Circulation

ORIGINAL RESEARCH ARTICLE

A Simple, Evidence-Based Approach to Help Guide Diagnosis of Heart Failure With Preserved Ejection Fraction

Editorial, see p 871

Yogesh N.V. Reddy, MD
Rickey E. Carter, PhD
Masaru Obokata, MD, PhD
Margaret M. Redfield, MD
Barry A. Borlaug, MD

BACKGROUND: Diagnosis of heart failure with preserved ejection fraction (HFpEF) is challenging in euvolemic patients with dyspnea, and no evidence-based criteria are available. We sought to develop and then validate noninvasive diagnostic criteria that could be used to estimate the likelihood that HFpEF is present among patients with unexplained dyspnea to guide further testing.

METHODS: Consecutive patients with unexplained dyspnea referred for invasive hemodynamic exercise testing were retrospectively evaluated. Diagnosis of HFpEF (case) or noncardiac dyspnea (control) was ascertained by invasive hemodynamic exercise testing. Logistic regression was performed to evaluate the ability of clinical findings to discriminate cases from controls. A scoring system was developed and then validated in a separate test cohort.

RESULTS: The derivation cohort included 414 consecutive patients (267 cases with HFpEF and 147 controls; HFpEF prevalence, 64%). The test cohort included 100 consecutive patients (61 with HFpEF; prevalence, 61%). Obesity, atrial fibrillation, age >60 years, treatment with ≥2 antihypertensives, echocardiographic E/e' ratio >9, and echocardiographic pulmonary artery systolic pressure >35 mm Hg were selected as the final set of predictive variables. A weighted score based on these 6 variables was used to create a composite score (H_2FPEF score) ranging from 0 to 9. The odds of HFpEF doubled for each 1-unit score increase (odds ratio, 1.98; 95% CI, 1.74–2.30; $P<0.0001$), with an area under the curve of 0.841 ($P<0.0001$). The H_2FPEF score was superior to a currently used algorithm based on expert consensus (increase in area under the curve of 0.169; 95% CI, 0.120–0.217; $P<0.0001$). Performance in the independent test cohort was maintained (area under the curve, 0.886; $P<0.0001$).

CONCLUSIONS: The H_2FPEF score, which relies on simple clinical characteristics and echocardiography, enables discrimination of HFpEF from noncardiac causes of dyspnea and can assist in determination of the need for further diagnostic testing in the evaluation of patients with unexplained exertional dyspnea.

Key Words: catheterization ■ exercise test ■ heart failure

Sources of Funding, see page 869

https://www.ahajournals.org/journal/circ

ORIGINAL RESEARCH ARTICLE

Clinical Perspective

What Is New?

- We show, using simple, universally available clinical and echocardiographic characteristics, that the probability that heart failure with preserved ejection fraction is present can be accurately estimated in the patient presenting with unexplained exertional dyspnea.

What Are the Clinical Implications?

- The H$_2$FPEF score enables providers and patients to estimate the probability of underlying heart failure with preserved ejection fraction.
- This allows more informed decision-making about the likelihood of disease and thus the yield of additional testing to confirm or refute the diagnosis in a Bayesian approach.

Exertional dyspnea may be caused by cardiac and noncardiac disorders. Among the cardiovascular causes, heart failure with preserved ejection fraction (HFpEF) is an increasingly common one characterized by pathological increases in cardiac filling pressures at rest or with exertion.[1–6] Decompensated patients with HFpEF typically display overt congestion on physical examination and chest radiography, and in this setting, the diagnosis is straightforward. However, compensated, euvolemic patients presenting with exertional dyspnea in the absence of overt clinical, radiographic, or biomarker evidence of congestion present a greater diagnostic challenge.

The reference standard to diagnose HFpEF in these patients is right-sided heart catheterization followed by invasive exercise testing if resting intracardiac pressures are normal.[7–10] Because of its invasive nature, technical complexity, and cost, this test is impractical for routine evaluation but is more logically reserved for situations in which diagnosis remains uncertain after less invasive test results are equivocal.[7] To make this determination, the probability of disease must first be estimated, allowing clinicians to decide whether disease is likely present or absent or intermediate, in which case more definitive testing is required. Currently, no data are available to guide this sort of Bayesian approach to the evaluation of unexplained dyspnea.

To fill this gap, we evaluated clinical data from consecutive patients in whom the diagnosis of HFpEF or a noncardiac cause of dyspnea was ascertained conclusively by invasive exercise testing to develop a scoring system that could be used in the diagnostic evaluation of HFpEF. We then validated this new scoring system in a separate cohort.

METHODS

This was a retrospective analysis of all consecutive patients undergoing invasive exercise testing for the evaluation of unexplained dyspnea between 2006 and 2016 at the Mayo Clinic in Rochester, MN. The data, analytical methods, and study materials will not be made available to other researchers for purposes of reproducing the results or replicating the procedure. Exclusion criteria included ejection fraction <50% (current or prior), significant valvular heart disease (greater than mild stenosis, greater than moderate regurgitation), pulmonary arterial hypertension, constrictive pericarditis, primary cardiomyopathies, or heart transplantation. All patients referred for hemodynamic catheterization were evaluated by Mayo Clinic staff cardiologists and concluded to have dyspnea not explainable by pulmonary disease on the basis of evaluations performed at the discretion of the referring physicians.

Patients with HFpEF were identified by elevated pulmonary capillary wedge pressure at rest (≥15 mm Hg) or during exercise (≥25 mm Hg).[7,8] Noncardiac dyspnea was defined as no evidence of a cardiac cause for dyspnea after exhaustive clinical evaluation, including normal rest and exercise hemodynamics. Data included in the study were authorized by the patient for use in research with informed consent, and the study was approved by the Mayo Clinic Institutional Review Board.

Clinical Evaluation

All patients were evaluated by a board-certified cardiologist. Medical history was determined from detailed manual chart review by 2 independent observers (Y.N.V.R. and M.O. with discrepancies arbitrated by B.A.B.). Hypertension was defined by treatment with antihypertensive medications, and the number of antihypertensive medications was quantified for each patient. Atrial fibrillation was determined from clinical history and ECG. Diabetes mellitus was defined by treatment with antidiabetic medications, fasting plasma glucose ≥126 mg/dL, or a hemoglobin A_{1c} ≥6.5 mg/dL. Prediabetes was defined as fasting plasma glucose between 100 and 126 mg/dL or a hemoglobin A_{1c} between 5.7 and 6.5 mg/dL. Laboratory values, including hemoglobin and creatinine, were obtained on the day of catheterization. NT-proBNP (N-terminal pro-B-type natriuretic peptide) levels were obtained from samples drawn within 3 months of the assessment. Echocardiography was performed according to American Society of Echocardiography guidelines and interpreted by Mayo Clinic staff cardiologists.[11] Details of the echocardiographic measurements performed are included in the online-only Data Supplement.

Ascertainment of Diagnosis

Subjects were studied on long-term medications in the fasted state and supine position using high-fidelity micromanometer catheters and directly measured O_2 consumption at rest and during supine cycle ergometry exercise to exhaustion, as previously described.[7,8] Pressures in the right atrium, pulmonary artery, and pulmonary artery wedge positions were measured at end expiration from electronically stored continuous recordings of pressure tracings digitized at 240 Hz. Systemic and mixed venous O_2 content were determined by blood sampling. Cardiac output was determined by the Fick method.

Statistical Analysis

Data are reported as mean and SD or median (25th–75th interquartile range). The χ^2, Wilcoxon rank-sum, or *t* test was used as appropriate to examine differences between cases with HFpEF and controls. Before the development of the final models, data in the development cohort were imputed with random forest imputation (missForest package version 1.4).[12] Two modeling strategies were considered. The primary analysis plans were to develop a multivariable logistic regression model that could be summarized with a simple additive score based on prior knowledge of HFpEF pathophysiology while allowing categorization of variables to be considered in the modeling process.[13] As an alternative to address the limitations of variable categorization, a model consisting of only continuous variables was also estimated. Second, 2 agnostic supplemental models were built as sensitivity analyses: (1) a classification and regression tree (CART) model to enable easier graphical representation with inclusion of higher-order interaction terms that would be complex to represent with the additive score, and (2) a fully agnostic multivariable logistic model.

Predictors for HFpEF were first analyzed with simple logistic regression to identify candidate variables that were significantly associated with disease status. For ease of clinical use, continuous variables that were significant were dichotomized with receiver-operating characteristic curves to identify optimal cut points for discrimination, which were rounded to the nearest clinically significant integer when applicable. Next, significant variables ($P<0.05$) were entered into multivariable logistic regression models to determine a final model.

Obesity[14,15] and atrial fibrillation[16,17] are known to be important in HFpEF pathogenesis, and these variables were a priori forced into the model. Additional variables that were significant on univariable analysis were added, with care taken to avoid clinically relevant collinearity. Once the full multivariable model was created, stepwise backward elimination was performed with the least significant variable removed 1 variable at a time, until all included model variables were statistically significant. A noninvasive prediction score was created with the variables and strength of association by β coefficients, as previously described.[13] In addition to the points-based score, a continuous model was built with the same variables.

Using this prediction score on a continuous scale, we then evaluated its diagnostic performance by the area under the receiver-operating characteristic curve (AUC), or c statistic. To evaluate the likelihood of the model to generalize to a new sample, Harrell optimism was calculated with 1000 bootstrap replicates,[18] and to evaluate incremental discrimination beyond existing criteria, we compared the AUCs from our derived scoring system with the current and prior consensus algorithms endorsed by the European Society of Cardiology guidelines (Figure I in the online-only Data Supplement) using the DeLong test.[5] Calibration of the predicted probabilities with the empirical probabilities for HFpEF was assessed with the Hosmer-Lemeshow goodness-of-fit test.

Two completely agnostic models were built as a sensitivity analysis. First, an agnostic multivariable logistic model included all significant predictors of HFpEF on univariable analysis, with stepwise backward regression using a probability to leave of 0.05. Second, random forest classifiers were constructed with the full list of candidate predictors in the development data to develop a CART model. Variable importance plots were used to begin subsetting the number of variables. The subset of variables was then used to train a CART model with the rpart package in R under the default tuning configuration. Both the resulting CART model and the prediction score were validated on an independent test cohort from patients. All tests were 2-sided, with a value of $P<0.05$ considered significant. Analyses were performed with JMP 13.0.0 (SAS Institute, Cary, NC) and R 3.4.1 (R Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Patients with HFpEF (n=267) were older and had higher body mass index and more hypertension, glucose intolerance, atrial fibrillation, NT-proBNP elevation, and renal dysfunction compared with patients with noncardiac causes of dyspnea (n=147; Table 1). They were more likely to have a pacemaker; QRS, QTc, and PR interval prolongation on ECG; and cardiomegaly on chest radiography. Two thirds of patients came from local practices served by the Mayo Clinic (n=273, 66%), with the remainder referred from tertiary academic centers. Of patients found to have HFpEF, 45% (n=121) displayed elevation in filling pressures only during exercise (early-stage HFpEF).

Transthoracic echocardiography revealed that patients with HFpEF were more likely to have diastolic dysfunction, with higher noninvasive estimates of filling pressure (higher E/e' ratio). Although the ejection fraction was similar, patients with HFpEF had subtle impairment in systolic function as evidenced by lower global longitudinal strain (Table 1). Estimated pulmonary artery pressure was higher and right ventricular dysfunction and dilatation were more common in HFpEF. Although group differences for many variables were highly significant, there was a substantial degree of overlap between the 2 groups (Table 1).

At cardiac catheterization, patients with HFpEF displayed higher ventricular filling pressures and pulmonary artery pressures and lower cardiac output compared with patients with noncardiac dyspnea, as expected (Table I in the online-only Data Supplement).

Univariable Predictors of HFpEF

Clinical, demographic, and echocardiographic criteria were evaluated as predictors of HFpEF in isolation (Table 2 and Table II in the online-only Data Supplement). Certain variables were highly specific for the presence of HFpEF, including grade II obesity (body mass index >35 kg/m^2; specificity, 88%), chronic kidney disease (≥stage 3, 90%), atrial fibrillation (96%), diabetes mellitus (88%), the presence of a pacemaker (99%), cardiomegaly on chest film (96%), mildly depressed ejection fraction of 50% to 54% (96%), E/e' >14 (89%), pulmonary artery systolic pressure >35 mm Hg (86%), NT-proBNP >450 pg/mL (85%), and the presence of right ventricular dysfunction.

Table 1. Baseline Characteristics

	Noncardiac Dyspnea (n=147)	HFpEF (n=267)	*P* Value
Age, y	56±15	68±11	<0.0001
Female, %	59	61	0.4
Body mass index, kg/m^2	28.2±5.4	33.0±7.4	<0.0001
Comorbidities			
Hypertension, %	53	86	<0.0001
Antihypertensive drugs, n	1.2±1.3	2.2±1.3	<0.0001
Impaired glucose tolerance any, %	29.9	54.7	<0.0001
Prediabetes	17.7	26.6	
Diabetes mellitus	12.2	28.1	
Atrial fibrillation any, %	4.1	34.4	<0.0001
Paroxysmal atrial fibrillation	3.4	17.2	
Permanent atrial fibrillation	0.7	17.2	
Hemoglobin, g/dL	12.9±1.3	12.2±1.5	<0.0001
Diuretic, %	23	48	<0.0001
NT-proBNP, pg/mL	122 (52–259)	384 (131–1111)	<0.0001
Creatinine, mg/dL	0.96±0.22	1.13±0.40	<0.0001
Glomerular filtration rate, mL·min^{-1}·1.73 m^{-2}	93±31	83±37	0.006
Kidney disease grade 3 or higher, %	10	26	<0.0001
ECG			
Pacemaker, %	0.7	12.7	<0.0001
QRS duration, ms	94±19	99±27	0.04
Left bundle-branch block,%	2	2	0.7
Left atrial enlargement, %	1	4	0.2
PR interval, ms	159±26	175±39	<0.0001
LV hypertrophy, %	2	3	0.7
QTc interval, ms	435±26	445±33	0.0009
Chest radiography			
Cardiomegaly, %	4	24	<0.0001
Pleural effusion, %	2	4	0.4
Echocardiography			
LV			
LV end diastolic dimension, mm	48±5	48±5	0.1
LV mass index, g/m^2	84±19	92±23	<0.0001
LV hypertrophy, %	12	26	0.0009
LV ejection fraction, %	63±5	63±6	0.7
LV global longitudinal strain, %	16.3±2.6	15.2±3.0	0.0001
LA volume index, mL/m^2	29±9	38±14	<0.0001
E/e' ratio	10±4	14±7	<0.0001
Septal e', cm/s	8±3	7±2	<0.0001
RV			
Pulmonary artery systolic pressure, mm Hg	30±5	38±12	<0.0001
TAPSE, mm	23±3	21±4	<0.0001
TV lateral s', cm/s	14±3	13±3	0.002
RV fractional area change, %	53±5	50±9	0.0003
RV basal diameter, mm	31±5	33±6	0.0003
RV mid diameter, mm	24±4	25±6	0.004
Visual RV dysfunction,%	6	22	<0.0001
Visual RV dilation, %	12	31	<0.0001

Values represent mean ± standard deviation. e' Indicates septal mitral annulus tissue relaxation velocity in early diastole; E/e', ratio of early diastolic mitral inflow velocity to septal mitral annulus tissue relaxation velocity; HFpEF, heart failure with preserved ejection fraction; LA, left atrial; LV, left ventricle; NT-proBNP, N-terminal pro-B-type natriuretic peptide; RV, right ventricle, TAPSE, tricuspid annular plane systolic excursion; and TV, tricuspid valve.

Derivation of the H$_2$FPEF Score

The variables identified through univariable screening were entered into a multivariable model (Table 3). Obesity (body mass index >30 kg/m^2), atrial fibrillation, age >60 years, treatment with ≥2 antihypertensive drugs, E/e' >9, and pulmonary artery systolic pressure >35 mm Hg were associated with HFpEF in combination (all *P*<0.05). A score was assigned to these 6 variables based on strength of association in logistic regression with HFpEF (atrial fibrillation, 3 points; obesity, 2 points; others, 1 each), creating an H$_2$FPEF score (heavy, 2 or more hypertensive drugs, atrial fibrillation, pulmonary hypertension [pulmonary artery systolic pressure>35 mm Hg], elder age>60, elevated filling pressures [E/e'>9]) ranging from 0 to 9 (Figure 1, top). The probability of HFpEF increased with increasing H$_2$FPEF score (Figure 1, bottom). Model-based probabilities closely matched the observed prevalence for each given score value, indicating good calibration (Figure 2).

The odds of having HFpEF increased by a factor of 2 for every 1-unit increase in the score (odds ratio, 1.98; 95% CI, 1.73–2.30). The H$_2$FPEF score provided strong discrimination of HFpEF from controls (AUC, 0.841; 95% CI, 0.802–0.881). The H$_2$FPEF score better discriminated HFpEF from noncardiac causes of dyspnea compared with widely used diagnostic algorithms based on expert consensus[4,5] (AUC comparison, 0.169 [95% CI, 0.120–0.217] versus 2016 European Society of Cardiology guidelines and 0.173 [95% CI, 0.132–0.215) versus 2007 European Society of Cardiology guidelines; both *P*<0.0001; Table III in the online-only Data Supplement). The use of NT-proBNP levels did not incrementally add diagnostic ability to the H$_2$FPEF score (Table 3).

Because the points-based score can result in loss of information as a result of dichotomization, we also evaluated the H$_2$FPEF score on a continual scale

Table 2. Univariate Predictors of HFpEF

	OR (95% CI)	β Estimate	AUC	Sensitivity, %	Specificity, %	*P* Value
Clinical						
Age >60 y	6.20 (3.96–9.69)	1.82	0.704	80	60	<0.0001
Body mass index >30 kg/m^2	3.46 (2.27–5.29)	1.24	0.651	65	65	<0.0001
Body mass index >35 kg/m^2	4.02 (2.23–7.07)	1.39	0.615	35	88	<0.0001
NT-proBNP >275 pg/mL	4.82 (3.06–7.59)	1.57	0.680	59	77	<0.0001
NT-proBNP >450 pg/mL	4.93 (3.00–8.40)	1.60	0.657	46	85	<0.0001
Chronic kidney disease grade 3 or higher	3.38 (1.88–6.47)	1.22	0.584	26	90	<0.0001
Any hypertension	5.33 (3.35–8.61)	1.67	0.664	86	47	<0.0001
Treatment with ≥2 antihypertensives	4.49 (2.94–6.94)	1.50	0.678	72	63	<0.0001
Atrial fibrillation, any	12.35 (5.69–302.41)	2.51	0.652	35	96	<0.0001
Paroxysmal atrial fibrillation	5.91 (2.51–17.36)	1.78	0.569	17	97	<0.0001
Permanent atrial fibrillation	30.39 (6.53–540.93)	3.41	0.583	17	99	<0.0001
Diabetes mellitus	2.80 (1.60–4.90)	1.03	0.579	28	88	0.0003
Prediabetes or diabetes mellitus	2.82 (1.84–4.33)	1.04	0.624	55	70	<0.0001
Pacemaker	21.30 (2.89–157.31)	3.06	0.560	13	99	<0.0001
Chest radiography						
Cardiomegaly	7.56 (3.45–19.97)	2.02	0.601	24	96	<0.0001
Pleural effusion	4.96 (1.70–21.08)	1.60	0.537	9	98	0.002
Cardiomegaly or pleural effusion	5.99 (3.05–13.21)	1.79	0.610	28	94	<0.0001
Echocardiogram						
Ejection fraction <55%	2.39 (0.88–6.47)	0.87	0.522	8	96	0.09
Global longitudinal strain <16%	2.10 (1.39–3.16)	0.74	0.591	62	56	0.0004
LV hypertrophy	2.55 (1.48–4.59)	0.94	0.570	26	88	0.0006
LA volume index >30 mL/m^2	5.65 (3.64–8.79)	1.73	0.704	70	71	<0.0001
E/e' ratio>9	5.23 (3.37–8.11)	1.65	0.687	78	59	<0.0001
E/e' ratio>13	5.20 (3.09–8.76)	1.65	0.661	46	86	<0.0001
Septal e' velocity <7, cm/s	2.90 (1.87–4.59)	1.07	0.619	48	76	<0.0001
Right atrial pressure>10 mm Hg	6.80 (2.05–22.58)	1.91	0.564	16	97	<0.0001
Pulmonary artery systolic pressure >35 mm Hg	5.05 (3.05–8.69)	1.61	0.657	46	86	<0.0001
RV fractional area change <48%	4.88 (2.78–8.59)	1.59	0.637	39	88	<0.0001
Tricuspid annular plane systolic excursion <21 mm	3.69 (2.29–5.94)	1.30	0.637	46	81	<0.0001
Visual RV dysfunction	4.26 (2.04–8.87)	1.45	0.578	22	94	<0.0001
Visual RV dilation	3.45 (1.96–6.09)	1.24	0.598	32	88	<0.0001

Cut points derived from receiver-operating curve analysis as shown in Table II in the online-only Data Supplement. AUC indicates area under the curve; e', septal mitral annulus tissue relaxation velocity in early diastole; E/e', ratio of early diastolic mitral inflow velocity to septal mitral annulus tissue relaxation velocity; HFpEF, heart failure with preserved ejection fraction; LA, left atrial; LV, left ventricular; NT-proBNP, N-terminal pro-B-type natriuretic peptide; OR, odds ratio; and RV, right ventricular.

(Figure II in the online-only Data Supplement). This resulted in a slightly better-performing model (AUC comparison, 0.022; 95% CI, 0.002–0.042; *P*=0.03; Table III in the online-only Data Supplement). In contrast to the points-based H$_2$FPEF model, the number of hypertension medicines did not remain predictive in the continuous model, so this variable was not included. Calibration remained robust using the continuous model with a goodness-of-fit *P*>0.1 (Figure III in the online-only Data Supplement). Findings for the models were upheld in the bootstrap (internal) validation, with an optimism-corrected AUC of 0.838 for the categorical model and 0.857 for the continuous model.

Table 3. Multivariable Predictors of HFpEF

	OR (95% CI)	β Estimate	*P* Value
Multivariable model (AICc, 393.72; AUC, 0.854; *P*<0.0001)			
Atrial fibrillation	4.59 (1.84–13.22)	1.52	0.0007
Body mass index >30 kg/m^2	2.90 (1.68–5.09)	1.07	0.0001
Age >60 y	2.12 (1.12–3.82)	0.75	0.01
Treatment with ≥2 antihypertensives	1.78 (1.04–3.02)	0.58	0.03
E/e' ratio >9	1.87 (1.07–3.26)	0.63	0.03
Pulmonary artery systolic pressure >35 mm Hg	1.74 (0.92–3.35)	0.55	0.09
Diabetes mellitus or prediabetes	1.67 (0.97–2.87)	0.51	0.06
LA volume index >30 mL/m^2	1.59 (0.88–2.88)	0.47	0.1
Chronic kidney disease stage 3 or greater	1.46 (0.66–3.30)	0.37	0.4
NT-proBNP >275 pg/mL	1.26 (0.66–2.41)	0.23	0.5
H$_2$FPEF score (AICc, 393.36; AUC, 0.841; *P*<0.0001)			
Body mass index >30 kg/m^2	3.10 (1.85–5.18)	1.13 (Score 2)	<0.0001
Atrial fibrillation	5.78 (2.28–14.62)	1.75 (Score 3)	<0.0001
Age >60 y	2.83 (1.65–4.84)	1.04 (Score 1)	0.0001
Treatment with ≥2 antihypertensives	1.99 (1.18–3.33)	0.69 (Score 1)	0.01
E/e' >9	2.15 (1.27–3.67)	0.77 (Score 1)	0.005
Pulmonary artery systolic pressure >35 mm Hg	2.05 (1.11–3.78)	0.72 (Score 1)	0.02

AICc indicates Akaike information criterion, corrected; AUC, area under the curve; E/e', ratio of early diastolic mitral inflow velocity to septal mitral annulus tissue relaxation velocity; HFpEF, heart failure with preserved ejection fraction; LA, left atrial; NT-proBNP, N-terminal pro-B-type natriuretic peptide; and OR, odds ratio.

Sensitivity Analyses

The agnostic CART model used a more nuanced diagnostic scheme because it allows empirically determined thresholds and interactions based on patterns in the data. As a result, the CART was slightly more predictive than the logistic regression–derived H$_2$FPEF score, with an AUC of 0.8831, an increase of 0.044 (*P*=0.002; Figure IV and Table III in the online-only Data Supplement). The agnostic logistic model based on automated stepwise backward selection of all predictors also verified discrimination similar to the H$_2$FPEF score (AUC, 0.857) and included the same variables, except that right ventricular fractional area change supplanted pulmonary artery systolic pressure as being predictive in the final agnostic logistic model (Table IV in the online-only Data Supplement).

Sensitivity analyses applying the H$_2$FPEF model restricted to local patients from the regional practice (AUC, 0.841) or patients with early-stage HFpEF (AUC, 0.814) demonstrated a performance similar to that of the overall cohort (Figure V in the online-only Data Supplement).

Validation in the Test Cohort

The test cohort included 100 consecutive patients (61 cases with HFpEF and 39 controls) whose baseline characteristics were similar to those of the derivation cohort (Table V in the online-only Data Supplement). Performance of the points-based H$_2$FPEF score (AUC, 0.886) and continuous variable–based score (AUC, 0.910) remained robust in this cohort (Table III in the online-only Data Supplement).

DISCUSSION

HFpEF accounts for half of heart failure hospitalizations, and in hospitalized patients, overt congestion is typically obvious from physical examination, chest radiography, and natriuretic peptide assays.[1] However, in outpatients with exertional dyspnea, overt congestion is often absent at rest, and the diagnosis may be challenging.[7,8] Right-sided heart catheterization, with exercise if resting filling pressures are normal, is the gold standard for HFpEF diagnosis but is not universally available, and noninvasive estimates of cardiac filling pressures lack sensitivity.[1–8] In this study, we derived and then validated a new score using clinical and echocardiographic variables that are widely available in clinical practice. In the derivation and test cohorts and in sensitivity analyses restricted to community-based patients and those with early-stage HFpEF, the H$_2$FPEF score effectively discriminated patients with HFpEF from a comparator population of patients with exertional dyspnea that was not caused by heart failure, ascertained with the gold standard of invasive hemodynamic exercise testing. Inclusion

	Clinical Variable	Values	Points
H_2	**H**eavy	Body mass index > 30 kg/m^2	2
	Hypertensive	2 or more antihypertensive medicines	1
F	Atrial **F**ibrillation	Paroxysmal or Persistent	3
P	**P**ulmonary Hypertension	Doppler Echocardiographic estimated Pulmonary Artery Systolic Pressure > 35 mmHg	1
E	**E**lder	Age > 60 years	1
F	**F**illing Pressure	Doppler Echocardiographic E/e' > 9	1
H_2FPEF score			Sum (0-9)

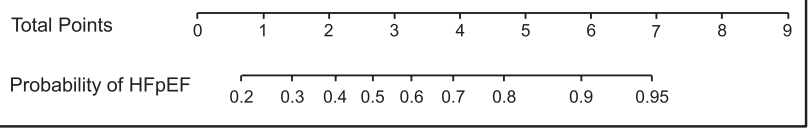

Figure 1. Description of the H_2FPEF score. Description of the H_2FPEF score and point allocations for each clinical characteristic (top), with associated probability of having heart failure with preserved ejection fraction (HFpEF) based on the total score as estimated from the model (bottom).

of this control group was crucial to our study design because it would not have otherwise been possible to judge the ability of clinical characteristics to estimate the likelihood of HFpEF without the ability to definitively identify or exclude disease on the basis of invasive criteria.

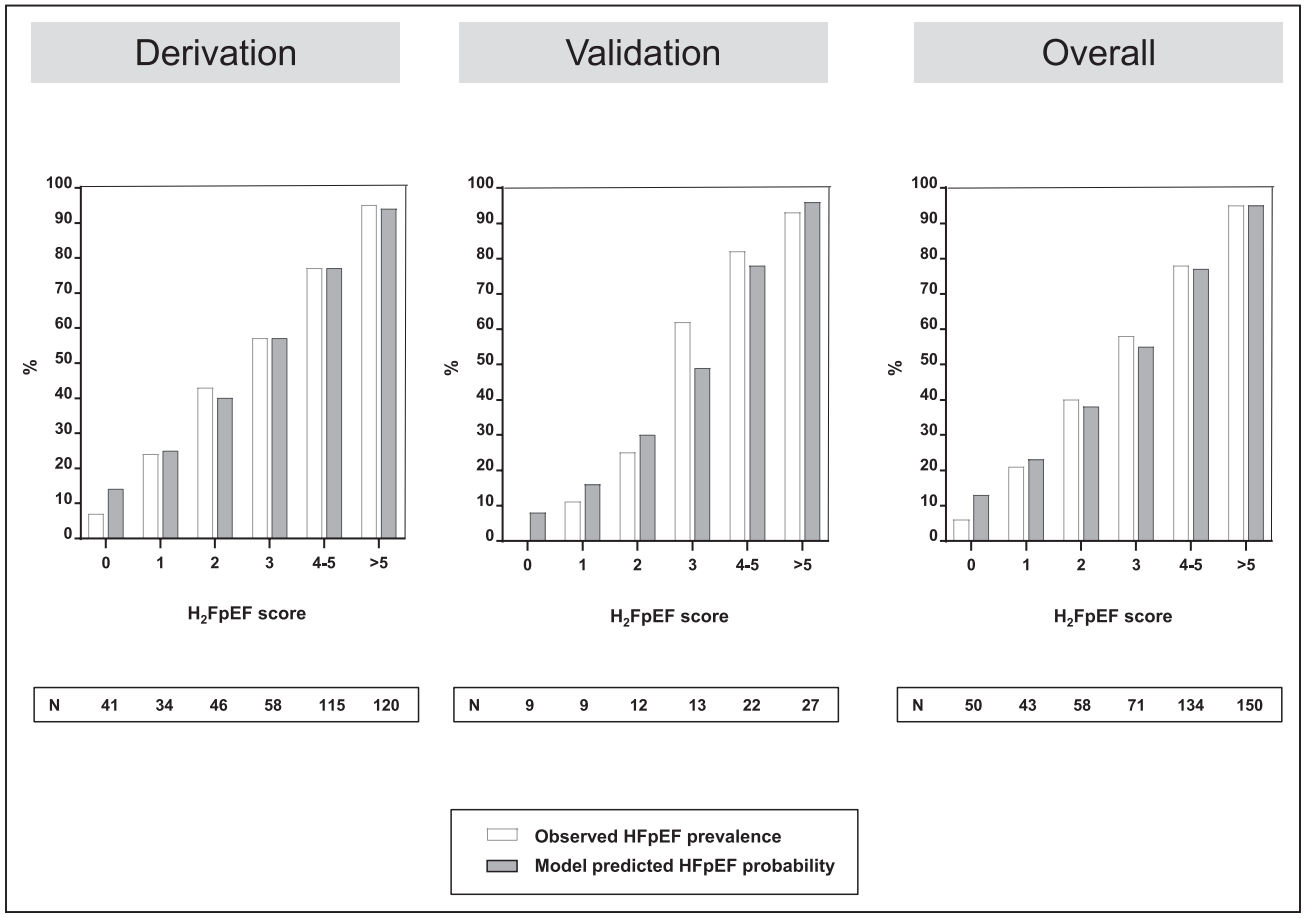

Figure 2. Calibration of the H_2FPEF score.
The Hosmer-Lemeshow goodness-of-fit test results using deciles of predicted probabilities were *P*=0.14, 0.53, and 0.18 for the derivation, validation, and pooled overall sample, respectively, indicating support for a properly calibrated model. HFpEF indicates heart failure with preserved ejection fraction.

Diagnostic algorithms for HFpEF used in practice and for entry into clinical trials are based on expert consensus opinion.[4,5] When these criteria were prospectively evaluated, specificity was robust but sensitivity was poor.[7] Therefore, HFpEF remains underdiagnosed in the community. In recent years, there has been increased use of invasive cardiopulmonary exercise testing to evaluate patients with exertional dyspnea, which is the gold standard to establish or refute the diagnosis of HFpEF.[7–10] Although this definitive approach has been shown to be cost-effective and safe,[19] its uniform application is not practical for all diagnostic evaluations, given the enormous number of patients in the community presenting with exertional dyspnea.

By establishing the probability of disease, the H_2FPEF score may be used to effectively rule out disease among patients with low scores (eg, 0 or 1), to establish the diagnosis with reasonably high confidence at higher scores (eg, 6–9), and to identify patients for whom additional testing is needed with intermediate scores (eg, 2–5). Rather than forcing a probabilistic diagnosis (HFpEF) into binary categories (present or absent), this Bayesian approach provides a framework that can be used to determine whether there is sufficient confidence in the working diagnosis or whether further evaluation is necessary based on the identified probability of disease. This system could be readily applied for diagnostic purposes in clinical care and research settings to help refine enrollment criteria for clinical trials. Although the categorical H_2FPEF score is easily calculated even at the bedside to rapidly estimate low or high probability of HFpEF, the more complex continual HFpEF calculator (online-only Data Supplement) can also be used to provide a more precise estimate of the probability of HFpEF in an individual when required for clinical use or in screening or research settings.

Selection of the Final Model

In this analysis, we examined complementary modeling strategies that strove to balance parsimony, ease of calculation, and discriminatory capabilities. Although we also considered more complex machine learning approaches, we finalized our models using multiple logistic regression analysis and the agnostic CART. Many of the candidate variables for the models were highly collinear, so multiple sets of variables were often found to be equally discriminatory. The final model reflected a combination of variables selected a priori because of their central role in HFpEF pathogenesis (eg, obesity and atrial fibrillation), as well as stepwise multivariable regression with systematic backward elimination to include only variables that were independently predictive of HFpEF in combination. This yielded the components of our final H_2FPEF score.

Sensitivity analyses using purely agnostic methods, including an unbiased logistic model, yielded nearly identical results, apart from the inclusion of right ventricular fractional area change in place of pulmonary artery systolic pressure (Table IV in the online-only Data Supplement). Because right ventricular fractional area change (a measure of right ventricular function) varies inversely with pulmonary artery pressure,[20] it is not surprising that both measures can discriminate HFpEF from noncardiac dyspnea. Because estimated pulmonary artery systolic pressure is a well-established marker of HFpEF[21] and is more commonly measured in practice, we chose to include it in the final model rather than right ventricular fractional area change, which is not part of the routine clinical echocardiogram in many centers.

The lack of a particular variable in the final model, such as NT-proBNP, should not be interpreted as revealing a lack of association with HFpEF. Rather, our data suggest that NT-proBNP may not add incremental information to clinical variables and echocardiography in diagnosing HFpEF among patients with unexplained dyspnea. This is in contrast to patients presenting with acute dyspnea that is present at rest, for which the diagnostic performance of the natriuretic peptides is well established.[22,23] Although the discrimination of cases and controls was slightly improved with the CART model and the continuous HFpEF score model, the differences were minor, and we propose that the simplicity of the H_2FPEF score system outweighs this difference because it improves the feasibility of applying this approach in everyday practice. However, if precise estimation of an individual patient's probability of underlying HFpEF is to be calculated, the more complex continuous variable version of the HFpEF score from our online calculator can be applied.

Association of Comorbidities With HFpEF

HFpEF is currently believed to be a systemic disorder driven in large part by comorbidities.[2,3] We observed that 2 comorbidities, obesity and atrial fibrillation, independently increase the probability that HFpEF is present. Severe hypertension identified by treatment with ≥2 antihypertensive drugs was another independent predictor. Diabetes mellitus is common in HFpEF, seen in 30% to 40%,[24] but the presence of abnormal glucose tolerance did not add incremental diagnostic value beyond obesity alone, supporting the emerging evidence of the importance of obesity as a cause of HFpEF.[14]

Limitations

NT-proBNP data were missing at random in 24% of patients because some cardiologists did not obtain this laboratory value during their evaluation. Therefore, imputation was performed to account for the missing data, which

may have affected the inclusion of NT-proBNP in the final model. However, a sensitivity analysis yielded similar results in the 76% of patients who had directly measured NT-proBNP, increasing our confidence in the imputation-derived values. This study was single-center, limiting its generalizability. There is referral bias in that all patients were referred for invasive testing, which may have inflated the prevalence of HFpEF. However, this analysis would not have been possible without the use of a gold standard assessment. Although this study was performed in a tertiary referral center, our practice also serves the local population, and sensitivity analysis restricted to local patients revealed that the H$_2$FPEF score performed similarly well in this subset (AUC, 0.841), increasing confidence in the generalizability of our results. Although discrimination was maintained in our separate validation cohort, external validation was not performed, and replication in other centers is necessary. Physical examination findings were not included in the models because there may be variability in examination skill and interpretation[25] and because overt congestion was absent in the patients included in this study, who were deemed to have indeterminate dyspnea after thorough evaluation by board-certified cardiologists on the basis of history, physical examination, and echocardiography. Therefore, the present results may not apply to patients with more frank evidence of tissue congestion, in whom testing beyond the history and physical examination may not be necessary to diagnose HFpEF. Assessment for lung disease was performed at the discretion of referring physicians and was not done in all patients. However, this reflects practice in the community, and the presence or absence of pulmonary disease is not relevant to the primary study goal of discriminating cardiac dyspnea (HFpEF) and noncardiac dyspnea.

CONCLUSIONS

The H$_2$FPEF score, which uses 6 clinical and echocardiographic characteristics that are universally obtained in the evaluation of patients with unexplained exertional dyspnea, enables robust discrimination of HFpEF from noncardiac causes of dyspnea at low and high scores while identifying patients at intermediate probability in whom additional testing is needed to refine the diagnosis.

ARTICLE INFORMATION

Received February 28, 2018; accepted May 2, 2018.

The online-only Data Supplement, podcast, and transcript are available with this article at https://www.ahajournals.org/doi/suppl/10.1161/circulationaha.118.034646.

Correspondence

Barry A. Borlaug, MD, Mayo Clinic and Foundation, 200 First Street SW, Rochester, MN 55905. Email borlaug.barry@mayo.edu

Affiliations

Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN (Y.N.V.R., M.O., M.M.R., B.A.B.). Department of Health Sciences Research, Mayo Clinic, Jacksonville, FL (R.E.C.).

Sources of Funding

Dr Borlaug is supported by R01 HL128526, R01 HL126638, U01 HL125205, and U10 HL110262 from the National Institutes of Health. Dr Reddy is supported by T32 HL007111 from the National Institutes of Health. Dr Carter is supported by R01 HL128526 and UL1 TR02377 from the National Institutes of Health. Dr Obokata is supported by a research fellowship from the Uehara Memorial Foundation, Japan.

Disclosures

None.

REFERENCES

1. Owan TE, Hodge DO, Herges RM, Jacobsen SJ, Roger VL, Redfield MM. Trends in prevalence and outcome of heart failure with preserved ejection fraction. *N Engl J Med*. 2006;355:251–259. doi: 10.1056/NEJMoa052256
2. Redfield MM. Heart failure with preserved ejection fraction. *N Engl J Med*. 2016;375:1868–1877. doi: 10.1056/NEJMcp1511175
3. Shah SJ, Kitzman DW, Borlaug BA, van Heerebeek L, Zile MR, Kass DA, Paulus WJ. Phenotype-specific treatment of heart failure with preserved ejection fraction: a multiorgan roadmap. *Circulation*. 2016;134:73–90. doi: 10.1161/CIRCULATIONAHA.116.021884
4. Paulus WJ, Tschöpe C, Sanderson JE, Rusconi C, Flachskampf FA, Rademakers FE, Marino P, Smiseth OA, De Keulenaer G, Leite-Moreira AF, Borbély A, Edes I, Handoko ML, Heymans S, Pezzali N, Pieske B, Dickstein K, Fraser AG, Brutsaert DL. How to diagnose diastolic heart failure: a consensus statement on the diagnosis of heart failure with normal left ventricular ejection fraction by the Heart Failure and Echocardiography Associations of the European Society of Cardiology. *Eur Heart J*. 2007;28:2539–2550. doi: 10.1093/eurheartj/ehm037
5. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, Falk V, González-Juanatey JR, Harjola VP, Jankowska EA, Jessup M, Linde C, Nihoyannopoulos P, Parissis JT, Pieske B, Riley JP, Rosano GMC, Ruilope LM, Ruschitzka F, Rutten FH, van der Meer P; ESC Scientific Document Group. 2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J*. 2016;37:2129–2200. doi: 10.1093/eurheartj/ehw128
6. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr, Drazner MH, Fonarow GC, Geraci SA, Horwich T, Januzzi JL, Johnson MR, Kasper EK, Levy WC, Masoudi FA, McBride PE, McMurray JJ, Mitchell JE, Peterson PN, Riegel B, Sam F, Stevenson LW, Tang WH, Tsai EJ, Wilkoff BL. 2013 ACCF/AHA guideline for the management of heart failure: executive summary: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2013;128:1810–1852. doi: 10.1161/CIR.0b013e31829e8807
7. Obokata M, Kane GC, Reddy YN, Olson TP, Melenovsky V, Borlaug BA. Role of diastolic stress testing in the evaluation for heart failure with preserved ejection fraction: a simultaneous invasive-echocardiographic study. *Circulation*. 2017;135:825–838. doi: 10.1161/CIRCULATIONAHA.116.024822
8. Borlaug BA, Nishimura RA, Sorajja P, Lam CS, Redfield MM. Exercise hemodynamics enhance diagnosis of early heart failure with preserved ejection fraction. *Circ Heart Fail*. 2010;3:588–595. doi: 10.1161/CIRCHEARTFAILURE.109.930701
9. Maron BA, Cockrill BA, Waxman AB, Systrom DM. The invasive cardiopulmonary exercise test. *Circulation*. 2013;127:1157–1164. doi: 10.1161/CIRCULATIONAHA.112.104463
10. Givertz MM, Fang JC, Sorajja P, Dimas V, Forfia PR, Kapur NK, Kern MJ, Naidu SS, Borlaug BA. Executive summary of the SCAI/HFSA clinical expert consensus document on the use of invasive hemodynamics for the diagnosis and management of cardiovascular disease. *J Card Fail*. 2017;23:487–491. doi: 10.1016/j.cardfail.2017.04.013

11. Lang RM, Badano LP, Mor-Avi V, Afilalo J, Armstrong A, Ernande L, Flachskampf FA, Foster E, Goldstein SA, Kuznetsova T, Lancellotti P, Muraru D, Picard MH, Rietzschel ER, Rudski L, Spencer KT, Tsang W, Voigt JU. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging*. 2015;16:233–270. doi: 10.1093/ehjci/jev014
12. Stekhoven DJ, Bühlmann P. MissForest–non-parametric missing value imputation for mixed-type data. *Bioinformatics*. 2012;28:112–118. doi: 10.1093/bioinformatics/btr597
13. Sullivan LM, Massaro JM, D'Agostino RB Sr. Presentation of multivariate data for clinical use: the Framingham study risk score functions. *Stat Med*. 2004;23:1631–1660. doi: 10.1002/sim.1742
14. Obokata M, Reddy YNV, Pislaru SV, Melenovsky V, Borlaug BA. Evidence supporting the existence of a distinct obese phenotype of heart failure with preserved ejection fraction. *Circulation*. 2017;136:6–19. doi: 10.1161/CIRCULATIONAHA.116.026807
15. Kitzman DW, Lam CSP. Obese heart failure with preserved ejection fraction phenotype: from pariah to central player. *Circulation*. 2017;136:20–23. doi: 10.1161/CIRCULATIONAHA.117.028365
16. Reddy YNV, Obokata M, Gersh BJ, Borlaug BA. High prevalence of occult heart failure with preserved ejection fraction among patients with atrial fibrillation and dyspnea. *Circulation*. 2018;137:534–535. doi: 10.1161/CIRCULATIONAHA.117.030093
17. Lam CS, Rienstra M, Tay WT, Liu LC, Hummel YM, van der Meer P, de Boer RA, Van Gelder IC, van Veldhuisen DJ, Voors AA, Hoendermis ES. Atrial fibrillation in heart failure with preserved ejection fraction: association with exercise capacity, left ventricular filling pressures, natriuretic peptides, and left atrial volume. *JACC Heart Fail*. 2017;5:92–98. doi: 10.1016/j.jchf.2016.10.005
18. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001;54:774–781.
19. Huang W, Resch S, Oliveira RK, Cockrill BA, Systrom DM, Waxman AB. Invasive cardiopulmonary exercise testing in the evaluation of unexplained dyspnea: insights from a multidisciplinary dyspnea center. *Eur J Prev Cardiol*. 2017;24:1190–1199. doi: 10.1177/2047487317709605
20. Melenovsky V, Hwang SJ, Lin G, Redfield MM, Borlaug BA. Right heart dysfunction in heart failure with preserved ejection fraction. *Eur Heart J*. 2014;35:3452–3462. doi: 10.1093/eurheartj/ehu193
21. Lam CS, Roger VL, Rodeheffer RJ, Borlaug BA, Enders FT, Redfield MM. Pulmonary hypertension in heart failure with preserved ejection fraction: a community-based study. *J Am Coll Cardiol*. 2009;53:1119–1126. doi: 10.1016/j.jacc.2008.11.051
22. Steinhart B, Thorpe KE, Bayoumi AM, Moe G, Januzzi JL Jr, Mazer CD. Improving the diagnosis of acute heart failure using a validated prediction model. *J Am Coll Cardiol*. 2009;54:1515–1521. doi: 10.1016/j.jacc.2009.05.065
23. Januzzi JL Jr, Chen-Tournoux AA, Christenson RH, Doros G, Hollander JE, Levy PD, Nagurney JT, Nowak RM, Pang PS, Patel D, Peacock WF, Rivers EJ, Walters EL, Gaggin HK; ICON-RELOADED Investigators. N-terminal pro-B-type natriuretic peptide in the emergency department: the ICON-RELOADED Study. *J Am Coll Cardiol*. 2018;71:1191–1200. doi: 10.1016/j.jacc.2018.01.021
24. Lindman BR, Dávila-Román VG, Mann DL, McNulty S, Semigran MJ, Lewis GD, de las Fuentes L, Joseph SM, Vader J, Hernandez AF, Redfield MM. Cardiovascular phenotype in HFpEF patients with or without diabetes: a RELAX trial ancillary study. *J Am Coll Cardiol*. 2014;64:541–549. doi: 10.1016/j.jacc.2014.05.030
25. From AM, Lam CS, Pitta SR, Kumar PV, Balbissi KA, Booker JD, Singh IM, Sorajja P, Reeder GS, Borlaug BA. Bedside assessment of cardiac hemodynamics: the impact of noninvasive testing and examiner experience. *Am J Med*. 2011;124:1051–1057. doi: 10.1016/j.amjmed.2011.05.034