JACC: ASIA VOL. 4, NO. 12, 2024

ORIGINAL RESEARCH

Predicting Mortality and Hospitalization in Heart Failure With Preserved Ejection Fraction by Using Machine Learning

Chieh-Yu Chang, MD,[a,*] Chun-Chi Chen, MD,[a,*] Ming-Lung Tsai, MD,[b] Ming-Jer Hsieh, MD, PHD,[a] Tien-Hsing Chen, MD,[c] Shao-Wei Chen, MD, PHD,[d,e] Shang-Hung Chang, MD, PHD,[a] Pao-Hsien Chu, MD,[a] I-Chang Hsieh, MD,[a] Ming-Shien Wen, MD,[a] Dong-Yi Chen, MD[a]

ABSTRACT

BACKGROUND Few studies have incorporated echocardiography and laboratory data to predict clinical outcomes in heart failure with preserved ejection fraction (HFpEF).

OBJECTIVES This study aimed to use machine learning to find predictors of heart failure (HF) hospitalization and cardiovascular (CV) death in HFpEF.

METHODS From the Chang Gung Research Database in Taiwan, 6,092 HFpEF patients (2,898 derivation, 3,194 validation) identified between 2008 and 2017 were followed until 2019. A random survival forest model, using 58 variables, was developed to predict the composite outcome of HF hospitalization and CV death.

RESULTS During 2.9-year follow-up, 37.7% of derivation and 36.0% of validation cohort patients experienced HF hospitalization or CV death. The study identified 15 predictive indicators, including age ≥65 years, B-type natriuretic peptide level ≥600 pg/mL, left atrium size ≥46 mm, atrial fibrillation, frequency of HF hospitalization within 3 years, body mass index <30 kg/m^2, moderate or severe mitral regurgitation, left ventricular (LV) posterior wall thickness of <10 or ≥13 mm, dysnatremia, LV end-diastolic dimension of <40 or ≥56 mm, uric acid level ≥7 mg/dL, triglyceride level of <70 or ≥200 mg/dL, blood urea nitrogen level ≥20 mg/dL, interventricular septum thickness of <11 or ≥20 mm, and glycated hemoglobin (HbA_{1c}) level of <6% or ≥8%. The random survival forest model demonstrated robust external generalizability with an 86.9% area under curve in validation.

CONCLUSIONS Machine learning identified 15 predictors of HF hospitalization and CV death in HFpEF patients, helping doctors identify high-risk individuals for tailored treatment. (JACC Asia. 2024;4:956-968) © 2024 The Authors. Published by Elsevier on behalf of the American College of Cardiology Foundation. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

From the [a]Division of Cardiology, Department of Internal Medicine, Chang Gung Memorial Hospital at Linkou, and Chang Gung University College of Medicine, Taoyuan, Taiwan; [b]Division of Cardiology, New Taipei Municipal TuCheng Hospital, Taiwan, and Chang Gung University College of Medicine, Taoyuan, Taiwan; [c]Division of Cardiology, Department of Internal Medicine, Chang Gung Memorial Hospital at Keelung, and Chang Gung University College of Medicine, Taoyuan, Taiwan; [d]Department of Thoracic and Cardiovascular Surgery, Chang Gung Memorial Hospital at Linkou, and Chang Gung University College of Medicine, Taoyuan, Taiwan; and the [e]Center for Big Data Analytics and Statistics, Department of Medical Research and Development, Chang Gung Memorial Hospital at Linkou, Taoyuan, Taiwan. *Drs Chang and Chen contributed equally to this work.
The authors attest they are in compliance with human studies committees and animal welfare regulations of the authors' institutions and Food and Drug Administration guidelines, including patient consent where appropriate. For more information, visit the Author Center.

Manuscript received April 3, 2024; revised manuscript received August 19, 2024, accepted September 8, 2024.

ISSN 2772-3747 https://doi.org/10.1016/j.jacasi.2024.09.003

Heart failure (HF) is a complex and life-threatening syndrome that profoundly affects quality of life and is a significant contributor to morbidity and mortality worldwide.[1] In contrast to heart failure with reduced ejection fraction (HFrEF), heart failure with preserved ejection fraction (HFpEF) is characterized by compromised diastolic function, resulting in elevated filling pressure and pulmonary congestion.[2] The prevalence of HFpEF is increasing; the condition accounts for nearly one-half of all HF cases.[3] Individuals with HFpEF are at increased risks of recurrent HF-related hospitalizations and poor clinical outcomes and have a 5-year survival rate of <50%, similar to that of patients with HFrEF; HFpEF is thus a substantial public health burden.[3] Due to the distinct pathophysiological mechanisms of HFpEF and HFrEF, the risk factors influencing clinical outcomes must be assessed separately for each condition.

Numerous risk calculation models are available for predicting the clinical outcomes of patients with HF; however, few specifically focus on patients with HFpEF.[4-7] The I-PRESERVE (Irbesartan in Heart Failure with Preserved Ejection Fraction Study) is one of the few predictive investigations focusing on patients with HFpEF.[8] The I-PRESERVE trial utilized multivariable Cox regression models to identify factors associated with the outcomes of HFpEF. Nevertheless, the study was limited because it excluded patients with moderate or severe renal dysfunction and because it did not perform validation in an independent cohort. In another study, data from the TOPCAT (Treatment of Preserved Cardiac Function Heart Failure with an Aldosterone Antagonist) trial were utilized to formulate models for predicting mortality and HF hospitalization of outpatients with HFpEF.[9] However, that study was limited by the lack of follow-up data and did not include time-to-event analysis. Additionally, neither the I-PRESERVE nor the TOPCAT trial studies included echocardiographic parameters in their prediction models. Consequently, a substantial gap regarding the prediction of clinical outcomes for patients with HFpEF remains in the published data.

Novel machine learning methods, such as random survival forest (RSF), can accommodate high-dimensional and nonlinear effects of variables, enabling identification of the effects of individual variables from large variable sets.[10] These methods enable the calibration and integrated analysis of factors encompassing demographic, clinical, biological, and echocardiographic variables to identify predictors associated with adverse clinical outcomes. The present study employed RSF to formulate and validate models predicting HF hospitalization and cardiovascular (CV) death in patients with HFpEF. The multi-institutional medical records database of the Chang Gung Memorial Hospitals (CGMH) in Taiwan was employed for data.

METHODS

DATA SOURCE. The data set for this study was sourced from the Chang Gung Research Database (CGRD). This electronic medical data set contains systematically gathered records of all emergency services and inpatient and outpatient visits within the CGMH system.[11] The CGMH system is a multi-institutional network comprising 7 branches across Taiwan, of which 4 are tertiary academic medical centers and 3 are teaching hospitals.[12] The CGRD provides comprehensive information encompassing demographic characteristics, electronic medical records, pharmacy dispensing details, echocardiography reports, laboratory results, discharge summaries, and nursing records. Disease diagnoses and procedures are documented using the International Classification of Diseases-9th Revision-Clinical Modification (ICD-9-CM) for cases before 2016 and the International Classification of Diseases-10th Revision-Clinical Modification (ICD-10-CM) for cases after 2016. The ICD-9-CM and ICD-10-CM codes used in this study are listed in Supplemental Table 1. The study received approval from the Ethics Review Board of CGMH (IRB No.202300950B0). All patient data were processed using anonymization and deidentification measures, and the requirement for informed consent was thus waived.

IDENTIFICATION OF HFpEF COHORT. In this study, individuals admitted to CGMH hospitals for acute HF between January 1, 2007, and December 31, 2017, were identified from the CGRD. If patients were admitted for HF 2 or more times during the inclusion period, the first admission was selected as the index HF hospitalization. The exclusion criteria were the absence of essential demographic information (age and sex), age under 20 years, missing left ventricular ejection fraction (LVEF) data during hospitalization, in-hospital mortality, and being previously lost to follow-up after discharge (Supplemental Figure 1). Patient baseline echocardiography data from the preceding 3 months of index admission were obtained from examination reports within the CGRD; these parameters are presented in **Table 1**. The present study included patients meeting the criteria for HFpEF, specifically, with baseline LVEF ≥50%,

ABBREVIATIONS AND ACRONYMS

BNP = B-type natriuretic peptides

BUN = blood urea nitrogen

CGRD = Chang Gung Research Database

CV = cardiovascular

HF = heart failure

HFpEF = heart failure with preserved ejection fraction

HFrEF = heart failure with reduced ejection fraction

LA = left atrium

LV = left ventricular

RSF = random survival forest

VIMP = variable importance

TABLE 1 Baseline Characteristics of Patients in the Derivation (Northern Branches) and Validation (Southern Branches) Cohorts

	Available Number	Total (N = 6,092)	Derivation Cohort (n = 2,898)	Validation Cohort (n = 3,194)	*P* Value	STD
Age, y	6,092	75.0 ± 12.8	75.3 ± 13.1	74.7 ± 12.4	0.092	0.04
Age, y	6,092				0.002	
18-49		272 (4.5)	134 (4.6)	138 (4.3)		0.01
50-64		956 (15.7)	445 (15.4)	511 (16.0)		−0.02
65-74		1,299 (21.3)	602 (20.8)	697 (21.8)		−0.03
75-84		2,267 (37.2)	1,037 (35.8)	1,230 (38.5)		−0.06
≥85		1,298 (21.3)	680 (23.5)	618 (19.3)		0.10
Male	6,092	2,611 (42.9)	1,268 (43.8)	1,343 (42.0)	0.179	0.03
Smoking	6,092	1,200 (19.7)	655 (22.6)	545 (17.1)	<0.001	0.14
Body mass index, kg/m^2	5,907	25.5 ± 5.9	25.7 ± 5.8	25.4 ± 5.9	0.021	0.06
Body surface area, m^2	5,907	1.65 ± 0.22	1.66 ± 0.23	1.65 ± 0.22	0.046	0.05
Vital sign at admission						
Heart rate, beats/min	6,091	84.8 ± 19.9	84.2 ± 20.2	85.5 ± 19.6	0.011	−0.06
Systolic blood pressure, mm Hg	6,091	138.0 ± 26.7	135.9 ± 26.0	139.9 ± 27.2	<0.001	−0.15
Diastolic blood pressure, mm Hg	6,091	75.7 ± 16.2	74.3 ± 15.7	77.0 ± 16.5	<0.001	−0.17
Number of HFH in the previous 3 y	6,092				<0.001	
0		4,309 (70.7)	2,441 (84.2)	1,868 (58.5)		0.59
1		1,529 (25.1)	355 (12.2)	1,174 (36.8)		−0.59
≥2		254 (4.2)	102 (3.5)	152 (4.8)		−0.06
Comorbidity						
Hypertension	6,092	4,940 (81.1)	2,318 (80.0)	2,622 (82.1)	0.036	−0.05
Diabetes mellitus	6,092	3,072 (50.4)	1,427 (49.2)	1,645 (51.5)	0.078	−0.05
Hyperlipidemia	6,092	2,504 (41.1)	1,168 (40.3)	1,336 (41.8)	0.227	−0.03
Peripheral arterial disease	6,092	670 (11.0)	357 (12.3)	313 (9.8)	0.002	0.08
Liver cirrhosis	6,092	335 (5.5)	153 (5.3)	182 (5.7)	0.474	−0.02
Atrial fibrillation	6,092	2,688 (44.1)	1,371 (47.3)	1,317 (41.2)	<0.001	0.12
Myocardial infarction	6,092	471 (7.7)	212 (7.3)	259 (8.1)	0.247	−0.03
Stroke	6,092	943 (15.5)	404 (13.9)	539 (16.9)	0.002	−0.08
Coronary artery disease	6,092	2,754 (45.2)	1,241 (42.8)	1,513 (47.4)	<0.001	−0.09
Chronic obstructive pulmonary disease	6,092	1,829 (30.0)	768 (26.5)	1,061 (33.2)	<0.001	−0.15
Malignancy	6,092	924 (15.2)	467 (16.1)	457 (14.3)	0.050	0.05
Previous coronary revascularization	6,092	571 (9.4)	191 (6.6)	380 (11.9)	<0.001	−0.18
Baseline echocardiography						
Left ventricular ejection fraction, %	6,092	65.6 ± 9.1	65.4 ± 9.1	65.8 ± 9.2	0.087	−0.04
Left ventricular end-diastolic diameter, mm	6,092	48.7 ± 7.7	48.5 ± 7.8	48.8 ± 7.6	0.184	−0.03
Left ventricular end-systolic diameter, mm	6,092	30.9 ± 6.6	30.9 ± 6.7	30.8 ± 6.6	0.574	0.01
Left ventricular end-diastolic volume, mL	6,092	115.0 ± 41.8	114.3 ± 42.0	115.6 ± 41.6	0.238	−0.03
Left ventricular end-systolic volume, mL	6,092	40.4 ± 20.7	40.6 ± 20.7	40.2 ± 20.6	0.470	0.02
Left atrium diameter, mm	6,062	43.6 ± 9.1	45.3 ± 9.6	42.1 ± 8.3	<0.001	0.35
Interventricular septum thickness, mm	6,086	12.7 ± 3.0	12.8 ± 3.2	12.7 ± 2.7	0.387	0.02
Left ventricular posterior wall thickness, mm	6,046	10.9 ± 2.4	11.2 ± 2.6	10.6 ± 2.2	<0.001	0.22
Left ventricular mass, g	6,046	225.6 ± 82.1	229.4 ± 86.5	222.2 ± 77.8	0.001	0.09
Left ventricular mass index, g/m^2	5,866	137.3 ± 49.1	139.1 ± 51.4	135.7 ± 46.9	0.007	0.07
E/A ratio	1,981	1.11 ± 0.71	1.15 ± 0.76	1.06 ± 0.62	0.004	0.14
Peak E-wave velocity, cm/s	1,705	90.5 ± 35.9	92.3 ± 37.4	88.1 ± 33.8	0.019	0.12
Peak A-wave velocity, cm/s	1,627	91.6 ± 32.1	90.6 ± 32.3	93.0 ± 31.8	0.145	−0.07
Relative wall thickness, mm	6,046	0.46 ± 0.13	0.47 ± 0.14	0.45 ± 0.12	<0.001	0.19
Tricuspid regurgitation peak gradient, mm Hg	5,735	32.0 ± 15.7	33.3 ± 16.4	30.8 ± 15.0	<0.001	0.15

Continued on the next page

receiving a diuretic during hospitalization, and having a B-type natriuretic peptide (BNP) level exceeding 35 pg/mL (for patients with sinus rhythm) or exceeding 105 pg/mL (for patients with atrial fibrillation). The index date was defined as the date of discharge after the index HF admission. The patient cohort was subsequently divided into 2 subcohorts—a derivation cohort and a validation cohort—on the

TABLE 1 Continued

	Available Number	Total (N = 6,092)	Derivation Cohort (n = 2,898)	Validation Cohort (n = 3,194)	*P* Value	STD
Mitral regurgitation severity	5,953				<0.001	
None		1,093 (18.4)	519 (18.1)	574 (18.6)		−0.01
Mild		3,095 (52.0)	1,537 (53.7)	1,558 (50.4)		0.07
Moderate		1,167 (19.6)	567 (19.8)	600 (19.4)		0.01
Severe		598 (10.0)	240 (8.4)	358 (11.6)		−0.11
Laboratory data						
B-type natriuretic peptide, pg/mL	6,092	623 (293, 1199)	579 (275, 1120)	672 (308, 1260)	<0.001	−0.13
Blood urea nitrogen, mg/dL	5,896	33.3 ± 24.2	32.9 ± 23.5	33.6 ± 24.9	0.283	−0.03
Creatinine, mg/dL	6,088	1.91 ± 1.85	1.83 ± 1.71	1.99 ± 1.97	0.001	−0.09
eGFR, mL/min/1.73 m^2	6,092	54.9 ± 33.8	56.2 ± 34.2	53.8 ± 33.4	0.007	0.07
Sodium, mEq/L	6,091	137.1 ± 5.3	137.7 ± 5.2	136.6 ± 5.5	<0.001	0.19
Potassium, mEq/L	6,091	3.98 ± 0.67	3.99 ± 0.67	3.98 ± 0.68	0.476	0.02
Hemoglobin, g/dL	6,089	11.1 ± 2.3	11.1 ± 2.4	11.1 ± 2.3	0.728	0.01
Hematocrit, %	6,089	33.7 ± 6.9	33.8 ± 6.9	33.6 ± 6.9	0.129	0.04
Total cholesterol, mg/dL	3,887	159.8 ± 42.7	158.0 ± 42.9	161.6 ± 42.3	0.010	−0.08
Low-density lipoprotein, mg/dL	3,844	75.0 ± 43.2	78.4 ± 43.3	71.4 ± 42.8	<0.001	0.16
High-density lipoprotein, mg/dL	3,664	42.0 ± 14.0	40.2 ± 13.5	43.9 ± 14.2	<0.001	−0.27
Triglyceride, mg/dL	3,800	114.1 ± 66.1	114.8 ± 68.5	113.3 ± 63.5	0.476	0.02
HbA_{1c}, %	3,839	6.7 ± 1.4	6.7 ± 1.4	6.7 ± 1.4	0.374	−0.03
Uric acid, mg/dL	3,070	7.7 ± 2.6	7.8 ± 2.7	7.6 ± 2.6	0.105	0.06
Neutrophil, $10^3/\mu L$	6,084	6.8 ± 3.8	6.7 ± 3.8	6.9 ± 3.8	0.022	−0.06
Platelet, $10^3/\mu L$	6,088	204.8 ± 85.4	203.0 ± 86.6	206.4 ± 84.4	0.120	−0.04
Lymphocyte, $10^3/\mu L$	6,084	1.48 ± 1.02	1.49 ± 1.05	1.47 ± 0.99	0.507	0.02
White blood cell, $10^3/\mu L$	6,089	9.1 ± 4.2	9.0 ± 4.3	9.2 ± 4.2	0.054	−0.05
Neutrophil-to-lymphocyte ratio	6,082	4.5 (2.7, 8.1)	4.4 (2.6, 7.8)	4.6 (2.7, 8.3)	0.021	−0.02
Platelet lymphocyte ratio	6,081	11.7 (7.2, 21.1)	11.4 (7.0, 20.0)	12.1 (7.3, 21.8)	0.003	−0.02
Hypoglycemic medications						
Thiazolidinedione	6,092	189 (3.1)	56 (1.9)	133 (4.2)	<0.001	−0.13
GLP1RA/ SGLT2 inhibitor	6,092	24 (0.4)	17 (0.6)	7 (0.2)	0.022	0.06
Other medication						
ACEI/ARB/ARNI	6,092	4,052 (66.5)	1,834 (63.3)	2,218 (69.4)	<0.001	−0.13
Beta-blocker	6,092	3,600 (59.1)	1,908 (65.8)	1,692 (53.0)	<0.001	0.26
Mineralocorticoid receptor antagonist	6,092	1,672 (27.4)	740 (25.5)	932 (29.2)	0.001	−0.08
Digoxin	6,092	990 (16.3)	609 (21.0)	381 (11.9)	<0.001	0.25
Amiodarone	6,092	844 (13.9)	273 (9.4)	571 (17.9)	<0.001	−0.25
Prior implantable cardioverter-defibrillator	6,092	267 (4.4)	104 (3.6)	163 (5.1)	0.004	−0.07
Length of hospital stay, d	6,092	15.0 ± 13.7	15.5 ± 15.0	14.5 ± 12.3	0.004	0.07

Values are mean ± SD, n (%), or median (25th percentile, 75th percentile).

ACEI = angiotensin-converting enzyme inhibitor; ARB = angiotensin receptor blocker; ARNI = angiotensin receptor-neprilysin inhibitor; E/A = E-wave velocity/A-wave velocity; eGFR = estimated glomerular filtration rate; GLP1RA = glucagon-like peptide-1 receptor agonist; HFH = heart failure hospitalization; SGLT2 = sodium-glucose cotransporter 2; STD = standardized difference.

basis of whether the branch of the CGMH was a northern or southern branch. The northern branches of the CGMH are the Keelung, Taipei, Linkou, and Taoyuan branches, whereas the southern branches of the CGMH are the Chiayi, Kaohsiung, and Fengshan branches.

MEASUREMENT OF COVARIATES. This study examined demographic factors, physiological indicators, comorbidities, coronary intervention history, baseline echocardiography, laboratory findings, presence of implantable cardioverter-defibrillators, and concurrent medications. Patient smoking history was documented in the CGRD nursing care subdatabase. BMI, physiological parameters, laboratory results, echocardiographic measurements, and medication usage for the 3 months preceding the index date were derived from CGRD medical records. Comorbidities were defined based on 2 outpatient diagnoses or any single inpatient diagnosis before index admission. These included hypertension, diabetes, dyslipidemia, peripheral arterial disease, liver cirrhosis, atrial fibrillation, myocardial infarction, stroke, coronary artery disease, and chronic obstructive pulmonary

disease. The frequency of HF admissions in the preceding 3 years was examined to assess HF severity.

OUTCOMES AND FOLLOW-UP. The primary clinical outcome for the risk prediction model was a composite of HF hospitalization and CV death. The HF event was defined as hospitalization with a discharge diagnosis of heart failure, which has been validated in previous claim-based studies.[13] Furthermore, we restricted the unscheduled HF hospitalization to cases where the patient required at least 1 treatment involving a diuretic, nitrate, or inotropic agent.[12] The date, location, and cause of a patient's death were recorded in the Taiwan Death Registry database, which can be linked to CGRD. CV death was defined as including fatalities resulting from acute myocardial infarction, sudden cardiac death, HF, stroke, CV procedures, CV hemorrhage, or other CV causes.[14] The ICD-9-CM and ICD-10-CM codes utilized for identifying baseline comorbidities and composite outcomes are outlined in Supplemental Table 1. Patients were followed from the index date (the day of their discharge from the index HF hospitalization) until the day of HF hospitalization or CV death within 5 years, the fifth year (for those who did not develop subsequent HF and were still at risk), or the end of the database cutoff period for this study (December 31, 2019), whichever came first.

RSF ANALYSIS. All variables presented in **Table 1** were used as candidate predictor variables, except for 3 echocardiographic parameters (peak E-wave, A-wave velocity, and E/A ratio) because of the lack of this information for more than two-thirds of the cohort. We also excluded body surface area, left ventricular (LV) end-diastolic volume, and LV end-systolic volume from the predictor variables because of their strong correlation with BMI, left ventricular end-diastolic dimension (LVEDD), and LV end-systolic dimension, respectively. The baseline characteristics of the patients after single expectation-maximization imputation of missing data are presented in Supplemental Table 2. The initial RSF model for HF hospitalization and CV death was constructed using 58 variables on the basis of the data for the derivation cohort.[15] For RSF analyses, we employed an iterative node-splitting process commencing at the root node to construct a binary survival tree. At each node, variables were randomly selected to determine the optimal split, with the aim of maximizing the disparity in survival between daughter nodes. The tree was systematically expanded to its maximum extent while ensuring that each terminal node encapsulated at least one distinct outcome. This tree construction procedure was performed 500 times. Subsequently, the cumulative hazard function for each tree was computed, and an ensemble hazard function was formulated by averaging across all trees. We employed variable importance (VIMP) to rank the 58 variables by their prognostic capability. After developing the initial RSF model, we systematically reduced the size of the model, using 5 fewer predictors each time (ie, 30, 25, 20, 15, 10, and 5 features) to obtain a model that can be easily applied clinically. The final RSF model had to have minimal loss of discriminatory performance compared with models with more predictors (eg, 15 vs 20 features) and have satisfactory performance (eg, AUC ≥85%). The final RSF model incorporated 15 variables selected in accordance with their VIMP.

PREDICTIVE BEHAVIOR OF THE RSF VARIABLES. The relationship between the risk of the HF hospitalization-CV death composite and each of the 15 selected variables was visualized using partial dependence plots. The steps for generating a partial dependence plot are as follows: 1) fit the RSF model using the training data and include 15 features; 2) choose the feature of interest (eg, age) and vary its value while keeping the other 14 features constant; 3) obtain the predicted value of the target variable for each value of the selected feature (eg, from 20 to 100 years); and 4) plot the feature values on the x-axis against the corresponding mean predicted values on the y-axis, then connect them with a line to visualize the relationship.[16,17] The advantage of partial dependency plots is their ability to illustrate the nonlinear effects of a variable on an outcome.[10] Using partial dependency plots from the final RSF model, we dichotomized the top 15 variables, considering both the observed trend in the plot and clinical relevance. Taking age as an example: the predicted survival rates begin to decrease rapidly starting around 60-65 years (Supplemental Figure 2A). Because the conventional definition of elderly is 65 years, we decided to dichotomize age at this threshold. For another instance, consider LVEDD: the predicted survival rates were relatively low before approximately 37 to 43 mm and after about 55 to 57 mm (Supplemental Figure 2J). From a clinical perspective, an LVEDD measurement of <40 or >56 mm indicates an abnormal heart size. Therefore, we decided to categorize LVEDD into 2 groups: 40 to 55.9 mm and <40 or >56 mm. Kaplan-Meier survival curves along with log-rank tests were conducted to evaluate the difference in risk of outcome among

the aforementioned dichotomized subgroups (Supplemental Figures 2A to 2O).

STATISTICAL ANALYSIS. Comparative analyses of the demographic characteristics and comorbidities of the derivation and validation cohorts were conducted using chi-square tests for categorical variables, independent-sample Student's *t*-tests for continuous variables, and the Mann-Whiney *U* test for skewed continuous variables (eg, BNP level). We assessed the performance of the prediction models by computing their AUCs. The AUCs for the initial RSF (58-variable) and final RSF (15-variable) models were evaluated and compared. The RSF model with 15 variables was then externally validated using data from the validation cohort. Additionally, patients with possible HFpEF mimics were excluded, and the RSF models were refitted to evaluate the robustness of the results. Subsequently, the prognostic value of predictors was investigated for the derivation and validation cohorts. In this analysis, the total population was divided into 4 groups in accordance with the number of predictors applicable to the patients (0-4, 5-6, 7-9, or ≥10). Cumulative event rates were estimated from Kaplan-Meier censoring estimates, and the log-rank test was used to compare groups' survival curves. Moreover, HRs (without adjusting for any covariates) from the Cox proportional hazards model were obtained by classifying the group of patients with lower risk (0-4 features) as the reference category. The RSF models were created using R software version 4.3.2 with the "randomForestSRC" package (R Foundation for Statistical Computing). To mitigate the effects of HFpEF mimics, we conducted a sensitivity analysis by excluding patients with cardiac amyloidosis, sarcoidosis, and hemochromatosis in the training cohort. We also excluded cases of severe mitral regurgitation (MR) caused by rheumatic heart disease, prolapse, chordae rupture, leaflets perforation, and mitral vegetation. The RSF analysis prediction was repeated using this refined training cohort. Other statistical analyses were conducted using SAS version 9.4 (SAS Institute).

RESULTS

PATIENT CHARACTERISTICS AND OUTCOMES. A total of 6,092 patients with a diagnosis of HFpEF were included in the analysis. Of these, 2,898 (47.6%) were in the derivation cohort, and 3,194 (52.4%) in the validation cohort. The baseline characteristics of the patients are presented in Table 1. The average age of the total study population at baseline was 75.0 ± 8.8 years; 57.1% of the patients were women. The mean LVEF was 65.6% ± 9.1%, and the median BNP level was 623.0 pg/mL. The mean follow-up duration was 2.9 ± 2.4 years, and the maximum follow-up duration was limited to 5 years. As presented in Table 2, by the end of the 5-year follow-up, 2,113 patients (34.7% of the cohort) had experienced either HF hospitalization or CV death. A total of 1,027 (35.4%) patients were hospitalized for HF, and 1,086 (34.0%) died of a CV disease or CV-related conditions. The all-cause mortality rate was 38.1% at the 2-year follow-up and 58.4% at the 5-year follow-up. Supplemental Table 3 presents additional baseline characteristics of the patients who experienced HF hospitalization or CV death and those who did not.

TABLE 2 Follow-Up Outcomes for Study Patients in the Derivation (Northern Branches) and Validation (Southern Branches) Cohorts

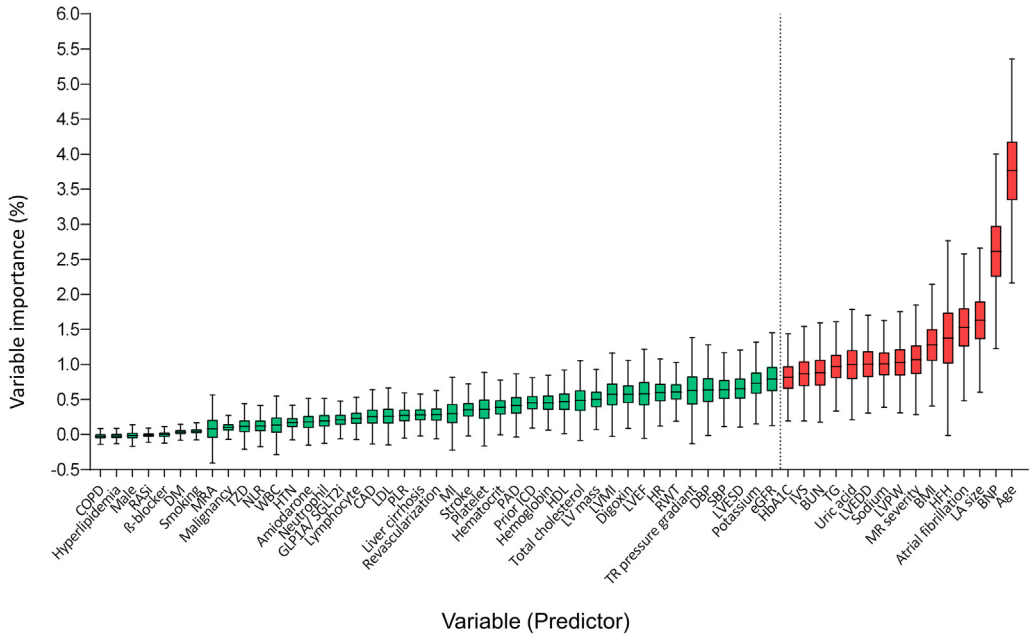
	Total (N = 6,092)	Derivation (n = 2,898)	Validation (n = 3,194)
Second-year follow-up			
Heart failure hospitalization	970 (15.9)	485 (16.7)	485 (15.2)
CV death	853 (14.0)	420 (14.5)	433 (13.6)
Composite of CV death and heart failure hospitalization	1,623 (26.6)	807 (27.8)	816 (25.5)
All-cause death	2,323 (38.1)	1,103 (38.1)	1,220 (38.2)
Fifth-year follow-up			
Heart failure hospitalization	1,153 (18.9)	570 (19.7)	583 (18.3)
CV death	1,303 (21.4)	626 (21.6)	677 (21.2)
Composite of CV death and heart failure hospitalization	2,113 (34.7)	1,027 (35.4)	1,086 (34.0)
All-cause death	3,559 (58.4)	1,669 (57.6)	1,890 (59.2)

Values are n (%).
CV = cardiovascular.

RSF MODEL AND SELECTION OF OUTCOME PREDICTORS. The initial RSF model was constructed using all 58 variables from the data of the derivation cohort (Figure 1). The VIMPs of these 58 variables are presented in Supplemental Table 4. Age was the most influential factor (VIMP: 3.76%), followed by BNP level (VIMP: 2.61%) and left atrium (LA) size (VIMP: 1.63%). Figure 2A illustrates the outstanding discriminatory performance of the initial RSF, which had an AUC of 90.0% (95% CI: 89.3%-90.7%). Table 3 presents the discriminatory performance of the more parsimonious models. The loss in AUC was not substantial for the models with 30 features (AUC: 88.6%), 25 features (AUC: 88.0%), 20 features (AUC: 87.3%), or 15 features (AUC: 85.6%), and the difference in the AUC between any 2 of these models that differed in their number of features by only 5 features was <2%. However, the AUC of the model with 10 features (82.8%) was nearly 3% lower than that with 15 features, and further reducing the number of features to

JACC: ASIA, VOL. 4, NO. 12, 2024
DECEMBER 2024:956–968

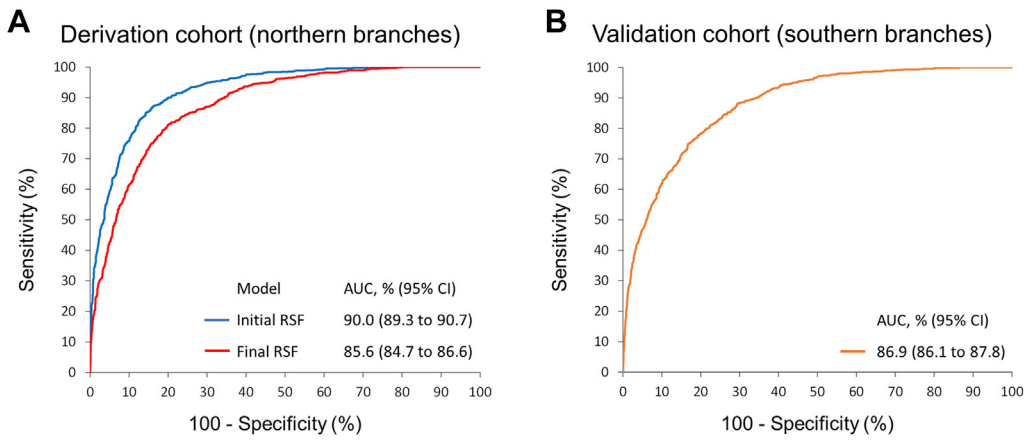
FIGURE 1 Importance of Variables



Variable importance distribution of the 58 variables in the initial random survival forest model. The 15 most strongly correlated variables were selected for the final random survival forest model. Data are derived from the random survival forest model on the basis of 1,000 bootstrap samples, and the wings of the plot represent the 95% CI of variable importance. BMI = body mass index; BNP = B-type natriuretic peptide; BUN = blood urea nitrogen; CAD = coronary artery disease; COPD = chronic obstructive pulmonary disease; DBP = diastolic blood pressure; DM = diabetes mellitus; eGFR = estimated glomerular filtration rate; GLP1RA = glucagon-like peptide-1 receptor agonist; HDL = high-density lipoprotein cholesterol; HbA_{1c} = glycated hemoglobin; HFH = heart failure hospitalization; HTN = hypertension; ICD = implantable cardioverter-defibrillator; IVS = interventricular septum; LA = left atrium; LDL = low-density lipoprotein; LV = left ventricular; LVEDD = left ventricular end-diastolic dimension; LVEF = left ventricular ejection fraction; LVESD = left ventricular end-systolic dimension; LVMI = left ventricular mass index; LVPW = left ventricular posterior wall; MI = myocardial infarction; MR = mitral regurgitation; MRA = mineralocorticoid receptor antagonist; NLR = neutrophil-lymphocyte ratio; PAD = peripheral arterial disease; PLR = platelet-lymphocyte ratio; RWT = relative wall thickness; SBP = systolic blood pressure; SGLT2i = sodium-glucose cotransporter 2 inhibitor; TG = triglyceride; TR = tricuspid regurgitation; TZD = thiazolidinedione; WBC = white blood cell count.

5 led to a further 4% reduction in the AUC (79.0%). Finally, using the VIMP values, the 15 most important variables (age, BNP level, LA size, atrial fibrillation, number of HF hospitalizations in the preceding 3 years, BMI, MR severity, LV posterior wall thickness, sodium level, LVEDD, uric acid level, triglyceride level, blood urea nitrogen [BUN] level, interventricular septum, and HbA_{1c} level) were selected for the construction of the risk model (**Figure 1**). In addition, the AUC values of all 58 RSF models according to the rankings of the initial RSF model were also provided (Supplemental Table 5). For the validation sample, the RSF model with 15 predictors exhibited excellent discriminatory performance; its AUC of 86.9% indicates generalizability (**Figure 2B**). Supplemental Figure 2 illustrates the partial dependency plots for the 15 predictive variables. After integrating the effects of all variables and adjusting for variable dependencies, each variable in the plots exhibited distinct nonlinear behavior. The factors predicting high risk of HF hospitalization and CV death were age ≥65 years, BNP level ≥600 pg/mL, LA size ≥46 mm, atrial fibrillation, frequency of HF hospitalization within 3 years, BMI <30 kg/m^2, moderate or severe MR, LV posterior wall thickness of <10 or ≥13 mm, hyponatremia (Na level <135 mEq/L) or hypernatremia (Na level ≥144 mEq/L), LVEDD of <40 or ≥56 mm, uric acid level ≥7 mg/dL, triglyceride level of <70 or ≥200 mg/dL, BUN level ≥20 mg/dL, interventricular septum thickness of <11 or ≥20 mm, and HbA_{1c} level of <6% or ≥8%. The identified thresholds were verified using Kaplan-

FIGURE 2 Discriminatory Performance of RSF Models



(A) Model containing all 58 variables (initial random survival forest [RSF]), and that containing 15 selected variables (final RSF) for the deviation cohort. (B) Estimated AUC of the final RSF model for the validation cohort.

Meier survival analysis on the basis of the data of the derivation cohort (Supplemental Figures 2A to 2O).

PROGNOSTIC IMPLICATIONS OF RISK FACTORS. A comparison of the 5-year cumulative rates of HF hospitalization and CV death revealed that the event rate increased with an increase in the number of predictors for the patients in both cohorts (**Figure 3**). In the derivation cohort, the rate of 5-year HF hospitalization or CV death was 18.3%, 28.1%, 40.7%, and 57.9% for patients with 0 to 4, 5 to 6, 7 to 9, and ≥10 risk factors, respectively (*P* trend of log-rank <0.001) (**Figure 3A**). The calibration results of the number of risk factors for HF hospitalization and CV death in the validation cohort indicated satisfactory predictive ability. For the validation cohort, the rate of HF hospitalization or CV death at the 5-year follow-up was 18.3%, 28.1%, 36.7%, and 43.6% for patients with 0 to 4, 5 to 6, 7 to 9, and ≥10 risk factors, respectively (*P* trend of log-rank <0.001) (**Figure 3B**).

TABLE 3 Performance of Random Survival Forest Models With Different Numbers of Predictors According to Variable Importance

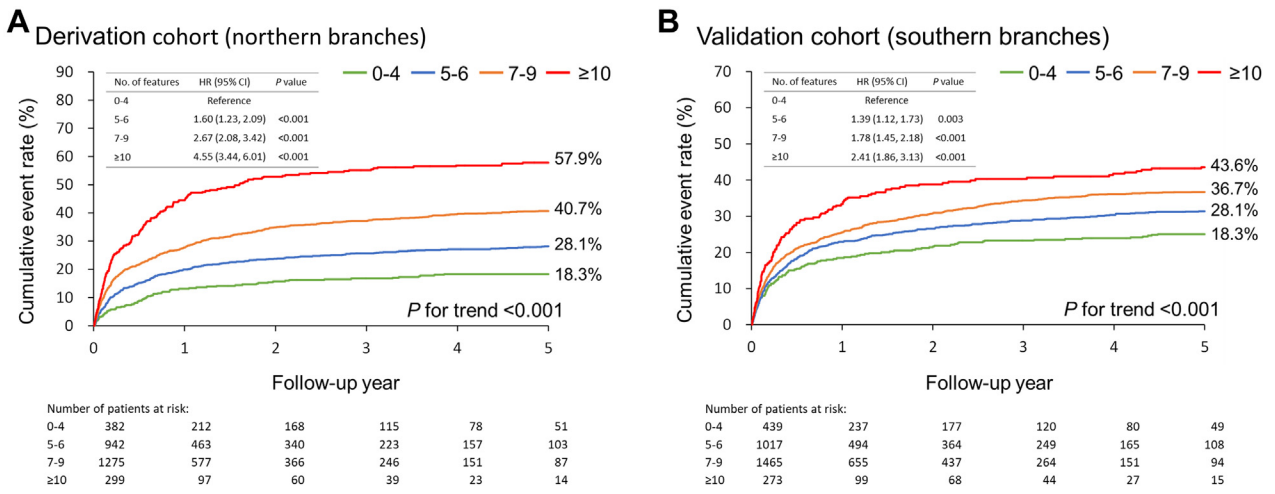
Feature Numbers	AUC (95% CI), %
All (58 features)	90.0 (89.3-90.7)
Top 30 features	88.6 (87.8-89.4)
Top 25 features	88.0 (87.2-88.8)
Top 20 features	87.3 (86.5-88.2)
Top 15 features	85.6 (84.7-86.6)
Top 10 features	82.8 (81.7-83.9)
Top 5 features	79.0 (77.8-80.2)

SENSITIVITY ANALYSIS. In total, 64 (2.2%) patients with possible HFpEF mimics or severe MR were excluded, including those with cardiac amyloidosis, sarcoidosis, hemochromatosis, and severe MR caused by rheumatic heart disease, prolapse, chordae rupture, leaflets perforation, or mitral vegetation. The results were highly consistent with the primary analysis, showing similar AUC values (85.4%) and identical rankings for the top 15 variables (Supplemental Table 6).

DISCUSSION

The present study employed machine learning to develop a model for predicting HF hospitalization and CV death in patients with HFpEF. Utilizing RSF, we identified 15 predictive markers, which comprised data on the patient's physical condition (age and BMI), echocardiography parameters (LA size, MR severity, LV posterior wall thickness, LVEDD, and interventricular septum thickness), laboratory data (BNP, sodium, uric acid, triglyceride, BUN, and HbA_{1c} levels), history of atrial fibrillation, and prior HF hospitalization. The predictive models exhibited robust performance (AUC: 85.6%), with high calibration and discriminatory ability (AUC: 86.9%). The predictive variables in the model can be obtained from clinical data routinely collected when managing patients with HFpEF. The risk factor scoring system can easily be incorporated into clinical practice through automated methodologies within electronic medical records, enabling clinicians to evaluate risks

FIGURE 3 Prognostic Implications of the 15 Relevant Predictors in the Cohorts



Cumulative incidence of the HF hospitalization–CV death composite outcome, stratified by the number of predictors in the (A) derivation and (B) validation cohorts. Variables based on the thresholds from the partial dependence plots were estimated using maximal log-rank statistics. The number of features refers to the number of 15 predictive factors. Abbreviations as in Figure 1.

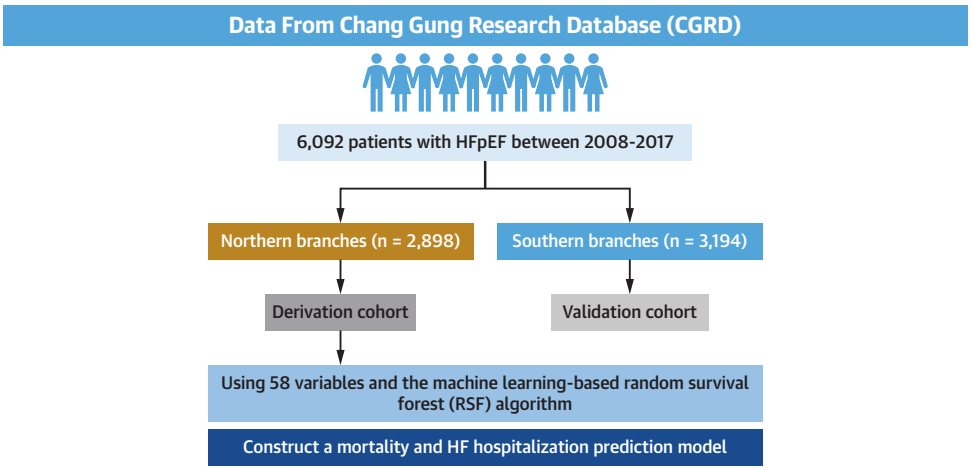
of HF hospitalization and CV death when providing care for patients with HFpEF (**Central Illustration**).

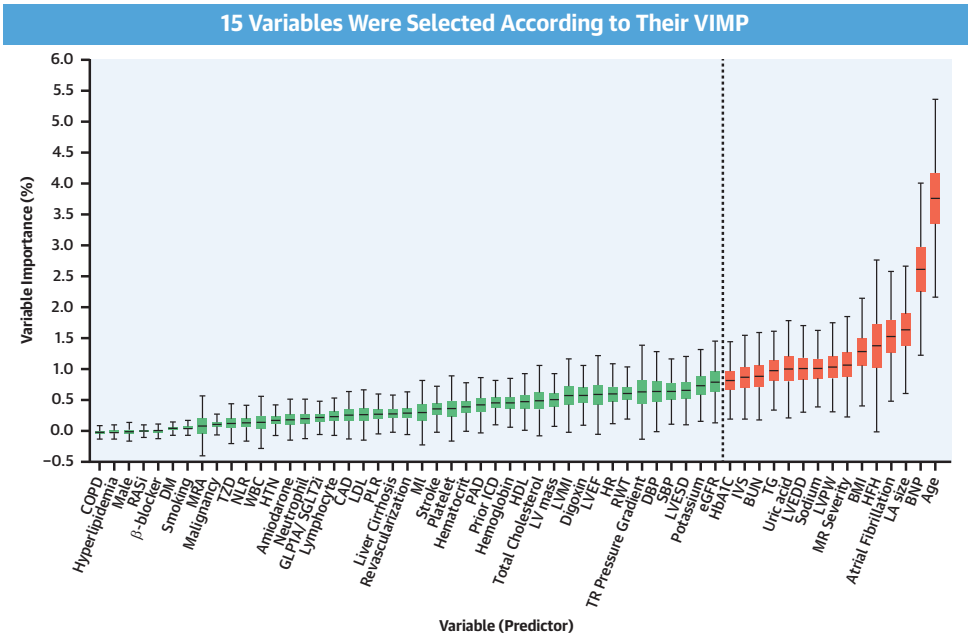
Few risk calculation models predict HFpEF clinical outcomes. The I-PRESERVE trial identified N-terminal proBNP level, age, prior HF hospitalization, and diabetes mellitus as strong independent predictors of all-cause mortality in HFpEF using multivariable Cox regression models.[8] Angraal et al,[9] using TOPCAT trial data, found that BUN level and BMI were strongly associated with mortality, and BUN level and time since previous HF hospitalization predicted HF hospitalization. Our model incorporated these risk factors and used partial dependency plots to illustrate their nonlinear effects on HF hospitalization and CV death. We refined earlier criteria, identifying age ≥65 years, BNP level ≥600 pg/mL, frequency of HF hospitalization within 3 years, BMI <30 kg/m^2, BUN level ≥20 mg/dL, and HbA_{1c} level of <6% or ≥8% as predictive markers of adverse clinical outcomes in HFpEF patients. This approach provides a more comprehensive and nuanced understanding of risk factors in HFpEF.

Our study differs from the PREDICT-HFpEF (Prognostic Models for Mortality and Morbidity in Heart Failure With Preserved Ejection Fraction) models, which used data from 6,263 DELIVER trial participants to develop predictors for composite outcomes in HFpEF.[18] PREDICT-HFpEF used multivariable analyses with a discrimination C statistic of 0.73 (95% CI: 0.71-0.75) in validation cohorts. However, it excluded patients with advanced chronic kidney disease and creatinine clearance under 25 mL/min/1.73 m^2, lacked important laboratory markers (sodium, uric acid, triglyceride, BUN), and omitted crucial echocardiographic parameters (left atrial size, MR, LVEDD) vital for HFpEF evaluation. In contrast, our study employed a machine learning-based RSF model, demonstrating higher external generalizability with an 86.9% AUC in validation. We included a comprehensive array of 58 variables, encompassing extensive laboratory data and echocardiographic parameters, to predict HFpEF outcomes based on real-world data. This approach offers a more thorough and widely applicable predictive model for HFpEF.

This study incorporated echocardiographic predictors related to heart morphology in HFpEF, including atrial and ventricular remodeling. LA dilatation, indicative of atrial remodeling, can manifest as atrial contractile dysfunction and fibrillation, marking advanced HFpEF. The present study demonstrated that an interventricular septum thickness of ≥20 or <11 mm, an LV posterior wall thickness of ≥13 or <10 mm, an LVEDD of ≥56 or <40 mm, and moderate or severe MR were strongly associated with HF hospitalization and CV death. HFpEF typically involves increased LV stiffness with concentric

CENTRAL ILLUSTRATION **Machine Learning for Mortality and Heart Failure Hospitalization Prediction in Heart Failure With Preserved Ejection Fraction**





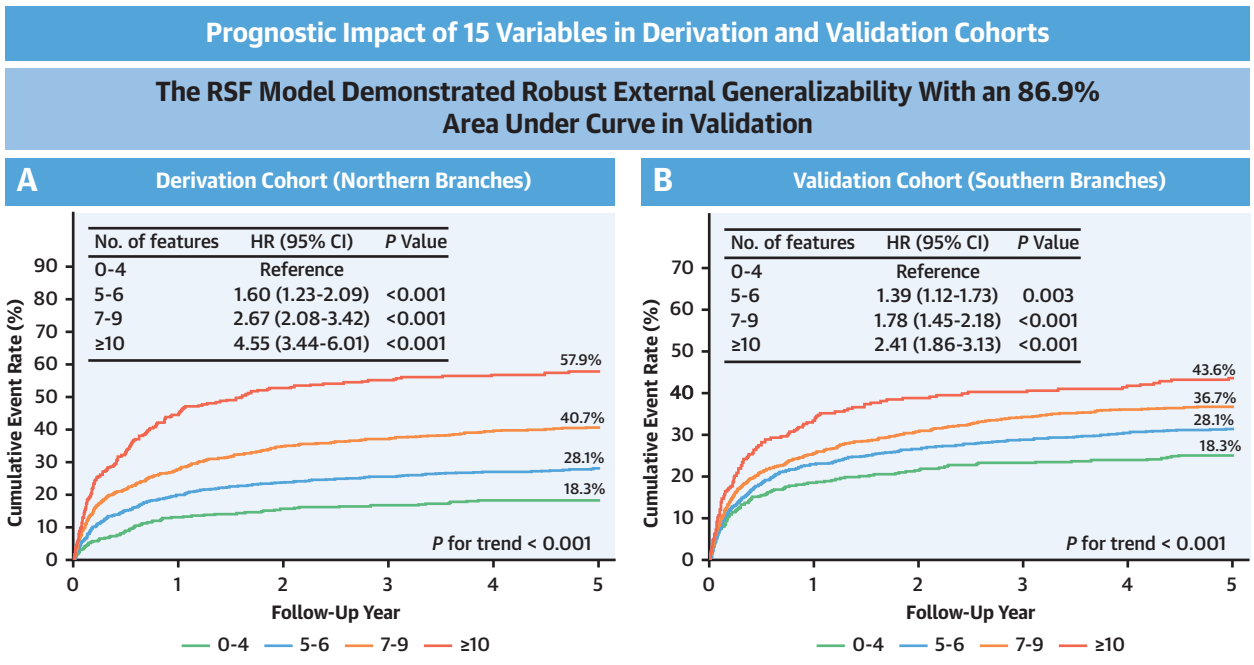
Predictors	
Age	≥65 years
BNP	≥600 pg/mL
LA size	≥46 mm
Atrial fibrillation	
Previous HF hospitalization	
BMI	<30 kg/m^2
Mitral regurgitation	Moderate/severe
LV posterior wall thickness	<10 or ≥13 mm
Serum sodium level	<135 or ≥145 mEq/L
LV end diastolic dimension	<40 or ≥56 mm
Uric acid	≥7 mg/dL
Triglyceride	<70 or ≥200 mg/dL
Blood urea nitrogen	≥20 mg/dL
IVS thickness	<11 or ≥20 mm
HbA1c	<6 % or ≥8 %

Chang C-Y, et al. JACC Asia. 2024;4(12):956–968.

Machine learning pinpointed 15 predictors of HF hospitalization and CV death in HFpEF patients, helping doctors identify high-risk individuals for tailored treatment. BMI = body mass index; BNP = B-type natriuretic peptide; CV = cardiovascular; HbA_{1c} = glycated hemoglobin; HF = heart failure; HFpEF = heart failure with preserved ejection fraction; LA = left atrial; LV = left ventricular; RSF = random survival forest; VIMP = variable importance

Continued on the next page

CENTRAL ILLUSTRATION Continued



Chang C-Y, et al. JACC Asia. 2024;4(12):956–968.

remodeling or hypertrophy linked to diastolic dysfunction.[19] Larger LV chambers are associated with higher BMI, valvular etiology, higher E/e' value, larger LA diameter, and more severe MR, all linked to higher rates of all-cause death and acute HF readmission.[20]

The present study also identified serum sodium and uric acid levels as risk factors. Both hyponatremia (<135 mEq/L) and hypernatremia (>145 mEq/L) were critical in identifying HFpEF patients at increased risk. Hyponatremia indicates adverse neurohormonal activation, whereas hypernatremia correlates with increased mortality risk.[21] A J-shaped relationship between serum sodium and long-term outcomes was observed in a larger study.[22] Additionally, uric acid levels ≥7 mg/dL predicted poor outcomes, aligning with previous studies showing elevated uric acid associated with HFpEF deterioration and increased risk of all-cause death and HF rehospitalization.[23,24] The study identified HbA_{1c} levels ≥8% or <6% and triglyceride levels ≥200 or <70 mg/dL as predictive markers for HF hospitalization and CV death in HFpEF patients. This supports evidence that HFpEF encompasses not only older adults with isolated hypertensive heart disease but also patients with multimorbidity, including diabetes, metabolic syndromes, and systemic inflammation.[25-27] These findings align with a study showing a J-shaped association between the triglyceride-glucose index and clinical outcomes in HFpEF patients.[28] Metabolic disturbances may lead to endothelial dysfunction, microvascular damage, diastolic impairment, increased proinflammatory cytokines, inflammatory pathway activation, and myocardial fibrosis promotion, all crucial to HFpEF development and progression.[28]

This study's strengths include its large, representative sample from multi-institutional data and longitudinal follow-up. The predictive model, based on real-world population data, offers greater clinical applicability than those from clinical trials. The CGRD provided comprehensive echocardiography data, often lacking in previous HFpEF predictive factor studies.[8,9] We employed machine learning, specifically RSF, to identify risk factors for HF hospitalization and CV death in HFpEF patients. Machine learning algorithms learn from sample data to perform tasks like classification or regression. Unlike conventional binary classification techniques, RSF is a nonparametric ensemble method designed for analyzing right-censored survival data.[29,30] This approach extends the traditional random forest method to accommodate time-to-event data analysis.[31]

STUDY STRENGTHS AND LIMITATIONS. First, although the study incorporated a range of variables—including the patients' physical condition, echocardiography parameters, laboratory data, and history of comorbidities—the database from which data were obtained lacked information on family history, lifestyle factors, and certain echocardiographic parameters such as e' and left atrial volume. The inclusion of these additional factors could potentially enhance the predictive accuracy of our model. Second, the prediction model was formulated using data from the northern-branch hospitals of CGMH and validated in an independent cohort from the southern-branch hospitals of CGMH. Future investigations are warranted to validate the predictive variables and assess their generalizability across diverse ethnic and racial groups. Third, the results were based on the imputed data set without missing values. However, the complete data set without any missing values accounts for only 37% of the training sample, which may not be representative of the entire study population.

CONCLUSIONS

We developed and validated a model predictive of HF hospitalization and CV death in patients with HFpEF. The RSF model contains 15 significant clinical predictors of HF hospitalization and CV death, namely age, BNP, LA size, atrial fibrillation, frequency of HF hospitalization within 3 years, BMI, MR, LV posterior wall thickness, dysnatremia, LVEDD, uric acid level, triglyceride level, BUN level, interventricular septum thickness, and HbA_{1c} level. The increasing incidence of adverse events with increasing numbers of predictive factors should assist physicians in identifying high-risk patients with HFpEF and tailoring treatment to their needs.

FUNDING SUPPORT AND AUTHOR DISCLOSURES

This study was supported by the Chang Gung Memorial Hospital (grant CORPG3N0341). The authors have reported that they have no relationships related to the contents of this paper to disclose.

ADDRESS FOR CORRESPONDENCE: Dr Dong-Yi Chen, Division of Cardiology, Department of Internal Medicine, Chang Gung Memorial Hospital at Linkou, Chang Gung University College of Medicine, Taoyuan, Taiwan, 5, Fuxing Street, Guishan District, Taoyuan City 33305, Taiwan. E-mail: dongyi70@gmail.com.

PERSPECTIVES

CLINICAL COMPETENCIES: HF hospitalization and CV death in patients with HFpEF can be estimated using 15 predictors: age, BNP level, LA size, atrial fibrillation, frequency of HF hospitalization within 3 years, BMI, MR severity, LV posterior wall thickness, dysnatremia, LVEDD, uric acid level, triglyceride level, BUN level, interventricular septum thickness, and HbA_{1c} level.

TRANSLATIONAL OUTLOOK: Our risk prediction model can help physicians identify high-risk patients with HFpEF and tailor treatments to meet their needs.

REFERENCES

1. Savarese G, Becher PM, Lund LH, et al. Global burden of heart failure: a comprehensive and updated review of epidemiology. *Cardiovasc Res*. 2023;118:3272–3287. https://doi.org/10.1093/cvr/cvac013

2. Redfield MM, Borlaug BA. Heart failure with preserved ejection fraction: a review. *JAMA*. 2023;329:827–838. https://doi.org/10.1001/jama.2023.2020

3. Shah KS, Xu H, Matsouaka RA, et al. Heart failure with preserved, borderline, and reduced ejection fraction: 5-year outcomes. *J Am Coll Cardiol*. 2017;70:2476–2486. https://doi.org/10.1016/j.jacc.2017.08.074

4. Peterson PN, Rumsfeld JS, Liang L, et al. A validated risk score for in-hospital mortality in patients with heart failure from the American Heart Association Get With The Guidelines program. *Circ Cardiovasc Qual Outcomes*. 2010;3:25–32. https://doi.org/10.1161/circoutcomes.109.854877

5. Pocock SJ, Ariti CA, McMurray JJ, et al. Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies. *Eur Heart J*. 2013;34:1404–1413. https://doi.org/10.1093/eurheartj/ehs337

6. Lagu T, Pekow PS, Shieh MS, et al. Validation and comparison of seven mortality prediction models for hospitalized patients with acute decompensated heart failure. *Circ Heart Fail*. 2016;9(8):e002912. https://doi.org/10.1161/circheartfailure.115.002912

7. Suzuki S, Yoshihisa A, Sato Y, et al. Clinical significance of Get With the Guidelines-Heart Failure Risk Score in patients with chronic heart failure after hospitalization. *J Am Heart Assoc*. 2018;7:e008316. https://doi.org/10.1161/jaha.117.008316

8. Komajda M, Carson PE, Hetzel S, et al. Factors associated with outcome in heart failure with preserved ejection fraction: findings from the Irbesartan in Heart Failure with Preserved Ejection Fraction Study (I-PRESERVE). *Circ Heart Fail*. 2011;4:27–35. https://doi.org/10.1161/circheartfailure.109.932996

9. Angraal S, Mortazavi BJ, Gupta A, et al. Machine learning prediction of mortality and hospitalization in heart failure with preserved ejection fraction. *JACC Heart Fail*. 2020;8:12–21. https://doi.org/10.1016/j.jchf.2019.06.013

10. Dietrich S, Floegel A, Troll M, et al. Random Survival Forest in practice: a method for modelling complex metabolomics data in time to event analysis. *Int J Epidemiol*. 2016;45:1406–1420. https://doi.org/10.1093/ije/dyw145

11. Tsai MS, Lin MH, Lee CP, et al. Chang Gung Research Database: a multi-institutional database consisting of original medical records. *Biomed J*. 2017;40:263–269. https://doi.org/10.1016/j.bj.2017.08.002

12. Chen DY, Chen CC, Tseng CN, et al. Clinical outcomes of Sacubitril/Valsartan in patients with acute heart failure: a multi-institution study. *EClinicalMedicine*. 2021;41:101149. https://doi.org/10.1016/j.eclinm.2021.101149

13. Cheng CL, Chien HC, Lee CH, Lin SJ, Yang YH. Validity of in-hospital mortality data among patients with acute myocardial infarction or stroke in National Health Insurance Research Database in Taiwan. *Int J Cardiol*. 2015;201:96–101. https://doi.org/10.1016/j.ijcard.2015.07.075

14. Hicks KA, Mahaffey KW, Mehran R, et al. 2017 Cardiovascular and stroke endpoint definitions for clinical trials. *Circulation*. 2018;137:961–972. https://doi.org/10.1161/circulationaha.117.033502

15. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. *Ann Appl Stat*. 2008;2:841–860, 820.

16. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat*. 2001:1189–1232.

17. Goldstein A, Kapelner A, Bleich J, Pitkin E. Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *J Comput Graph Stat*. 2015;24:44–65.

18. McDowell K, Kondo T, Talebi A, et al. Prognostic models for mortality and morbidity in heart failure with preserved ejection fraction. *JAMA Cardiol*. 2024;9:457–465. https://doi.org/10.1001/jamacardio.2024.0284

19. Zawadzka MM, Grabowski M, Kapłon-Cieślicka A. Phenotyping in heart failure with preserved ejection fraction: a key to find effective treatment. *Adv Clin Exp Med*. 2022;31:1163–1172. https://doi.org/10.17219/acem/149728

20. Ogawa S, Nagatomo Y, Takei M, et al. Impact of left ventricular chamber size on outcome in heart failure with preserved ejection fraction. *Int Heart J*. 2022;63:62–72. https://doi.org/10.1536/ihj.21-486

21. Deubner N, Berliner D, Frey A, et al. Dysnatraemia in heart failure. *Eur J Heart Fail*. 2012;14:1147–1154. https://doi.org/10.1093/eurjhf/hfs115

22. Patel YR, Kurgansky KE, Imran TF, et al. Prognostic significance of baseline serum sodium in heart failure with preserved ejection fraction. *J Am Heart Assoc*. 2018;7(12):e007529. https://doi.org/10.1161/jaha.117.007529

23. Kobayashi Y, Omote K, Nagai T, et al. Prognostic value of serum uric acid in hospitalized heart failure patients with preserved ejection fraction (from the Japanese Nationwide Multicenter Registry). *Am J Cardiol*. 2020;125:772–776. https://doi.org/10.1016/j.amjcard.2019.12.003

24. Nishino M, Egami Y, Kawanami S, et al. Lowering uric acid may improve prognosis in patients with hyperuricemia and heart failure with preserved ejection fraction. *J Am Heart Assoc*. 2022;11:e026301. https://doi.org/10.1161/jaha.122.026301

25. Hahn VS, Petucci C, Kim MS, et al. Myocardial metabolomics of human heart failure with preserved ejection fraction. *Circulation*. 2023;147:1147–1161. https://doi.org/10.1161/circulationaha.122.061846

26. Borlaug BA, Sharma K, Shah SJ, Ho JE. Heart failure with preserved ejection fraction: JACC scientific statement. *J Am Coll Cardiol*. 2023;81:1810–1834. https://doi.org/10.1016/j.jacc.2023.01.049

27. Lawson CA, Tay WT, Bernhardt L, et al. Association between diabetes, chronic kidney disease, and outcomes in people with heart failure from Asia. *JACC: Asia*. 2023;3:611–621. https://doi.org/10.1016/j.jacasi.2023.03.005

28. Zhou Q, Yang J, Tang H, et al. High triglyceride-glucose (TyG) index is associated with poor prognosis of heart failure with preserved ejection fraction. *Cardiovasc Diabetol*. 2023;22:263. https://doi.org/10.1186/s12933-023-02001-4

29. Motwani M, Dey D, Berman DS, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *Eur Heart J*. 2017;38:500–507.

30. Molenaar MA, Bouma BJ, Asselbergs FW, et al. Explainable machine learning using echocardiography to improve risk prediction in patients with chronic coronary syndrome. *Eur Heart J Digit Health*. 2024;5:170–182.

31. Pezel T, Sanguineti F, Garot P, et al. Machine-learning score using stress CMR for death prediction in patients with suspected or known CAD. *JACC Cardiovasc Imaging*. 2022;15:1900–1913. https://doi.org/10.1016/j.jcmg.2022.05.007

KEY WORDS cardiovascular death, heart failure hospitalization, heart failure with preserved ejection fraction, machine learning, prediction model, random survival forest

APPENDIX For supplemental figures and tables, please see the online version of this paper.

Go to http://www.acc.org/jacc-journals-cme to take the CME/MOC/ECME quiz for this article.