


Heart failure with preserved ejection fraction phenogroup classification using machine learning

Atsushi Kyodo¹, Koshiro Kanaoka¹, Ayaka Keshi¹, Maki Nogi¹, Kazutaka Nogi¹, Satomi Ishihara¹, Daisuke Kamon¹, Yukihiro Hashimoto¹, Yasuki Nakada¹, Tomoya Ueda¹, Ayako Seno¹, Taku Nishida¹, Kenji Onoue¹, Tsuneari Soeda¹, Rika Kawakami¹, Makoto Watanabe¹, Toshiyuki Nagai², Toshihisa Anzai² and Yoshihiko Saito^{1*} 

¹Department of Cardiovascular Medicine, Nara Medical University, Kashihara, Japan; and ²Department of Cardiovascular Medicine, Hokkaido University Graduate School of Medicine, Sapporo, Japan

Abstract

Aims Heart failure (HF) with preserved ejection fraction (HFpEF) is a complex syndrome with a poor prognosis. Phenotyping is required to identify subtype-dependent treatment strategies. Phenotypes of Japanese HFpEF patients are not fully elucidated, whose obesity is much less than Western patients. This study aimed to reveal model-based phenomapping using unsupervised machine learning (ML) for HFpEF in Japanese patients.

Methods and results We studied 365 patients with HFpEF (left ventricular ejection fraction >50%) as a derivation cohort from the Nara Registry and Analyses for Heart Failure (NARA-HF), which registered patients with hospitalization by acute decompensated HF. We used unsupervised ML with a variational Bayesian–Gaussian mixture model (VBGMM) with common clinical variables. We also performed hierarchical clustering on the derivation cohort. We adopted 230 patients in the Japanese Heart Failure Syndrome with Preserved Ejection Fraction Registry as the validation cohort for VBGMM. The primary endpoint was defined as all-cause death and HF readmission within 5 years. Supervised ML was performed on the composite cohort of derivation and validation. The optimal number of clusters was three because of the probable distribution of VBGMM and the minimum Bayesian information criterion, and we stratified HFpEF into three phenogroups. Phenogroup 1 ($n = 125$) was older (mean age 78.9 ± 9.1 years) and predominantly male (57.6%), with the worst kidney function (mean estimated glomerular filtration rate 28.5 ± 9.7 mL/min/1.73 m²) and a high incidence of atherosclerotic factor. Phenogroup 2 ($n = 200$) had older individuals (mean age 78.8 ± 9.7 years), the lowest body mass index (BMI; 22.78 ± 3.94), and the highest incidence of women (57.5%) and atrial fibrillation (56.5%). Phenogroup 3 ($n = 40$) was the youngest (mean age 63.5 ± 11.2) and predominantly male (63.5 ± 11.2), with the highest BMI (27.46 ± 5.85) and a high incidence of left ventricular hypertrophy. We characterized these three phenogroups as atherosclerosis and chronic kidney disease, atrial fibrillation, and younger and left ventricular hypertrophy groups, respectively. At the primary endpoint, Phenogroup 1 demonstrated the worst prognosis (Phenogroups 1–3: 72.0% vs. 58.5% vs. 45%, $P = 0.0036$). We also successfully classified a derivation cohort into three similar phenogroups using VBGMM. Hierarchical and supervised clustering successfully showed the reproducibility of the three phenogroups.

Conclusions ML could successfully stratify Japanese HFpEF patients into three phenogroups (atherosclerosis and chronic kidney disease, atrial fibrillation, and younger and left ventricular hypertrophy groups).

Keywords Heart failure with preserved ejection fraction; Machine learning; Unsupervised clustering

Received: 9 April 2022; Revised: 5 January 2023; Accepted: 13 March 2023

*Correspondence to: Yoshihiko Saito, Department of Cardiovascular Medicine, Nara Medical University, 840 Shijo-cho, Kashihara, Nara 634-8522, Japan. Tel: +81-744-22-3051 (ext. 3411); Fax: +81-744-22-9726. Email: saitonaramed@gmail.com

Introduction

Heart failure (HF) with preserved ejection fraction (HFpEF) has a complex aetiology and has been increasing in multiple ethnicities with a variety of lifestyles and related phenotypes.¹ Most previous studies have failed to show effective treatment for HFpEF other than sodium–glucose cotransporter 2 (SGLT2) inhibitors; thus, a ‘one size fits all’ concept is not acceptable in HFpEF management.^{2,3} Many cardiologists have recognized the need for subgroup-dependent therapies.^{4,5}

Recently, phenotyping using machine learning (ML), which makes it possible to clarify complex phenotypes from data patterns in multi-dimensional datasets, has been applied in this field.⁶ Prior studies have succeeded in stratifying HFpEF into several phenotypes in multiple cohorts that included a small number of Asian but not Japanese patients.^{7–12}

The distribution of patient population characteristics was significantly different between Western and Eastern countries, especially in terms of mean body mass index (BMI) and the incidence of obesity. In the recent HFpEF worldwide registry, the Asian patient population demonstrated lower BMI, a frequent incidence of atrial fibrillation (AF), poor kidney function, and a history of HF admission.¹³ The difference in obesity and disease distribution might result in differences in phenotyping results and cardiovascular event rates. Therefore, we considered that HFpEF phenomapping in Japanese and Asian populations is needed to understand their pathology and optimal treatment.

This study aimed to develop model-based phenotyping using unsupervised clustering for HFpEF in Japanese people.

Methods

Study population

This study included two HF cohorts: the Nara Registry and Analyses for Heart Failure (NARA-HF) study and the Japanese Heart Failure Syndrome with Preserved Ejection Fraction Registry (JASPER). The NARA-HF study was used as a derivation model, JASPER, and the composite cohort combining NARA-HF and JASPER was used as a validation cohort.

The NARA-HF study is a dynamic prospective cohort study that conformed to the principles outlined in the Declaration of Helsinki and was approved by the ethics committee of Nara Medical University (Approval No. 624). All participants provided written informed consent. All 1448 patients recruited met the Framingham criteria for emergency admission due to acute decompensated HF (ADHF) (either acute new-onset or acute-on-chronic HF) between January 2007 and December 2018. The inclusion and exclusion criteria have been previously reported.¹⁴ For the HFpEF cohort, we excluded patients with left ventricular ejection fraction (LVEF) <50% at dis-

charge ($n = 852$). From the remaining 596 patients, we excluded 231 patients because of in-hospital death ($n = 24$), haemodialysis at discharge ($n = 78$), right HF ($n = 12$), constrictive pericarditis ($n = 1$), adult congenital heart disease ($n = 9$), severe valvular disease ($n = 71$), and estimated glomerular filtration rate (eGFR) <15 mL/min/1.73 m² ($n = 36$). The remaining 365 patients satisfied the European Society of Cardiology guideline definition of HFpEF at admission.¹⁵ Finally, 365 HFpEF patients from NARA-HF cohort were analysed who had discharge data and 5 year mortality rates (Figure 1).

The details of JASPER have been previously described.¹⁶ A total of 534 hospitalized patients with HFpEF from 15 university or teaching hospitals were registered from November 2011 to March 2015. Patients with missing variables ($n = 230$) and an eGFR < 15 mL/min/1.73 m² at discharge ($n = 31$) were excluded. Finally, we selected 273 patients for the validation cohort (Figure 1).

We calculated eGFR using the following formula: eGFR (mL/min/1.73 m²) in men = $194 * \text{serum creatinine level (mg/dL)}^{-1.094} * \text{age}^{-0.287}$ and eGFR (mL/min/1.73 m²) in women = $(194 * \text{serum creatinine level (mg/dL)}^{-1.094} * \text{age}^{-0.287}) * 0.739$.¹⁷

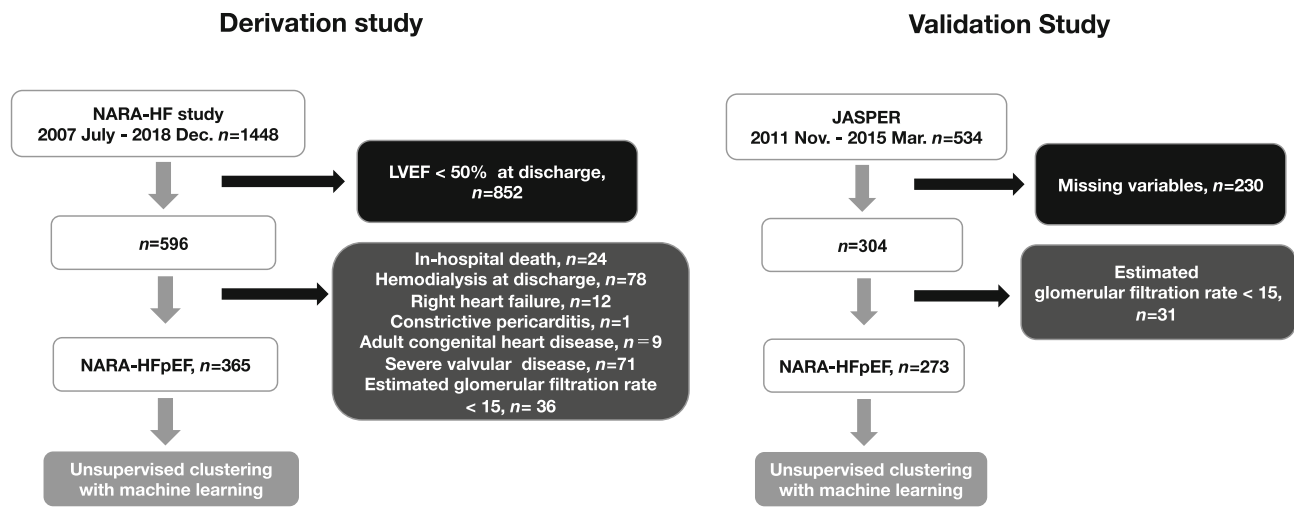
Clinical phenogroup assignment

A variational Bayesian–Gaussian mixture model (VBGMM) was used to determine clusters of clinical phenotypes. VBGMM is a statistical technique used in the ML domain for an unknown number of clusters.¹⁸ The core idea is based on VBGMM, which allows the computation of the posterior distribution of the Gaussian clusters for each instance. In contrast to other ML learning algorithms, in which the number of clusters is predefined, the proper number is estimated via the sparseness of the Dirichlet before the mixture weights.¹⁹ We adopted VBGMM because this method has been recommended when the number of clusters is unknown or when the number of data is <10 000 in the ML domain. The second advantage is that redundant clusters are identified due to the small expected values of the mixing coefficients and are therefore eliminated from the original cluster. Therefore, redundant clusters are less likely to be generated, and clustering with less noise might be possible.¹⁹ In medical biology, VBGMM is currently used in image recognition and is known to represent clusters.^{19–22}

The VBGMM algorithm used 18 singular-value continuous and 6 binary discrete variables. These clinical variables were selected because of their general clinical use and ease of routine practice, considering known associations with adverse outcomes in HFpEF, a missing values rate of <20%, and a correlation coefficient of <80%. The selected variables were documented in Supporting Information, Table S1.

In this study, the optimal number of clusters was three as suggested by the VBGMM in the probability distribution

Figure 1 Study flow. HFpEF, heart failure with preserved ejection fraction; JASPER, Japanese Heart Failure Syndrome with Preserved Ejection Fraction Registry; LVEF, left ventricular ejection fraction; NARA-HF, Nara Registry and Analyses for Heart Failure.



derived from automatic relevance determination (Supporting Information, *Figure S1A*). To validate the probability distribution of VBGMM, the Bayesian information criterion (BIC) was calculated. The minimum BIC score of between two and nine clusters also predicted that the optimal number of clusters was three (Supporting Information, *Figure S1B*). Additionally, we performed principal component analysis (PCA), a singular-value decomposition (SVD) identifying an orthogonal change in the dataset. PCA is a traditional dimensionality reduction method frequently used in the ML domain. Supporting Information, *Figure S2* plots the two-dimensional space of the first two principal components and is colour coded by three phenogroups. PCA illustrated some overlaps between the three phenogroups. Therefore, we decided that the optimal number of phenogroups was three.

To validate the results of the phenomapping of NARA-HF, we performed three different validations: (i) matching the rate of stratification using another unsupervised ML algorithm, (ii) applying VBGMM to another HFpEF cohort as an external validation, and (iii) ensuring the reproducibility with supervised ML.

(i) We applied hierarchical clustering to NARA-HF as another unsupervised ML algorithm. Hierarchical clustering is a representative unsupervised algorithm and has been used in previous HFpEF phenomapping.^{7,12,23} The advantage of hierarchical clustering is that it does not require a pre-specified number of clusters. Only the 18 standardized continuous variables from the 24 used in VBGMM were selected to apply hierarchical clustering. We stratified NARA-HF patients into three phenogroups by hierarchical clustering with Ward's method based on the dendrogram (Supporting Information, *Figure S3*). The three phenogroups were labelled as Phenogroups A–C. The matching rate between VBGMM and hierarchical clustering was analysed.

(ii) VBGMM was also adopted using JASPER as the derivation, divided into three phenogroups, and their characteristics were confirmed.

(iii) We applied the supervised ML method as a random forest (RF) method, which is generally used in the ML classification domain. The RF method builds decision trees and uses their majority vote for classification. The RF algorithm was trained by the NARA-HF cohort and VBGMM phenomapping labels (RF-NARA). To validate the reproducibility of the NARA-HF phenomapping, NARA-HF and JASPER combined cohorts were used as the composite cohort and used as the test dataset, classified using RF-NARA and VBGMM, and the three phenotypes mapped by VBGMM were compared with those conducted by RF-NARA by accuracy score and F1-measure score, calculated by $(2 \times \text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$.

VBGMM, hierarchical clustering, and RF were performed in Python (Version 3.6.5), scikit-learn package 0.19.1, NumPy package 1.14.3, pandas 0.23.0, scipy, and matplotlib 2.2.2 in the Jupyter Notebook (4.4.0). Before analysis, continuous missing values were imputed for ML using SVD in JMP 14.3.0 because the ML algorithm did not accept missing values. All missing data were considered missing at random. Given a matrix with missing values, the missing entries were imputed using a low-rank SVD approximation. SVD provides better imputation for small and large datasets compared with other algorithms.²⁴

Outcomes of interest

In the NARA-HF study, the primary endpoint was defined as all-cause death and HF readmission within 5 years after registration. As a secondary outcome, cardiovascular death was

defined as death due to HF, myocardial infarction, sudden death, stroke, or vascular diseases such as aortic dissection. We defined all-cause mortality as the primary endpoint because the cardiovascular death in elderly HF patients was hard to distinguish from other causes of death affected by their comorbidities. We checked the medical records to determine vital status and cause of death. When this information was unavailable in the medical records, we contacted patients or their families. If the patient had died, we interviewed his or her family about the institute where he or she died. We then asked the physician and confirmed the patient's cause of death. In the validation cohort of JASPER, we defined the primary and secondary endpoints as described above for 2 years because of differences in study design and mean follow-up duration.

Statistical analysis

Statistical analyses were performed using JMP Version 14.3.0 (SAS Institute, Cary, NC, USA). All values are expressed as mean \pm standard deviation or as median with an inter-quartile range for continuous variables and counts and percentages for categorical variables. Continuous variables were compared using parametric one-way analysis of variance or the non-parametric Kruskal–Wallis test based on the normality of a variable's distribution. Categorical data

were evaluated using Pearson's χ^2 test. Statistical significance was set at $P < 0.05$.

Results

Classification of heart failure with preserved ejection fraction in derivation study

We divided the derivation cohort into three phenogroups (1–3) on the basis of the VBGMM probability distribution, the minimum BIC, and PCA plot (Supporting Information, Figures S1–S3).

A comparison of the baseline characteristics in NARA-HF among the three clusters is shown in Tables 1 and 2. Phenogroup 1 was older (mean age 78.9 \pm 9.1 years) and predominantly male and had a high incidence of hypertension, diabetes mellitus (DM), hyperlipidaemia, old myocardial infarction, and anaemia at discharge. Also, Phenogroup 1 showed the highest levels of brain natriuretic peptide (BNP) and C-reactive protein (CRP) among the three phenotypes. This phenogroup also had worse kidney function. The feature of Phenogroup 1 is characterized by atherosclerotic vascular diseases and related organ damage such as old myocardial infarction and chronic kidney disease (CKD). Phenogroup 2 had

Table 1 Baseline characteristics comparison between the three phenogroups in the derivation study

	Phenogroup 1 (n = 125)	Phenogroup 2 (n = 200)	Phenogroup 3 (n = 40)	P value
Demographics on admission				
Age (years)	78.9 \pm 9.1	78.8 \pm 9.7	63.5 \pm 11.2	<0.0001
Male, n (%)	72 (57.6)	85 (42.5)	23 (57.5)	0.0164
Clinical characteristics on admission				
BMI at admission (kg/m ²)	23.40 \pm 3.60	22.78 \pm 3.94	27.46 \pm 5.85	<0.0001
HR (b.p.m.)	86.06 \pm 26.04	92.97 \pm 31.62	103.50 \pm 29.73	<0.0001
SBP (mmHg)	149.06 \pm 31.08	147.29 \pm 32.68	175.78 \pm 40.26	<0.0001
DBP (mmHg)	79.20 \pm 20.76	79.70 \pm 19.98	108.62 \pm 27.18	<0.0001
NYHA, n (%)				<0.0001
II	12 (10)	25 (13)	3 (8)	
III	69 (55)	90 (45)	11 (28)	
IV	44 (35)	85 (43)	26 (65)	
Medical history on admission				
Hypertension, n (%)	116 (92.8)	145 (72.5)	32 (80.0)	<0.0001
Diabetes mellitus, n (%)	68 (54.4)	74 (37.0)	16 (40.0)	0.0079
Hyperlipidaemia, n (%)	61 (48.8)	55 (27.5)	21 (52.5)	<0.0001
History of myocardial infarction, n (%)	50 (40.0)	11 (5.5)	9 (22.5)	<0.0001
Atrial fibrillation, n (%)	49 (39.2)	113 (56.5)	16 (40.0)	0.005
Anaemia at discharge, n (%)	53 (42.5)	45 (22.5)	1 (2.5)	<0.0001
Echocardiographic data on admission				
LVEF (%)	60.4 \pm 10.5	60.3 \pm 12.1	54.5 \pm 14.5	0.0156
LVDd (mm)	48.8 \pm 0.66	46.7 \pm 0.53	49.1 \pm 1.19	0.0273
LVDs (mm)	32.9 \pm 0.65	31.1 \pm 0.53	35.7 \pm 1.17	0.0009
Left ventricular hypertrophy, n (%)	40 (32)	19 (9.5)	19 (47.5)	<0.0001
IVST (mm)	11.6 \pm 2.3	10.3 \pm 1.76	13.2 \pm 3.8	<0.0001
PWT (mm)	11.2 \pm 1.9	10.2 \pm 1.7	12.5 \pm 2.7	<0.0001
LAD (mm)	45.4 \pm 8.7	44.8 \pm 8.8	43.0 \pm 6.3	0.429

BMI, body mass index; DBP, diastolic blood pressure; HR, heart rate; IVST, interventricular septal thickness; LAD, left arterial diameter; LVDd, left ventricular end-diastolic diameter; LVDs, left ventricular end-systolic diameter; LVEF, left ventricular ejection fraction; NYHA, New York Heart Association functional classification; PWT, posterior wall thickness; SBP, systolic blood pressure.

Table 2 Laboratory findings and clinical characteristics at discharge

	Phenogroup 1 (n = 125)	Phenogroup 2 (n = 200)	Phenogroup 3 (n = 40)	P value
HbA1c (%)	6.15 ± 0.95	6.10 ± 1.04	6.09 ± 1.03	0.9194
BNP (pg/dL)	412.1 ± 406.4	183 ± 147.1	212.7 ± 177.4	<0.0001
CRP (mg/dL)	1.07 ± 1.86	0.71 ± 1.01	0.56 ± 0.72	0.0285
Haemoglobin (g/dL)	10.31 ± 1.56	11.39 ± 1.89	13.64 ± 1.73	<0.0001
Serum renin (pg/mL)	6.47 ± 7.99	7.51 ± 13.0	9.31 ± 14.2	0.5332
Serum aldosterone (mg/dL)	109.7 ± 9.5	102.5 ± 7.3	145.0 ± 14.3	0.0316
Serum sodium (mmol/L)	139.1 ± 2.8	138.3 ± 4.3	140.2 ± 2.2	0.0010
Creatinine (mg/dL)	1.79 ± 0.54	0.94 ± 0.31	0.96 ± 0.29	<0.0001
Blood urea nitrogen (mg/mL)	41.3 ± 17.6	23.1 ± 9.1	20.4 ± 8.9	<0.0001
eGFR	28.5 ± 9.7	55.6 ± 19.3	59.8 ± 18.9	<0.0001
Total protein (g/dL)	6.53 ± 0.61	6.68 ± 0.76	7.17 ± 0.54	<0.0001
Serum albumin (g/dL)	3.51 ± 0.45	3.61 ± 0.44	4.03 ± 0.30	<0.0001
Vital signs				
Heart rate (b.p.m.)	68.8 ± 10.8	73.2 ± 12.9	67.6 ± 8.0	0.0007
Systolic blood pressure (mmHg)	118.5 ± 16.5	111.5 ± 14.7	122.4 ± 16.3	<0.0001
Diastolic pressure (mmHg)	62.6 ± 10.4	61.4 ± 9.9	67.9 ± 9.0	0.0011
Medication at discharge, n (%)				
ACE inhibitor	59 (47.2)	97 (48.5)	20 (50.0)	0.9469
Angiotensin receptor blockers	58 (46.4)	71 (35.5)	15 (37.5)	0.1424
Any RAS blocker	107 (85.6)	165 (82.5)	33 (82.5)	0.575
MRA	34 (27.2)	79 (39.5)	15 (37.5)	0.0733
Calcium channel blocker	69 (55.2)	67 (33.5)	14 (35.0)	0.0004
Diuretics	109 (87.2)	154 (77.0)	25 (62.5)	0.0024
Digitalis	7 (5.6)	18 (9.0)	2 (5.0)	0.4381
Antiplatelet therapy	29 (23.3)	23 (11.5)	8 (20.0)	0.0176
Any anticoagulation therapy	49 (39.2)	96 (48.0)	16 (40.)	0.2561
Statins	46 (36.8)	45 (22.5)	17 (42.5)	0.0038
Diabetes drugs	50 (40)	49 (24.5)	8 (20)	0.0266
SGLT2 inhibitor	1 (0.008)	2 (0.01)	0 (0)	0.8146
Insulin user	10 (8.0)	10 (5.0)	1 (2.5)	0.3407

ACE, angiotensin-converting enzyme; BNP, brain natriuretic peptide; CRP, C-reactive protein; eGFR, estimated glomerular filtration rate; HbA1c, glycated haemoglobin; MRA, mineralocorticoid receptor antagonist; RAS, renin-angiotensin system; SGLT2, sodium-glucose cotransporter 2.

a higher incidence of women and older individuals and the lowest BMI among the phenogroups. Phenogroup 2 also had the highest incidence of AF and the lowest incidence of atherosclerotic factors. The features of Phenogroup 2 were summarized as AF and aged women (mean age 78.8 ± 9.7 years). Phenogroup 3 was the youngest and predominantly male. BMI was the highest. Echocardiography revealed the highest incidence of left ventricular hypertrophy (LVH) (47.5%). The characteristics of Phenogroup 3 were as follows: youngest, obese, and LVH. A summary of the characteristics of the three phenotypes in the derivation cohort is illustrated in *Figure 2*. We named the three phenogroups as the atherosclerosis and CKD group, the AF group, and the younger and LVH group.

Prognostic relationship between clinical phenotypes and patient outcome

As shown in *Table 3*, Phenogroup 1 demonstrated a significantly worse prognosis of the primary composite endpoint compared with the other two phenotypes (Phenogroups 1–3: 72.0% vs. 58.5% vs. 45%, $P = 0.0036$). As secondary endpoints, the incidences of all-cause death, cardiovascular

death, or HF readmission within the 5 years were also significantly higher in Phenogroup 1 than in the other two groups (*Table 3*). When comparing Phenogroups 2 and 3, the occurrence of all-cause death tended to be higher in Phenogroup 2 than Phenogroup 3. In contrast, the rates of cardiovascular death and HF readmission were similar between the two phenogroups despite older age and worse kidney function in Phenogroup 2. The Kaplan–Meier survival curves of the primary composite endpoint and each secondary endpoint for 5 years are illustrated in *Figure 3*.

Hierarchical clustering in NARA-HF

The characteristics of the three phenogroups stratified by hierarchical clustering are documented in Supporting Information, *Table S2* and are summarized as follows: Phenogroup A was older with predominantly atherosclerotic factors and CKD, Phenogroup B was predominantly female with AF, and Phenogroup C was younger with relatively high BMI. These characteristics were similar to the VBGMM phenomapping results. From the above matching features, we determined that VBGMM Phenogroups 1–3 corresponded to the hierarchical clustering Phenogroups A–C, respectively. Similar to

Figure 2 Characteristics summary of the three phenogroups in the Nara Registry and Analyses for Heart Failure. AF, atrial fibrillation; BMI, body mass index; CKD, chronic kidney disease; DM, diabetes mellitus; HL, hyperlipidaemia; HR, heart rate; HT, hypertension; LV, left ventricle; LVH, left ventricular hypertrophy.

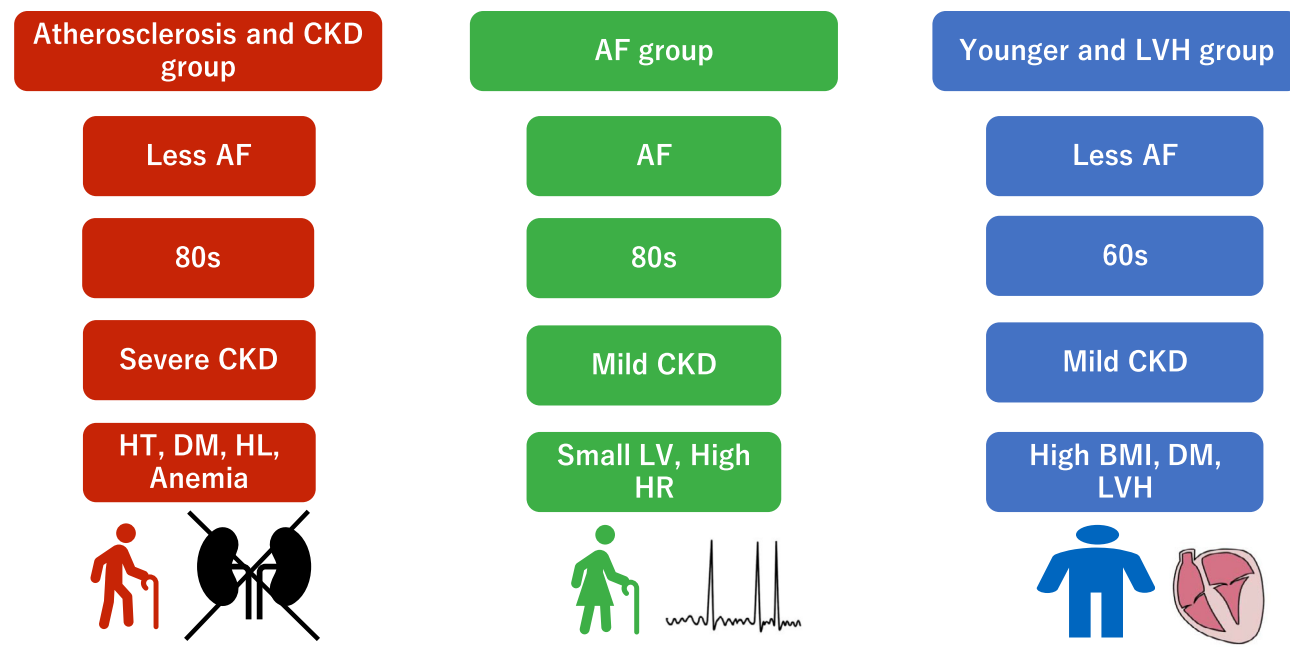


Table 3 Comparison of 5 year clinical outcomes between the three phenogroups in the derivation cohort

5 year clinical outcome	Phenogroup 1 (n = 119)	Phenogroup 2 (n = 184)	Phenogroup 3 (n = 39)	P value
Primary endpoint, n (%)	90 (72.0)	117 (58.5)	18 (45)	0.0036
Secondary endpoint				
All-cause death, n (%)	62 (51.7)	88 (44.0)	8 (20)	0.0021
Cardiovascular death, n (%)	35 (28.0)	39 (19.5)	5 (12.5)	0.0467
Heart failure re-hospitalization, n (%)	53 (42.4)	51 (25.5)	12 (30.0)	0.0061

Primary endpoint included all-cause death and heart failure re-hospitalization.

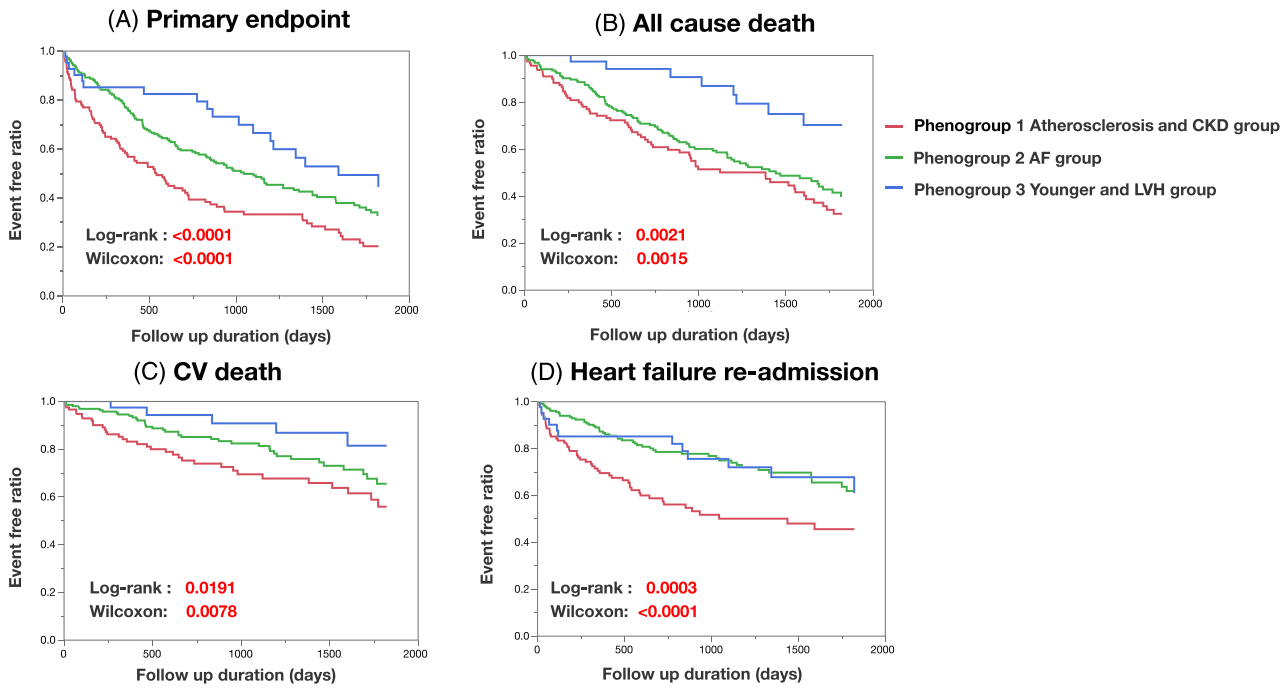
the previous report that compared the result of orthogonal two unsupervised ML algorithms, the matching rate between these two algorithms is illustrated in Supporting Information, *Figure S4*. The percentage of matched individuals in VBGMM (% cluster in VBGMM) of Phenogroup 1 and Phenogroup A was 69.6%, of Phenogroup 2 and Phenogroup B was 49.0%, and of Phenogroup 3 and Phenogroup C was 77.5%. However, Phenogroup 2 and Phenogroup C did not show complete agreement regarding the prevalence of LVH, which may have caused the difference between the clustering algorithms. Phenogroup C included only 29.3% of Phenogroup 3 and 40.6% of Phenogroup 2.

External validation of the phenogroups in the JASPER cohort

For external validation, we validated our results using data from the JASPER cohort. Among the 534 patients with HFpEF

in the JASPER study, 273 patients were analysed after excluding 230 patients with missing values for 24 variables used for clustering (Supporting Information, *Table S1*) and 31 patients with $eGFR < 15 \text{ mL/min/1.73 m}^2$ (*Figure 1*). The major differences in baseline characteristics between the derivation and validation cohorts are shown in Supporting Information, *Table S3*. After phenomapping by VBGMM using the same 24 variables as in the derivation analyses, we succeeded in dividing the patients into three phenogroups similar to those in the derivation cohort. The characteristics of each phenogroup are presented in *Table 4*. Phenogroup 1 had a higher incidence of older aged patients, men, and atherosclerotic risk factors and the highest rate of old myocardial infarction and poor kidney function. Phenogroup 2 had the oldest age, lowest BMI, and highest prevalence of women and AF. Phenogroup 3 had the youngest age and highest BMI, relatively better kidney function, and LVH. The characteristics of these three phenogroups in the validation cohort were similar to those in the derivation cohort of the NARA-HF study,

Figure 3 Kaplan–Meier survival curves of primary and secondary endpoints in the Nara Registry and Analyses for Heart Failure: (A) primary endpoint, (B) all-cause death, (C) cardiovascular death, and (D) heart failure readmission. AF, atrial fibrillation; CKD, chronic kidney disease; LVH, left ventricular hypertrophy.



which was clearly shown by the correlation heat map of each variable (*Figure 4*). The variables of sex, age, BMI, heart rate, blood pressure at admission, atherosclerotic risk factors, AF, left ventricular wall thickness, and kidney function showed a similar correlation coefficient pattern between the two cohorts. However, laboratory findings such as CRP, serum total protein (TP), and serum albumin (Alb) did not clarify the correspondence between the two cohorts. However, among the three phenogroups in the validation cohort, the phenogroup of AF showed a significantly worse prognosis of 2 year primary and secondary endpoints than the phenogroup of atherosclerosis and CKD (Supporting Information, *Table S4* and *Figure S5*), which were different from those in the derivation cohort.

Supervised machine learning validation

We validated our results of the present unsupervised phenomapping for NARA-HF and JASPER by supervised ML, an RF algorithm. The composite cohort combining NARA-HF and JASPER was classified into three phenogroups as the derivation by VBGMM. Subsequently, the composite cohort was also classified into three phenogroups by RF-NARA, which was trained by the derivation phenomapping result of NARA-HF. The accuracy and F1-measure scores comparing the three phenogroup labels obtained by VBGMM and

RF-NARA to the composite cohort were 0.845 and 0.785, respectively.

Discussion

We stratified HFpEF into three phenogroups based on standard clinical variables using unsupervised ML with the NARA-HF study as a derivation cohort in Japan. Next, we validated the phenomapping of NARA-HF using another unsupervised ML algorithm and a supervised one. Finally, using another Japanese multicentre HFpEF registry, the JASPER study, as a validation cohort, we succeeded in stratifying three phenogroups similar to those in the derivation cohort. Both cohorts included patients who were admitted to the hospital for ADHF. Therefore, the patients were much older than those enrolled in randomized control trials (RCTs), such as the I-Preserved trial and the TOPCAT trial, and the BMI was much lower in Japanese cohorts than in Western populations.^{2,8}

About the clustering variables

Phenotyping using ML was dependent on the variables used. In the present study, we used standard variables such as BMI, atherosclerotic risk factors, kidney function, AF, LVH, and BNP

Table 4 Clinical characteristics of the three phenotypes in the validation cohort

	Phenogroup 1 (n = 126)	Phenogroup 2 (n = 74)	Phenogroup 3 (n = 73)	P value
Demographics on admission				
Age (years)	78.2 ± 7.9	82.6 ± 6.6	70.7 ± 11.7	<0.0001
Male, n (%)	80 (63.5)	24 (34.3)	30 (41.1)	<0.0001
BMI (kg/m ²)	24.6 ± 3.6	21.6 ± 4.4	25.2 ± 6.0	<0.0001
Medical history on admission				
Diabetes, n (%)	70 (55.6)	24 (32.4)	15 (20.6)	<0.0001
Hypertension, n (%)	113 (89.7)	49 (66.2)	49 (67.1)	<0.0001
Hyperlipidaemia, n (%)	83 (65.9)	16 (21.6)	21 (28.8)	<0.0001
Hyperuricaemia, n (%)	68 (53.9)	26 (35.1)	20 (27.4)	0.0015
Atrial fibrillation, n (%)	70 (55.6)	57 (77.0)	48 (65.8)	0.0088
History of myocardial infarction, n (%)	27 (21.4)	1 (1.4)	0 (0)	<0.0001
Echocardiographic data on admission				
LVDD (mm)	50.3 ± 6.0	44.3 ± 5.6	43.6 ± 6.5	<0.0001
LVDs (mm)	33.6 ± 6.0	27.8 ± 4.2	27.6 ± 6.5	<0.0001
LVEF (%)	59.5 ± 6.7	61.0 ± 7.8	61.2 ± 7.9	0.2004
IVST (mm)	10.4 ± 1.7	9.9 ± 1.9	12.3 ± 3.1	<0.0001
PWT (mm)	10.3 ± 1.4	9.8 ± 1.4	11.8 ± 2.7	<0.0001
LAD (mm)	47.2 ± 9.2	46.4 ± 10.7	44.8 ± 8.7	0.2813
Laboratory findings at discharge				
HbA1c (%)	6.4 ± 1.1	6.0 ± 1.0	5.9 ± 0.9	0.2201
BNP (pg/dL)	180.1 ± 160.1	256.1 ± 289.5	212.9 ± 249.8	0.0730
CRP (mg/dL)	0.40 ± 0.53	0.94 ± 1.04	0.30 ± 0.34	<0.0001
Haemoglobin (g/dL)	11.2 ± 1.7	10.5 ± 1.5	13.2 ± 1.9	<0.0001
Creatinine (mg/dL)	1.42 ± 0.47	1.15 ± 0.37	0.81 ± 0.21	<0.0001
Blood urea nitrogen (mg/dL)	34.9 ± 1.3	30.6 ± 1.7	19.6 ± 1.6	<0.0001
eGFR at discharge (mL/min/1.73 m ²)	38.1 ± 13.2	43.0 ± 15.3	64.2 ± 11.9	<0.0001
Total protein (g/dL)	7.0 ± 0.6	6.4 ± 0.6	6.9 ± 0.6	<0.0001
Serum albumin (mg/dL)	3.8 ± 0.4	3.4 ± 0.4	3.9 ± 0.4	<0.0001
Vital signs at discharge				
Heart rate (b.p.m.)	64.1 ± 10.4	71.9 ± 11.9	66.9 ± 10.1	<0.0001
Systolic blood pressure (mmHg)	118.8 ± 14.9	109.6 ± 14.5	112.1 ± 14.4	<0.0001
Diastolic pressure (mmHg)	60.2 ± 11.8	61.0 ± 10.0	66.5 ± 10.1	0.0004

BMI, body mass index; BNP, brain natriuretic peptide; CRP, C-reactive protein; eGFR, estimated glomerular filtration rate; HbA1c, glycated haemoglobin; IVST, interventricular septal thickness; LAD, left arterial diameter; LVDD, left ventricular end-diastolic diameter; LVDs, left ventricular end-systolic diameter; LVEF, left ventricular ejection fraction; PWT, posterior wall thickness.

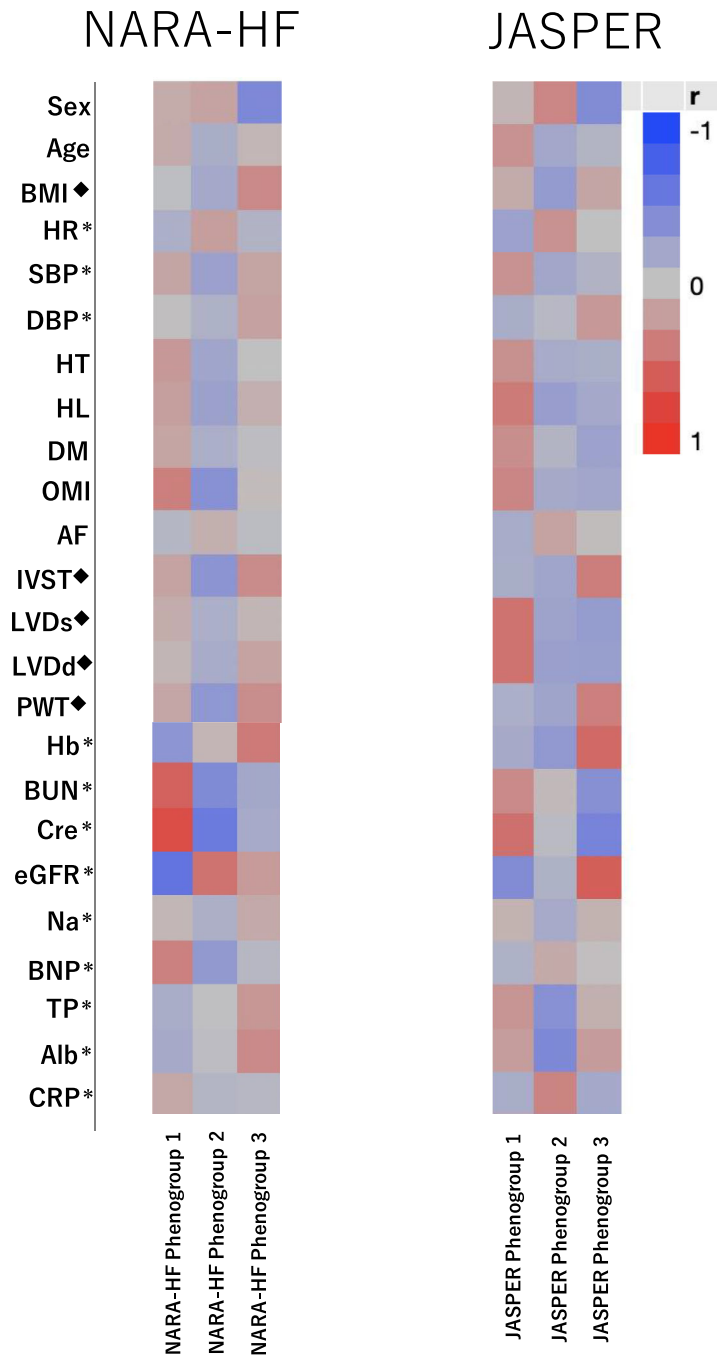
levels, so that the present phenotyping can be widely applied. Consequently, the three phenogroups stratified in the present study were characterized as atherosclerosis and CKD, AF, and younger and LVH, providing insight on the development of HFpEF. Although the present phenotyping was not entirely consistent with previous studies, these characteristics were commonly observed in each phenogroup in earlier reports in Western countries,^{7–12} in which patients were more obese than Japanese patients. However, when comparing the phenotype results of the JASPER and NARA-HF studies, the trends in mean values of CRP, TP, and Alb were not completely consistent. These results might not be important variables for the phenomapping of HFpEF using clinical data.

Phenogroup 2 (AF group)

In the phenogroup characterized by AF, the patients were older women and frequently had CKD in both our cohorts and several previous studies' populations, whereas BMI was much lower in patients in Japanese studies than in

those in Western countries.^{7–9} Shah *et al.* first reported a phenogroup characterized by CKD and AF in ML phenomapping.⁷ Cohen *et al.* also presented an older, with stiff arteries, small left ventricles, diastolic function, and AF group.⁸ Additionally, Jones *et al.* also reported a distinct HFpEF phenogroup, which showed characteristics of diastolic dysfunction in haemodynamic analysis, similar to Cohen *et al.*'s AF group.¹² The AF group we presented here was very similar to previously reported phenogroups, especially in terms of frequency of AF, lower BMI, older age, and poorer renal function than the other groups. Given that AF is almost equally observed in patients with higher and lower BMI in the Japanese AF cohort,²⁵ it is possible that AF contributes to the development of HFpEF with or without obesity. Considering that AF occurs more frequently in men than in women, it is notable that the incidence of women was higher in this phenogroup of HFpEF.²⁵ Additionally, a phenogroup with AF, predominantly elderly women, and mildly CKD was frequently observed in previous Western studies. Therefore, the phenogroup with elderly, AF, and CKD may be a common HFpEF phenotype worldwide, although their BMI was slightly different.

Figure 4 Heat mapping of clinical variables in derivation and validation studies. AF, atrial fibrillation; Alb, serum albumin; BMI, body mass index; BNP, brain natriuretic peptide; BUN, blood urea nitrogen; Cre, serum creatinine level; CRP, C-reactive protein; DBP, diastolic blood pressure; DM, diabetes mellitus; eGFR, estimated glomerular filtration rate; Hb, haemoglobin; HL, hyperlipidaemia; HR, heart rate; HT, hypertension; IVST, interventricular septal thickness; JASPER, Japanese Heart Failure Syndrome with Preserved Ejection Fraction Registry; LVDd, left ventricular end-diastolic diameter; LVDs, left ventricular end-systolic diameter; Na, serum sodium level; NARA-HF, Nara Registry and Analyses for Heart Failure; OMI, old myocardial infarction; PWT, posterior wall thickness; SBP, systolic blood pressure; TP, serum total protein.



Phenogroups 1 (atherosclerosis and CKD group) and 3 (younger and LVH group)

Phenogroups with a higher incidence of atherosclerotic risk factors have also been identified in previous reports. Kao *et al.* reported phenogroups with a high rate of atherosclerotic factors, such as coronary artery disease, DM, and CKD, in patients enrolled in the I-Preserved trial, and these phenogroups with a high risk of atherosclerosis showed the worst outcome.⁹ Cohen *et al.* proposed that the phenogroup with high-risk factors for atherosclerosis and obesity had the highest mortality in HFpEF phenotyping using the TOPCAT trial.⁸ Thus, a phenogroup similar to Phenogroup 1 in the present study was also classified in previous Western cohorts of HFpEF. This phenotype would be similar to one recently reported by Hahn *et al.*¹¹ and Jones *et al.*¹² that is characterized by a higher N-terminal pro-BNP value and relatively low ejection fraction (EF). However, the characteristics of our atherosclerotic phenogroup were not completely consistent with those of atherosclerotic and obesity phenotypes in previous reports, such as the prevalence of DM. For example, Cohen *et al.*'s Phenogroup 3 (obese, diabetic, with advanced symptoms) had a DM frequency of 88%, whereas our equivalent Phenogroup 3 had a DM frequency of 40%.⁸ Subgroup C presented by Kao *et al.* (elderly patients with a high prevalence of atherosclerotic factor and CKD) also had a DM frequency of 100%, whereas our Phenogroup 1 had a frequency of 68%.⁹ This may be explained by the difference in lifestyle between Western countries and Japan, which could be represented by the proportion of obesity between our study and previous reports. Generally, obesity is a stronger risk factor for atherosclerotic disease in Western countries than in Japan and other Asian countries. In the present study, patients in the atherosclerosis and CKD phenogroup were not obese, and their cardiac geometry was not hypertrophied. However, patients in the corresponding phenogroups in the I-Preserved and TOPCAT trials were obese and had LVH.^{8,9} Our patients in Phenogroup 3 (younger and LVH) were more obese and had the most severe LVH compared with the other two phenogroups. The phenogroup of atherosclerosis in Western countries might include a subgroup of patients who would be classified as Phenogroup 3 if in Japan.

In previous phenotyping of HFpEF using RCTs, there was a phenogroup with relatively normal left ventricular geometry, with a lower proportion of atherosclerotic risk factors and AF.⁸ In the present study, we were unable to identify this phenogroup. A possible explanation may be the difference in patient recruitment between our study and previous studies. The NARA-HF and JASPER studies are registries of patients with ADHF at unexpected admission to hospitals, whereas the I-Preserved and TOPCAT trials are RCTs that recruited outpatient patients with HFpEF.

Possible treatment strategies by phenotype

Recently, SGLT2 inhibitors were reported to be drugs useful for reducing a composite of cardiovascular death and HF hospitalization in patients with HFpEF.³ However, given that HFpEF is a complex syndrome consisting of different subgroups classified with some common characteristics, phenotype-related treatment strategies should be proposed in the future. Considering the present results, the subgroup of patients characterized with AF would possibly be suitable candidates for treatment with ablation,^{26,27} those characterized with atherosclerosis and CKD would be for SGLT2 inhibitors,²⁸ and those characterized with LVH would be for sacubitril-valsartan²⁹ or mineralocorticoid receptor antagonists.⁸

Limitations

The present study had several limitations. First, the NARA-HF study is a single-centre, relatively small study; therefore, regional and clinical decision bias might be related to the phenotypes and clinical outcomes. Second, because the appropriate number of clusters varies depending on the dataset and clustering method, it is difficult to determine whether the number of clusters currently presented is truly optimal. Third, the selection bias of variables for phenomapping might have affected the results.

Fourth, we could not completely exclude the patients with secondary cardiomyopathy and recovered EF who improved their LVEF during hospitalization in the NARA-HF study. Fifth, we used only a few echo parameters in our analysis, because many echo parameters had 20% or more missing values. This may have affected the present phenomapping result. Sixth, we imputed some missing variables. The largest missing variable was Alb at discharge (deficit rate was 19.5%). Segar *et al.*¹⁰ excluded variables with 10% or more missing values, and the missing values were imputed using the missForest package. Kao *et al.*⁹ also imputed missing data for 67% of patients, using 20 multiply imputed datasets. Compared to past studies, the percentages of imputed variables in the current study might be acceptable. However, there might be some bias depending on the imputation algorithm. Finally, the primary composite event rate in the derivation study did not match that in the validation study. This could have resulted from the difference in kidney function and age of each phenogroup between the two cohorts. In the NARA-HF study, the incidence of worse CKD was higher than that in the JASPER study.

In conclusion, we identified three HFpEF phenotypes with different clinical characteristics in Japanese patients with HFpEF, who had lower BMI and incidence of obesity than Western patients. The feature values and comorbidities of each phenogroup such as atherosclerotic risk factors, age,

sex, and AF in the present study were similar to those of earlier studies conducted in Western countries, whereas BMI and the rate of obesity did not correspond to these earlier studies. Further studies with a much larger sample size are necessary to further clarify the precise phenotypes in HFpEF, a complex syndrome.

Acknowledgements

We thank Yoko Wada, Yuki Kamada, and Rika Nagao for their support during data collection.

Conflict of interest

Y.S. has received research funds from Otsuka Pharmaceutical Co., Ltd., Ono Pharmaceutical Co., Ltd., Takeda Pharmaceutical Co., Ltd., Daiichi Sankyo Co., Ltd., Mitsubishi Tanabe Pharma Corporation, Bristol-Myers Squibb Company, Actelion Pharmaceuticals Japan Ltd., Kyowa Kirin Co., Ltd., Kowa Pharmaceutical Co. Ltd., Shionogi & Co., Ltd., Dainippon Sumitomo Pharma Co., Ltd., Teijin Pharma Ltd., Chugai Pharmaceutical Co., Ltd., Eli Lilly Japan K.K., Nihon Medi-Physics Co., Ltd., Novartis Pharma K.K., Pfizer Japan Inc., and Fuji Yakuhin Co., Ltd.; research expenses from Novartis Pharma K.K., Roche Diagnostics K.K., Amgen Inc., Bayer Yakuhin, Ltd., Astellas Pharma Inc., and Actelion Pharmaceuticals Japan Ltd.; speakers' bureau/honorarium from Alnylam Japan K.K., AstraZeneca K.K., Otsuka Pharmaceutical Co., Ltd., Kowa Pharmaceutical Co. Ltd., Daiichi Sankyo Co., Ltd., Mitsubishi Tanabe Pharma Corporation, Tsumura & Co., Teijin Pharma Ltd., Toa Eiyo Ltd., Nippon Shinyaku Co., Ltd., Nippon Boehringer Ingelheim Co., Ltd., Novartis Pharma K.K., Bayer Yakuhin, Ltd., Pfizer Japan Inc., Bristol-Myers Squibb Company, and Mochida Pharmaceutical Co., Ltd.; and consultation fees from Ono Pharmaceutical Co., Ltd. and Novartis Pharma K.K. The remaining authors have no conflicts of interest to report.

References

- 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.
- Anand IS, Rector TS, Cleland JG, Kuskowski M, McKelvie RS, Persson H, McMurray JJ, Zile MR, Komajda M, Massie BM, et al. Prognostic value of baseline plasma amino-terminal pro-brain natriuretic peptide and its interactions with irbesartan treatment effects in patients with heart failure and preserved ejection fraction findings from the I-PRESERVE trial. *Circ Heart Fail.* 2011; **4**: 569–577.
- Anker SD, Butler J, Filippatos G, Ferreira JP, Bocchi E, Böhm M, Brunner-La Rocca H-P, Choi D-J, Chopra V, Chuquiure-Valenzuela E, et al. Empagliflozin in heart failure with a preserved ejection fraction. *N Engl J Med.* 2021; **385**: 1451–1461.
- Senni M, Paulus WJ, Gavazzi A, Fraser AG, Díez J, Solomon SD, Smiseth OA, Guazzi M, Lam CSP, Maggioni AP, et al. New strategies for heart failure with preserved ejection fraction: the importance of targeted therapies for heart failure

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. List of selected variables.

Table S2. Patient characteristics of the three phenogroups stratified by hierarchical clustering in NARA-HF.

Table S3. Major differences between NARA-HF study, JASPER, Sanjiv et al. and TOPCAT.

Table S4. Comparison of clinical outcomes after 2 years between the three phenogroups in the validation cohort.

Figure S1. Representative probability distribution proposed by variational Bayesian Gaussian mixture and Bayesian information criterion analysis for the identification of the optimal number of phenogroups. (A) Representative probability distribution proposed by VBGMM. Blue bar shows that the number of clusters indicates the possibility of classification. The highest blue bar was shown in three clusters. (B) The blue bar represents the BIC of the number of clusters. The minimum BIC was documented in three. Abbreviations: VBGMM, variational Bayesian Gaussian mixture; BIC, Bayesian information criterion.

Figure S2. Illustration of the distance between three phenogroups and their overlap in principal component analysis. PC1 and PC2 are principal components of their dataset. The purple dot is Phenogroup 1, the yellow dot is Phenogroup 2 and the blue dot is Phenogroup 3.

Figure S3. The cluster dendrogram of the patients in NARA-HF using the hierarchical clustering algorithm. Green lines indicate Phenogroup A, red lines indicate Phenogroup B, and blue lines indicate Phenogroup C. The horizontal axis illustrates the patients-index. The vertical axis illustrates the threshold of the dendrogram.

Figure S4. Patient overlap between phenogroups stratified by VBGMM and hierarchical clustering.

Figure S5. Kaplan–Meier survival curve of the primary endpoint in JASPER. JASPER, Japanese Heart Failure Syndrome With Preserved Ejection Fraction; CV, cardiovascular; CKD, chronic kidney disease; AF, atrial fibrillation; LVH, left ventricular hypertrophy.

- phenotypes. *Eur Heart J*. 2014; **35**: 2797–2811d.
5. Ahmad T, Pencina MJ, Schulte PJ, O'Brien E, Whellan DJ, Piña IL, Kitzman DW, Lee KL, O'Connor CM, Felker GM. Clinical implications of chronic heart failure phenotypes defined by cluster analysis. *J Am Coll Cardiol*. 2014; **64**: 1765–1774.
 6. Ahlqvist E, Storm P, Käräjämäki A, Martinell M, Dorkhan M, Carlsson A, Vikman P, Prasad RB, Aly DM, Almgren P, Wessman Y, Shaat N, Spégel P, Mulder H, Lindholm E, Melander O, Hansson O, Malmqvist U, Lernmark Å, Lahti K, Forsén T, Tuomi T, Rosengren AH, Groop L. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol*. 2018; **6**: 361–369.
 7. Shah SJ, Katz DH, Selvaraj S, Burke MA, Yancy CW, Gheorghiane M, Bonow RO, Huang CC, Deo RC. Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation*. 2015; **131**: 269–279.
 8. Cohen JB, Schrauben SJ, Zhao L, Basso MD, Cvijic ME, Li Z, Yarde M, Wang Z, Bhattacharya PT, Chirinos DA, et al. Clinical phenogroups in heart failure with preserved ejection fraction: detailed phenotypes, prognosis, and response to spironolactone. *JACC Heart Fail*. 2020; **8**: 172–184.
 9. Kao DP, Lewsey JD, Anand IS, Massie BM, Zile MR, Carson PE, McKelvie RS, Komajda M, McMurray JJ, Lindenfeld JA. Characterization of subgroups of heart failure patients with preserved ejection fraction with possible implications for prognosis and treatment response. *Eur J Heart Fail*. 2015; **17**: 925–935.
 10. Segar MW, Patel KV, Ayers C, Basit M, Tang WHW, Willett D, Berry J, Grodin JL, Pandey A. Phenomapping of patients with heart failure with preserved ejection fraction using machine learning-based unsupervised cluster analysis. *Eur J Heart Fail*. 2019: 1–11.
 11. Hahn VS, Knutsdottir H, Luo X, Bedi K, Margulies KB, Haldar SM, Stolina M, Yin J, Khakoo AY, Vaishnav J, Bader JS, Kass DA, Sharma K. Myocardial gene expression signatures in human heart failure with preserved ejection fraction. *Circulation*. 2021; **143**: 120–134.
 12. Jones E, Randall EB, Hummel SL, Cameron DM, Beard DA, Carlson BE. Phenotyping heart failure using model-based analysis and physiology-informed machine learning. *J Physiol*. 2021; **599**: 4991–5013.
 13. Tromp J, Claggett BL, Liu J, Jackson AM, Jhund PS, Køber L, Widimský J, Boytsov SA, Chopra VK, Anand IS, et al. Global differences in heart failure with preserved ejection fraction: the PARAGON-HF trial. *Circ Heart Fail*. 2021: 468–477.
 14. Nakada Y, Kawakami R, Matsushima S, Ide T, Kanaoka K, Ueda T, Ishihara S, Nishida T, Onoue K, Soeda T, Okayama S, Watanabe M, Okura H, Tsuchihashi-Makaya M, Tsutsui H, Saito Y. Simple risk score to predict survival in acute decompensated heart failure: A₂B score. *Circ J*. 2019; **83**: 1019–1024.
 15. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumgartner H, Böhm M, Burri H, Butler J, Celutkiene J, Chioncel O, et al. 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: developed by the Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure of the European Society of Cardiology (ESC) with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J*. 2021; **42**: 3599–3726.
 16. Nagai T, Yoshikawa T, Saito Y, Takeishi Y, Yamamoto K, Ogawa H, Anzai T. Clinical characteristics, management, and outcomes of Japanese patients hospitalized for heart failure with preserved ejection fraction—a report from the Japanese Heart Failure Syndrome with Preserved Ejection Fraction (JASPER) registry. *Circ J*. 2018; **82**: 1534–1545.
 17. Matsuo S, Imai E, Horio M, Yasuda Y, Tomita K, Nitta K, Yamagata K, Tomino Y, Yokoyama H, Hishida A, et al. Revised equations for estimated GFR from serum creatinine in Japan. *Am J Kidney Dis [Internet]*. 2009; **53**: 982–992.
 18. Corduneanu A, Bishop CM. Variational Bayesian model selection for mixture distributions. In *Artificial Intelligence and Statistics*. Burlington, MA: Morgan Kaufmann Publishers; 2001.
 19. Loukas C, Sgouros NP. Multi-instance multi-label learning for surgical image annotation. *Int J Med Robot Comput Assist Surg*. 2020; **16**: 1–12.
 20. Loukas C, Nikiteas N, Schizas D, Georgiou E. Shot boundary detection in endoscopic surgery videos using a variational Bayesian framework. *Int J Comput Assist Radiol Surg*. 2016; **11**: 1937–1949.
 21. Zhang X, Ma J, Cheng Z, Huang S, Ge SS, Lee TH. Trajectory generation by chance-constrained nonlinear MPC with probabilistic prediction. *IEEE Trans Cybern*. 2021; **51**: 3616–3629.
 22. Ma X, Yang X, Fan R. Screening of miRNA target genes in coronary artery disease by variational Bayesian Gaussian mixture model. *Exp Ther Med*. 2019: 2129–2136.
 23. Murray EM, Greene SJ, Rao VN, Sun JL, Alhanti BA, Blumer V, Butler J, Ahmad T, Mentz RJ. Machine learning to define phenotypes and outcomes of patients hospitalized for heart failure with preserved ejection fraction: findings from ASCEND-HF. *Am Heart J*. 2022; **254**: 112–121.
 24. Mandel JSP. A comparison of six methods for missing data imputation. *J Biom Biostat*. 2015; **06**: 1–6.
 25. Akao M, Chun YH, Wada H, Esato M, Hashimoto T, Abe M, Hasegawa K, Tsuji H, Furuue K. Current status of clinical background of patients with atrial fibrillation in a community-based survey: the Fushimi AF Registry. *J Cardiol [Internet]*. 2013; **61**: 260–266.
 26. Packer DL, Mark DB, Robb RA, Monahan KH, Bahnson TD, Poole JE, Noseworthy PA, Rosenberg YD, Jeffries N, Mitchell LB, Flaker GC, Pokushalov E, Romanov A, Bunch TJ, Noelker G, Ardashv A, Revishvili A, Wilber DJ, Cappato R, Kuck KH, Hindricks G, Davies DW, Kowey PR, Naccarelli GV, Reiffel JA, Piccini JP, Silverstein AP, al-Khalidi HR, Lee KL, for the CABANA Investigators. Effect of catheter ablation vs antiarrhythmic drug therapy on mortality, stroke, bleeding, and cardiac arrest among patients with atrial fibrillation: the CABANA randomized clinical trial. *JAMA*. 2019; **321**: 1261–1274.
 27. Packer DL, Piccini JP, Monahan KH, Al-Khalidi HR, Silverstein AP, Noseworthy PA, Poole JE, Bahnson TD, Lee KL, Mark DB. Ablation versus drug therapy for atrial fibrillation in heart failure: results from the CABANA trial. *Circulation*. 2021; **143**: 1377–1390.
 28. Heerspink HJL, Stefánsson BV, Correa-Rotter R, Chertow GM, Greene T, Hou F-F, Mann JFE, McMurray JJV, Lindberg M, Rossing P, et al. Dapagliflozin in patients with chronic kidney disease. *N Engl J Med*. 2020; **383**: 1436–1446.
 29. Januzzi JL, Prescott MF, Butler J, Felker GM, Maisel AS, McCague K, Camacho A, Pinã IL, Rocha RA, Shah AM, et al. Association of change in N-terminal pro-B-type natriuretic peptide following initiation of sacubitril-valsartan treatment with cardiac structure and function in patients with heart failure with reduced ejection fraction. *JAMA*. 2019; **322**: 1085–1095.