Heart failure with preserved ejection fraction A clustering approach to a heterogenous syndrome

Florian Schrub, Emmanuel Oger, Auriane Bidaut, Camilla Hage, Marion Charton, Jean Claude Daubert, Chrtistophe Leclercq, Cecila Linde, Lars Lund, Erwan Donal

Heart failure with preserved ejection fraction: A clustering approach to a heterogenous syndrome

Abbreviated title: A clustering approach to HFpEF

Florian Schrub[a], Emmanuel Oger[b,c], Auriane Bidaut[a], Camilla Hage[d], Marion Charton[a], Jean Claude Daubert[a], Chrtistophe Leclercq[a,b,c], Cecila Linde[d], Lars Lund[d], Erwan Donal[a,b,c,*]

[a] *Department of Cardiology & CIC-IT U 804, Rennes University Hospital, 35000 Rennes, France*

[b] *Rennes Clinical Investigation Centre, INSERM CIC-0203, CHU de Rennes, 35033 Rennes, France*

[c] *LTSI, University of Rennes 1, 35042 Rennes, France*

[d] *Heart and Vascular Theme, Karolinska University Hospital, Karolinska Institutet, 171 76 Stockholm, Sweden*

* Corresponding author at: Cardiology Department, Rennes University Hospital, 2 rue Henri le Guilloux, 35000 Rennes, France.

E-mail address: erwan.donal@chu-rennes.fr (E. Donal).

Summary

Background. – Heart failure with preserved ejection fraction (HFpEF) is a complex syndrome at the crossroads of multiple co-morbidities; there is no valid treatment for this condition. Defining new phenotypes could play a role in improving treatment and prognosis.

Aim. – To identify groups with different pathophysiologies by applying a clustering approach to a multicentric cohort of patients with HFpEF.

Methods. – A total of 538 patients from the multicentre KaRen study were included. Accurate clinical, biological and ultrasound data are available, with a mean follow-up of 28 months. Based on a clustering analysis, the population was separated into groups based on 55 variables, comparing distribution of deaths and hospitalizations between groups.

Results. – Three clusters were identified from 356 analysable patients (mean age 76.1 ± 9.31 years; 43.5% men): cluster 1 (n = 128) comprised overweight, relatively young men at high cardiovascular risk, in sinus rhythm, with altered renal function; cluster 2 (n = 134) comprised women, most of whom had conserved left ventricular function; cluster 3 (n = 94) had the highest incidence of mitral regurgitation, atrial remodelling and rhythm disorders. There were no significant differences, only a trend towards early mortality in cluster 3.

Conclusions. – Clustering analysis seems to be effective at individualizing subgroups with different physiopathologies in HFpEF. The clinical relevance of these phenotypes needs to be studied, and may concern treatment strategy more than prognostic differences.

Résumé

Contexte. – L'insuffisance cardiaque à fraction d'éjection préservée (ICFEp) reste un syndrome complexe au carrefour de multiples comorbidités, sans traitement validé. La définition de nouveaux phénotypes est une piste pour améliorer le traitement et le pronostic.

Objectif. – Notre objectif est d'appliquer une approche par « clustering » sur une cohorte multicentrique de patients ICFEp afin d'identifier des groupes de différentes physiopathologies.

Méthodes. – 538 patients de l'étude multicentrique KaRen ont été inclus. La conception de l'étude a été publiée précédemment. Les données cliniques, biologiques et échographiques précises sont disponibles, sur un suivi moyen de 28 mois. A partir d'une analyse par « cluster », la population a été

séparée en groupes sur la base de 55 variables, en étudiant la distribution des décès et des hospitalisations.

Résultats. – Trois clusters ont été identifiés sur 356 patients analysables (âge moyen = 76,1 ± 9,31 ans ; 43,5 % d'hommes) : cluster 1 (*n* = 128) avec des hommes, plus jeunes, en surpoids, à haut risque cardiovasculaire, en rythme sinusal, de fonction rénale altérée ; cluster 2 (*n* = 134) avec des femmes, une fonction ventriculaire gauche la plus préservée ; cluster 3 (*n* = 94) avec l'incidence la plus haute de régurgitation mitrale, de remodelage auriculaire et de troubles du rythme. Il n'y pas de différence significative, hormis une tendance à la mortalité précoce dans le groupe 3.

Conclusions. – L'analyse par « clustering » semble être efficace pour individualiser des sous-groupes de différentes physiopathologies. L'impact clinique de ces trois phénotypes reste à démontrer.

KEYWORDS

Heart failure;

Preserved ejection fraction;

Clustering;

Phenotype;

KaRen substudy

MOTS CLÉS

Insuffisance cardiaque ;

Fraction d'éjection préservée ;

Clustering ;

Phenotypage ;

Sous étude KaRen

Abbreviations: 2D, two-dimensional; AF, atrial fibrillation; EF, ejection fraction; HF, heart failure; HFpEF, heart failure with preserved ejection fraction; HFrEF, heart failure with reduced ejection fraction; LA, left atrial.

Background

Within the broad spectrum of “heart failure” (HF), heart failure with preserved ejection fraction (HFpEF) remains a diagnostic and therapeutic challenge. The prevalence is continuing to rise, with HFpEF representing 50% of HF cases [1, 2] – mostly older women [3] with multiple co-morbidities such as hypertension, atrial fibrillation (AF), obesity and kidney disease [3].

Higher ejection fraction (EF) does not mean better prognosis. Shah et al. [2] found the same poor 5-year outcome in the three categories of HF (preserved EF, mid-range EF and altered EF), with a mortality rate of 75% and a rehospitalization rate of approximately 83%. Cause of death is the main difference between HF categories, with a twofold higher rate of non-cardiovascular death among patients with HFpEF, mostly linked to the fragility of this older and co-morbid population with more infection and cancer [4].

First presented as “diastolic dysfunction” [5], the latest guidelines define HFpEF as an association of “clinical signs and symptoms”, “functional and morphological abnormalities” and “modification of biomarkers” [6]; this contrasts with heart failure with reduced ejection fraction (HFrEF), which is mainly defined by an EF < 40% [6].

Physiopathology is determined by a complex interaction between many factors, such as systolic dysfunction and deformation impairment, chronotropic incompetence and impaired contractility reserve, left atrial (LA) dysfunction and AF, arterial stiffening and abnormal ventricular-arterial coupling, autonomic imbalance and endothelial dysfunction [7]. Synergy between systemic inflammation and coronary microvascular endothelial dysfunction appears to be the theory unifying the different cardiovascular and extracardiovascular factors, leading to dysfunctional myocardial signalling and multiorgan alterations [8]. Instead of being a precise and well defined “disease”, it seems rather to be a complex “syndrome” with various clinical presentations [9, 10].

All previous clinical trials involving drugs have failed to have a significant impact on HFpEF [9]; one reason for this that has been highlighted by different authors in the literature is the vast heterogeneity of the population and the physiopathological mechanisms [10].

Instead of focusing on one target, a new statistical approach known as “clustering” or “phenomapping” has shown its relevance in cardiology in identifying subgroups with similar characteristics. In chronic HF, Ahmad et al. [11] found four specific clusters with specific responses to exercise training and different prognoses. In Japan, three subtypes of acute HF were highlighted,

corresponding to previous physiopathological concepts [12]. In addition, two recent papers used the same statistical technique at rest [13] and during exercise [14] in HFpEF, with consistent findings.

Diagnosis of HF according to the guidelines has been previously validated [15], and many substudies have explored different dimensions and phenotypes [16].

The aims of our study were: (1) to apply a clustering approach to the KaRen study cohort to identify specific subgroups; and (2) to study the physiopathological, clinical and prognostic meaning of the phenotypes.

Methods

KaRen was a prospective multicentre international observational study of patients with HFpEF. The design of the study has been published previously [17], and the primary echocardiographic characteristics have also been reported [18]. The KaRen study included patients with HF symptoms admitted to the emergency ward. The following inclusion criteria were used: (1) acute presentation to the hospital with clinical signs and symptoms of HF, according to the Framingham criteria; (2) B-type natriuretic peptide > 100 pg/mL or N-terminal prohormone of B-type natriuretic peptide > 300 pg/mL; and (3) left ventricular EF > 45% by echocardiography within the first 72 hours. The measurements were performed according to previously established guidelines [6]. All three inclusion criteria were verified within 72 hours of presentation. Enrollment occurred during the initial visit, provided that any exclusion criteria had been ruled out. Information concerning clinical history, co-morbidities, clinical signs and symptoms, standard biology and treatments was collected prospectively during a scheduled hospital visit 4–8 weeks after initiating treatment for acute HF.

Echocardiographic methods and definitions

All patients were followed and reassessed during a visit scheduled 4–8 weeks after the acute visit, including an extensive echocardiographic assessment [19]. Doppler echocardiography was performed using a ViVid 7 echo-platform (GE VingMed, Horten, Norway). The acquisitions were standardized to images of the left and right heart. Thereafter, all examinations were analysed at the core laboratory in Rennes. Each measurement was performed three times and averaged. The echocardiography reader was blinded to the clinical history of any patient [18]. This KaRen substudy included all patients with

analysable trasnthoracic echocardiography at the visit at 4–8 weeks, according to the European Association of Cardiovascular Imaging/American Society of Echocardiography recommendations [20].

Analysis of left ventricular longitudinal deformation was conducted for each patient using the loops from the apical four-, three- and two-chamber views. According to the regional thickness of each segment, the region of interest was adapted to systematically include the endocardial and epicardial borders. The regional adaptation of the size of the region of interest was possible on the EchoPAC version we were using (EchoPAC B13; GE Healthcare, Horten, Norway). The same approach was used for the right ventricle, focusing on the free wall in the apical four-chamber view.

Follow-up

Patients were included between May 2007 and December 2011, and were followed prospectively from the initial hospitalization until November 2012. Vital status was assessed by clinical visit, telephone contact or, in Sweden, by the Swedish National Patient Register and Population Register. [20]

Study endpoints

The primary outcome was all-cause mortality or first hospitalization for HF. Hospitalization for HF was defined as admission to the hospital for any length of time, including day care, with either HF treatment or HF as the primary reason for admission. The secondary outcome was time to all-cause mortality [20]. Because this variable was measured from the visit at 4–8 weeks, the index date and start of follow-up were defined as the date of the visit at 4–8 weeks and data collection.

Statistical analysis

Before analysis, missing data were imputed using the SVDimpute function within the impute package in R. The percentage of missing values for features ranged from 0% to 28% (for LA volume index).

Before the cluster analysis was performed, the Gower dissimilarity measure was chosen to measure the closeness between each observation; it implies standardization, which is set to range for interval and ordinal variables. A hierarchical cluster analysis was conducted in PROC CLUSTER (SAS, version 9.4; SAS Institute, Cary, NC, USA) using two-stage density linkage, specifying five neighbours for k-nearest neighbour density estimation. All clustering was performed blinded to clinical outcome data.

Hierarchical cluster analysis is an attempt to classify data of previously unknown structure into discrete groups. Each observation begins in a cluster by itself. The two closest clusters are merged to form a new cluster that replaces the two old clusters. Merging of the two closest clusters is repeated until only one cluster is left. The various clustering methods differ in how the distance between two clusters is computed. The two-stage density linkage is a modification of density linkage that ensures that all points are assigned to modal clusters before the modal clusters are permitted to join. In the first stage, disjoint modal clusters are formed. Two clusters are joined only if at least one of the two clusters has fewer members than the number specified (in our analysis, five). At the end of the first stage, each point belongs to one modal cluster. In the second stage, the modal clusters are hierarchically joined by single linkage. The final number of clusters can exceed one when there are wide gaps between the clusters or when the smoothing parameter is small.

Once clusters were defined, we compared differences in demographic, clinical and echocardiographic characteristics between clusters, using the χ^2 test (or Fisher's exact test, when appropriate) for categorical variables and analysis of variance (or the Kruskal-Wallis test, when appropriate) for continuous variables. Fifty-five variables were finally included in the analysis (see Table 1 for details).

For outcomes analyses, we used Kaplan-Meier survival estimates to determine the independent association between clusters and outcomes.

Results

The overall KaRen population has been described previously [21]. Cluster analysis included 356 patients with evaluable echocardiography, including 55 variables to identify three clusters. The mean age was 76.1 ± 9.31 years and 56.5% were female; a large majority of patients (79.4%) had hypertension and 45.4% had renal insufficiency. Sinus rhythm was present at enrollment in 56.3%. Subclinical left ventricular dysfunction was observed, with a mid-altered two-dimensional (2D) strain peak (–14.6 ± 3.96%) and moderate LA remodelling, with a mean LA volume index of 49.4 ± 18 mL/m². Distribution of chronic pulmonary obstructive disease, peak of tricuspid regurgitation and E/e' ratio were not significantly different between the clusters.

Patient characteristics overall, and within the different clusters are shown in Table 2.

Cluster 1 (*n* = 128)

This cluster can be summed up as the “cardiovascular risk group” with the highest rates of hypertension (89.1%) and diabetes (60.2%), despite a relatively young age (73.8 ± 10.3 years), and the highest proportion of males (60.2%). Obesity was more prevalent (mean body mass index 31.2 ± 6.55 kg/m^2) and renal inufficiency was more frequent (60.2%) than in the other clusters, with a lower concentration of haemoglobin (118 ± 21 g/L). Fewer rhythm disorders occurred than in the other clusters, with 88.3% in sinus rhythm at enrolment. Concerning echocardiography variables, left ventricular remodelling was predominant (interventricular septal thickness 12.0 ± 2.53 mm), with moderate left ventricular deformation (mean 2D strain peak –14.0 ± 3.47%). Right ventricular function was preserved, with a mean tricuspid annulus plan systolic excursion of 18.8 ± 4.74 mm, and we observed the smallest LA dilatation (45.7 ± 2.4 mL/m^2) and the lowest rate of severe mitral regurgitation (2.3%) in this cluster.

Cluster 2 (*n* = 134)

This cluster had the highest proportion of females (68.7%) and the lowest rates of diabetes (11.9%) and hypertension (73.1%). A sinus rhythm electrocardiogram was less frequent (67.2%) than in cluster 1. Renal function was most preserved in this cluster (only 35.1% had renal insufficiency), and it was the only group with subnormal left ventricular systolic function, with a mean 2D strain peak of –17.3 ± 3.05% in association with a preserved left ventricular EF. The results were concordant with the mean population values for LA dilatation, mitral regurgitation severity and other echocardiographic variables.

Cluster 3 (*n* = 94)

This cluster had the oldest population (mean age 78.3 ± 6.92 years), with a majority of females (61.7%); the rates of diabetes (23.4%) and hypertension (75.5%) were lower than in cluster 1, but higher than in cluster 2. These patients had the highest rate of rhythm disorders, with only 12.8% in sinus rhythm. Mitral regurgitation was more frequent: five times the rate in cluster 1 for class III–IV (10.6% vs 2.3%) and twice the rate in cluster 1 for class II (22.3% vs 10.9%). Left ventricular systolic and diastolic dysfunction were altered (mean 2D strain peak –11.6 ± 2.87%), with severe LA enlargment (LA volume index 54.3 ± 15.7 mL/m^2).

Cluster association and clinical outcomes

Fig. 1, Fig. 2 and Fig. 3 compare outcomes between the different clusters. There were no statistical differences between the three clusters for the primary endpoint.

We found a tendency towards a worse long-term outcome for patients in cluster 2. When death was the only criterion analysed, cluster 3 showed higher rates of short- and mid-term mortality (40% vs 22% and 23% in clusters 1 and 2, respectively, at 500 days of follow-up), but no difference in long-term survival (35% vs 66% and 63% in clusters 1 and 2, respectively, at 1000 days of follow-up).

Discussion

HFpEF: Heterogenous and complex remodelling with fibrosis and energetic impairement

From diastolic dysfunction to the impact of multiple co-morbidities, a new paradigm [22] has emerged, presenting HFpEF as the final expression of a complex interaction between multiorgan dysfunction, neurohormonal activation and haemodynamic stress, linked by the same pathological disorder, systemic inflammation and endothelial dysfunction, resulting in myocardial stiffness and hypertrophy.

Clustering approach

New methods are needed to be more discriminating in future trials and to help to improve the specific diagnosis of patients. Clustering is a new statistical method belonging to the category of non-supervised analysis, without predefined classes. The goal is to sort objects into different subgroups, “the clusters”, and each cluster includes objects with similarities to each other, and dissimilarities to objects outside the subgroup. The hierarchical method is based on progressive step-by-step aggregation of patients, until one specific class is found.

Following the path opened by Shah et al. [13] and Przewlocka-Kosmala et al. [14], a cluster analysis technique was applied to the 356 patients with HFpEF from the KaRen cohort in order to individualize phenotypic groups with specific patterns of co-morbidites. Three clusters, summarized in Table 3, were discovered, based on multiple clinical and paraclinical variables, selected for their “physiological” meaning and the absence of high statistical redundance between each other.

Metabolic syndrome and young HFpEF

Cluster 1 included the youngest, mostly male patients from the population, with a high rate of cardiovascular co-morbidities. These findings resonate with previous literature. In the RELAX trial [23], diabetes, hypertension, obesity and renal dysfunction were predominant in young men. In the study by Zacharias et al. [24], younger patients hospitalized for acute HF were more likely to be male, obese and have a history of diabetes and chronic kidney disease than older patients with HFpEF. This specific phenotype may be related to a well-known clinical entity, the “metabolic syndrome” mediated by a pathological insulin cycle. In synergy with radical oxygen species, insulin resistance enhances diastolic dysfunction, leading to HF [25]. This hormonal disorder stands at the crossroads of many co-morbidities, as both a cause and a consequence of them, thus increasing the development of HFpEF functional impairment, but also its risk factors. Renal dysfunction seems to be part of another “vicious cycle” with HFpEF and the other cardiovascular risk factors

Cluster 1 patients may benefit from adapted therapies. Aerobic training and calorie restriction may help with recovery of insulin sensitivity, loss of weight and restoration of an efficient energy cycle. Kitzman et al. found an significant improvement in peak oxygen consumption (VO_2) in patients with obesity and HFpEF [26]. From another perspective, sodium-glucose cotransporter inhibitors effectively reduced HF hospitalization and cardiovascular death in patients with non-insulin-dependent diabetes. Interference with calcium homeostasis, improvement in mitochondrial function, decrease in inflammation and the production of reactive oxygen species and restoration of nitric oxide formation represent some of the mechanisms behind these effects, and each of them is involved in HFpEF development [27].

Mitral regurgitation: An atrial remodelling endpoint

Cluster 3 included the oldest patients, with the worst left ventricular systolic and diastolic dysfunction, the highest degree of LA remodelling and moderate-to-severe mitral regurgitation, and with fewer patients in sinus rhythm, because of the loss of atrioventricular coupling. A continuous process of inflammation, microvascular and diastolic dysfunction, increasing left ventricular filling pressure and LA stress means that the patient with HFpEF is more dependent on LA contraction than control subjects [7]. Mediated by LA fibrosis, LA dilatation and altered LA reservoir function lead to high filling pressure and AF. HFpEF and AF are known to influence each other’s development and persistence,

and to worsen LA mechanics, thus leading to altered geometry and dynamic function of the mitral annulus and leaflets [28], causing functional mitral regurgitation. More than a risk factor, mitral regurgitation may represent the severity of an evolved HFpEF form. Kajimoto et al. [29] found a prognostic association between mitral regurgitation, even at a mild grade, and outcomes for HFpEF, whereas this prognostic association is observed for moderate-to-severe mitral regurgitation in HFrEF. This sequence of LA remodelling-AF-functional mitral regurgitation in cluster 3 may explain the tendency towards short- and mid-term mortality, whereas there is no significant difference in long-term outcome; it can be linked to the older age of this subgroup, but also to a cumulative impact of LA size or dysfunction, AF and mitral regurgitation [29]. A therapeutic strategy may need to take into account the abnormal mitral subvalvular and valvular system, with the objectives of reducing congestion and treating risk factors, or invasive treatment may be required, despite controversial data on the subject.

Cluster 2 seems to a be a “previous” stage of cluster 3, with younger patients, less diabetes and lower rates of mitral regurgitation and arrhythmia. Global preserved left ventricular function may be a marker of less inflammation and better systolic-diastolic dynamics, slowing evolution of LA dysfunction. This clinical phenotype can be similar to “exercise intolerance” previously presented by Shah [10], and may benefit from a sport practice programme. Sacubitril-valsartan may represent another effective therapy, by reducing left ventricular filling pressure and LA stress, restoring fluid balance and vasodilation function or improving ventricular-arterial coupling.

Clusters in the literature

Table 4 presents some of the main studies using clustering analysis in HF. No specific phenotype seems to perfectly fit every study, but some similarities can be noted. Cluster 3 in the studies by Shah et al. [13] and Horiuchi et al. [12] and our cluster 3 share the same high level of cardiac remodelling and AF in older patients. Obesity and diabetes coexist in each study, supporting the insulin resistance concept. The main difference concerns the youngest part of the population, with a predominance of men in our study, and of women in the phenomapping study. The relative homogeneity of the E/e’ ratio, tricupsid regurgitation and right ventricular function in our analysis may be explained the absence of statistical prognostic value in the study.

Another explanation may be the harmonious management of patients in the KaRen study, without any specific intervention adapted to a particular phenotype. Indeed, in another trial, Ahmad et al. [11]

found a significant prognostic difference between four clusters extracted from a HFrEF cohort, when an exercise training programme was applied to the population. Testing a specific intervention in different phenotypes may highlight positive results in terms of morbimortality for some subgroups previously defined.

Study limitations

A complete and analysable echocardiographic recording was unfortunately only available for 356/539 patients (66%). No significant prognostic difference was found between clusters. At first glance, the conclusion appears to be only negative, with prognostic heterogeneity still to be proven. However, from another point of view, this statement reinforces our misunderstanding of this pathology and the lack of effective treatments.

Conclusions

HFpEF is a heteregenous syndrome, with complex entanglement between multiple co-morbidites and cellular and molecular remodelling, all sharing the same initial phenomenon – systemic inflammation. The lack of effective therapy influences current and perhaps future trials, targeting the approach of “selected therapy” for “specific phenotype”.

Cluster analysis applied to a well-known HFpEF cohort seems to be effective at individualizing new phenotypes with different characteristics, emphasizing a common physiopathology. Further studies are needed to prove the clinical relevance of these results, and their implications for the therapeutic strategy in patients.

Sources of funding

None.

Disclosure of interest

The authors declare that they have no conflicts of interest concerning this article.

References

[1] Owan TE, Hodge DO, Herges RM, Jacobsen SJ, Roger VL, Redfield MM. Trends in prevalence and outcome of heart failure with preserved ejection fraction. N Engl J Med 2006;355:251-9.

[2] Shah KS, Xu H, Matsouaka RA, et al. Heart Failure With Preserved, Borderline, and Reduced Ejection Fraction: 5-Year Outcomes. J Am Coll Cardiol 2017;70:2476-86.

[3] Dunlay SM, Roger VL, Redfield MM. Epidemiology of heart failure with preserved ejection fraction. Nat Rev Cardiol 2017;14:591-602.

[4] Vaduganathan M, Patel RB, Michel A, et al. Mode of Death in Heart Failure With Preserved Ejection Fraction. J Am Coll Cardiol 2017;69:556-69.

[5] Paulus WJ, Tschope C, Sanderson JE, et al. How to diagnose diastolic heart failure: a consensus statement on the diagnosis of heart failure with normal left ventricular ejection fraction by the Heart Failure and Echocardiography Associations of the European Society of Cardiology. Eur Heart J 2007;28:2539-50.

[6] Ponikowski P, Voors AA, Anker SD, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur Heart J 2016;37:2129-200.

[7] Borlaug BA. The pathophysiology of heart failure with preserved ejection fraction. Nat Rev Cardiol 2014;11:507-15.

[8] Franssen C, Chen S, Unger A, et al. Myocardial Microvascular Inflammatory Endothelial Activation in Heart Failure With Preserved Ejection Fraction. JACC Heart Fail 2016;4:312-24.

[9] Senni M, Paulus WJ, Gavazzi A, et al. New strategies for heart failure with preserved ejection fraction: the importance of targeted therapies for heart failure phenotypes. Eur Heart J 2014;35:2797-815.

[10] Shah SJ. Matchmaking for the optimization of clinical trials of heart failure with preserved ejection fraction: no laughing matter. J Am Coll Cardiol 2013;62:1339-42.

[11] Ahmad T, Pencina MJ, Schulte PJ, et al. Clinical implications of chronic heart failure phenotypes defined by cluster analysis. J Am Coll Cardiol 2014;64:1765-74.

[12] Horiuchi Y, Tanimoto S, Latif A, et al. Identifying novel phenotypes of acute heart failure using cluster analysis of clinical variables. Int J Cardiol 2018;262:57-63.

[13] Shah SJ, Katz DH, Selvaraj S, et al. Phenomapping for novel classification of heart failure with preserved ejection fraction. Circulation 2015;131:269-79.

[14] Przewlocka-Kosmala M, Marwick TH, Dabrowski A, Kosmala W. Contribution of Cardiovascular Reserve to Prognostic Categories of Heart Failure With Preserved Ejection Fraction: A Classification Based on Machine Learning. J Am Soc Echocardiogr 2019;32:604-15 e6.

[15] Persson H, Donal E, Lund LH, et al. Importance of structural heart disease and diastolic dysfunction in heart failure with preserved ejection fraction assessed according to the ESC guidelines - A substudy in the Ka (Karolinska) Ren (Rennes) study. Int J Cardiol 2019;274:202-7.

[16] Lund LH, Donal E, Oger E, et al. Association between cardiovascular vs. non-cardiovascular co-morbidities and outcomes in heart failure with preserved ejection fraction. Eur J Heart Fail 2014;16:992-1001.

[17] Donal E, Lund LH, Linde C, et al. Rationale and design of the Karolinska-Rennes (KaRen) prospective study of dyssynchrony in heart failure with preserved ejection fraction. Eur J Heart Fail 2009;11:198-204.

[18] Donal E, Lund LH, Oger E, et al. Value of exercise echocardiography in heart failure with preserved ejection fraction: a substudy from the KaRen study. Eur Heart J Cardiovasc Imaging 2016;17:106-13.

[19] Cosyns B, Garbi M, Separovic J, Pasquet A, Lancellotti P, Education Committee of the European Association of Cardiovascular Imaging. Update of the echocardiography core syllabus of the European Association of Cardiovascular Imaging (EACVI). Eur Heart J Cardiovasc Imaging 2013;14:837-9.

[20] Nagueh SF, Smiseth OA, Appleton CP, et al. Recommendations for the Evaluation of Left Ventricular Diastolic Function by Echocardiography: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. J Am Soc Echocardiogr 2016;29:277-314.

[21] Donal E, Lund LH, Oger E, et al. Baseline characteristics of patients with heart failure and preserved ejection fraction included in the Karolinska Rennes (KaRen) study. Arch Cardiovasc Dis 2014;107:112-21.

[22] Paulus WJ, Tschope C. A novel paradigm for heart failure with preserved ejection fraction: comorbidities drive myocardial dysfunction and remodeling through coronary microvascular endothelial inflammation. J Am Coll Cardiol 2013;62:263-71.

[23] Lewis GA, Schelbert EB, Williams SG, et al. Biological Phenotypes of Heart Failure With Preserved Ejection Fraction. J Am Coll Cardiol 2017;70:2186-200.

[24] Zacharias M, Joffe S, Konadu E, et al. Clinical epidemiology of heart failure with preserved ejection fraction (HFpEF) in comparatively young hospitalized patients. Int J Cardiol 2016;202:918-21.

[25] Fontes-Carvalho R, Ladeiras-Lopes R, Bettencourt P, Leite-Moreira A, Azevedo A. Diastolic dysfunction in the diabetic continuum: association with insulin resistance, metabolic syndrome and type 2 diabetes. Cardiovasc Diabetol 2015;14:4.

[26] Kitzman DW, Brubaker P, Morgan T, et al. Effect of Caloric Restriction or Aerobic Exercise Training on Peak Oxygen Consumption and Quality of Life in Obese Older Patients With Heart Failure With Preserved Ejection Fraction: A Randomized Clinical Trial. JAMA 2016;315:36-46.

[27] Uthman L, Baartscheer A, Schumacher CA, et al. Direct Cardiac Actions of Sodium Glucose Cotransporter 2 Inhibitors Target Pathogenic Mechanisms Underlying Heart Failure in Diabetic Patients. Front Physiol 2018;9:1575.

[28] Pai RG, Varadarajan P, Tanimoto M. Effect of atrial fibrillation on the dynamics of mitral annular area. J Heart Valve Dis 2003;12:31-7.

[29] Kajimoto K, Sato N, Takano T, Investigators of the Acute Decompensated Heart Failure Syndromes Registry. Functional mitral regurgitation at discharge and outcomes in patients hospitalized for acute decompensated heart failure with a preserved or reduced ejection fraction. Eur J Heart Fail 2016;18:1051-9.

Figure legends

Figure 1. Stacked bar chart of outcomes by clusters. CV: cardiovascular.

Figure 2. Kaplan-Meier survival free of cardiovascular (CV) hospitalization or death.

Figure 3. Kaplan-Meier survival curve.

Table 1 Variables that served as features for the clustering analysis, including clinical variables, physical characteristics, laboratory data and echocardiographic variables.

Domain	Variables
Demographics	Age, sex
Physical characteristics	Body mass index, sinus rhythm
Medical history	Renal insufficiency, COPD, arterial hypertension, diabetes mellitus
Laboratory	Haemoglobin
Echocardiography	
Left heart structure	LV end-diastolic volume, LV end-systolic volume, LV end-diastolic diameter, LV end-systolic diameter, septal wall thickness, LV mass, LV fractional shortening, LA volume, LA diameter
LV systolic function	LVEF
	LV pre-ejection time interval, LV ejection time duration
LV diastolic function	Mitral inflow characteristics (E-wave deceleration time, end-diastolic mitral inflow A -wave, E/A ratio, IVRT, mitral inflow velocity time integral, mitral inflow duration/cycle length)
	Tissue Doppler characteristics (e′ velocities, LV septal e' velocity, E/e′ ratio)
	E-wave/LV end-diastolic peak
Right heart structure	RV diameter
RV function	RV fractional area change, TAPSE, RV shortening

	RV pre-ejection time interval, interventricular time delay, RV ejection time duration
Haemodynamics	Stroke volume, velocity time integral of the flow in the LVOT, cardiac index
	Mitral regurgitation
	Aortic valve calcifications, mean pressure gradient across aortic valve, maximal aortic gradient, aortic valve effective area, aortic valve regurgitation
	Tricuspid regurgitation, systolic peak velocity at the tricuspid annulus
	Septo-lateral delay DTI
	Time between QRS and aortic valve closure
	2D strain, peak of LV deformation in the basal part of the septum
	2D strain, mean of the peak of LV deformation in apical four-chamber view
	2D strain, mean of the peak of LV deformation in the 16 LV segments
	LV systolic peak in strain rate
	LV end-diastolic peak in strain rate

2D: two-dimensional; COPD: chronic obstructive pulmonary disease; DTI: Doppler tissue imaging; IVRT: isovolumic relaxation time; LV: left ventricular; LVEF: left ventricular ejection fraction; LVOT: left ventricular outflow tract; RV: right ventricular; TAPSE: tricuspid annular plane systolic excursion.

Table 2 Patient characteristics across clusters.

Characteristics	Missing values	Total (n = 356)	Cluster 1 (n = 128)	Cluster 2 (n = 134)	Cluster 3 (n = 94)	P
Male sex	0	155 (43.5)	77 (60.2)	42 (31.3)	36 (38.3)	< 0.0001
Age (years)	0	76.1 ± 9.31	73.8 ± 10.3	76.7 ± 9.36	78.3 ± 6.92	0.0010
Body mass index (kg/m²)	17 (4.77)	28.5 ± 6.48	31.2 ± 6.55	28.6 ± 6.17	28.5 ± 5.81	0.0009
Haemoglobin (g/L)	8 (2.25)	123 ± 20	118 ± 21	125 ± 17	126 ± 18	0.0048
Renal insufficiency	6 (1.68)	159 (45.4)	77 (60.2)	47 (35.1)	35 (37.2)	< 0.0001
Sinus rhythm at enrolment	33 (9.27)	182 (56.3)	113 (88.3)	90 (67.2)	12 (12.8)	< 0.0001
COPD	0	45 (12.6)	14 (10.9)	19 (14.2)	12 (12.8)	0.73
Hypertension	1 (0.28)	282 (79.4)	114 (89.1)	98 (73.1)	71 (75.5)	0.0033
Diabetes mellitus	0	115 (32.3)	77 (60.2)	16 (11.9)	22 (23.4)	< 0.0001
Mitral regurgitation	16 (4.49)					0.0002
0		74 (21.8)	42 (32.8)	22 (16.4)	10 (10.6)	
1		186 (54.7)	69 (53.9)	80 (59.7)	53 (56.4)	
2		59 (17.3)	14 (10.9)	24 (17.9)	21 (22.3)	
3–4		21 (6.18)	3 (2.3)	8 (6.0)	10 (10.6)	
RV pre-ejection time interval (ms)	28 (7.86)	84.2 ± 24.1	80.1 ± 24.3	82.4 ± 22.5	92.3 ± 20.5	0.0003

LV pre-ejection time interval (ms)	2.25 (8)	83.1 ± 27.2	79.2 ± 26.8	77.1 ± 24.4	96.9 ± 25.7	< 0.0001
TAPSE (mm)	1.97 (7)	17.2 ± 4.74	18.8 ± 4.74	17.9 ± 4.35	14.1 ± 3.55	< 0.0001
IVST (mm)	4.21 (15)	11.6 ± 2.25	12.0 ± 2.53	11.3 ± 1.85	11.3 ± 2.08	0.0109
Tricuspid regurgitation velocity (m/s)	13.8 (49)	2.87 ± 0.64	2.79 ± 0.60	2.92 ± 0.63	2.92 ± 0.51	0.13
LA volume index (mL/m^2)	28.4 (101)	49.4 ± 18.0	45.7 ± 12.4	49.5 ± 16.3	54.3 ± 15.7	0.0001
E/e' ratio	6.74 (24)	12.9 ± 6.00	13.5 ± 5.58	12.8 ± 6.05	12.0 ± 5.67	0.17
E-wave/LV end-diastolic peak ratio	10.9 (39)	5.10 ± 11.6	3.09 ± 4.36	4.07 ± 6.78	9.31 ± 18.4	< 0.0001
LA diameter (mm)	10.1 (36)	45.5 ± 6.66	45.0 ± 5.82	44.7 ± 6.38	47.4 ± 6.52	0.0027
LV deformation[a]/apical four-chamber view (%)	2.25 (8)	–14.6 ± 3.96	–14.0 ± 3.47	–17.3 ± 3.05	–11.6 ± 2.87	< 0.0001

Data are expressed as number (%) or mean ± standard deviation. COPD: chronic obstructive pulmonary disease; IVST: interventricular septal thickness; LA: left atrial; LV: left ventricular; RV: right ventricular; TAPSE: tricuspid annulus plan systolic excursion.

[a] Two-dimensional strain, mean peak.

Table 3 Main features of the clusters.

Cluster 1	Young, male, history of hypertension, diabetes mellitus, impaired renal function, obesity, sinusal rhythm, anaemia
Cluster 2	Female, preserved LV systolic function, less renal dysfunction, moderate atrial remodelling
Cluster 3	Elderly, female, rhythm disorders, midly impaired LV systolic function (GLS), atrial remodelling, mitral regurgitation

GLS : global longitudinal strain; LV: left ventricular.

Table 4 Clustering in heart failure research.

Study	Pathology	Cluster 1	Cluster 1	Cluster 1	Cluster 1	Prognosis
Shah et al. [13] (*n* = 397)	Chronic HFpEF	Young, female, low BNP, low cardiac remodelling/dysfunction	Obesity, diabetes mellitus, OSA, worst relaxation, high pulmonary pressure	Oldest, CKD, AF, high cardiac remodelling (mass, LA dilatation, E/e'), worst RV function	-	Cluster 1 > 2 > 3
Horiuchi et al. [12] (*n* = 345)	Acute HF	"Vascular": hypertension, diabetes mellitus, mildly impaired systolic function, high LV mass	"Cardiac": male, CKD, BNP, impaired diastolic and systolic function	"HFpEF": oldest, anaemia, AF, HFpEF or HFmrEF	-	Cluster 1 > 2 and 3
Ahmad et al. [11] (*n* = 1619)	Chronic HFrEF	Eldest, caucasian, ICM, co-morbidites	Young, obese, less ICM, low BNP, African, hospitalization	Caucasian, ICM, angina, hospitalization	Caucasian, female, non-ICM, low rates of co-morbidities	Death, ++ cluster 1; hospitalization, +++ clusters 1 & 3
Schrub et al. (*n* = 356)	Chronic HFpEF	Young, male, obesity, hypertension, CKD, anaemia, LV mass, low atrial remodelling	Female, preserved LV function, moderate atrial remodelling	Eldest, female, AF, midly impaired systolic function, atrial remodelling, mitral regurgitation	-	No statistical difference

AF: atrial fibrillation; BNP: B-type natriuretic peptide; CKD: chronic kidney disease; HF: heart failure; HFmrEF: heart failure with mid-range ejection fraction; HFpEF: heart

failure with preserved ejection fraction; HFrEF: heart failure with reduced ejection fraction; ICM: ischaemic heart disease; LA: left atrial; LV: left ventricular; OSA: obstructive sleep apnoea; RV: right ventricular.

Figure 1

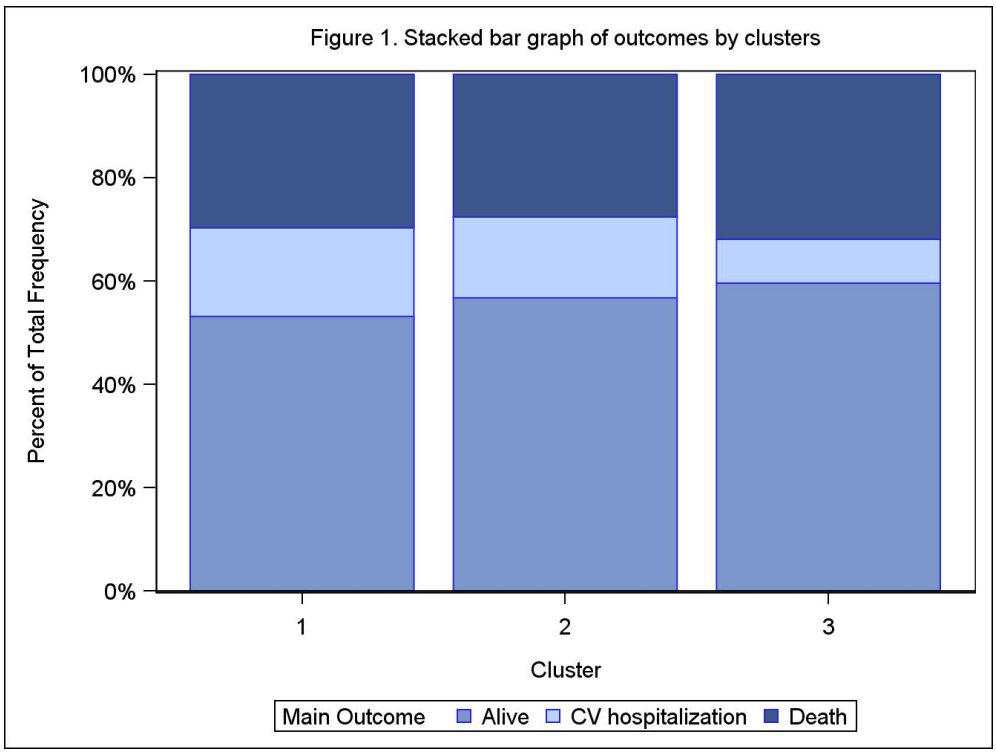


Figure 1. Stacked bar graph of outcomes by clusters
Percent of Total Frequency
100%
80%
60%
40%
20%
0%
1
2
3
Cluster
Main Outcome
Alive
CV hospitalization
Death

Figure 2

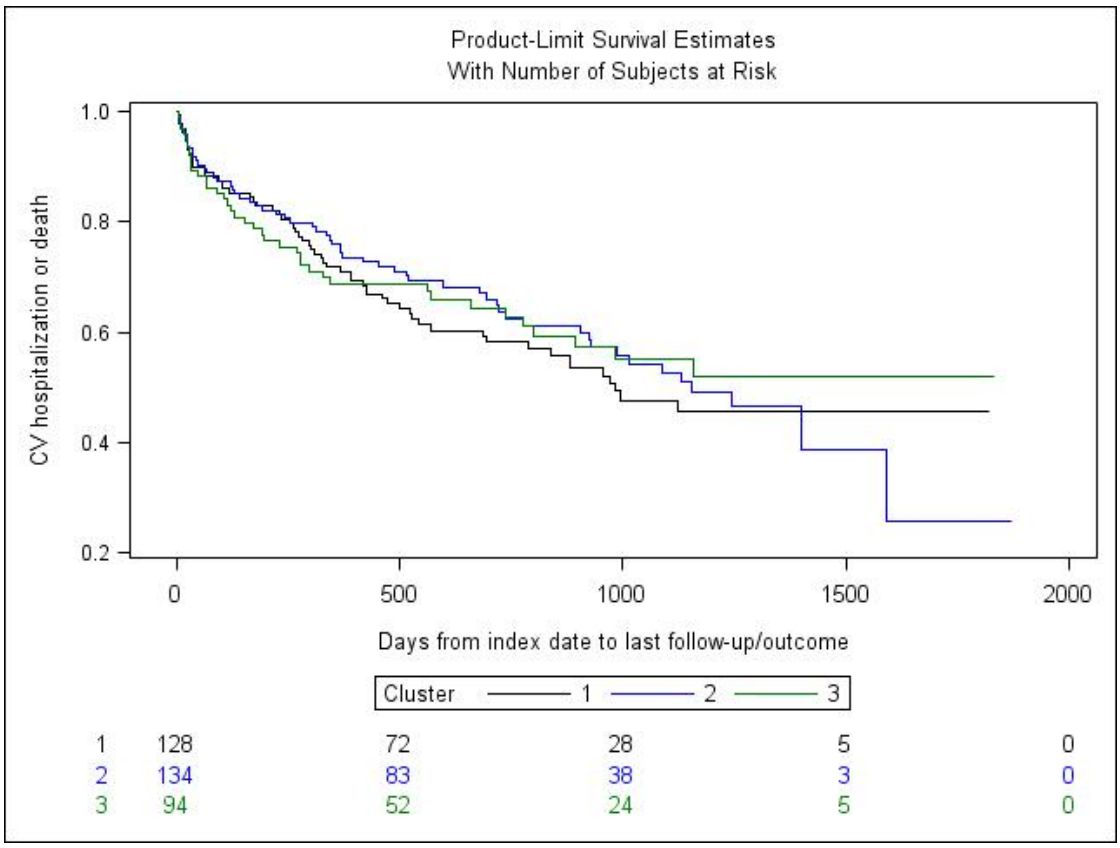
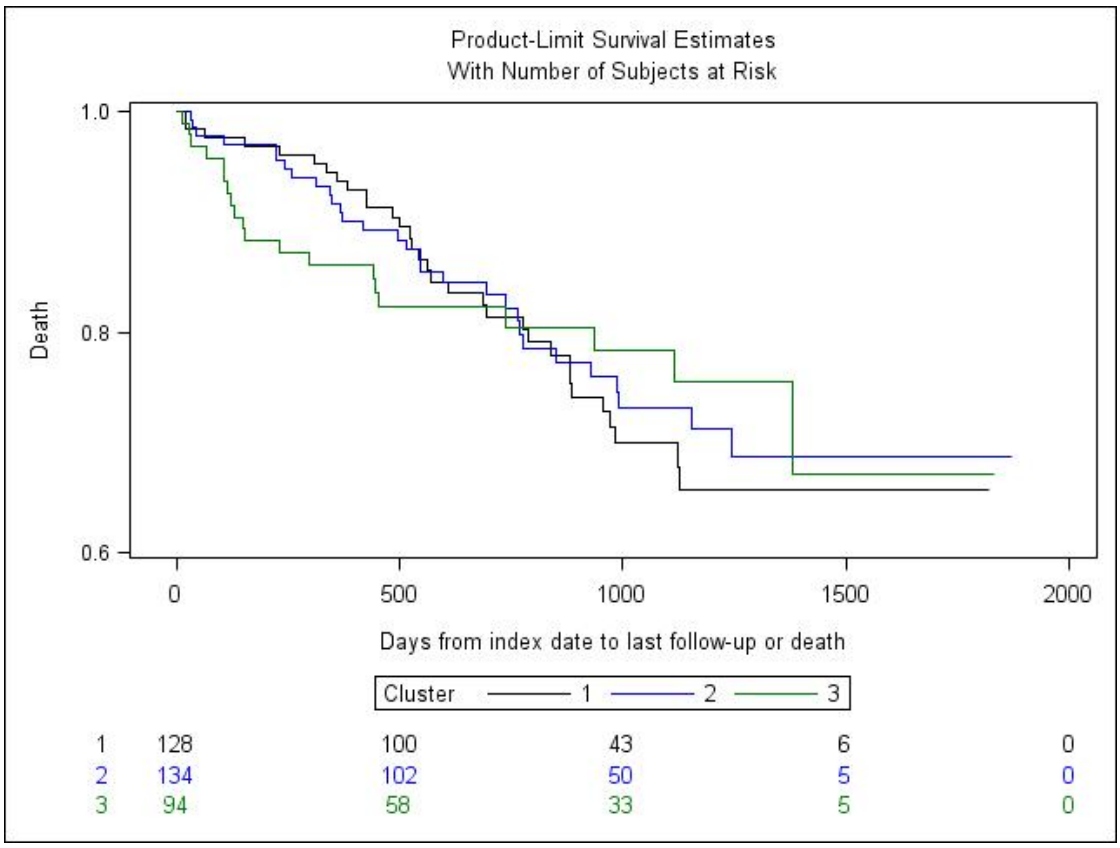

Product-Limit Survival Estimates
With Number of Subjects at Risk
CV hospitalization or death
1.0
0.8
0.6
0.4
0.2
0
500
1000
1500
2000
Days from index date to last follow-up/outcome
Cluster 1 2 3
1 128 72 28 5 0
2 134 83 38 3 0
3 94 52 24 5 0

Figure 3


Product-Limit Survival Estimates
With Number of Subjects at Risk
Death
1.0
0.8
0.6
0
500
1000
1500
2000
Days from index date to last follow-up or death
Cluster 1 2 3
1 128 100 43 6 0
2 134 102 50 5 0
3 94 58 33 5 0