



Discovering Distinct Phenotypical Clusters in Heart Failure Across the Ejection Fraction Spectrum: a Systematic Review

Claartje Meijs^{1,2} · M. Louis Handoko³ · Gianluigi Savarese⁴ · Robin W. M. Vernooij^{1,5} · Ilonca Vaartjes¹ · Amitava Banerjee⁶ · Stefan Koudstaal⁷ · Jasper J. Brugts⁸ · Folkert W. Asselbergs^{6,9} · Alicia Uijl^{1,4,9} 

Accepted: 27 June 2023
© The Author(s) 2023

Abstract

Review Purpose This systematic review aims to summarise clustering studies in heart failure (HF) and guide future clinical trial design and implementation in routine clinical practice.

Findings 34 studies were identified ($n = 19$ in HF with preserved ejection fraction (HFpEF)). There was significant heterogeneity in variables and techniques used. However, 149/165 described clusters could be assigned to one of nine phenotypes: 1) young, low comorbidity burden; 2) metabolic; 3) cardio-renal; 4) atrial fibrillation (AF); 5) elderly female AF; 6) hypertensive-comorbidity; 7) ischaemic-male; 8) valvular disease; and 9) devices. There was room for improvement on important methodological topics for all clustering studies such as external validation and transparency of the modelling process.

Summary The large overlap between the phenotypes of the clustering studies shows that clustering is a robust approach for discovering clinically distinct phenotypes. However, future studies should invest in a phenotype model that can be implemented in routine clinical practice and future clinical trial design.

Keywords Heart failure · Machine learning · Clustering · Phenotyping · Precision medicine

Introduction

Heart failure (HF) is a heterogeneous, chronic syndrome with high morbidity and high mortality, with 10–20% of patients rehospitalised for HF within 1 year and less than 50% of patients surviving 5 years after diagnosis [1, 2]. The prevalence of HF is only expected to increase with an aging general population [3]. Left ventricular ejection fraction (EF) plays a central role in the diagnosis, prognosis,

and treatment indication for patients with HF. The European Society of Cardiology (ESC) differentiates EF between HF with reduced EF (HFrEF; $EF \leq 40\%$), HF with mildly reduced EF (HFmrEF; $EF 41–49\%$), and HF with preserved EF (HFpEF; $EF \geq 50\%$) [4].

At both ends of the EF spectrum there are limitations in the treatment of patients, which indicates there could be potential for personalisation of care. Treatment of HF follows a “one-size-fits-all” approach, with four main

✉ Alicia Uijl
a.uijl@amsterdamumc.nl

¹ Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

² Helmholtz Zentrum München GmbH - German Research Center for Environmental Health, Institute of Computational Biology, Neuherberg, Germany

³ Department of Cardiology, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam Cardiovascular Sciences, Amsterdam, the Netherlands

⁴ Division of Cardiology, Department of Medicine, Karolinska Institutet, Stockholm, Sweden

⁵ Department of Nephrology and Hypertension, University Medical Centre Utrecht, Utrecht University, Utrecht, The Netherlands

⁶ Health Data Research UK London, Institute for Health Informatics, University College London, London, UK

⁷ Department of Cardiology, Green Heart Hospital, Gouda, the Netherlands

⁸ Department of Cardiology, Thoraxcenter, Erasmus MC University Medical Center, Rotterdam, The Netherlands

⁹ Department of Cardiology, Amsterdam University Medical Centers, University of Amsterdam, Amsterdam, The Netherlands

treatments that should be considered for patients with HFrEF. However, with this multitude of evidence-based therapies, an aging population and multimorbidity the management of these patients is complicated [5]. Currently, prioritisation or sequencing of guideline directed medical therapy is lacking, yet personalisation of treatment strategies could be an option for these patients [6].

Only sodium-glucose co-transporter 2 inhibitor (SGLT2i) have demonstrated benefit in patients with HFpEF [7, 8]. Overall, there have been disappointing neutral trial results for patients with HFpEF [9]. The inconclusive trial results in patients with HFpEF might be a consequence of increased underlying heterogeneity in patients with higher LVEF. Yet, there could be subgroups of patients that would benefit from some therapies. This indicates that personalisation is a key concept that could be implemented across the EF spectrum.

Given the high variation in pathophysiology, symptoms, and comorbidities among HF patients, there is significant potential for personalized care. To address the above-mentioned issues, there has been a surge of studies that aimed to describe the heterogeneity of HF patients in a more multidimensional manner, using clustering to characterize phenotypical subgroups.

Unsupervised clustering analysis is a machine learning algorithm that can classify patients according to patient characteristics. Cluster analysis is especially suitable for subgroup discovery when dealing with unknown and complex relationships between variables, as these relationships do not have to be pre-specified to be modelled correctly. A series of clustering studies has been instigated since Shah et al. in 2015 used clustering, which they termed “phenomapping”, to identify clusters of patients with HFpEF [10]. The hypothesis is that increased patient heterogeneity could lead to dilution of beneficial treatment effects.

There is a wide variety of clustering studies in HF, using different clustering methods, identifying variables and HF populations, which makes it difficult to compare these studies. To date, several reviews have discussed clustering, in particular in HFpEF, yet results have not yet been synthesised in a systematic review [11–13]. This systematic review aims to examine and compare the methodology and results of clustering studies that are performed in patients with HF. A comprehensive summary of the clustering studies can shed light on the utility of clustering for patients with HF and the usefulness of corresponding phenotype cluster models, and could help shape future research on treatment personalisation for patients with HF.

Methods

The review protocol was previously specified and registered in PROSPERO (CRD42022362925). The Preferred Reporting Items for Systematic Reviews and Meta-Analyses

(PRISMA) was used to ensure transparent reporting of review methods.

Eligibility Criteria

Randomised clinical trials and observational studies (cross-sectional, cohorts, registries, and electronic health records) reporting on unsupervised clustering analysis in HF were considered for inclusion. Patients had to be diagnosed with HF, HFrEF, HFmrEF or HFpEF, subpopulations of HF were excluded (e.g. patients with HF and diabetes or patients with HF and destination therapy left ventricular assist devices). Studies were also excluded if the aim of the article was not to define and describe phenotypes within patients with HF or if the analysis did not include unsupervised clustering methods. Clustering studies based on symptoms were excluded. Studies were excluded if they were review articles or case reports. Only studies conducted after 1 January 2010 were considered for synthesis to include contemporary studies on HF and machine learning techniques. The language was restricted to English or Dutch.

Literature Search

We included relevant search terms for HF, including HFrEF, HFmrEF, HFpEF. In addition, we searched for clustering methods, general terms such as “machine learning” and “clustering analysis” were combined with specific clustering methods such as “latent class analysis”, “hierarchical clustering” and “phenomapping”. Last, we included the outcome of clustering methods such as “clusters”, “phenogroups” and “subgroups”. MeSH terms that were relevant were included. All searches were combined using the Boolean Operators “AND” and “OR”. The search was conducted in two databases: PubMed and EMBASE. The search strategy was conducted on 13 October 2022. A detailed search strategy can be found in Supplementary Table S1.

Final consensus on eligibility, based on title/abstract and full text screening, was reached by two independent reviewers (CM and AU) using the Rayyan web tool.

Data Extraction and Synthesis

Data was extracted from the included articles according to the following characteristics: 1) general information (year of publication, author, data source), 2) study characteristics (sample size, age and sex distribution), 3) characteristics of clustering (method, number of variables, number of clusters, external validation), 4) data on outcome (identifying variables for each cluster, morbidity and mortality outcomes). A proposed qualitative cluster framework was

created summarizing similarities and differences between the cluster models. This framework was developed based on phenotype patterns that we could identify across the different clustering studies. Within the framework characteristic frequency was quantified. Additionally, the characteristics and proportions and prognoses of most common clusters are discussed. Data was extracted by one reviewer (CM) and checked by a second reviewer (AU).

Quality Assessment

To assess quality of the clustering studies, the methodology of all studies was compared. We consulted the scoping review of Hond et al. [14•], and two practical guidelines on clustering to create a comparison structure that contains most crucial aspects of unsupervised learning (Supplementary Table S2) [14•, 15, 16]. The methodology comparison is structured into three phases: 1) preparation, collection, and checking of the data, 2) development of the model, and 3) validation of the model.

Results

Literature Search

A total of 1097 studies were identified in PubMed and EMBASE, of which 472 studies were excluded as duplicates. Studies ($n = 625$) were screened on title/abstract and 52 were selected for full-text review. Of these, 18 were excluded based on wrong methods (i.e. supervised

clustering or prediction modelling), wrong study population (i.e. also including non-HF participants) or a missing description of phenotypes (i.e. missing outcome). In total, 34 studies were included in the systematic review (Fig. 1).

Study Characteristics

We found 34 eligible clustering studies that were performed between 2012 and 2022, and used varying datatypes, clustering methods, and sample sizes (Table 1) [10, 17–33, 34•, 35–38, 39•, 40–42, 43•, 44–49]. Clustering techniques that were used included hierarchical clustering ($n = 14$) [10, 22, 24, 25, 27, 28, 31, 34•, 36, 39•, 40, 45, 47, 49], LCA ($n = 10$) [17, 21, 26, 32, 33, 35, 37, 43•, 46, 48], PAM ($n = 5$) [19, 29, 30, 34•, 38], k-means clustering ($n = 5$) [23, 34•, 41, 42, 44], and model-based clustering ($n = 3$) [18, 20, 34•]. Dataset sizes ranged from 103 patients to 318,384 patients. Datatypes varied between registry-based data ($n = 6$) [19, 26, 27, 42, 43•, 46], cohort data ($n = 7$) [20, 22, 24, 28, 30, 38, 41], EHR data ($n = 9$) [10, 23, 29, 31, 34•, 44, 47–49], and trial data ($n = 12$) [17, 18, 21, 25, 32, 33, 35–37, 39•, 40, 45], using varying variable types for the clustering such as clinical variables ($n = 31$), echocardiographic variables ($n = 7$) [10, 18, 20, 22, 23, 40, 49], biomarkers ($n = 4$) [24, 28, 38, 41], hemodynamic parameters ($n = 1$) [23], and demographic variables ($n = 1$) [27]. The number of variables used for analysis also varied between 8 and 415, and the number of clusters discovered ranged between 2 to 15.

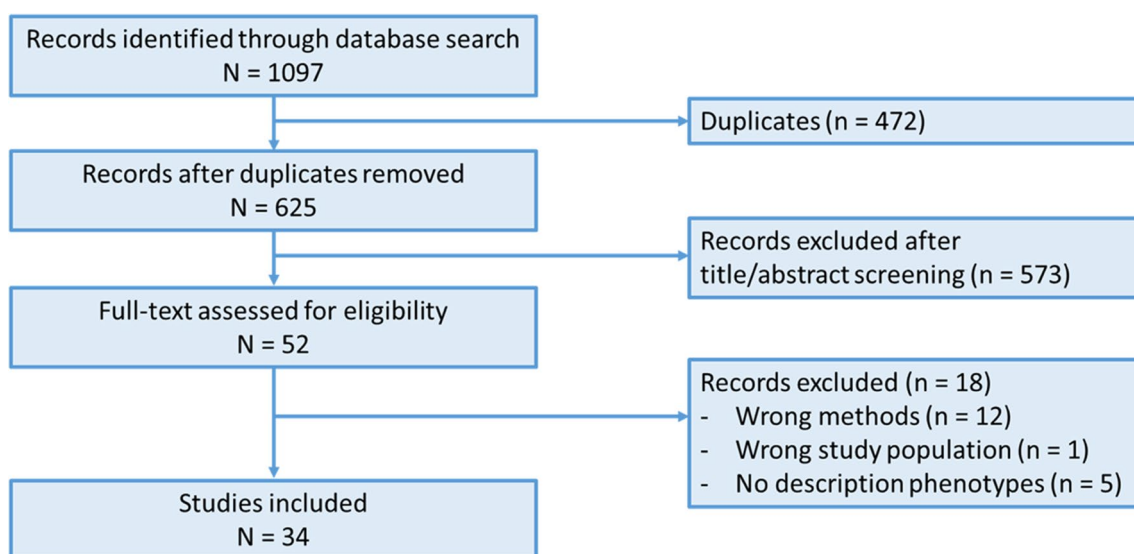


Fig. 1 Study flow of the literature search and study selection

Table 1 Summary of key characteristics of clustering studies that were performed in HF patients for phenotype discovery

	Shah [10]	Kao [17]	Segar [18]	Arévalo-Lorido [19]	Hedman [20]	Cohen [21]	Schrub [22]	Harada [23]	Stienen [24]	Gu [25]	Ujii [26]
HF type	HFpEF	HFpEF	HFpEF	HFpEF	HFpEF	HFpEF	HFpEF	HFpEF	HFpEF	HFpEF	HFpEF
Publication year	2015	2015	2019	2020	2020	2020	2020	2020	2020	2020	2021
Dataset	HFpEF Program	I-PRESERVE	TOPCAT	DICUMAP	KaRen	TOPCAT	KaRen	Single center	Media-DHF	Single center	SwedeHF
Data type	EHR	Trial	Trial	Registry	Cohort	Trial	Cohort	EHR	Cohort	Trial	Registry
Variable type	Clinical and echocardiographic	Clinical	Clinical and echocardiographic	Clinical	Clinical and echocardiographic	Clinical	echocardiographic	Clinical, hemodynamic parameters, and echocardiographic	Biomarkers	Clinical	Clinical
Number of patients	397	4113	654	103	320	1765	356	350	392	970	6909
Mean/median age	65	NA	71	79	76	69	76	77	74	70	80
% Female	62	60	48	60	56	52	57	55	64	42	52
Clustering method	Hierarchical clustering	LCA	Model-based clustering	PAM	Model based clustering	LCA	Hierarchical clustering	K-means clustering	Hierarchical clustering	Hierarchical clustering	LCA
Number of variables	67	11	61	23	43	8	55	37	415 (349)	11	10
Number of clusters	3	6	3	5	6	3	3	4	2	3	5
Outcome	All-cause mortality, CV-hospitalization and non-CV hospitalization	Event-free survival, all-cause mortality, and CV-hospitalization	All-cause mortality and HF hospitalization	1-year hospitalization and mortality	All-cause mortality and HF-hospitalization	6-year HF hospitalization or all-cause mortality	All-cause mortality and HF hospitalization	Cardiac events	CV death and CV hospitalizations	5-year all-cause mortality and HF hospitalization	All-cause mortality, CV mortality, non-CV mortality, HF hospitalization
External validation	None	Prognosis and treatment response validation with external validation cohort	Outcome validation in RELAX cohort	None	None	None	None	Outcome validation with internal validation cohort	None	None	CHECK-HF

Table 1 (continued)

HF type	Casebeer [27]	Woolley [28]	Nouraei [29]	Perry [30]	Fayol [31]	Murray [32]	Choy [33]	Banerjee [34•]	Kao [35]	Ahmad [36]	Ferreira [37]	Tromp [38]
Publication year	2021	2021	2021	2021	2022	2022	2022	2022 (pre-print)	2012	2014	2018	2018
Dataset	MAPD	BIOSTAT-CHF	Single center	Single center	Single center	ASCEND-HF	TOPCAT	THIN	BEST	HF-ACTION	EMPHASIS-HF	BIOSTAT-CHF
Data type	Registry	Cohort	EHR	Cohort	EHR	Trial	Trial	EHR	Trial	Trial	Trial	Cohort
Variable type	Demographic and clinical	Biomarkers	Clinical	Clinical	Clinical	Clinical	Clinical	Clinical	Clinical	Clinical	Clinical	Biomarkers
Number of patients	1515	429	196	889	928	812	1540	188,799	1121	1619	2279	1802
Mean/median age	73	77	77	61	75	74	NA	NA	NA	59	NA	68
% Female	54	45	56	47	43	51	NA	NA	33	28	24	24
Clustering method	Hierarchical clustering	Hierarchical clustering	PAM	PAM	Hierarchical agglomerative clustering	LCA	LCA	K-means, hierarchical, PAM, mixture modelling	LCA	Hierarchical clustering	LCA	PAM
Number of variables	9	363	18	11	15	11	10	87	7	45 (13)	18	92
Number of clusters	3	6	6	7	3	4	3	5	6	4	4	6
Outcome	Treatment and 1-year hospitalization	2-year all-cause mortality and hospitalization	CV mortality	Mortality	All-cause death	30-day hospitalization and mortality	CV mortality, aborted cardiac arrest, or hospitalization for HF	Risk of non-fatal CV diseases, all-cause hospitalization	1-year mortality	Hospitalization and mortality risk	composite of CV death and HF hospitalization	All-cause mortality and HF hospitalisation
External validation	None	None	None	None	None	None	None	CPRD and UK Biobank	Validation of outcomes using the MOCHA population	None	EPHESUS trial	Scottish independent dataset

Table 1 (continued)

	Karwath [39•]	Bouali [40]	de Lange [41]	Ahmad [42]	Tromp [43•]	Nagamine [44]	Gevaert [45]	Gutela[46]	Uszko-Lencer [47]	Zheng [48]	Zhou [49]
HF type	HFrEF	HFrEF	HFrEF	all HF	all HF	all HF	all HF	all HF	all HF	all HF	all HF
Publication year	2021	2022	2022	2018	2018	2020	2021	2021	2022	2022	2022
Dataset	BB-meta-HF	Single-center	Bio-SHIFT	SwedeHF	ASIAN-HF	Single Center	PACT-HF	OLDW	Single-center	Single-center	Single-center
Data type	Trial	Trial	Cohort	Registry	Registry	EHR	Trial	Registry	EHR	EHR	EHR
Variable type	Clinical	Clinical and Echocardiographic	Biomarker	Clinical	Clinical	Clinical and symptoms	Clinical	Clinical	Clinical	Clinical	Clinical and echocardiographic
Number of patients	15659	108	250	44886	6480	25952	1693	318384	603	4063	562
Mean/median age	64	66	68	76	62	60	77	73	65	69	78
% Female	24	22	26	38	27	43	49	51	29	66	46
Clustering method	Hierarchical clustering	Hierarchical clustering	K-means clustering	Random Forest and k-means (partly supervised)	LCA	K-means clustering	Agglomerative hierarchical clustering	LCA	Hierarchical clustering	LCA	Agglomerative hierarchical clustering
Number of variables	10	17	92	86 (8)	16	NA	14	12	8	12	NA
Number of clusters	12	2	3	4	5	15	6	5	5	4	5
Outcome	All-cause mortality	Major adverse cardiac events (MACEs) or death, and RV-remodelling	Event-free survival time	1-year survival	1-year all-cause death or HF hospitalization	Mortality and cardiac events	1-year all-cause death or all-cause rehospitalization	Hospitalization and mortality	Hospital admissions	1-year HF rehospitalization and all-cause mortality	All-cause mortality, CV mortality
External validation	None	None	none	None	None	None	None	None	None	None	None

EHR Electronic health record; *LCA* Latent class analysis; *PAM* Partitioning around medoids; *CV* Cardiovascular; *HF* Heart failure

Methodology Comparison

Below, we will discuss a few of the trends that could be observed within the three phases of the quality assessment (Table 2 and Supplementary Table S3).

Preparation, Collection, and Checking of the Data

In over half of the studies the generalizability and representativeness of the participants is evaluated ($n=24$). However, only rarely sample size requirements are discussed ($n=2$) [36, 41]. Still, most studies exceeded the threshold of 100 participants for each discovered subgroup ($n=29$). Description of missingness ranged from not mentioning missing values at all ($n=11$) [18, 19, 23, 25, 30, 31, 35, 36, 43•, 44, 48] to reporting percentage of missing for each variable ($n=12$) [10, 20–22, 24, 26, 29, 37, 39•, 40, 46, 47], but also several studies have only given a very global description of the missingness in the dataset usually limited to which variables passed a specific threshold of missingness ($n=11$) [17, 27, 28, 32, 33, 34•, 38, 41, 42, 45, 49]. From the studies that describe handling of missingness they either performed complete case analysis ($n=9$) [26, 32, 33, 36–38, 39•, 43•, 46] or imputation ($n=13$) [10, 17, 18, 20, 22, 24, 29, 34•, 40–42, 48, 49].

Development of the Model

Although most studies often described how they selected the number of clusters and helped the reader to interpret the clustering model with either visual aid or with an explanation ($n=27$), only a small part of the studies provided a

description of the advantages and pitfalls of their chosen clustering technique ($n=11$) [10, 18, 19, 23, 32, 33, 38, 39•, 42, 45, 47]. What is noteworthy is that especially regarding modelling transparency the studies showed low quality, because only rarely the clustering algorithm are being shared ($n=9$) [17, 23, 26, 34•, 37, 42, 43•, 46, 48], and the code or pipeline was never provided. When it comes to feature selection, part of the studies used all variables available or seem to have used all variables available, as they do not mention feature selection ($n=18$), some studies select features a priori (i.e., based on clinician perspective, literature or general availability of the variable in the clinic) ($n=9$) [17, 21, 24, 26, 27, 35, 38, 39•, 47], and other studies use computational approaches to select features (e.g., select features using PCA or correlation coefficient) ($n=7$) [10, 28, 33, 34•, 36, 42, 44].

Validation of the Model

In total, eight studies validated their results in an external validation dataset (Table 2) [17, 18, 23, 26, 34•, 35, 37, 38]. Two of the studies that performed external validation did this with a dataset that was either a subset from the same original dataset or within a dataset that was from the same country, time period, and healthcare setting as the development cohort [10, 34•]. The other studies used external data from different time period, place, or healthcare setting [17, 18, 26, 35, 37, 38]. In the external validation, it was found that phenotypes in the validation cohort had similar outcomes or similar group sizes, depending on whether follow-up data was available.

Table 2 Methodology comparison that summarizes crucial aspects of unsupervised clustering

	HFpEF														HFREF						All HF																		
	Shah[10]	Kao[17]	Segar[18]	Arévalo-Lordó[19]	Hedman[20]	Cohen[21]	Schnab[22]	Harada[23]	Stienen[24]	Gou[25]	Ujj[26]	Casebeer[27]	Woolley[28]	Nouraei[29]	Perry[30]	Faye[31]	Murrey[32]	Choy[33]	Banerjee[34]	Kao[35]	Ahmed[36]	Ferreira[37]	Tromp[38]	Karwath[39]	Bouali[40]	de Lange[41]	Ahmed[42]	Tromp[43]	Nigamini[44]	Gevaert[45]	Guille[46]	Usako[47]	Lencer[48]	Zheng[49]	Zhou[49]				
1. Preparation, collection and checking of the data																																							
1.1 Sample size																																							
1.2 Representativeness																																							
1.3 Data quality																																							
1.4 Data pre-processing																																							
2. Development of the model																																							
2.1 Model selection and interpretability																																							
2.2 Feature selection																																							
2.3 Internal validation																																							
2.4 Measures to reduce risk of overfitting																																							
2.5 Transparency modelling process																																							
3. Validation of the model																																							
3.1 External validation																																							
3.2 Generalizability																																							

Legend. Structure is based on a scoping review and two practical guides for unsupervised learning [14•, 15, 16]. Green = all requirements for that methodology item have been met, yellow = part of requirements for that methodology item have been met, orange = none of the requirements for that methodology item have been met. Supplementary Table S1 provides a more elaborate description on all methodology requirements, supplementary Table S2 provides a more detailed explanation on the requirement assessment. Unmet requirements can be the result of either not (adequately) performing specific analysis steps or the lack of reporting on details of the analysis steps

Phenotype Comparison

Of the 165 described phenotypes, 149 could be assigned to a proposed qualitative framework of nine phenotypes that transcended studies and EF subtypes (Table 3): 1) young-low comorbidity burden phenotype ($n = 32$); 2) metabolic phenotype ($n = 29$); 3) cardio-renal phenotype ($n = 19$); 4) AF phenotype ($n = 17$); 5) elderly female AF phenotype ($n = 16$); 6) hypertensive-comorbidity phenotype ($n = 14$); 7) ischaemic-male phenotype ($n = 16$); 8) valvular disease phenotype ($n = 2$); and 9) a devices phenotype ($n = 4$). The prevalence of the phenotype characteristics of these nine phenotypes are quantified in Table 4.

Young-low Comorbidity Burden Phenotype

The young low comorbidity burden phenotype could be assigned in 17/19 studies in HFpEF, 6/7 studies in HFrEF and 7/8 studies in all HF patients. This cluster is characterised by a lower comorbidity burden and younger age, with in addition obesity ($n = 10$), lower NT-proBNP levels ($n = 7$) and milder HF symptoms ($n = 7$). To some extent lower NYHA ($n = 4$) and smoking ($n = 4$) is reported for this phenotype. Sex is not reported consistently; six studies mention more males while five studies mention more females.

Table 3 Phenotype key characteristics for each study

		HFpEF									
		Shah[10]	Kao[17]	Segar[18]	Arévalo-Lorido[19]	Hedman[20]	Cohen[21]	Schrub[22]	Harada[23]	Stienen[24]	Gu[25]
Number of clusters		3	6	3	5	6	3	3	4		3
1	Youngest, low BNP, least comorbidities, least electric and myocardial remodeling		Male, younger, higher alcohol use, less comorbidities	Diastolic dysfunction, lowest burden of symptoms	diabetes, dyslipidemia, AF	Young, male, obese, milder HF, low NT-proBNP	Younger, smoking, low comorbidity burden	Female, lower comorbidity burden, subnormal left ventricular systolic function	Younger, mild symptoms, LV relaxation abnormality, male, higher BMI	Younger, low comorbidity burden, less advanced HF	Youngest, lower comorbidity burden, more male, more smoking, lower NYHA, lower BNP
2	Obesity, diabetes, OSAS, hyperlipidemia, hypertension		Female, younger, anaemia, less comorbidities	Obesity, diabetes, high comorbidity burden, high NT-proBNP, high NYHA	diabetes, peripheral artery disease, CKD, obese, hypertension	Younger, hypertension, CAD, CKD, anaemia, diabetes	Obesity, diabetes, worse renal function, worse NYHA, more edema	Male, younger, hypertension, diabetes, obesity, CKD	Older, renal dysfunction, female, RV and LV function preserved	Older, diabetes, AF, worse renal function, worse NYHA	Male, IHD, diabetes, hypertension, overweight, dyslipidemia, higher BNP
3	Oldest, CKD, AF, highest BNP, worst RV function		Obesity, diabetes, hyperlipidemia, CAD, CKD	Intermediate comorbidity burden, hypertensive	CKD, dyslipidemia, obese	older AF, COPD, CKD, AF, anaemia, high NYHA and NT-proBNP	older female, AF, lower renal function	Older, female, AF, mitral regurgitation, severe LA enlargement	AF, advanced biventricular diastolic dysfunction		Oldest, female, more AF
4			Female, diabetes, hyperlipidaemia, obesity and CKD		Male, AF, low NYHA	Male, hypertension, AF, pacemakers			Old age, renal dysfunction, female, AF, RV afterload mismatch		
5			Female, oldest, lean, AF, valvular disease, CKD, anaemia		female, hypertension, low NYHA, stroke	older female, lean, hypertension, AF					
6			Male, lean, AF, CAD, alcohol use			older female, hypertension, CAD					

		HFpEF								
		Ujjl[26]	Casebeer[27]	Woolley[28]	Nouraei[29]	Perry[30]	Fayol[31]	Murray[32]	Choy[33]	Banerjee[34]
Number of clusters		5	3	6	6	7	3	4	3	5
1	Young, low comorbidity burden		Older, female, low comorbidity burden	Young, obesity, low comorbidity burden, lowest NTproBNP	Female, low comorbidity	No comorbidities	Lowest rate of comorbidities, amyloidosis	Younger, male, black, obesity, diabetes	Young, low comorbidity burden, smoking	Young, low comorbidity burden, obesity
2	Younger, obesity, diabetes, hypertension, IHD		Younger, diabetes, obesity, IHD, OSAS, smoking, dyslipidemia	CKD, diabetes, overweight, higher NYHA	Female, hypertension, diabetes, dyslipidemia	Diabetes, younger, lower eGFR	Obesity, dyslipidaemia, diabetes, COPD	Older, white/asian, AF, high BNP, low blood pressure, high heart rate	Young, diabetes, obesity, CKD	Obesity, medium comorbidity burden, low CVD
3	older, female, IHD, hypertension, AF, CKD, worse NYHA		Male, AF, CKD, dyslipidemia	older AF, hypertension, CKD	Males, CAD, hypertension, OSAS, diabetes, overweight	AF	CAD, male, MI	Older, female, white/asian, AF, high BNP	older, AF, CKD, pacemaker, hypothyroidism	Obesity, high rate of prescribed medications, high rate of comorbidities, and CVD
4	AF, hypertension			COPD, IHD, anaemia, smoking	Males, older, AF, hypertension, higher BNP, RV size and LV mass index larger	AF, CRT, hypertension		Older, vascular disease, hypertension, low heart rate		AF, valve disease, obesity, CKD
5	older female, lean, AF, hypertension				Female, older, AF, hypertension	older wide QRS, Lean				older female, hypertension, low prevalence CVD
6					older male, hypertension, CAD, dyslipidemia	CRT, Younger				
7						IHD, male				

Table 3 (continued)

Number of clusters	HFrEF						
	Kao[35] 6	Ahmad[36] 4	Ferreira[37] 4	Tromp[38] 6	Karwath[39] 11	Bouali[40] 2	de Lange[41] 3
1	young, obesity, anemia, non-caucasian	Youngest, obesity, less comorbidities, less advanced HF, lowest NT-proBNP	Younger, less comorbidities, lower BMI, higher eGFR, lower LVEF	Youngest, obesity, less comorbidities, lower NYHA, mild HF, lowest NT-proBNP	Younger, MI, Male, high BMI	Lower comorbidity burden, less advanced pathology, lower NT-pro-BNP levels, lower NYHA	Older, longer duration of HF, lower eGFR, diabetes, hypertension, AF, higher baseline biomarker levels, worst prognosis
2	Female, obesity, diabetes, hypertension, hyperlipidemia, anemia, CKD	Ischemic cardiomyopathy, angina symptoms, PCI/CABG, ICD/CRT use.	Female, hypertension, diabetes, AF, high BMI	Oldest, anemia, CKD, highest NT-proBNP	Younger, lower LVEF, high HR, high BMI, higher NYHA	Higher comorbidity burden, AF, advanced pathology, higher NT-pro-BNP levels	Younger, shorter duration of HF, larger slopes for most biomarkers, MI/PCI/CABG
3	male, obesity, hypertension, hyperlipidemia	Young, most women, less comorbidities, mild HF, ICD/CRT use	Younger, male, CKD, anemia, hypertension, diabetes, PCI/CABG, AF	High NT-proBNP, AF, elevated heart rate, anemia, higher NYHA	AF, older higher LVEF		Younger, less advanced HF, more often cardiomyopathy, lower NYHA, ICD/CRT
4	older male, hypertension, CKD, non-caucasian	older, men, smoking, ischemic cardiomyopathy, higher NYHA, AF, CKD, COPD, ICD/CRT use, highest NT-proBNP	Male, hypertension, diabetes, PCI/CABG, high BMI, COPD	Hypertension, lower NT-proBNP, IHD	AF, older		
5	older, male, AF, valve disease, pacemaker			Ischemic, less edema	AF, Younger, low MI, high BMI		
6	female, LBBB, hyperlipidemia, caucasian			Low levels of CHIT1, anemia	AF, high MI, male, low LVEF, low HR, low BMI		
7					AF, Younger, low MI, male, low LVEF, low BMI		
8					older female, higher LVEF, hypertension, higher NYHA		
9					older MI, higher LVEF, lower NYHA		
10					older MI, high LVEF		
11					low LVEF, low		

Number of clusters	all HF							
	Ahmad[42] 4	Tromp[43] 5	Nagamine[44] 15	Gevaert[45] 6	Gulea[46] 5	Uszko-Lencer[47] 5	Zheng[48] 4	Zhou[49] 5
1	Youngest, male, ischaemic cardiomyopathy, smoking, low comorbidity burden, devices, HFrEF	Low comorbidity burden, female, HFrEF, eccentric hypertrophy	AF, dyspnea	Younger, low comorbidity burden	Low comorbidity burden, Younger, CAD	Low comorbidity burden, obesity, low NT-proBNP, lower NYHA	Younger, hyperlipidaemia, obesity, mild HF, diabetes	Younger, less comorbidities, lower LVEF
2	Male, younger, obesity, hypertension, diabetes, dyslipidaemia, HFrEF	Obesity, hypertension, diabetes, HFpEF, concentric remodelling	Hypertensive heart disease, obesity, hyperlipidaemia, female	Younger, highest weight, diabetes, hypertension, OSAS	Diabetes, obesity, hypertension, younger	Male, older ischemic, lower LVEF, diabetes, obesity, smoking	hypertension, CKD, previous MI, diabetes, anaemia, lung disease, various CVD, high NT-proBNP	AF, hyperlipidaemia, IHD
3	Older, lean, CKD, AF, aortic stenosis, COPD, prior MI, highest NT-proBNP.	Female, diabetes, low BMI, HFpEF, concentric hypertrophy, CKD	Hypertension, angina, history MI, male	oldest, female, AF	Diabetes, CAD, PAD, CVA, AF, COPD, CKD	Older CKD, anaemia, more hospital admissions, higher NT-proBNP, higher CRP	older AF, valvular heart disease, lung disease, high NT-proBNP	Oldest, CKD, CV comorbidities
4	Oldest, lean, female, hypertension, AF, stroke/tia, high NT-proBNP, HFpEF	Older, AF, stroke, female, HFpEF, concentric remodelling	Acute MI, ACS, CAD, stenosis, male	COPD, hypertension	Anaemia, CKD, hypertension	Younger, high LVEF, anxiety and depression, obesity	previous MI, valvular heart disease, ischaemic, lung disease, male, higher NYHA	Higher LVEF, lower LV parameters, female, older, AF
5		Male, CAD, ischemic, HFrEF, eccentric hypertrophy	Unstable angina, CAD, ACS, male	CAD, male, coronary revascularization, high NT-proBNP	Ischemic, CAD, PAD, CVA, AF, COPD	Female, lower glucose and better eGFR, lean, osteoporosis		hypertension, hyperlipidaemia, diabetes, IHD, PAD
6			Angina, IHD, cerebrovascular diseases, Cardiac surgery, male	Valvular heart disease, older, high NT-proBNP				
7			Aortic valve disease/stenosis/insufficiency					
8			Congenital heart defects					
9			NICU, congenital heart disease					
10			Decompensated CHF, pulmonary hypertension, pulmonary embolism, male					
11			Dilated cardiomyopathy, dyspnea, mitral valve insufficiency, male					
12			Cerebrovascular disease					
13			Hypertrophic cardiomyopathy, mitral valve dysfunction, diastolic dysfunction					
14			Isolated cardiomyopathy, female					
15			Pediatric cardiomyopathy, myocarditis, cardiac arrhythmias					

Table 3 (continued)

Phenotypes that show comparable characteristics are highlighted in the same colour. Yellow: young-low comorbidity phenotype; light orange: diabetic-obesity phenotype; dark orange: cardio-renal phenotype; red: AF phenotype; purple (dark letters): old female phenotype; purple (white letters): hypertensive phenotype; dark violet: ischaemic-male phenotype; dark blue: valvular disease phenotype; black: devices phenotype, grey: other. *AF* Atrial fibrillation; *BMI* Body mass index; *BNP* B-type natriuretic peptide; *CABG* Coronary artery bypass graft; *CAD* Coronary artery disease; *CHF* Congestive heart failure; *CHIT1* Chitotriosidase; *CKD* Chronic kidney disease; *COPD* Chronic obstructive pulmonary disease; *CRT* Cardiac resynchronization therapy; *CVA* Cerebral vascular accident; *CVD* Cardiovascular disease; *eGFR* Estimated glomerular filtration rate; *HF* Heart failure; *ICD* Implantable cardioverter-defibrillator; *IHD* Ischemic heart disease; *LA* Left atrial; *LV* Left ventricle; *LBBB* Left bundle branch block; *LVEF* Left ventricle ejection fraction; *MI* Myocardial infarction; *NICU* Neonatal intensive care unit; *NYHA* New York heart association; *OSAS* Obstructive sleep apnea syndrome; *PAD* Peripheral artery disease; *PCI* Percutaneous coronary intervention; *RV* Right ventricle

Metabolic Phenotype

The metabolic phenotype could be assigned in 17/19 studies in HFpEF, 3/7 studies in HFrEF and 5/8 studies in all HF patients. Patients in this phenotype more often have obesity or are overweight, and have diabetes and hypertension. In addition, younger age ($n = 10$); CKD ($n = 8$) and an imbalance of lipids ($n = 11$) are often reported. Several studies observed some form of ischaemia (IHD $n = 4$; CAD

$n = 3$). Sex is not reported consistently; five studies mention more males while four studies mention more females.

Cardio-renal Phenotype

The cardio-renal phenotype could be assigned to 10/19 studies in HFpEF, 3/7 studies in HFrEF and 6/8 studies in all HF patients. Patients clustered in this phenotype had CKD or worse renal function, were older and more often

Table 4 Frequency of phenotype characteristics of the nine most common phenotypes

Phenotype	Young - Low comorbidity burden	Metabolic	Cardio - Renal	AF	Elderly - Female - AF
Times phenotype was found in cluster studies	32	29	19	17	16
Percentage occurrence phenotype s per HF category	89% of studies in HFpEF, 86% of studies in HFrEF, 88% of studies all HF patients	89% of studies in HFpEF, 43% in studies in HFrEF, 63% of studies in all HF patients	53% of studies in HFpEF, 43% of studies in HFrEF, 75% of studies in all HF patients	42% of studies in HFpEF, 57% of studies in HFrEF, 25% of studies in all HF patients	42% of studies in HFpEF, 14% of studies in HFrEF, 63% of studies in all HF patients
	Characteristic Count	Characteristic Count	Characteristic Count	Characteristic Count	Characteristic Count
	Low comorbidity burden 26 Young 23 Obesity 10 Mild HF 7 Lower NT-proBNP 7 Male 6 Female 5 Smoking 4 Lower NYHA 4 Lower LVEF 3 Anaemia 2 Less advanced pathology 2 HFrEF 2 Abnormal LV relaxation 1 Amyloidosis 1 Coronary artery disease 1 Devices 1 Diabetes 1 Diastolic dysfunction 1 Eccentric Hypertrophy 1 Higher alcohol use 1 Higher eGFR 1 Higher heart rate 1 Higher NYHA 1 Hyperlipidaemia 1 Ischaemic cardiomyopathy 1 Lower BMI 1 Myocardial infarction 1 Non-Caucasian 1 Older 1 Subnormal LV systolic function 1	Diabetes 27 Obesity 20 Hypertension 16 Younger 10 CKD 8 Dyslipidaemia 6 Hyperlipidaemia 5 Overweight/higher BMI 5 Male 5 Female 4 OSAS 4 Ischaemic heart disease 4 AF 3 lower eGFR 3 Coronary artery disease 3 Anaemia 2 Older 2 Smoking 2 Black 1 Concentric remodelling 1 COPD 1 HFpEF 1 HFrEF 1 High comorbidity burden 1 High CVD 1 High prescription rate 1 Higher baseline biomarkers 1 Higher BNP 1 Higher NYHA 1 Longer duration HF 1 low CVD 1 Lower LVEF 1 Lower NYHA 1 Medium comorbidity burden 1 More oedema 1 Peripheral artery disease 1	CKD 17 Older 13 AF 11 Anaemia 7 Hypertension 5 Diabetes 5 Higher NT-proBNP 5 Female 4 COPD 3 Lean/low BMI 3 Higher NYHA 3 Cardiovascular disease 2 Male 2 Myocardial infarction 2 Obesity 2 Valvular disease 2 Worse renal function 2 Aortic stenosis 1 Concentric hypertrophy 1 Coronary artery disease 1 CVA 1 Dyslipidaemia 1 HFpEF 1 Higher BNP 1 Higher CRP 1 Hypothyroidism 1 Ischaemic heart disease 1 Lung disease 1 Non-Caucasian 1 Pacemaker 1 PCI/CABG 1 Peripheral artery disease 1 RV/LV function preserved 1 Valvular disease 1 Worse RV function 1 Younger 1	AF 17 Male 8 Hypertension 4 Lean/low BMI 3 Higher NT-proBNP 3 Advanced pathology 2 CKD 2 Dyslipidaemia 2 Higher comorbidity burden 2 Less myocardial infarction 2 Lower LVEF 2 Older 2 Pacemaker 2 Younger 2 advanced biventricular diastolic dysfunction 1 Anaemia 1 Coronary artery disease 1 CRT 1 Dyspnoea 1 Higher alcohol use 1 Higher BMI 1 Higher BNP 1 Higher heart rate 1 Higher NYHA 1 Hyperlipidaemia 1 Ischaemic heart disease 1 Larger LV mass index 1 Larger RV size 1 Lower heart rate 1 Lower NYHA 1 More myocardial infarction 1 Valvular disease 1	Older 16 AF 16 Female 12 Hypertension 4 Caucasian/Asian 2 HFpEF 2 Higher BNP 2 Higher LVEF 2 Higher NT-proBNP 2 Lean/low BMI 2 Stroke 2 CAD 1 Concentric remodelling 1 Higher heart rate 1 Lower blood pressure 1 lower eGFR 1 Lower LV parameters 1 Lung disease 1 Mitral regurgitation 1 Renal dysfunction 1 RV afterload mismatch severe LA enlargement 1 Valvular disease 1

Table 4 (continued)

Phenotype	Hypertensive - Comorbidity burden		Ischaemic - Male		Valvular disease		Devices	
Times phenotype was found in cluster studies	14		16		2		4	
Percentage occurrence phenotypes per HF category	37% of studies in HFpEF, 43% of studies in HF rEF, 50% of studies in all HF patients		10% of studies in HFpEF, 57% of studies in HF rEF, 63% of studies in all HF patients		0% of studies in HFpEF, 0% of studies in HF rEF, 25% of studies in all HF patients		5% of studies in HFpEF, 29% of studies in HF rEF, 0% of studies in all HF patients	
	Characteristic	Count	Characteristic	Count	Characteristic	Count	Characteristic	Count
	Hypertension	14	Male	9	Valvular disease	2	ICD/CRT	4
	Female	5	Ischaemic heart disease	6	Aortic valve stenosis	1	Younger	3
	Older	5	Coronary artery disease	6	Aortic valve insufficiency	1	Mild HF	2
	Ischaemic heart disease	4	Myocardial infarction	6	Higher NT-proBNP	1	Older	2
	COPD	3	Angina	4	Older	1	Ischaemic cardiomyopathy	2
	Anaemia	2	Coronary revascularisation	3			AF	1
	Male	2	ASC	2			CKD	1
	Hyperlipidaemia	2	Cerebrovascular disease	2			COPD	1
	Diabetes	2	Higher NYHA	2			Female	1
	CKD	1	Higher LVEF	2			Higher NT-proBNP	1
	Coronary artery disease	1	Acute MI	1			Higher NYHA	1
	Dyslipidaemia	1	AF	1			Lean	1
	Higher BMI	1	Cardiac surgery	1			Low comorbidity burden	1
	Higher LVEF	1	COPD	1			Lower NYHA	1
	Higher NYHA	1	Eccentric remodelling	1			Male	1
	Intermediate comorbidity burden	1	HF rEF	1			Smoking	1
	Lower CVD	1	Higher NT-proBNP	1			Wide QRS	1
	Lower heart rate	1	Hypertension	1				
	Lower NT-proBNP	1	ICD/CRT	1				
	Lower NYHA	1	Ischaemic cardiomyopathy	1				
	Obesity	1	Larger slopes biomarkers	1				
	PCI/CABG	1	Less oedema	1				
	Peripheral artery disease	1	Lower BMI	1				
	Smoking	1	Lower LVEF	1				
	Stroke	1	Lower NYHA	1				
	Vascular disease	1	Lung disease	1				
			Peripheral artery disease	1				
			Shorter duration HF	1				
			Stenosis	1				
			Valvular disease	1				
			Younger	1				

AF Atrial fibrillation; BMI Body mass index; BNP B-type natriuretic peptide; CABG Coronary artery bypass graft; CAD Coronary artery disease; CHF Congestive heart failure; CHIT1 Chitotriosidase; CKD Chronic kidney disease; COPD Chronic obstructive pulmonary disease; CRT Cardiac resynchronization therapy; CVA Cerebral vascular accident; CVD Cardiovascular disease; eGFR Estimated glomerular filtration rate; HF Heart failure; ICD Implantable cardioverter-defibrillator; IHD Ischemic heart disease; LA Left atrial; LV Left ventricle; LBBB Left bundle branch block; LVEF Left ventricle ejection fraction; MI Myocardial infarction; NICU Neonatal intensive care unit; NYHA New York heart association; OSAS Obstructive sleep apnea syndrome; PAD Peripheral artery disease; PCI Percutaneous coronary intervention; RV Right ventricle

had AF. Also more often reported were anemia ($n = 7$), hypertension ($n = 5$) and diabetes ($n = 5$). Several CVDs are observed in this phenotype, including myocardial infarction, valvular disease and coronary artery disease. Sex is not reported consistently; two studies mention more males while four studies mention more females.

AF Phenotype

The AF phenotype could be assigned in 8/19 studies in HFpEF, 4/7 studies in HF rEF and 2/8 studies in all HF patients. This phenotype mainly includes patients with AF. Male sex is more reported ($n = 8$) as well as hypertension ($n = 4$). There are inconsistencies between clusters assigned to this phenotype, some studies report younger patients ($n = 2$) whereas others report older patients ($n = 2$).

Elderly Female AF Phenotype

The older female phenotype could be assigned in 8/19 studies in HFpEF, 1/7 studies in HF rEF and 5/8 studies in all HF patients. Patients in this phenotype are elderly,

have AF and are more often female. In addition, hypertension ($n = 4$), higher BNP/NT-proBNP ($n = 4$) and HFpEF ($n = 2$) are reported.

Hypertensive-comorbidity Phenotype

The hypertensive-comorbidity phenotype could be assigned in 7/19 studies in HFpEF, 3/7 studies in HF rEF and 4/8 studies in all HF patients. Patients clustered to this phenotype have hypertension as main comorbidity. In addition, older age ($n = 5$); IHD ($n = 4$) and COPD ($n = 3$) are often reported. Several studies reported anemia, hyperlipidaemia or diabetes (all $n = 2$). Sex is not reported consistently; two studies mention more males while five studies mention more females.

Ischaemic-male Phenotype

The ischaemic-male phenotype could be assigned in 2/19 studies in HFpEF, 4/7 studies in HF rEF and 5/8 studies in all HF patients. Patients assigned to this phenotype more often have ischaemic heart disease, CAD or previous myocardial infarction.

In addition, angina ($n=4$); revascularisation ($n=3$) are more often reported. Several studies also reported higher NYHA ($n=2$). Nine studies reported more males in this phenotype.

Valvular Phenotype

The valvular phenotype could be assigned in 2/8 studies in all HF patients and no studies specifically in patients with HFpEF or HFrEF. Patients assigned to this phenotype more often have valvular disease as main comorbidity. Few other characteristics are reported.

Devices Phenotype

The devices phenotype could be assigned in 1/19 studies in HFpEF, 2/7 studies in HFrEF and no studies in all HF patients. Patients assigned to this phenotype more often have implantable devices such as ICD or CRT. In addition, they have milder HF ($n=2$) and ischaemic cardiomyopathy ($n=2$). Age is not reported consistently; 3 studies mention younger patients while 2 studies mention older patients.

Prognosis

The young-low comorbidity phenotype most often had the best prognosis compared to the other subgroups (Table 5). However, this trend was not present in the studies performed on HFrEF patients, where their outcomes were mostly

intermediate. The group with the worst outcomes was the cardio-renal phenotype, and this trend can be seen across the EF spectrum. The AF phenotype and male-ischaemic phenotype mostly had intermediate prognosis, a trend that was also present in all EF categories. The prognoses of the metabolic phenotype and hypertensive phenotype were highly variant in relation to the other phenotypes, however for the metabolic phenotype it seems that their prognosis in patients with HFpEF is worse than in patients with HFrEF. There was not enough data on the prognosis of the valvular disease phenotype and devices phenotype to discover any trends.

Discussion

In this systematic review we examined 34 clustering studies in patients with HF, of which 19 studies were exclusively performed in patients with HFpEF. [10, 17–33, 34, 35–38, 39, 40–42, 43, 44–49] Methodologies and phenotypes showed major heterogeneity in the study designs, including the types and sizes of the datasets, clustering algorithms, and selected variables. None of the clustering studies fulfilled all components of the quality assessment, however the degree of methodological limitations differed between the studies. Especially model validation was lacking, only eight studies performed external validation. There was a large overlap in clusters found in the studies, and we identified nine commonly described phenotypes: young-low comorbidity burden; metabolic; cardio-renal; AF; elderly female AF; hypertensive-comorbidity; ischaemic-male; valvular disease; and devices.

Table 5 Prognosis of the nine most common phenotypes

Phenotypes	HFpEF			HFrEF			all HF		
	Best Survival (%)	Intermediate survival (%)	Worst survival (%)	Best survival (%)	Intermediate survival (%)	Worst survival (%)	Best survival (%)	Intermediate survival (%)	Worst survival (%)
Young-Low Comorbidity Burden	12 (75.0)	3 (18.8)	1 (6.3)	2 (28.6)	3 (42.9)	2 (28.6)	5 (83.3)	1 (16.7)	0 (0.0)
Metabolic	4 (22.2)	5 (27.8)	9 (50.0)	2 (50.0)	1 (25.0)	1 (25.0)	2 (50.0)	2 (50.0)	0 (0.0)
Cardio-Renal	1 (10.0)	4 (40.0)	5 (50.0)	1 (33.3)	0 (0.0)	2 (66.7)	0 (0.0)	0 (0.0)	5 (100.0)
AF	4 (44.4)	3 (33.3)	2 (22.2)	2 (33.3)	1 (16.7)	3 (50.0)	0 (0.0)	1 (100.0)	0 (0.0)
Elderly - Female - AF	0 (0.0)	5 (62.5)	3 (37.5)	0 (0.0)	2 (100.0)	0 (0.0)	1 (20.0)	4 (80.0)	0 (0.0)
Hypertensive - Comorbidity	0 (0.0)	4 (57.1)	3 (42.9)	1 (33.3)	1 (33.3)	1 (33.3)	0 (0.0)	1 (50.0)	1 (50.0)
Ischaemic - Male	1 (50.0)	1 (50.0)	0 (0.0)	1 (20.0)	2 (40.0)	2 (40.0)	0 (0.0)	3 (75.0)	1 (25.0)
Valvular disease	NA	NA	NA	NA	NA	NA	0 (0.0)	0 (0.0)	1 (100.0)
Devices	1 (50.0)	0 (0.0)	1 (50.0)	2 (66.7)	1 (33.3)	0 (0.0)	NA	NA	NA

AF Atrial fibrillation; HFpEF Heart failure with preserved ejection fraction; HFrEF Heart failure with reduced ejection fraction; HF Heart failure

Qualitative Phenotype Framework

Based on the clustering studies we created a qualitative phenotype framework consisting of 9 phenotypes. Two phenotypes were most consistently seen in all clustering studies: the young-low comorbidity burden phenotype and metabolic phenotype. To explain the young-low comorbidity burden phenotype, we hypothesize that part of these patients might have BNP deficiency syndrome as proposed by Shah et al. in 2015 [10]. At least 10 cluster studies reported obesity in this phenotype and it is known that obesity can influence the BNP clearance through higher neprilysin levels and increased renal filtration [50]. Furthermore, especially in the studies in HFpEF, it could indicate that patients in this phenotype have recovered HF after treatment with guideline recommended therapy. Another potential reason for this phenotype could be that these patients simply have less severe/advanced HF, which is in line with the better prognosis trends seen in this phenotype.

For the metabolic phenotype, both obesity and diabetes are prone to occur in HF patients. Obesity has been shown to be associated with adverse hemodynamic changes that predispose to cardiac remodelling and ventricular dysfunction and thus HF, also in the absence of other comorbidities [51]. In addition, diabetes has also been shown to be independently associated with an increased risk of HF, cardiovascular mortality, and HF hospitalization [52]. This phenotype appears to cluster more around patients with HFpEF compared to those with HFrEF, yet can still be found across the EF spectrum. This confirms the notion that HFpEF pathophysiology is more driven by metabolic disturbances and an inflammatory burden [53].

There were three phenotypes that all had AF as one of the main three components, these were the cardio-renal, AF and elderly-female phenotypes. These were included in the qualitative framework as three phenotypes as there were distinct differences between clusters with regards to presence across the EF spectrum and prognosis. Several studies have shown the close relation between AF and HF [54]. What is unique about the AF phenotype is that these patients often are of intermediate age and can be both male or female, in contrast to the elderly-female-AF phenotype. This phenotype also appeared across the EF spectrum, yet there could be differences in the pathophysiology of this phenotype. For example, the prognosis of this phenotype appears to be worse in patients with HFrEF, whereas it appears better in patients with HFpEF. It is proposed that in HFrEF, AF may be a consequence of the HF, whereas in HFpEF, both ventricular and atrial myopathy may develop in parallel [54, 55]. And indeed, several studies reported changes in left ventricle and left atrium parameters. Which can be seen in the elderly-AF-female phenotype, which was more prevalent in patients with HFpEF. The proportion of patients with HF and

concomitant AF increases with age, which is very likely observed in this phenotype [56].

In the cardio-renal phenotype we observe the bidirectional interaction between kidney function and HF [57]. Previous studies have shown that CKD is more common in patients with HFpEF, yet might play a larger role in the prognosis of patients with HFrEF [58, 59]. In this review we consistently show a worse survival regardless of EF. In addition, several studies reported anaemia, which could be a consequence of the presence of CKD [60]. Anaemia also independently contributes to worse prognosis in HF [61]. Studies have shown that treating anaemia in HF patients is associated with improvements of NYHA class, symptoms and HF hospitalisations [62, 63].

Two phenotypes occurred more frequently in the studies investigating HFrEF patients: The ischaemic-male and devices phenotypes. Ischaemia is one of the underlying cause of HF, and more often in men, where it is the main cause of HF [64, 65]. Patients in this cluster could have a variety of previous ischaemic diseases in their underlying disease pathology for HFrEF [66].

ICDs are implanted in patients with HF that are at risk for sudden cardiac death or all-cause mortality according to the recommendations in the guidelines as both primary and secondary prevention [4]. Primary prevention is targeted to those patients that have symptomatic HF (NYHA class II-III) of an ischaemic aetiology and LVEF $\leq 35\%$. Across the studies, this phenotype occurred mainly in HFrEF patients and was also seen in one study with recovered HF patients. Personalisation in ICD placement is a current unmet need [67]. Clustering could potentially play a role in this personalisation.

The hypertensive-comorbidity phenotype was characterised by the absence of obesity and diabetes and presence of comorbidities such as COPD and IHD. There are several diagnostic challenges in COPD and HF as clinical symptoms can be overlapping [68, 69]. Different characteristics could potentially be used to better define this phenotype, such as biomarkers or echocardiographic parameters.

The valvular disease phenotype was a specific phenotype related to hospitalised inpatients described in two studies in EHR data (in- and outpatients) and one based on a clinical trial (tertiary care or quaternary care) [44, 45]. Valve disease is a known aetiology for HF with a very poor prognosis, with the three main diseases aortic stenosis, aortic regurgitation and mitral insufficiency [4].

Are we There Yet? Precision Medicine for HF

There was significant overlap in the clustering outcomes between the various studies [10, 17–33, 34•, 35–38, 39•, 40–42, 43•, 44–49]. Yet, differences between the clustering studies still exist. This indicates that there is a lack of precision at least to some extent in the subgroups based on clustering.

We hypothesize that the differences in phenotype descriptions could be due to differences between the data sources, as phenotypes characteristics are relative to their patient population. This limits the reproducibility and generalisability of the cluster models to other patient populations and use in routine clinical care. A potential solution could be readjusting or fine-tuning the current models using site specific information to increase the generalisability.

In addition, it is important to underline that we grouped the clusters based on reported characteristics to one of the nine phenotypes. It could be that there are unreported characteristics that would categorize a cluster to a different phenotype if they were known.

Implementation and Future Perspective

Ideally, this systematic review could identify one or multiple clustering studies of sufficient quality for implementation in clinical trials or in clinical practice. Due to the high heterogeneity and absence of a gold standard, this is not possible. Nonetheless, the findings of this review suggest that clustering is a suitable and fruitful approach for capturing the underlying heterogeneity of patients with HF.

Future studies should be aware of the methodological caveats of clustering research and take this into account when performing these studies. Efforts should be directed towards improving the development and validation of clustering as machine learning model [70]. Validating existing models could potentially lead to a more precise, valid and reliable phenotyping model that could be implemented in clinical trial design or as a decision tool in daily clinical practice.

Most importantly, we found that clusters transcend across the EF spectrum, indicating that clusters might not be limited to heterogeneity in HFpEF, but could also play a role in HFrfEF. There were significant differences based on prognosis that are worth to be explored further. In addition, it has not yet been investigated whether patients could change between clusters over time. Longitudinal data is necessary to uncover any transitions over time.

Current studies should therefore be considered as hypothesis generating. In the future it would be potentially be possible to investigate differences in prognosis and treatment benefit in clinical trials. Differences in prognosis could be used to guide future trial inclusion to optimise and enrich clinical trials. Moreover, patients in trials could be stratified based on clustering models to see whether there are different treatment effects. Currently, there is limited data on treatment heterogeneity across clusters and should be studied further. Investigating this would mean a step forward towards finding beneficial treatment options or strategies on subgroup patient level and in the future on individual patient level.

Strengths and Limitations

One of the strengths of this systematic review is that both results and methodology of the clustering studies were compared. This gives context to the results and can provide nuance in the discussion on the reliability and validity of the clustering studies. In addition, the large amount of clustering studies and the heterogeneity of their study designs increase the meaningfulness of their similarities regarding their phenotype models. This enables us to quantify the degree of certainty to some extent regarding phenotype characteristics. Moreover, this systematic review showed a general overview of the requirements of unsupervised clustering. In cardiology, clustering is an increasingly common technique for subgroup discovery, and a basic understanding of the strengths, limitations, and pitfalls of clustering can help facilitate critical evaluation of these studies. Furthermore, at this moment we are the first systematic review that has performed rigorous review methodology and that gives an elaborate overview of both the discovered phenotypes and methodology.

In this systematic review, we developed a methodological quality assessment, as current tools suitable for clustering meta-analysis are non-existent and validated guidelines on reviewing clustering studies are lacking. Therefore, a more descriptive approach has been used. To compare methodologies, the quality assessment was based on a scoping review of Hond et al. [14•], and two practical guidelines on clustering [14•, 15, 16]. It is important to note that we could not always distinguish between worse performance of a study or only lacking to report certain aspects.

The issue of generalizability across different ethnic or social economic backgrounds in patients with HF has been debated. Some have argued that these differences may lead to variations in the presentation and treatment of HF that require separate subgroup analyses [71]. Currently, there is not enough evidence to support biological differences between different populations, therefore, it may be deemed appropriate to generalize our findings to other populations. Indeed, three studies in an Asian population presented comparable phenotypes as to those with other ethnicities [43•, 48, 49].

Lastly, we grouped the clusters according to the reported characteristics for each study. It could be that there are other underlying characteristics that would change the phenotype assignment.

Conclusions

There were many differences between the clustering studies regarding the sizes and types of the datasets, variable selection, and algorithms, but they yielded comparable phenotypes which implies that clustering is a fruitful approach

for phenotype discovery. Specifically, of the 165 phenotypes that were described, 149 could be assigned to one of nine most common phenotypes: young-low comorbidity burden; metabolic; cardio-renal; AF; elderly female AF; hypertensive-comorbidity; ischaemic-male; valvular disease; and a devices phenotype. These phenotypes are not limited to a particular EF, but rather transcended across the EF spectrum. Comparing methodologies of the studies showed that there was still room for improvement on topics concerning validity and reliability, especially regarding external validation. These methodological aspects limit the current implementation into clinical practice and effort should be directed towards improving the clinical utility of cluster analysis. Altogether, this systematic review is hypothesis generating and lays the groundwork for future research into a more precise and reliable phenotype model that can serve as a stratification and decision tool in clinical trial design and personalised medicine for patients with HF.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11897-023-00615-z>.

Funding This work has received support from the EU/EFPIA Innovative Medicines Initiative 2 Joint Undertaking BigData@Heart grant n° 116074. **FA** is supported by UCL Hospitals NIHR Biomedical Research Centre. **IV** is supported by the Dutch Heart Foundation, as part of “Facts and Figures”. **RV** and **MLH** are supported by the Dutch CardioVascular Alliance (2020B008 RECONNECT). **MLH** is supported by the Dutch Heart Foundation (NHS; 2020T058). **AB** is supported by National Institute for Health and Care Research (NIHR), British Medical Association, UK Research and Innovation, European Union, and is trustee of the South Asian Health Foundation and Long COVID SOS.

Declarations

Conflict of Interest **CM, RV, IV, SK, FA, AU** have nothing to disclose. **MLH** received educational/speaker/consultancy fees from Novartis, Boehringer Ingelheim, Daiichi Sankyo, Vifor Pharma, AstraZeneca, Bayer, MSD, and Quin; all not related to this work.

JJB received research grants and/or speaker fees from Vifor, Abbott, Boehringer Ingelheim, Bayer and Novartis outside the submitted work. **AB** received research grants from AstraZeneca; outside the submitted work. **GS** reports grants and personal fees from Vifor, grants and non-financial support from Boehringer Ingelheim, personal fees from Società Prodotti Antibiotici, grants and personal fees from AstraZeneca, personal fees from Roche, Servier, GENESIS, Cytokinetics, Medtronic, grants from Novartis, Boston Scientific, PHARMACOSMOS, Merck, Bayer, outside the submitted work.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in

the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

Papers of particular interest, published recently, have been highlighted as:

• Of importance

1. Crespo-Leiro MG, Anker SD, Maggioni AP, Coats AJ, Filippatos G, Ruschitzka F, et al. European Society of Cardiology Heart Failure Long-Term Registry (ESC-HF-LT): 1-year follow-up outcomes and differences across regions. *Eur J Heart Fail.* 2016;18:613–25.
2. Taylor CJ, Ordóñez-Mena JM, Roalfe AK, Lay-Flurrie S, Jones NR, Marshall T, et al. Trends in survival after a diagnosis of heart failure in the United Kingdom 2000–2017: population based cohort study. *BMJ.* 2019;364:l223.
3. Groenewegen A, Rutten FH, Mosterd A, Hoes AW. Epidemiology of heart failure. *Eur J Heart Fail.* 2020;22:1342–56.
4. McDonagh T, Metra M, Adamo M, Gardner R, Baumbach A, Böhm M, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J.* 2021;42:3599–726.
5. Beezer J, Al Hatrushi M, Husband Slovaki A, Kurdi A, Forsyth P. Polypharmacy definition and prevalence in heart failure: a systematic review. *Heart Fail Rev.* 2021;27:465–92.
6. Rao VN, Fudim M, Savarese G, Butler J. Polypharmacy in heart failure with reduced ejection fraction: progress. *Not Problem Am J Med.* 2021;134:1068–70.
7. Anker SD, Butler J, Filippatos G, Ferreira JP, Bocchi E, Böhm M, et al. Empagliflozin in heart failure with a preserved ejection fraction. *N Engl J Med.* 2021;385:1451–61.
8. Solomon SD, McMurray JJV, Claggett B, de Boer RA, DeMets D, Hernandez AF, et al. Dapagliflozin in heart failure with mildly reduced or preserved ejection fraction. *N Engl J Med.* 2022;387:1089–98.
9. Borlaug BA. Evaluation and management of heart failure with preserved ejection fraction. *Nat Rev Cardiol.* 2020;17:559–73.
10. Shah SJ, Katz DH, Selvaraj S, Burke MA, Yancy CW, Gheorghiu M, et al. Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation.* 2015;131:269–79.
11. Heinzel FR, Shah SJ. The future of heart failure with preserved ejection fraction : Deep phenotyping for targeted therapeutics. *Herz.* 2022;47:308–23.
12. Galli E, Bourg C, Kosmala W, Oger E, Donal E. Phenomapping heart failure with preserved ejection fraction using machine learning cluster analysis: prognostic and therapeutic implications. *Heart Fail Clin.* 2021;17:499–518.
13. Sun J, Guo H, Wang W, Wang X, Ding J, He K, et al. Identifying novel subgroups in heart failure patients with unsupervised machine learning: a scoping review. *Front Cardiovasc Med.* 2022;9:895836.
14. • de Hond AAH, Leeuwenberg AM, Hooft L, Kant IMJ, Nijman SWJ, van Os HJA, et al. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. *NPJ Digit Med Nat Res.* 2022;5:2–2. **This work identifies actionable guidelines for the development, evaluation and implementation of artificial intelligence prediction models.**
15. Kassambara A. Multivariate analysis 1: practical guide to cluster analysis in R. *STHDA* 2017;1–187.

16. Schreiber JB. Latent class analysis: an example for reporting results. *Res Soc Adm Pharm.* 2017;13:1196–201.
17. Kao DP, Lewsey JD, Anand IS, Massie BM, Zile MR, Carson PE, et al. Characterization of subgroups of heart failure patients with preserved ejection fraction with possible implications for prognosis and treatment response. *Eur J Heart Fail.* 2015;17:925–35.
18. Segar MW, Patel KV, Ayers C, Basit M, Tang WHW, Willett D, et al. Phenomapping of patients with heart failure with preserved ejection fraction using machine learning-based unsupervised cluster analysis. *Eur J Heart Fail.* 2020;22:148–58.
19. Arévalo-Lorido JC, Carretero-Gómez J, Aramburu-Bodas O, Grau-Amoros J, Torres-Cortada G, Camafort-Babkowski M. Blood pressure, congestion and heart failure with preserved ejection fraction among patients with and without type 2 diabetes mellitus. A cluster analysis approach from the observational registry DICUMAP. *High Blood Press Cardiovasc Prev Adis.* 2020;27:399–408.
20. Hedman ÅK, Hage C, Sharma A, Brosnan MJ, Buckbinder L, Gan L-MM, et al. Identification of novel pheno-groups in heart failure with preserved ejection fraction using machine learning. *Heart.* 2019;106:342–9.
21. Cohen JB, Schrauben SJ, Zhao L, Basso MD, Cvijic ME, Li Z, et al. Clinical phenogroups in heart failure with preserved ejection fraction: detailed phenotypes, prognosis, and response to spironolactone. *JACC Hear Fail.* 2020;8:172–84.
22. Schrub F, Oger E, Bidaut A, Hage C, Charton M, Daubert JC, et al. Heart failure with preserved ejection fraction: a clustering approach to a heterogeneous syndrome. *Arch Cardiovasc Dis.* 2020;113:381–90.
23. Harada D, Asanoi H, Noto T, Takagawa J. Different pathophysiology and outcomes of heart failure with preserved ejection fraction stratified by K-means clustering. *Front Cardiovasc Med.* 2020;7:607760.
24. Stienen S, Ferreira JP, Kobayashi M, Preud'homme G, Dobre D, Machu JL, et al. Enhanced clinical phenotyping by mechanistic bioprofiling in heart failure with preserved ejection fraction: insights from the MEDIA-DHF study (The Metabolic Road to Diastolic Heart Failure). *Biomarkers.* 2020;25:201–11.
25. Gu J, Pan J-A, Lin H, Zhang J-F, Wang C-Q. Characteristics, prognosis and treatment response in distinct phenogroups of heart failure with preserved ejection fraction. *Int J Cardiol.* 2021;323:148–54.
26. Uijl A, Savarese G, Vaartjes I, Dahlström U, Brugts JJ, Linssen GCMM, et al. Identification of distinct phenotypic clusters in heart failure with preserved ejection fraction. *Eur J Heart Fail.* 2021;23:973–82.
27. Casebeer A, Horter L, Hayden J, Simmons J, Evers T. Phenotypic clustering of heart failure with preserved ejection fraction reveals different rates of hospitalization. *J Cardiovasc Med.* 2021;22:45–52.
28. Woolley RJ, Ceelen D, Ouwerkerk W, Tromp J, Figarska SM, Anker SD, et al. Machine learning based on biomarker profiles identifies distinct subgroups of heart failure with preserved ejection fraction. *Eur J Heart Fail.* 2021;23:983–91.
29. Nouraei H, Rabkin SW. A new approach to the clinical sub-classification of heart failure with preserved ejection fraction. *Int J Cardiol.* 2021;331:138–43.
30. Perry A, Loh F, Adamo L, Zhang KW, Deych E, Foraker R, et al. Unsupervised cluster analysis of patients with recovered left ventricular ejection fraction identifies unique clinical phenotypes. *PLoS One.* 2021;16:e0248317.
31. Fayol A, Wack M, Livrozet M, Carves JBJ-B, Domengé O, Vermersch E, et al. Aetiological classification and prognosis in patients with heart failure with preserved ejection fraction. *ESC Hear Fail.* 2022;9:519–30.
32. Murray E, Greene S, Rao V, Sun J-L, Alhanti B, Blumer V, et al. Unsupervised machine learning to define acute hfpef phenotypes: findings from ascend-hf. *J Card Fail.* 2022;28:S10–1.
33. Choy M, Liang W, He J, Fu M, Dong Y, He X, et al. Phenotypes of heart failure with preserved ejection fraction and effect of spironolactone treatment. *ESC Hear Fail.* 2022;9:2567–75.
34. Banerjee A, Chen S, Dashtban M, Pasa L, Thygesen JH, Fatemifar G, et al. Identifying subtypes of heart failure with machine learning: external, prognostic and genetic validation in three electronic health record sources with 320,863 individuals. *medRxiv.* 2022. **This work performs extensive validation steps, including internal validation through comparison of the outcomes of four different clustering methods, external validation with comparing outcomes in a validation dataset, and genetic validation through association with polygenic risk scores.**
35. Kao DP, Wagner BD, Robertson AD, Bristow MR, Lowes BD. A personalized BEST: characterization of latent clinical classes of nonischemic heart failure that predict outcomes and response to bucindolol. *PLoS One.* 2012;7:e48184.
36. Ahmad T, Pencina MJ, Schulte PJ, O'Brien E, Whellan DJ, Piña IL, et al. Clinical implications of chronic heart failure phenotypes defined by cluster analysis. *J Am Coll Cardiol.* 2014;64:1765–74.
37. Ferreira JP, Duarte K, McMurray JJV, Pitt B, Van Veldhuisen DJ, Vincent J, et al. Data-driven approach to identify subgroups of heart failure with reduced ejection fraction patients with different prognoses and aldosterone antagonist response patterns. *Circ Hear Fail.* 2018;11:e004926.
38. Tromp J, Ouwerkerk W, Demissei BG, Anker SD, Cleland JG, Dickstein K, et al. Novel endotypes in heart failure: effects on guideline-directed medical therapy. *Eur Heart J.* 2018;39:4269–76.
39. Karwath A, Bunting KV, Gill SK, Tica O, Pendleton S, Aziz F, et al. Redefining β -blocker response in heart failure patients with sinus rhythm and atrial fibrillation: a machine learning cluster analysis. *Lancet.* 2021;398:1427–35. **This work shows that a cluster analysis can be developed and applied to beta-blocker trial data where there is a different response to beta-blockers in different clusters of heart failure patients. They find that for one cluster with patients with atrial fibrillation, beta-blockers reduce mortality significantly.**
40. Bouali Y, Galli E, Paven E, Laurin C, Arnaud H, Oger E, et al. Impact of sacubitril/valsartan on systolic heart failure: right heart location and clustering analysis. *Adv Clin Exp Med.* 2022;31:109–19.
41. de Lange I, Petersen TB, de Bakker M, Akkerhuis KM, Brugts JJ, Caliskan K, et al. Heart failure subphenotypes based on repeated biomarker measurements are associated with clinical characteristics and adverse events (Bio-SHiFT study). *Int J Cardiol.* 2022;364:77–84.
42. Ahmad T, Lund LH, Rao P, Ghosh R, Warier P, Vaccaro B, et al. Machine learning methods improve prognostication, identify clinically distinct phenotypes, and detect heterogeneity in response to therapy in a large cohort of heart failure patients. *J Am Heart Assoc E.* 2018;7:e008081.
43. Tromp J, Tay WT, Ouwerkerk W, Teng T-HK, Yap J, MacDonald MR, et al. Correction: multimorbidity in patients with heart failure from 11 Asian regions: a prospective cohort study using the ASIAN-HF registry. *PLoS Med.* 2018;15:e1002541. **This work shows cluster analysis on all HF patients in a diverse, Asian population, and they have detailed description of the found clusters beyond clinical characteristics. This study evaluates the role of ethnicity and ejection fraction regarding HF phenotypes.**
44. Nagamine T, Gillette B, Pakhomov A, Kahoun J, Mayer H, Burghaus R, et al. Multiscale classification of heart failure phenotypes by unsupervised clustering of unstructured

- electronic medical record data. *Sci Rep.* 2020;10:21340 (Nature Research).
45. Gevaert AB, Tibebe S, Mamas MA, Ravindra NG, Lee SF, Ahmad T, et al. Clinical phenogroups are more effective than left ventricular ejection fraction categories in stratifying heart failure outcomes. *ESC Hear Fail.* 2021;8:2741–54.
 46. Gulea C, Zakeri R, Quint JK. Model-based comorbidity clusters in patients with heart failure: association with clinical outcomes and healthcare utilization. *BMC Med.* 2021;19:9.
 47. Uszko-Lencer N, Janssen DJAA, Gaffron S, Vanfleteren L, Janssen E, Werter C, et al. Clustering based on comorbidities in patients with chronic heart failure: an illustration of clinical diversity. *ESC Hear Fail.* 2021;9:614–26.
 48. Zheng C, Han L, Tian J, Li J, He H, Han G, et al. Hierarchical management of chronic heart failure: a perspective based on the latent structure of comorbidities. *ESC Hear Fail.* 2022;9:595–605.
 49. Zhou X, Nakamura K, Sahara N, Asami M, Toyoda Y, Enomoto Y, et al. Exploring and identifying prognostic phenotypes of patients with heart failure guided by explainable machine learning. *Life (Basel, Switzerland).* 2022;12:776.
 50. Reinmann M, Meyer P. B-type natriuretic peptide and obesity in heart failure: a mysterious but important association in clinical practice. *Cardiovasc Med.* 2020;1:2020.
 51. Alpert MA, Lavie CJ, Agrawal H, Aggarwal KB, Kumar SA. Obesity and heart failure: epidemiology, pathophysiology, clinical manifestations, and management. *Transl Res.* 2014;164:345–56.
 52. Dauriz M, Targher G, Laroche C, Temporelli PL, Ferrari R, Anker S, et al. Association between diabetes and 1-Year adverse clinical outcomes in a multinational cohort of ambulatory patients with chronic heart failure: results from the ESC-HFA Heart Failure Long-Term Registry. *Diabetes Care.* 2017;40:671–8.
 53. Schiattarella GG, Rodolico D, Hill JA. Metabolic inflammation in heart failure with preserved ejection fraction. *Cardiovasc Res.* 2021;117:423.
 54. Sartipy U, Dahlström U, Fu M, Lund LH, Dahlstrom U, Fu M, et al. Atrial fibrillation in heart failure with preserved, mid-range, and reduced ejection fraction. *JACC Heart Fail.* 2017;5:565–74.
 55. Packer M, Lam CSP, Lund LH, Redfield MM. Interdependence of atrial fibrillation and heart failure with a preserved ejection fraction reflects a common underlying atrial and ventricular myopathy. *Circulation.* 2020;141:4–6.
 56. Zafrir B, Lund LH, Laroche C, Ruschitzka F, Crespo-Leiro MG, Coats AJS, et al. Prognostic implications of atrial fibrillation in heart failure with reduced, mid-range, and preserved ejection fraction: a report from 14 964 patients in the European Society of Cardiology Heart Failure Long-Term Registry. *Eur Heart J.* 2018;39:4277–84.
 57. Schefold JC, Filippatos G, Hasenfuss G, Anker SD, Von Haehling S. Heart failure and kidney dysfunction: epidemiology, mechanisms and management. *Nat Rev Nephrol.* 2016;12:610–23.
 58. Lofman I, Szummer K, Dahlstrom U, Jernberg T, Lund LH, Löfman I, et al. Associations with and prognostic impact of chronic kidney disease in heart failure with preserved, mid-range, and reduced ejection fraction. *Eur J Heart Fail.* 2017;19:1606–14.
 59. Löfman I, Szummer K, Evans M, Carrero JJ, Lund LH, Jernberg T. Incidence of, associations with and prognostic impact of worsening renal function in heart failure with different ejection fraction categories. *Am J Cardiol.* 2019;124:1575–83.
 60. Babitt JL, Lin HY. Mechanisms of anemia in CKD. *J Am Soc Nephrol.* 2012;23:1631–4.
 61. Go AS, Yang J, Ackerson LM, Lepper K, Robbins S, Massie BM, et al. Hemoglobin level, chronic kidney disease, and the risks of death and hospitalization in adults with chronic heart failure. *Circulation.* 2006;113:2713–23.
 62. Anker SD, Comin Colet J, Filippatos G, Willenheimer R, Dickstein K, Drexler H, et al. Ferric carboxymaltose in patients with heart failure and iron deficiency. *N Engl J Med.* 2009;361:2436–48.
 63. Ponikowski P, Kirwan BA, Anker SD, McDonagh T, Dorobantu M, Drozd J, et al. Ferric carboxymaltose for iron deficiency at discharge after acute heart failure: a multicentre, double-blind, randomised, controlled trial. *Lancet.* 2020;396:1895–904.
 64. Vedin O, Lam CSP, Koh AS, Benson L, Teng THK, Tay WT, et al. Significance of ischemic heart disease in patients with heart failure and preserved, midrange, and reduced ejection fraction: a nationwide cohort study. *Circ Heart Fail.* 2017;10:e003875.
 65. Postigo A, Martínez-Sellés M. Sex influence on heart failure prognosis. *Front Cardiovasc Med.* 2020;7:616273.
 66. Lee DS, Gona P, Vasani RS, Larson MG, Benjamin EJ, Wang TJ, et al. Relation of disease pathogenesis and risk factors to heart failure with preserved or reduced ejection fraction: Insights from the framingham heart study of the national heart, lung, and blood institute. *Circulation.* 2009;119:3070–7.
 67. Schrage B, Lund LH, Benson L, Dahlström U, Shadman R, Linde C, et al. Predictors of primary prevention implantable cardioverter-defibrillator use in heart failure with reduced ejection fraction: impact of the predicted risk of sudden cardiac death and all-cause mortality. *Eur J Heart Fail.* 2022;24:1212–22.
 68. Hawkins NM, Petrie MC, Jhund PS, Chalmers GW, Dunn FG, McMurray JJV. Heart failure and chronic obstructive pulmonary disease: diagnostic pitfalls and epidemiology. *Eur J Heart Fail.* 2009;11:130.
 69. Caruana L, Petrie MC, Davie AP, McMurray JJ. Do patients with suspected heart failure and preserved left ventricular systolic function suffer from “diastolic heart failure” or from misdiagnosis? A prospective descriptive study. *BMJ.* 2000;321:215–8.
 70. Banerjee A, Chen S, Fatemifar G, Zeina M, Lumbers RT, Mielke J, et al. Machine learning for subtype definition and risk prediction in heart failure, acute coronary syndromes and atrial fibrillation: systematic review of validity and clinical utility. *BMC Med.* 2021;19:1–14 (BioMed Central).
 71. Khariton Y, Nassif ME, Thomas L, Fonarow GC, Mi X, DeVore AD, et al. Health status disparities by sex, race/ethnicity, and socioeconomic status in outpatients with heart failure. *JACC Heart Fail.* 2018;6:465–73.

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