic studies were performed using a standard ultrasound machine with a 2.5-MHz probe. Standard images were obtained in the parasternal and apical views using two-dimensional Doppler and M-mode imaging. Echocardiographic parameters were measured according to the recommendations of the American Society of Echocardiography and the European Association of Cardiovascular Imaging<sup>32,33</sup>.

LVEF was measured using the modified Quinones method or calculated using the modified Simpson's method, employing apical four- and two-chamber views in patients with regional wall motion abnormalities<sup>34,35</sup>. LVMI was calculated based on left ventricular diameter and body surface area. The LAVI was measured using the prolate ellipsoid method or biplane-modified Simpson method and adjusted for body surface area. Mitral inflow velocities were recorded using pulsed-wave Doppler from the apical window, with a 1- to 3-mm sample volume placed between the tips and mitral leaflets during diastole. The following variables were obtained: peak velocity of early (E) and late (A) diastolic filling. Mitral annular velocity was measured by tissue Doppler imaging using the pulsed-wave Doppler mode. Early diastolic ( $e'$ ) and systolic ( $S'$ ) velocities of the mitral annulus were measured from the apical four-chamber view, with a 2- to 5-mm sample volume placed at the septal corner of the mitral annulus. The ratio of the early diastolic mitral inflow velocity to the early diastolic mitral annular velocity ( $E/e'$ ) was calculated. TR jet velocity was also obtained, if measurable, to estimate PASP using continuous wave Doppler. PASP was calculated using the modified Bernoulli equation as  $4 \times (\text{peak TR velocity})^2$  plus the estimated right atrial pressure. Right atrial pressure was estimated by measuring the diameter of the inferior vena cava and its response to inspiration. Among the 992 patients, 79 (8.0%) had missing data for TR velocity and PASP and were presumed to have PASP within the normal range.

## Data collection

We retrospectively collected data on participants' baseline characteristics, including demographic data, medical history, current medications, laboratory data, and echocardiographic parameters through a review of medical records. Hypertension was defined as a blood pressure of ≥ 140/90 mmHg or use of antihypertensive medication. The estimated glomerular filtration rate (eGFR) was calculated using the Chronic Kidney Disease Epidemiology Collaborative equation<sup>36</sup>. Chronic kidney disease was defined as an eGFR < 60 mL/min/1.73 m<sup>2</sup>. Serum NT-proBNP was measured by an electrochemiluminescence immunoassay using Cobas e601 analyzer (Roche, Basel, Switzerland).

The World Health Organization defines overweight as a BMI ≥ 25 kg/m<sup>2</sup> and obesity as a BMI ≥ 30 kg/m<sup>237</sup>. However, Asian populations including Koreans, have an increased risk of diabetes and cardiovascular disease even at lower BMI levels; therefore, different criteria for obesity are used for Asians, with as BMI of ≥ 25 kg/m<sup>2</sup> being classified as obesity<sup>38,39</sup>. When calculating the H<sub>2</sub>FPEF score, we used BMI criteria of > 30 kg/m<sup>2</sup> for the main analysis, consistent with a previous study<sup>14</sup>, and also performed supplementary analyses using BMI cut-off of 27–25 kg/m<sup>2</sup>. We evaluated the change in the H<sub>2</sub>FPEF score and concordance between HFpEF diagnostic algorithms according to these BMI cut-off changes.

## Statistical analyses

Categorical variables are reported as frequencies (percentages) and were compared using Pearson's chi-squared or Fisher's exact tests. Continuous variables were tested for normality using the Shapiro–Wilk normality test. Continuous variables with normal distribution are expressed as means ± standard deviation and were compared using Student's *t*-test. Non-normally distributed continuous variables are presented as medians with interquartile

ranges and were compared using the Mann–Whitney *U* test. One-way analysis of variance or the Kruskal–Wallis test was used to compare continuous variables across the three groups.

We used Cohen's kappa coefficients to assess concordance among the three classifications<sup>40</sup>. Concordance was defined as poor (0–0.20), fair (0.21–0.40), moderate (0.41–0.60), good (0.61–0.80), and optimal (0.81–1.00). When patients were categorized into three groups—confirmed HFpEF, suspected HFpEF, and no HFpEF—Cohen's weighted kappa statistic was used. For the categorization of patients into two groups, HFpEF and non-HFpEF, Cohen's unweighted kappa statistic was used. All tests were two-tailed, and *P* value < 0.05 was considered statistically significant.

## Data availability

The data used to support the findings of this study are available from the corresponding author (Seok-Min Kang) upon request.

Received: 23 December 2024; Accepted: 6 May 2025

Published online: 01 July 2025

## References

- Shahim, B., Kapelios, C. J., Savarese, G. & Lund, L. H. Global public health burden of heart failure: an updated review. *Cardiac Fail. Rev.* **9**, 1 (2023).
- Conrad, N. et al. Temporal trends and patterns in heart failure incidence: a population-based study of 4 million individuals. *Lancet* **391**, 572–580 (2018).
- Clark, K. A. et al. Trends in heart failure hospitalizations in the US from 2008 to 2018. *J. Card. Fail.* **28**, 171–180 (2022).
- Lee, C. J. et al. Heart failure statistics 2024 update: A report from the Korean society of heart failure. *Int. J. Heart Fail.* **6**, 56–69 (2024).
- van Riet, E. E. et al. Epidemiology of heart failure: the prevalence of heart failure and ventricular dysfunction in older adults over time. A systematic review. *Eur. J. Heart Fail.* **18**, 242–252 (2016).
- Tsao, C. W. et al. Heart disease and stroke statistics—2023 update: a report from the American heart association. *Circulation* **147**, e93–e621 (2023).
- Ho, J. E. et al. Differential clinical profiles, exercise responses, and outcomes associated with existing HFpEF definitions. *Circulation* **140**, 353–365 (2019).
- McDonagh, T. A. et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) With the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur. Heart J.* **42**, 3599–3726 (2021).
- Heidenreich, P. A. et al. 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American college of cardiology/american heart association joint committee on clinical practice guidelines. *J. Am. Coll. Cardiol.* **79**, e263–e421 (2022).
- Cho, J. Y. et al. Korean society of heart failure guidelines for the management of heart failure: definition and diagnosis. *Korean Circulation J.* **53**, 195–216 (2023).
- Ho, J. E., Redfield, M. M., Lewis, G. D., Paulus, W. J. & Lam, C. S. Deliberating the diagnostic dilemma of heart failure with preserved ejection fraction. *Circulation* **142**, 1770–1780 (2020).
- Borlaug, B. A., Nishimura, R. A., Sorajja, P., Lam, C. S. & Redfield, M. M. Exercise hemodynamics enhance diagnosis of early heart failure with preserved ejection fraction. *Circulation Heart Fail.* **3**, 588–595 (2010).
- Obokata, M. et al. Role of diastolic stress testing in the evaluation for heart failure with preserved ejection fraction: a simultaneous invasive-echocardiographic study. *Circulation* **135**, 825–838 (2017).
- Reddy, Y. N., Carter, R. E., Obokata, M., Redfield, M. M. & Borlaug, B. A. A simple, evidence-based approach to help guide diagnosis of heart failure with preserved ejection fraction. *Circulation* **138**, 861–870 (2018).
- Pieske, B. 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.* **40**, 3297–3317 (2019).
- Segar, M. W., Patel, K. V., Berry, J. D., Grodin, J. L. & Pandey, A. Generalizability and implications of the H2FPEF score in a cohort of patients with heart failure with preserved ejection fraction: insights from the TOPCAT trial. *Circulation* **139**, 1851–1853 (2019).
- Sepehrvand, N. et al. External validation of the H2F-PEF model in diagnosing patients with heart failure and preserved ejection fraction. *Circulation* **139**, 2377–2379 (2019).
- Barandiarán Aizpurua, A. et al. Validation of the HFA-PEFF score for the diagnosis of heart failure with preserved ejection fraction. *Eur. J. Heart Fail.* **22**, 413–421 (2020).
- Parcha, V. et al. Diagnostic and prognostic implications of heart failure with preserved ejection fraction scoring systems. *ESC Heart Fail.* **8**, 2089–2102 (2021).
- Seo, Y., Ishizu, T., Ieda, M., Ohte, N. & Investigators, J. L. Clinical usefulness of the HFA-PEFF diagnostic scoring system in identifying late elderly heart failure with preserved ejection fraction patients. *Circ. J.* **1**, 1 (2020).
- Ouwkerk, W. et al. Heart failure with preserved ejection fraction diagnostic scores in an Asian population. *Eur. J. Heart Fail.* **22**, 1737–1739 (2020).
- Tada, A. et al. Performance of the H2FPEF and the HFA-PEFF scores for the diagnosis of heart failure with preserved ejection fraction in Japanese patients: a report from the Japanese multicenter registry. *Int. J. Cardiol.* **342**, 43–48 (2021).
- Sanders-van Wijk, S. et al. The HFA-PEFF and H2FPEF scores largely disagree in classifying patients with suspected heart failure with preserved ejection fraction. *Eur. J. Heart Fail.* **23**, 838 (2021).
- Nikrowitsch, J. et al. Applying the ESC 2016, H2FPEF, and HFA-PEFF diagnostic algorithms for heart failure with preserved ejection fraction to the general population. *ESC Heart Fail.* **8**, 3603–3612 (2021).
- Anker, S. D. et al. Empagliflozin in heart failure with a preserved ejection fraction. *N. Engl. J. Med.* **385**, 1451–1461 (2021).
- Solomon, S. D. et al. Dapagliflozin in heart failure with mildly reduced or preserved ejection fraction. *N. Engl. J. Med.* **387**, 1089–1098 (2022).
- van Riet, E. E. et al. Prevalence of unrecognized heart failure in older persons with shortness of breath on exertion. *Eur. J. Heart Fail.* **16**, 772–777 (2014).
- Tromp, J. et al. Multimorbidity in patients with heart failure from 11 Asian regions: a prospective cohort study using the ASIAN-HF registry. *PLoS Med.* **15**, e1002541 (2018).
- Yang, M. et al. Geographical variation in patient characteristics and outcomes in heart failure with mildly reduced and preserved ejection fraction. *Eur. J. Heart Fail.* **26**, 1788–1803 (2024).
- Wang, X. et al. Effects of Dapagliflozin in patients in Asia: A post hoc subgroup analysis from the DELIVER trial. *JACC Asia.* **4**, 108–118 (2024).

31. Kittleston, M. M. et al. 2023 ACC expert consensus decision pathway on management of heart failure with preserved ejection fraction: A report of the American College of Cardiology Solution Set Oversight Committee. *J. Am. Coll. Cardiol.* **81**, 1835–1878 (2023).
32. Lang, R. M. et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American society of echocardiography and the European association of cardiovascular imaging. *J. Am. Soc. Echocardiogr.* **28**, 1–39 (2015).
33. Nagueh, S. F. 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. *Eur. J. Echocardiography.* **17**, 1321–1360 (2016).
34. Quiñones, M. A. et al. A new, simplified and accurate method for determining ejection fraction with two-dimensional echocardiography. *Circulation* **64**, 744–753 (1981).
35. Folland, E. et al. Assessment of left ventricular ejection fraction and volumes by real-time, two-dimensional echocardiography. A comparison of cineangiographic and radionuclide techniques. *Circulation* **60**, 760–766 (1979).
36. Levey, A. S. et al. A new equation to estimate glomerular filtration rate. *Ann. Intern. Med.* **150**, 604–612 (2009).
37. Organization, W. H. *Obesity: Preventing and Managing the Global Epidemic: Report of a WHO Consultation* (2000).
38. Organization, W. H. *The Asia-Pacific Perspective: Redefining Obesity and Its Treatment* (2000).
39. Haam, J. H. et al. Diagnosis of obesity: 2022 update of clinical practice guidelines for obesity by the Korean society for the study of obesity. *J. Obes. Metabolic Syndrome.* **32**, 121 (2023).
40. McHugh, M. L. Interrater reliability: the kappa statistic. *Biochem. Med. (Zagreb).* **22**, 276–282 (2012).

## Author contributions

MY, JO, CL, and SK conceptualized and designed the study. MY and JO conducted data analysis under the supervision of CL and SK. MY and JO drafted the manuscript, while CL, and SK contributed to the interpretation of findings, provided critical feedback, and revised the manuscript. All authors have approved the final version of the manuscript.

## Funding

This work was Supported by grant no 18-2023-0015 from the SNUBH Research Fund.

## Declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-01537-7>.

**Correspondence** and requests for materials should be addressed to S.-M.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025