ESC European Heart Journal - Digital Health (2026) 7, ztag142
https://doi.org/10.1093/ehjdh/ztag142

ORIGINAL ARTICLE

Automated assessment of echocardiograms to aid diagnosis of heart failure with preserved ejection fraction

Constantijn Sebastiaan Venema [1,†], Max Venner [2,†], Anouk Achten [2], Wouter Ouwerkerk[3,4], Anne Raafs [2,5], Sanne Mourmans [2], Maurits Sikking [2], Hans-Peter Brunner-La Rocca [2], Stephane Heymans [2], Vanessa P.M. van Empel [2], Yoran Hummel[6], Justin Ezekowitz [7], Elke Hoendermis[1], Adriaan A. Voors [1], Joanna J. Wykrzykowska [1], Christian Knackstedt [2], Jasper Tromp [1,4,8,*†‡], and Jerremy Weerts [2,‡]

[1]Department of Cardiology and Cardiothoracic Surgery, Heart Centre, University of Groningen, University Medical Centre Groningen, Hanzeplein 1, 9713 GZ Groningen, The Netherlands; [2]Department of Cardiology, Cardiovascular Research Institute Maastricht (CARIM), Maastricht University Medical Centre (MUMC+), Maastricht, The Netherlands; [3]Department of Dermatology, University of Amsterdam Medical Centre, Amsterdam, Netherlands; [4]National Heart Centre, 5 Hospital Drive, Singapore 169609, Singapore; [5]Department of Cardiology, Catharina Hospital, Eindhoven, The Netherlands; [6]Us2.ai, 72 Circular Road, #03-01, Singapore 049426, Singapore; [7]University of Alberta, 116 St and 85 Ave, Edmonton, Alberta T6G 2R3, Canada; and [8]Saw Swee Hock School of Public Health, National University of Singapore, and the National University Health System, Tahir Foundation Building, 12 Science Drive 2, #10-01, Singapore 117549, Singapore

Received 12 February 2026; revised 13 May 2026; accepted 10 August 2026; online publish-ahead-of-print 4 September 2026

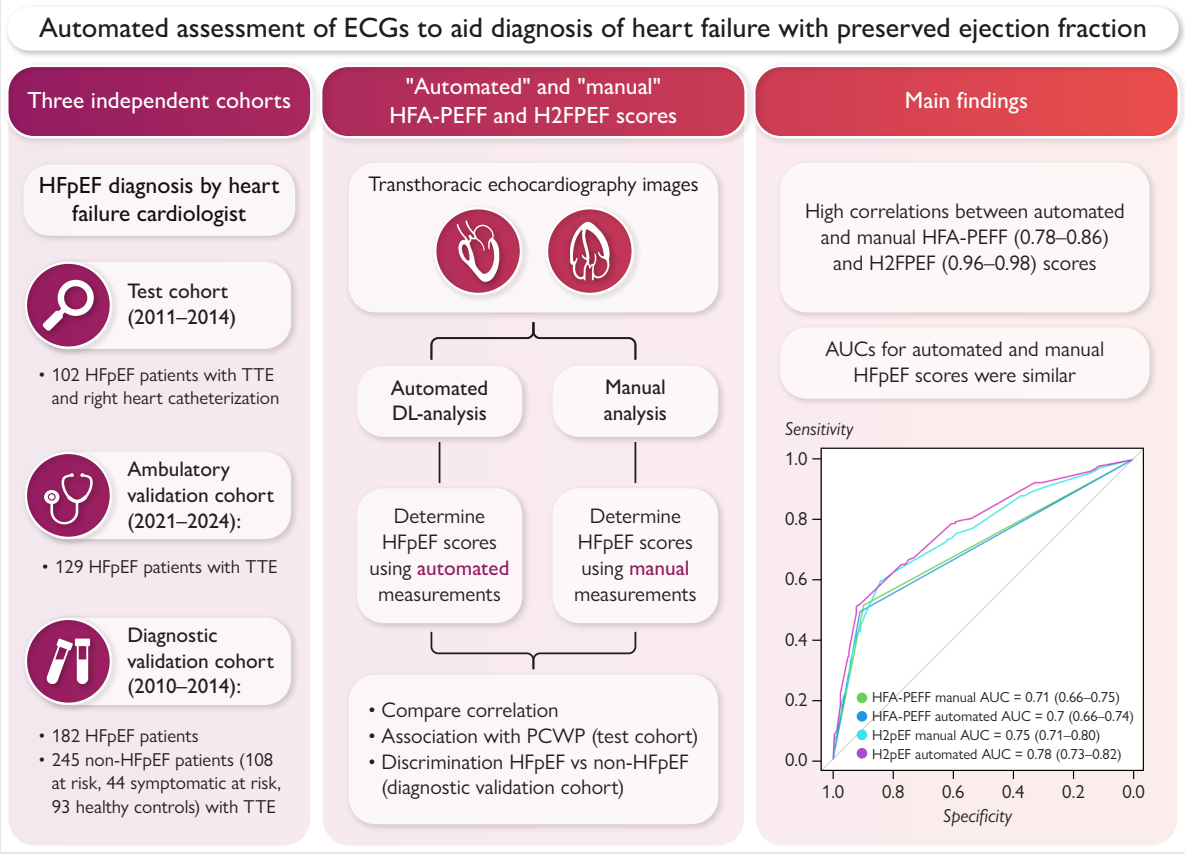
Aims Diagnosis of heart failure with preserved ejection fraction (HFpEF) using the HFA-PEFF and H_2FPEF scores remains challenging in clinical practice and relies on echocardiographic assessment. We aimed to determine whether diagnostic scoring based on automated deep learning interpretation of echocardiograms performs similar to manual measurements in diagnosing HFpEF.

Methods and results We analysed echocardiograms using an automated deep learning algorithm and manually in three cohorts: a test cohort (102 HFpEF patients diagnosed by right heart catheterization and echocardiography), an ambulatory validation cohort (129 HFpEF patients), and a diagnostic validation cohort (n = 427, of which 182 HFpEF and 245 non-HFpEF patients). We evaluated correlations between automated and manual HFA-PEFF and H_2FPEF scores across cohorts, their correlation with pulmonary capillary wedge pressures (PCWP), and compared diagnostic accuracy using the area under the curve (AUC). Automated and manual measurements showed good agreement across cohorts, with good correlations between HFA-PEFF (0.78–0.86) and H_2FPEF (0.96–0.98) scores and similar correlations with PCWP. One in five patients with high-likelihood HFpEF based on manual HFA-PEFF scores was classified as intermediate-likelihood by automated scores due to lower estimated left atrial volumes, without consistent interaction with atrial fibrillation. Areas under the curve for automated HFA-PEFF and H_2FPEF scores did not consistently differ from manual scores {0.70 [95% confidence interval (CI): 0.66–0.74] vs. 0.71 [95% CI: 0.66–0.75] and 0.78 [95% CI: 0.73–0.82] vs. 0.75 [95% CI: 0.71–0.80], respectively}.

* Corresponding author. Tel: 65 6516 4988, Email: jasper_tromp@nus.edu.sg
†These authors contributed equally to the study.
‡The last two authors contributed equally to the study.
© The Author(s) 2026. Published by Oxford University Press on behalf of the European Society of Cardiology.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Conclusion The HFA-PEFF and H_2FPEF scores based on automated and manual echocardiographic analysis showed similar diagnostic accuracy, suggesting automated HFpEF diagnosis using deep learning analysis of echocardiograms is feasible.

Graphical Abstract



Keywords Heart failure with preserved ejection fraction ● Diagnosis ● Artificial intelligence ● Echocardiography ● HFpEF ● eHealth

Background

Establishing the diagnosis of heart failure with preserved ejection fraction (HFpEF) remains challenging.[1] To aid the diagnosis in a patient with suspected HFpEF, two multiparametric scoring systems (the HFA-PEFF and H_2FPEF scores) have been developed to determine the likelihood of HFpEF.[2–6] Both scores include echocardiographic measurements, but these measurements are time-consuming and subject to interobserver variability.[7–10] It has been demonstrated that automated deep learning evaluation of echocardiographic images is associated with reduced measurement variability and has been validated in patients with heart failure.[9,11] Implementing automated echocardiographic measurements may therefore facilitate the application of diagnostic HFpEF scores in clinical practice by saving time and resources.

Aims

We hypothesized that using automated echocardiographic measurements to determine HFpEF scores yields similar diagnostic accuracy as manual measurements. Thus, we aimed to determine whether diagnostic HFpEF scoring based on fully automated deep learning interpretation of echocardiographic images is accurate and feasible compared with manual interpretation.

Table 1 Baseline characteristics of the three study cohorts

Demographics	Test cohort (*n* = 102)	Ambulatory validation cohort (*n* = 129)	Diagnostic validation cohort (*n* = 427)
Age (years), mean (SD)	73.61 (8.74)	81.4 (8.55)	67.0 (11.4)
Women (%)	68 (68.7%)	93 (72.1%)	212 (49.5%)
NYHA II	39 (39.4%)	65 (50.4%)	93 (21.7%)
NYHA III	58 (58.6%)	53 (41.1%)	52 (12.1%)
Body mass index (kg/m^2), mean (SD)	28.43 (5.71)	29.6 (6.06)	30.2 (6.9)
Medical history			
Atrial fibrillation	47 (47.5%)	83 (64.3%)	113 (26.4%)
Diabetes mellitus	31 (31.3%)	28 (21.7%)	NA
Hypertension	65 (65.7%)	61(47.3%)	271 (63.3%)
COPD	16 (16.2%)	11 (8.5%)	NA
Myocardial infarction	15 (15.2%)	14 (10.9%)	NA
NT-proBNP (pmol/L), median (IQR)	103 (54–230)	74 (39–177)	15.7 (5.5–62.4)

COPD, chronic obstructive pulmonary disease; NA, not available; NYHA, New York Heart Association; NT-proBNP, N-terminal pro-B-type natriuretic peptide.

Table 2 Yield and correlation between manual and automated HFA-PEFF and H_2FPEF scores and for selected echocardiographic parameters.

	Test cohort (*n* = 102)				Ambulatory validation cohort (*n* = 129)				Diagnostic validation cohort (*n* = 427)			
	Yield manual	Yield automated	r	RMSE	Yield manual	Yield automated	r	RMSE	Yield manual	Yield automated	r	RMSE
HFA-PEFF	NA	NA	0.78	0.79	NA	NA	0.80	0.88	NA	NA	0.86	0.98
H_2FPEF	NA	NA	0.97	0.79	NA	NA	0.98	0.41	NA	NA	0.96	0.62
LVEF	100	74	0.53	9.67	64.3	98.4	0.59	7.03	94.9	98.8	0.53	8.80
e′ septal	99	97.9	0.84	1.44	100	100	0.93	0.77	86.9	94.2	0.04	4.67
e′ lateral	99	97.9	0.83	2.03	100	100	0.93	1.06	87.1	94.9	0.12	3.60
E/e′ mean	97.9	85.4	0.92	2.84	100	100	0.94	1.4	79.7	93.7	0.26	6.52
LVMI	94.8	70.8	0.62	27.6	100	100	0.76	13.8	97.9	96.7	0.73	21.94
RWT	97.9	74	0.43	0.15	100	100	0.46	0.1	36.7	98.6	0.13	0.41
LVPWd	99	77.1	0.46	2.18	100	100	0.60	1.99	98.8	98.6	0.51	8.43
LAVI	66.7	74	0.74	15.22	100	100	0.86	12.51	92.5	57.0	0.74	8.74
PASP	76	77.1	0.85	36.21	100	100	0.79	7.97	45.6	74.3	0.65	12.61

LVEF was measured using Simpson's biplane method if both apical four-chamber (A4C) and apical two-chamber (A2C) views were available, or using A4C or A2C if both were not available. Manual relative wall thickness (RWT) was calculated as 2*LV posterior wall thickness/LV internal end-diastolic diameter. Manual PASP was calculated as 4*(tricuspid regurgitation peak jet velocity)^2.

e′ lateral, early diastolic tissue velocity at the lateral mitral annulus; e′ septal, early diastolic tissue velocity at the septal mitral annulus; LAVI, left atrial volume index; LVEF, left ventricular ejection fraction; LVMI, left ventricular mass index; LVPWd, left ventricular posterior wall thickness in diastole; E/e′, ratio of early mitral inflow velocity to early diastolic tissue velocity; NA, not applicable; PASP, pulmonary artery systolic pressure; RMSE, root mean squared error; RWT, relative wall thickness.

Methods

Study design

For this study, we used data from three cohorts: a test, an ambulatory, and a diagnostic validation cohort. For each cohort, all HFpEF diagnoses were established as a clinical diagnosis by heart failure cardiologists according to the leading European Society of Cardiology HF guidelines.[1] All participants provided written informed consent, and all cohorts were approved by local Ethics Review Committees.

Cohorts

The test cohort consisted of 102 patients with symptomatic HFpEF who were included in a randomized trial investigating the effect of sildenafil on pulmonary artery pressure between October 2011 and September 2014 (ClinicalTrials.gov, NCT01726049).[12] Patients underwent right heart catheterization with measurement of pulmonary capillary wedge pressure (PCWP) and simultaneous transthoracic echocardiography. Heart failure with preserved ejection fraction was defined as signs and symptoms of HF [New York Heart Association (NYHA) functional class II-IV], despite HF

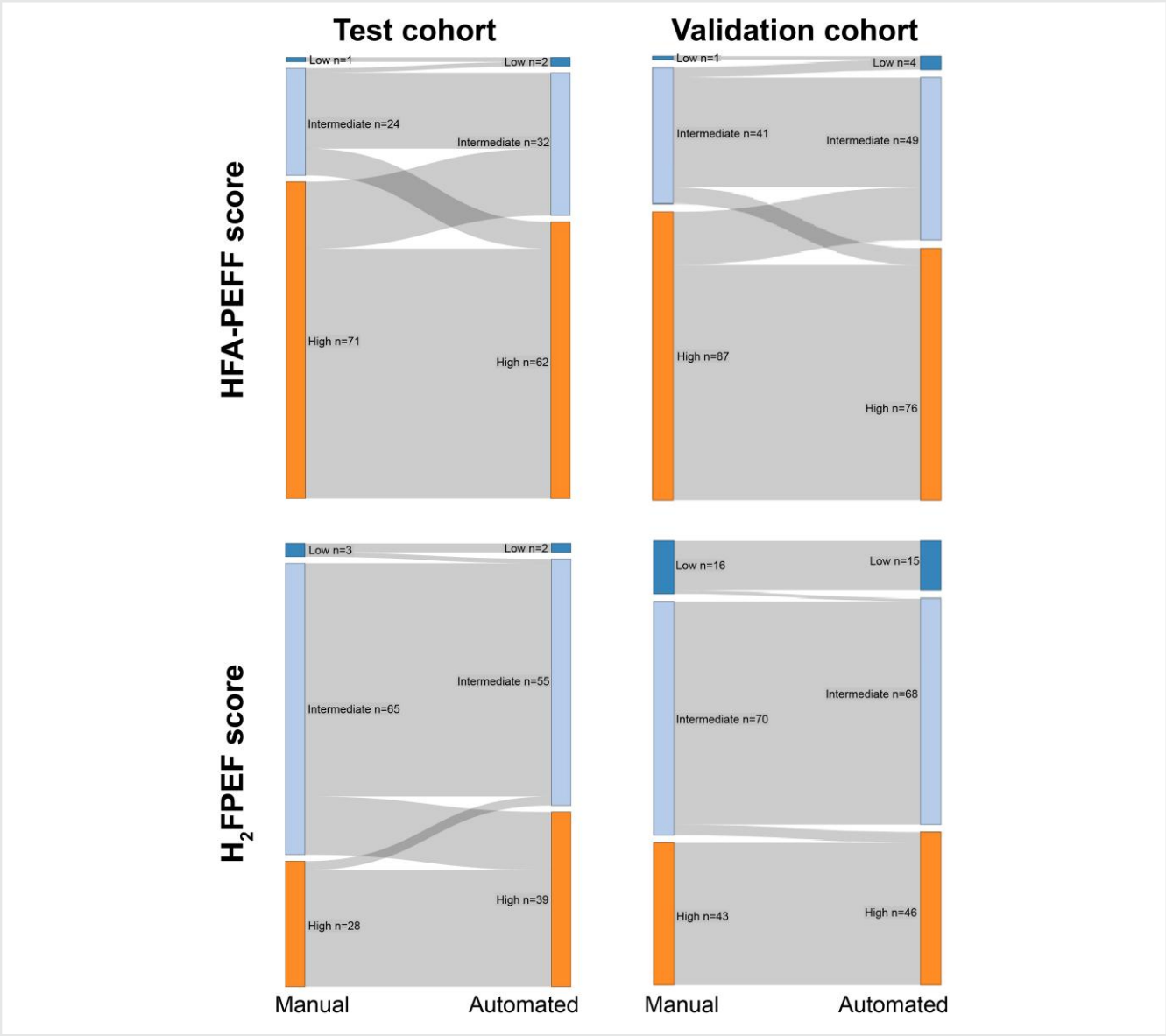


Figure 1 Sankey plots showing the classification of patients into a low, intermediate, or high probability of HFpEF using manual and automated HFA-PEFF score (upper panel) and the H_2FPEF score (lower panel) in the test and ambulatory validation cohorts. HFpEF, heart failure with preserved ejection fraction.

therapy, and a left ventricular ejection fraction (LVEF) ≥ 45%, mean pulmonary artery pressure (PAP) > 25 mmHg, and mean PCWP > 15 mmHg.

The ambulatory validation cohort comprised 129 patients with HFpEF from a dedicated HFpEF outpatient clinic between October 2021 and May 2024 with all evaluated echocardiographic parameters available by both methods (ClinicalTrials.gov, NCT04976348).[13] The diagnostic workflow has been detailed elsewhere.[14]

Finally, the diagnostic validation cohort was derived from the Alberta Heart Failure Etiology and Analysis Research Team (HEART) study and consisted of 182 patients with HFpEF and 245 patients without HFpEF, of which 108 patients at risk of HFpEF, 44 patients at risk of HFpEF with symptoms of other diseases, and 93 age- and sex-matched healthy controls (Alberta Heart Failure Etiology and Analysis Research Team; ClinicalTrials.gov, NCT02052804).[15] Participants were prospectively recruited between 2010 and 2014 and adjudicated into each group by two blinded investigators.

Echocardiographic analysis and calculation of diagnostic scores

Certified sonographers manually analysed all echocardiograms from the three cohorts according to local standards. In addition, all echocardiograms were processed using an automated, validated deep learning workflow (Us2.ai, version 2.2, Singapore). These algorithms automatically identify the correct 2D or Doppler view, annotate and segment structures, and provide echocardiographic measurements.[11]

Subsequently, the HFA-PEFF and H_2FPEF scores were determined in all cohorts using manual echocardiographic measurements (further referred to as manual scores) and deep learning-derived

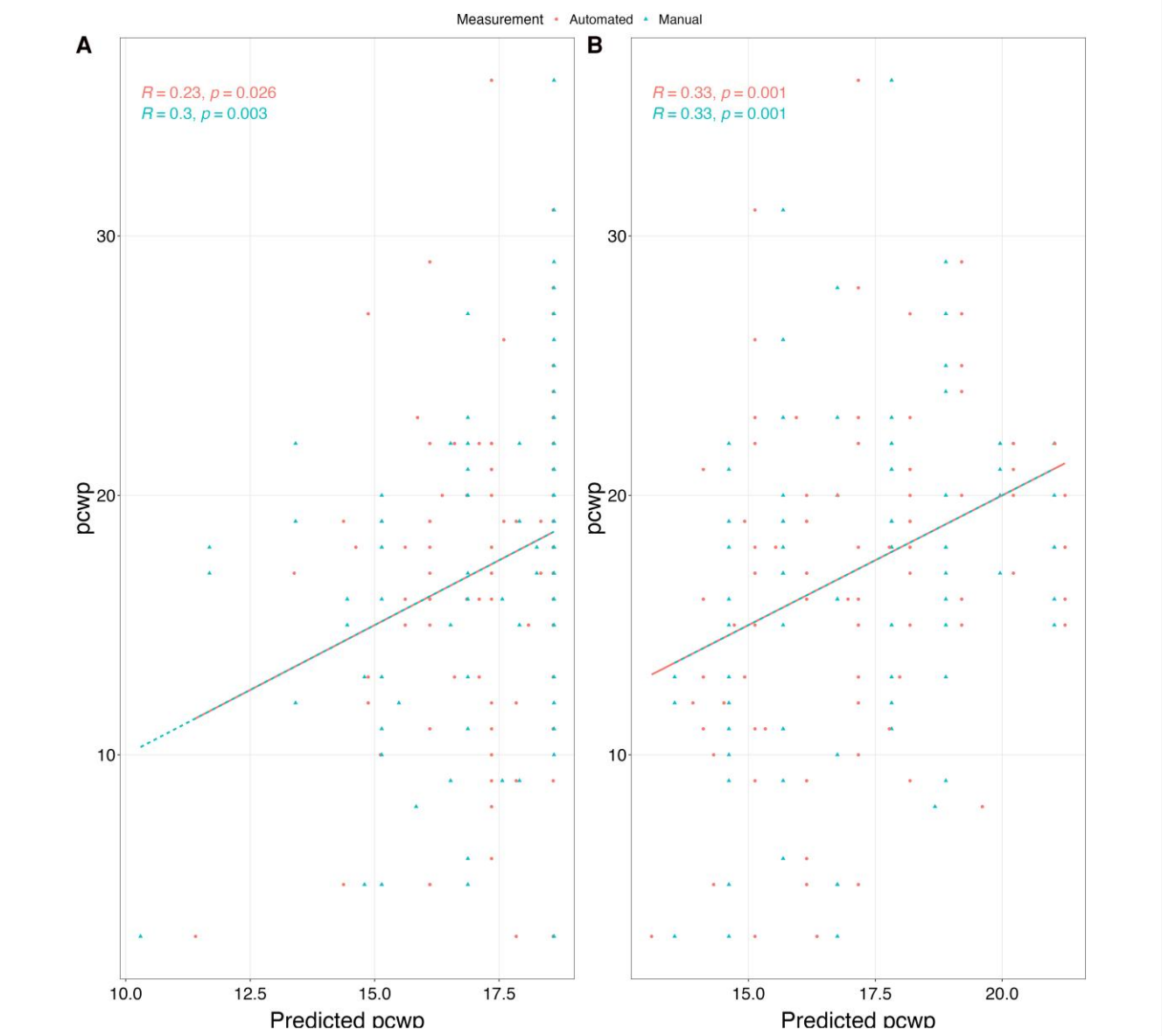


Figure 2 Predicted PCWP vs. actual PCWP for manual and automatically assessed variables included in the HFA-PEFF score (*A*) and the H_2FPEF score (*B*).

measurements (further referred to as automated scores). In case of missing data in the score components, data were imputed five times using Multiple Imputation by Chained Equations with predictive mean matching.[16] All echocardiographic, demographic, and medical history variables were used as predictors for imputation. Derived variables such as ratios were computed after imputation. The scores were then averaged over the imputed datasets.

Statistical analysis

The yield of manual and automated echocardiographic measurements was determined in each cohort, and correlations were assessed using Pearson's correlation coefficient. The differences between manual and automated measurements were quantified using the root mean square error (RMSE) and using Bland–Altman bias. The differences in classification between manual and automated HFA-PEFF and H_2FPEF scores were visualized using Sankey plots. Explorative interaction analyses were performed between scores, measurements, and atrial fibrillation status. In addition, we determined the correlation of manual and automated measurements with invasive PCWP in the test cohort. Finally, we evaluated the ability of automated and manual scores to discriminate between HFpEF (n = 182) and non-HFpEF (n = 245) in the diagnostic validation cohort by comparing the areas under the curve (AUCs) of the receiver operating characteristics (ROC) curves.[17,18] We determined 95% confidence intervals (CI) for paired AUC differences using DeLong's method.[19] Statistical analyses were performed using R version 4.5 (R Core Team, 2025), and two-sided P-values < 0.05 were considered statistically significant.

Results

Baseline characteristics of the three cohorts are presented in *Table 1*. In the test and ambulatory validation cohorts, most patients were female, overweight, and symptomatic, and

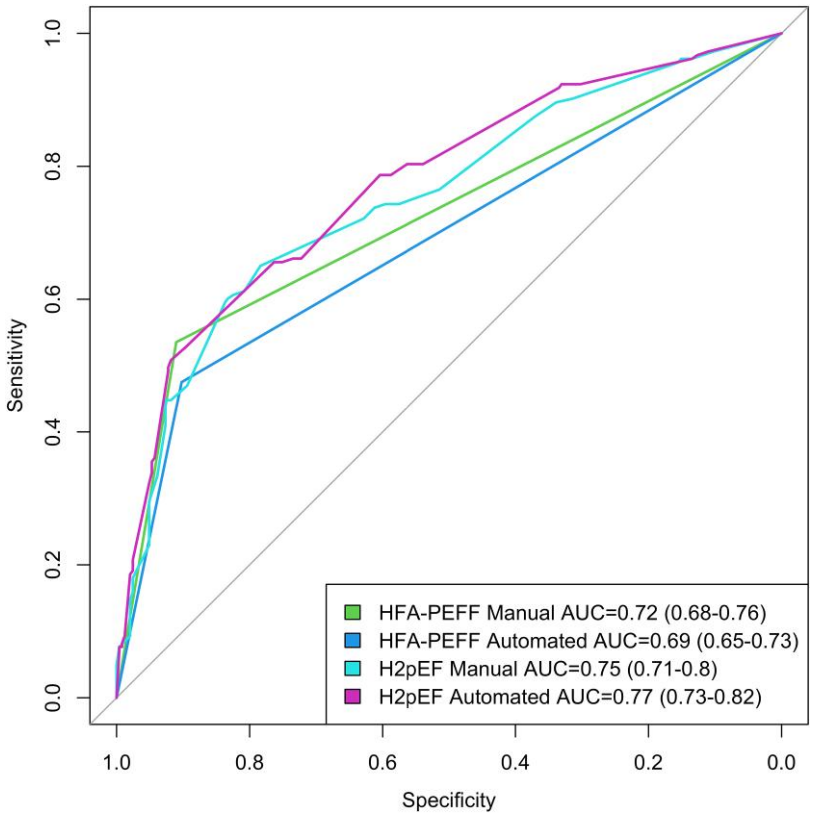


Figure 3 Receiver operating characteristic curves showing the ability of manual and automated HFpEF scores to discriminate between HFpEF and controls in the diagnostic validation cohort. AUCs of manual and automated HFA-PEFF scores did not differ significantly {AUC 0.72 [95% confidence interval (CI): 0.68–0.76] vs. 0.69 [95% CI: 0.65–0.73], $P = 0.05$}. The AUC of manual H_2FPEF scores was significantly lower than automated H_2FPEF scores [0.75 (95% CI: 0.71–0.80) vs. 0.77 (95% CI: 0.73–0.82), $P = 0.01$]. AUC, area under the curve; HFpEF, heart failure with preserved ejection fraction.

they commonly had a history of atrial fibrillation and hypertension.

Manual and automated echocardiographic measurement and diagnostic scores

Table 2 depicts the yield of manual and automated echocardiographic measurements and their correlation. Missing measurements were related to poor image quality. The proportion of cases in which the algorithm returned all echocardiographic components of both scores were 42%, 82%, and 49% in the test cohort, ambulatory validation cohort, and diagnostic validation cohort, respectively. Manual and automated echocardiographic measurements showed reasonable to high correlation coefficients ranging from 0.43 to 0.93 (*Table 2*). Similarly, manual and automated HFpEF scores showed good agreement in all cohorts with correlation coefficients ranging between 0.78 and 0.86 for the HFA-PEFF score and between 0.96 and 0.98 for the H_2FPEF score (*Table 2*).

Figure 1 shows the re-classification of patients into low, intermediate, and high risk of HFpEF using manual and automated HFpEF scores. In the test cohort, 21% of 71 patients with high-likelihood HFpEF based on manual HFA-PEFF scores were classified as intermediate likelihood using automated HFA-PEFF scores (18.6% of 87 patients in the ambulatory validation cohort). This was mainly due to smaller left atrial volume index (LAVI) estimations by the automated vs. manual echocardiographic measurement [i.e. ambulatory validation cohort LAVI 38 (32; 49) vs. 47 (40; 60) mL/m^2, respectively; Bland–Altman bias −9.9 (−11.3–8.6)], which did not consistently differ by atrial fibrillation status (P-interaction 0.63, 0.57, and <0.001, in the test, ambulatory, and diagnostic cohort, respectively).

Association with pulmonary capillary wedge pressure

Figure 2 shows the association of manual and automated HFA-PEFF and H_2FPEF scores with PCWP. The correlations of automated HFA-PEFF and H_2FPEF scores with PCWP were not significantly different from manual scores [$r = 0.23$ vs. $r = 0.30$ ($P = 0.25$) and $r = 0.33$ vs. $r = 0.33$ ($P = 0.93$), respectively]. Of the individual echocardiographic measurements, both manual and automated E/e′ showed a modest correlation with PCWP [$r = 0.24$ ($P = 0.021$) and $r = 0.33$ ($P = 0.003$), respectively; Supplementary material online, *Table S1*].

Classification of heart failure with preserved ejection fraction using manual and automated scores

Figure 3 depicts the ROC curves for manual and automated HFpEF scores to discriminate HFpEF from non-HFpEF in the diagnostic validation cohort. The AUCs for manual and automated HFA-PEFF were similar [0.71 (95% CI: 0.66–0.75) vs. 0.70 (95% CI: 0.66–0.74), difference 0.00 (95% CI: −0.03–0.04), $P = 0.78$], but manual H_2FPEF scores were significantly lower than automated H_2FPEF scores [0.75 (95% CI: 0.71–0.80) vs. 0.78 (95% CI: 0.73–0.82), difference −0.02 (95% CI: −0.04 to −0.01), $P = 0.01$]. Sensitivity, specificity, and other metrics of automated and manual HFpEF scores are provided in Supplementary material online, *Table S2*.

Conclusions

The main findings of this study were (i) HFA-PEFF and H_2FPEF scores determined using manual and automated echocardiographic measurements were comparable, (ii) manual and automated HFA-PEFF and H_2FPEF scores showed similar correlations with PCWP, and (iii) diagnostic accuracy of manual and automated HFA-PEFF and H_2FPEF scores to discriminate HFpEF from controls was similar. These findings suggest that facilitating HFpEF diagnosis using automated deep learning analysis of echocardiograms is feasible.

Diagnosing HFpEF in ambulatory patients can be challenging as dyspnoea and cardiac functional abnormalities may only manifest during exertion. Diagnostic scores were developed to assist in diagnosing patients suspected of having HFpEF. However, the practical application of these scores becomes complicated when there is disagreement between the predicted likelihoods of HFpEF.[20] As the HFA-PEFF and H_2FPEF scores contain different echocardiographic parameters, consistent and accurate measurements are critical to minimize disagreement. Our findings indicated that automated and manual measurements were comparable, without a consistent effect of atrial fibrillation. Manual and automated HFA-PEFF and H_2FPEF scores showed similar but poor correlation with resting PCWP, although PCWP may be normal in resting conditions. This is in line with previous studies demonstrating that echocardiographic variables only show modest correlation with invasive hemodynamics.[21,22] Finally, the ability of automated HFpEF scores to differentiate HFpEF was similar to that of manual scores. These findings imply that using automated echocardiographic analysis to determine HFpEF scores may facilitate and potentially accelerate the diagnostic process.

The current study underlines the promising application of artificial intelligence technologies to aid the diagnostic process. However, a limitation of this study is that echocardiograms from the cohorts were assessed according to local guidelines and not in a dedicated core lab, although this better resembles clinical practice. Nonetheless, echocardiograms from the ambulatory validation cohort that were conducted more recently compared with the other cohorts more often returned complete measurements by the algorithm. This indicates that newer and higher quality images improve performance of the automated software. Also, it should be noted that the AUC may have been inflated due to incorporation bias and inclusion of healthy controls in the non-HFpEF group. Moreover, a notable discrepancy between manual and automated measurements was observed for LA volume, leading to numerically more patients classified with intermediate instead of high risk for HFpEF according to the HFA-PEFF score. In such patients, manual reassessment of LA volume may influence diagnostic decisions and necessitate fewer additional diagnostics. Our findings, therefore, warrant prospective validation to assess whether automated HFpEF scores improve diagnostic consistency and accuracy.

In conclusion, this study demonstrated that facilitating the assessment of HFpEF scores using a fully automated deep learning echocardiography workflow yields similar results compared with manual scores. It additionally reduces the effects of interobserver variability and improves efficiency. In case of discrepancy between scores or intermediate likelihood, further investigation remains warranted, as per guideline recommendations.[3]

Supplementary material

Supplementary material is available at *European Heart Journal – Digital Health*.

Author contributions

Constantijn Sebastiaan Venema (Data curation, Investigation, Methodology, Visualization [supporting], Writing—original draft [lead]), Max Venner (Data curation, Formal analysis [supporting], Writing—original draft [equal]), Anouk Achten (Data curation, Writing—review & editing [supporting]), Wouter Ouwerkerk (Conceptualization, Data curation, Writing—review & editing [supporting], Formal analysis [lead]), Anne Raafs (Formal analysis, Investigation, Writing—review & editing [supporting]), Sanne Mourmans (Methodology, Writing—review & editing [supporting]), Maurits Sikking (Investigation, Writing—review & editing [supporting]), Hans-Peter Brunner-La Rocca (Conceptualization, Investigation, Methodology, Writing—review & editing [supporting]), Stephane Heymans (Investigation, Writing—review & editing [supporting]), Vanessa P.M. Van Empel (Conceptualization, Investigation, Writing—review & editing [supporting]), Yoran Hummel (Data curation, Software, Writing—review & editing [supporting]), Justin Ezekowitz (Data curation, Investigation, Writing—review & editing [supporting]), Elke Hoendermis (Data curation, Methodology, Writing—review & editing [supporting]), Adriaan A. Voors (Supervision, Writing—review & editing [supporting]), Joanna J. Wykrzykowska (Supervision, Writing—review & editing [supporting]), Christian Knackstedt (Conceptualization, Investigation, Methodology, Writing—review & editing [supporting]), Jasper Tromp (Conceptualization, Supervision [lead], Investigation, Methodology, Project administration, Writing—review & editing [supporting]), and Jerremy Weerts (Conceptualization, Data curation, Methodology, Supervision, Visualization, Writing—review & editing [supporting])

Funding

Us2.ai provided software (Us2v2) to the researchers to investigate the objective of this study. Funding was provided by Alberta Innovates – Health Solutions Interdisciplinary Team Grant to Alberta Heart Failure Etiology and Analysis Research Team (Alberta HEART), grant no. AHFMR ITG 200801018.

Conflict of interest: The employer of AAV received consultancy fees and/or research support from Adrenomed, Anacardio, AstraZeneca, Bayer AG, BMS, Boehringer Ingelheim, Cardurion, Corteria, EliLilly, Merck, Moderna, Novartis, Novo Nordisk, Rycarma, and SalubrisBio. J.W. reports personal fees from Corvia Medical, CSL Vifor, and Boehringer Ingelheim,

outside the submitted work. J.T. is supported by the National University of Singapore Start-up grant, the tier 1 grant from the Ministry of Education and the CS-IRG New Investigator Grant from the National Medical Research Council; has received research support from AstraZeneca and consulting or speaker fees from Daiichi Sankyo, Boehringer Ingelheim, Roche Diagnostics, and Us2.ai; and owns patent US-10702247-B2 unrelated to the present work. Y.M.H. is employee of Us2.ai. The other authors report no disclosures.

Data availability

All relevant data are included in the paper, and its supplementary material is online. The article's data may be shared upon a reasonable request to the corresponding author.

References

1. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, *et al.* 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J* 2021;**42**:3599–3726.
2. Reddy YNV, Carter RE, Obokata M, Redfield MM, Borlaug BA. A simple, evidence-based approach to help guide diagnosis of heart failure with preserved ejection fraction. *Circulation* 2018;**138**:861–870.
3. Pieske B, Tschöpe C, de Boer RA, Fraser AG, Anker SD, Donal E, *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**:3297–3317.
4. Barandiarán Aizpurua A, Sanders-van Wijk S, Brunner-La Rocca H-P, Henkens M, Heymans S, Beussink-Nelson L, *et al.* Validation of the HFA-PEFF score for the diagnosis of heart failure with preserved ejection fraction. *Eur J Heart Fail* 2020;**22**:413–421.
5. Ouwerkerk W, Tromp J, Jin X, Jaufeerally F, Yeo PSD, Leong KTG, *et al.* Heart failure with preserved ejection fraction diagnostic scores in an Asian population. *Eur J Heart Fail* 2020;**22**:1737–1739.
6. Selvaraj S, Myhre PL, Vaduganathan M, Claggett BL, Matsushita K, Kitzman DW, *et al.* Application of diagnostic algorithms for heart failure with preserved ejection fraction to the community. *JACC Heart Fail* 2020;**8**:640–653.
7. Nagueh SF, Abraham TP, Aurigemma GP, Bax JJ, Beladan C, Browning A, *et al.* Interobserver variability in applying American Society of Echocardiography/European Association of Cardiovascular Imaging 2016 guidelines for estimation of left ventricular filling pressure. *Circ Cardiovasc Imaging* 2019;**12**:e008122.
8. Morbach C, Gelbrich G, Breunig M, Tiffe T, Wagner M, Heuschmann PU, *et al.* Impact of acquisition and interpretation on total inter-observer variability in echocardiography: results from the quality assurance program of the STAAB cohort study. *Int J Cardiovasc Imaging* 2018;**34**:1057–1065.
9. Tromp J, Bauer D, Claggett BL, Frost M, Iversen MB, Prasad N, *et al.* A formal validation of a deep learning-based automated workflow for the interpretation of the echocardiogram. *Nat Commun* 2022;**13**:6776.
10. Weerts J, Amin H, Barandiarán Aizpurua A, Gevaert AB, Handoko ML, Dauw J, *et al.* Webtool to enhance the accuracy of diagnostic algorithms for HFpEF: a prospective cross-over study. *ESC Heart Fail* 2023;**10**:3493–3503.
11. Tromp J, Seekings PJ, Hung C-L, Iversen MB, Frost MJ, Ouwerkerk W, *et al.* Automated interpretation of systolic and diastolic function on the echocardiogram: a multicohort study. *Lancet Digit Health* 2022;**4**:e46-e54.
12. Hoendermis ES, Liu LCY, Hummel YM, van der Meer P, de Boer RA, Berger RMF, *et al.* Effects of sildenafil on invasive haemodynamics and exercise capacity in heart failure patients with preserved ejection fraction and pulmonary hypertension: a randomized controlled trial. *Eur Heart J* 2015;**36**:2565–2573.
13. Henkens MTHM, Weerts J, Verdonschot JAJ, Raafs AG, Stroeks S, Sikking MA, *et al.* Improving diagnosis and risk stratification across the ejection fraction spectrum: the Maastricht cardiomyopathy registry. *ESC Heart Fail* 2022;**9**:1463–1470.
14. Weerts J, Barandiarán Aizpurua A, Patel RB, van de Bovenkamp AA, Amin H, Sanders-van Wijk S, *et al.* A simplified approach for diagnosing heart failure with preserved ejection fraction by assessing the left atrium and natriuretic peptides. *Eur J Heart Fail* 2025;**27**:2897–2902.
15. Ezekowitz JA, McAlister FA, Howlett J, Alemayehu W, Paterson I, Belenkie I, *et al.* A prospective evaluation of the established criteria for heart failure with preserved ejection fraction using the Alberta HEART cohort. *ESC Heart Fail* 2018;**5**:19–26.
16. van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations inR. *J Stat Softw* 2011;**45**:1–67.
17. Fawcett T. An introduction to ROC analysis. *Pattern Recognit Lett* 2006;**27**:861–874.
18. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, *et al.* pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011;**12**:77.
19. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;**44**:837–845.
20. Sanders-van Wijk S, Barandiarán Aizpurua A, Brunner-La Rocca H-P, Henkens MTHM, Weerts J, Knackstedt C, *et al.* The HFA-PEFF and H2 FPEF scores largely disagree in classifying patients with suspected heart failure with preserved ejection fraction. *Eur J Heart Fail* 2021;**23**:838–840.
21. Nauta JF, Hummel YM, van der Meer P, Lam CSP, Voors AA, van Melle JP. Correlation with invasive left ventricular filling pressures and prognostic relevance of the echocardiographic diastolic parameters used in the 2016 ESC heart failure guidelines and in the 2016 ASE/EACVI recommendations: a systematic review in patients with heart failure with preserved ejection fraction. *Eur J Heart Fail* 2018;**20**:1303–1311.
22. Smiseth OA, Morris DA, Cardim N, Cikes M, Delgado V, Donal E, *et al.* Reviewers: this document was reviewed by members of the 2018-2020 EACVI scientific documents committee. Multimodality imaging in patients with heart failure and preserved ejection fraction: an expert consensus document of the European association of cardiovascular imaging. *Eur Heart J Cardiovasc Imaging* 2022;**23**:e34–e61.