ORIGINAL RESEARCH

Identification of novel pheno-groups in heart failure with preserved ejection fraction using machine learning

Åsa K Hedman [1,2] Camilla Hage,[3,4] Anil Sharma,[2] Mary Julia Brosnan,[5] Leonard Buckbinder,[5] Li-Ming Gan,[6,7] Sanjiv J Shah,[8] Cecilia M Linde,[3,4] Erwan Donal [9] Jean-Claude Daubert,[10] Anders Mälarstig [1,2] Daniel Ziemek,[11] Lars Lund[3,4]

► http://dx.doi.org/10.1136/heartjnl-2019-316030

Check for updates

ABSTRACT

Objective Heart failure with preserved ejection fraction (HFpEF) is a heterogeneous syndrome. We aimed to derive HFpEF phenotype-based groups ('phenogroups') based on clinical and echocardiogram data using machine learning, and to compare clinical characteristics, proteomics and outcomes across the phenogroups.

Methods We applied model-based clustering to 32 echocardiogram and 11 clinical and laboratory variables collected in stable condition from 320 HFpEF outpatients in the Karolinska-Rennes cohort study (56% female, median 78 years (IQR: 71–83)). Baseline proteomics and the composite end point of all-cause mortality or heart failure (HF) hospitalisation were used in secondary analyses.

Results We identified six phenogroups, for which significant differences in the prevalence of concomitant atrial fibrillation (AF), anaemia and kidney disease were observed (p<0.05). Fifteen out of 86 plasma proteins differed between phenogroups (false discovery rate, FDR<0.05), including biomarkers of HF, AF and kidney function. The composite end point was significantly different between phenogroups (log-rank p<0.001), at short-term (100 days), mid-term (18 months) and longer-term follow-up (1000 days). Phenogroup 2 was older, with poorer diastolic and right ventricular function and higher burden of risk factors as AF (85%), hypertension (83%) and chronic obstructive pulmonary disease (30%). In this group a third experienced the primary outcome to 100 days, and two-thirds to 18 months (HR (95% CI) versus phenogroups 1, 3, 4, 5, 6: 1.5 (0.8–2.9); 5.7 (2.6–12.8); 2.9 (1.5–5.6); 2.7 (1.6–4.6); 2.1 (1.2–3.9)).

Conclusions Using machine learning we identified distinct HFpEF phenogroups with differential characteristics and outcomes, as well as differential levels of inflammatory and cardiovascular proteins.

INTRODUCTION

Heart failure (HF) is a growing public health pandemic, with HF with preserved ejection fraction (HFpEF, defined as left ventricular ejection fraction (LVEF) ≥45%–50%[1]) constituting up to half of the HF population and increasing.[2] The HFpEF syndrome is multifactorial and heterogeneous with regards to clinical characteristics and prognosis, and the aetiology and pathophysiology of the disease are still poorly understood.[3] In addition, the heterogeneity of the disease poses significant challenges to drug development and clinical trial design and HFpEF trials to date have failed.[4–6]

In HFpEF, both cardiovascular and non-cardiovascular comorbidities are common and thought to contribute to disease progression and adverse outcomes.[7–9] One model posits that age and comorbidities may contribute to systemic inflammation, resulting in coronary microvascular endothelial dysfunction and inducing cardiomyocyte and interstitial changes and impaired relaxation and increased stiffness.[10 11]

Unsupervised clustering for the identification of patient subpopulations have been successfully applied in other diseases such as cancer[12 13] and diabetes.[14] While clinical translation to date has been limited, tests for risk stratification in breast cancer (using the PAM50 panel) are the results of such efforts. In HF, two previous studies suggest that clustering of patients with HFpEF based on clinical data and biochemical data can reveal subgroups of patients at greater risk of clinical end points.[15 16] These data warrant replication in European HF populations. In addition, previous studies have not investigated biomarkers across patient clusters beyond traditional routine chemistry. Therefore, we aimed to (1) Cluster patients with HFpEF in the multicentre Karolinska-Rennes (KaRen) Study based on a wide range of clinical information including clinical, laboratory and echocardiogram (echo) characteristics using unsupervised machine learning approaches. (2) Identify proteomic correlates of the phenotype-based groups (from now on referred to as phenogroups). (3) Assess associations between phenogroups and outcomes. (4) Optimise a model for phenotyping of new patients.

METHODS
Study participants and design

The KaRen Study, a prospective observational multicentre cohort study, and baseline characteristics of patients have previously been described.[17] The study conformed to the Declaration of Helsinki, and all patients provided written informed consent. Patients and the public were not involved in the

design, conduct, reporting or dissemination of this research. The study enrolled patients with HFpEF (n=539) with acute HF, LVEF ≥45% and elevated natriuretic peptides (NT-proBNP >300 ng/L or BNP >100 ng/L). At outpatient follow-up in stable state (4–8 weeks after hospitalisation), 98 clinical parameters (routine clinical and laboratory variables and detailed echo (n=320)) and cardiovascular plasma proteins (available in the Karen proteomics substudy,[18] n=76) were assessed. Echo data were collected as previously described (details in online supplementary methods).[19] The measurement of proteins has previously been described.[18] Briefly, a total of 92 plasma proteins were measured using the Olink ProSeek Cardiovascular I (CVD I) panel at Olink Proteomics (Uppsala, Sweden). Following quality control 86 proteins remained for analyses (details in online supplementary methods). Patients were followed for the composite end point of time to all-cause mortality and HF hospitalisation. Meta-analysis Global Group in Chronic Heart Failure (MAGGIC) Risk Score of patients were constructed as in Pocock et al, 2013.[20]

Phenotype-based clustering of patients with HFpEF

To identify HFpEF phenogroups we applied clustering to data from patients in the KaRen Study. Quantitative biomarkers (echo, routine clinical and laboratory variables; online supplementary table S1) remaining after preprocessing steps to remove variables with high missingness (≥30%) and highly correlated variables (details in online supplementary methods) were included in clustering. Model-based clustering of standardised variables (mean=0, SD=1) was performed in R using the *Mclust* function in the mclust package,[21] with default settings, and the optimal model and number of clusters determined by the maximum Bayesian Information Criteria. Comorbidities and additional variables were used to characterise the resulting phenogroups (online supplementary tables S3-6 and visualised in figure 1, online supplementary figures S1-3). This approach is analogous to previous work by Shah et al[15] (further information in online supplementary methods). To evaluate if detailed echo can provide fine-mapping of patients beyond routinely collected clinical and laboratory parameters and to provide a more balanced stratification of patients, we clustered patients separately on (1) Clinical and laboratory variables. (2) Echo variables, from which composite phenogroups were constructed.

Associations of clinical, laboratory, echo and proteomic parameters with pheno-groups

Differences in patient characteristics between the phenogroups were tested using analysis of variance (F-test) for continuous traits and Fisher's exact tests[22] for categorical traits. For plasma proteins, models were additionally adjusted for age, sex and mean protein level per individual, to control for protein differences related to these rather than phenogroup, and significance was determined using FDR (further details in online supplementary methods).

Associations with the composite outcome

Cox proportional hazards models were fitted to investigate the association between phenogroups and time-to-event (composite end point), using the *coxph* function in the R package survival. Kaplan-Meier curves were constructed using the survminer package in R. We examined differential outcome at an early (within 100 days), mid-range (18 months) and long-term (1000 days) follow-up time point.

Assigning patients to phenogroups

Classifiers were built on the standardised clinical variables used in clustering for assigning patients with HFpEF to phenogroups, evaluating three multinomial classification methods (Elastic Net, Neural Networks and Naive Bayes, in the R package caret[23]). A nested cross-validation procedure (repeated five times) was used to avoid overfitting, in which four-fifths of the samples ('training set') were used in optimisation of model parameters (using 10-fold cross-validation), and a fifth of the samples ('test set') were used in assessing prediction performance (further details in online supplementary methods). Prediction performance was assessed as the ability of the model to predict samples in the test set to the correct phenogroup (by multiclass AUC (Area under the ROC Curve)). In the final model (Elastic Net) optimisation of model parameters was performed on all data. Furthermore, we validated the performance of the complete procedure: clustering patients to derive phenogroups; training an Elastic Net model for assigning new patients to phenogroups; and assessing the difference in the composite outcome between the newly assigned phenogroups (further details in online supplementary methods).

RESULTS

Phenotype-based clustering of patients with HFpEF in the Karen Study

We identified three clusters based on 32 echo variables and two clusters based on 11 laboratory and routine clinical variables, respectively (further details in online supplementary results), from which six composite phenogroups were derived. We found distinct differences in comorbidities (with significant difference in prevalence of atrial fibrillation (AF), chronic kidney disease (CKD) and anaemia (p<0.001)), clinical and laboratory variables, and cardiac structure across phenogroups, and common features shared across all phenogroups (online supplementary tables S3-6 and visualised in figure 1, online supplementary figures S1-3). The six phenogroups revealed some distinct and some inconsistent patterns that may be generalised as in table 1, which also shows the corresponding event rates (more details in the outcomes section).

Two phenogroups (5 and 6) were consistent with the older female patient with HFpEF common in clinical practise, where one had more hypertension (80%) and AF (96%) and lower body mass index (BMI) (mean=27±5), and the other more coronary artery disease (CAD, 40%) and left-sided functional changes (online supplementary table S6). Phenogroup 2 had distinctly more severe HF, with the greatest degree of diastolic dysfunction (at least 30% of patients with a grade II or higher) and the worst right ventricular function, and the highest event rates. Although phenogroup 1 was younger, it also had more cardiovascular risk factors; more left-sided changes and had progressed to CKD (67%), and had the second highest event rates. Phenogroup 4 was male but otherwise similar to the female groups 5 and 6, with hypertension (75%), left atrial enlargement and AF (90%). Furthermore, this group had the largest proportion of pacemakers (25%) and previous myocardial infarction (MI, 21%), indicating ischaemic aetiology of HF. Finally, phenogroup 3 had mild HF and low risk, and HF symptoms may be explained by obesity (48%) and deconditioning, and may be similar to the types of patients included in HFpEF trials, that is, younger, milder HF (or even absent HF) and distinctly lower event rates, which may also explain the failure to demonstrate a treatment effect in these trials. Some unexpected or inconsistent findings include the relatively high LVEF (mean=63%) in group 1 given
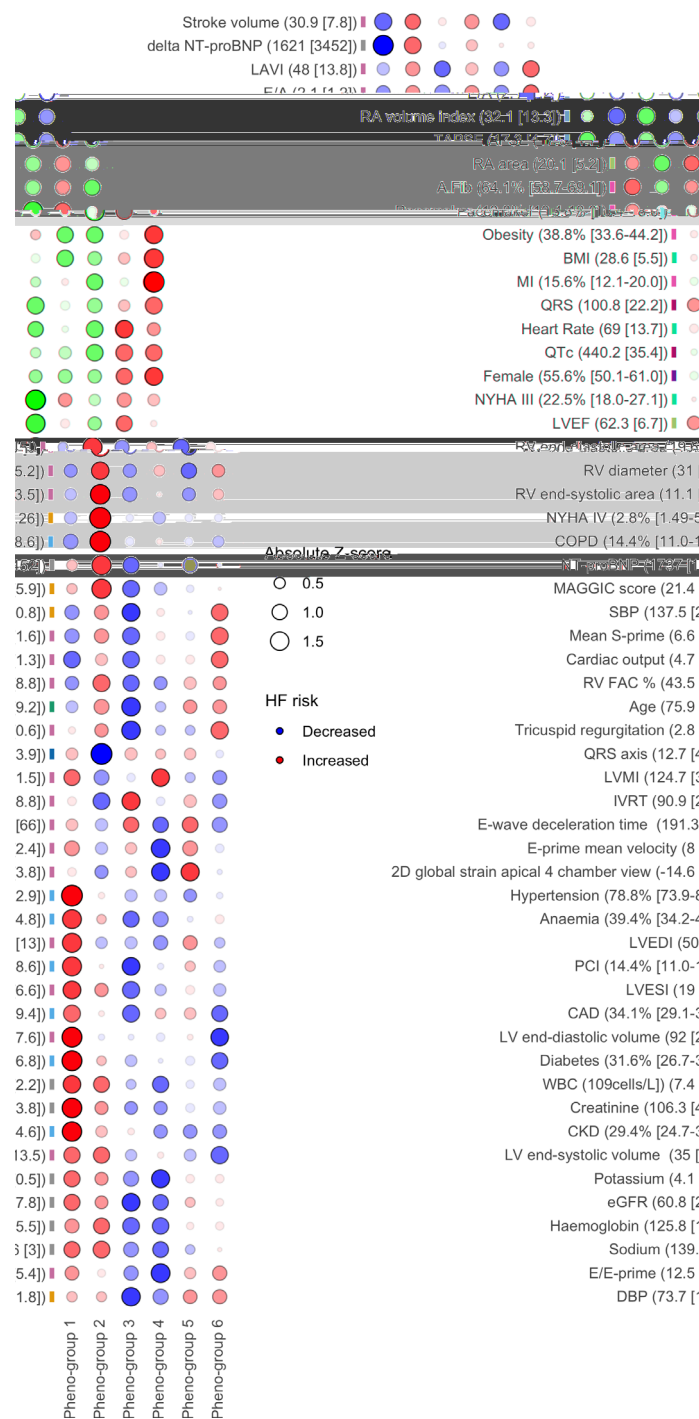
Figure 1 Visualisation of descriptive characteristics of patients in phenogroups. Phenotypes contained within online supplementary tables S3-4 are listed on y-axis and the six phenogroups are listed on the x-axis visualised in the plot. Pre z-scores for each phenotype based on either phenogroup means (for continuous variables) or proportions (for categorical variables). Bullets are sized by the absolute z-score, and coloured according to a priori knowledge on HF with red (increase) representing more severe HF and/or worse prognosis according to a priori knowledge, and blue (decrease) representing the opposite, with α levels reflecting the absolute deviation from the mean, to highlight stronger deviations. For clarity, where both low and high values may lead to increase in HF risk (eg, BMI) we have opted to only indicate higher values in red in the figure. For each variable label the population mean and SD or the percentage with 95% CIs are included as a reference. Phenotypes are ordered based on hierarchical clustering, to better display patterns relating to phenogroup differences and similarities. A.Fib atrial fibrillation; BMI, body mass index, CAD, coronary artery disease; COPD, chronic obstructive pulmonary disease; CKD, chronic kidney disease; DBP diastolic blood pressure; E/A ratio of mitral peak velocity of early filling to mitral peak velocity of late filling; E/E-prime ratio of mitral peak velocity of early filling to early diastolic mitral annular velocity; eGFR estimated glomerular filtration rate; FAC fractional area change; HF heart failure; IVRT isovolumic relaxation time; LAVI left atrial volume index; LVEDI left ventricular end-diastolic volume index; LVEF, left ventricular ejection fraction; LVESI left ventricular end-systolic volume index; LVMI left ventricular mass index; MAGGIC, Meta-analysis Global Group in Chronic Heart Failure; MI, myocardial infarction; NYHA, New York Heart Association; PCI percutaneous coronary intervention; RA right atrial; RV right ventricular; SBP systolic blood pressure; TAPSE tricuspid annular plane systolic excursion; WBC white blood cells.

Table 1 Clinical interpretation and generalised characterisation of the six phenogroups

Phenogroup (n)	Generalised characteristics	Primary events per 100 patient-years*
1 (n=30)	Younger, risk factors such as hypertension, CAD and diabetes, larger left ventricular volumes, and CKD	33
2 (n=40)	Older, less hypertrophy but worse left ventricular and right ventricular systolic function and more severe HF, and notably high prevalence of COPD	40
3 (n=48)	Younger, male, obesity, less left and right diastolic and systolic function, and milder HF	11
4 (n=52)	Male, hypertension, left ventricular hypertrophy, left atrial enlargement and AF	19
5 (n=101)	Older, female, hypertension, CAD, large left side but good right ventricular function	21
6 (n=49)	Older, female, low BMI, hypertension, AF and right-sided failure	27

*Derived from events to mid-term time point at 18 months.

AF, atrial fibrillation; BMI, body mass index; CAD, coronary artery disease; CKD, chronic kidney disease; COPD, chronic obstructive pulmonary disease; HF, heart failure.

the extent of left ventricular remodelling, the distinctly high prevalence of chronic obstructive pulmonary disease (COPD, 30%) in group 2, which also had the worst HF status, and the less than expected CKD (45%) in group 2, given the extent of RV and presumably backward failure.

Comparing the prognostic HF marker NT-proBNP across phenogroups we found the highest levels at initial hospitalisation for patients in phenogroup 2, and their NT-proBNP levels remained high at the follow-up visit (online supplementary figure S4, online supplementary table S3), in line with the highest proportion of patients in New York Heart Association class IV and poorer prognosis in these patients.

Differential proteomics profiles across phenogroups

Using targeted plasma protein data in a subset of Swedish KaRen patients (n=76 (6; 5; 24; 21; 17; 3 in the phenogroups, respectively)), we compared protein profiles across phenogroups observing significant differences in levels of 15 out of 86 proteins (FDR<0.05, table 2). Levels of these 15 proteins (after adjustments) across phenogroups are shown in online supplementary figure S5. In agreement with phenogroup characteristics, phenogroup 1 had the highest levels of proteins previously positively associated with CKD and HF incidence (FGF-23, PlGF, TRAIL-R2, U-PAR, online supplementary table S7), whereas HF-specific soluble ST2 (sST2) was highest in phenogroup 2, with the most severe HF. In addition, 10 of the proteins (including sST2) were directly associated with phenotypical variables and comorbidities (online supplementary results).

Differential composite end point between phenogroups

The incidence rate of the composite end point was significantly different (log rank p<0.001) between the phenogroups to the 18 months follow-up, with phenogroup 2 having the highest and phenogroup 3 having the lowest event rates (figure 2, online supplementary table S9). There was also a significant difference at the short-term (at 100 days) and at the longer-term follow-up (at 1000 days). Notably, 39% of patients in phenogroup 2 had an event already at 100 days. Investigating associations with the composite end point using Cox proportional hazards models, phenogroup 2 was 83% more likely than phenogroup 3 to have an event to 18 months (HR 5.73 (95% CI 2.57 to 12.77), p<0.001; online supplementary table S10). Higher risk for phenogroup 2 at 18 months (p<0.001) was also observed in comparison with phenogroup 5 (HR=2.69 (95% CI 1.58 to 4.57)).

In line with these findings, the MAGGIC Score, an established mortality risk score for patients with HF,[24] was significantly different across phenogroups (p<0.001; figure 3). Higher MAGGIC Score, predicting increased risk of 3-year mortality, correlated with belonging to phenogroups with the highest event rate, with patients in phenogroup 2 having the highest average score and risk, and phenogroup 3 the lowest.

Assigning patients into phenogroups

We assessed how well we could assign patients into phenogroups, using multinomial classifiers trained on the same clinical

Table 2 Plasma proteins differentially expressed between heart failure with preserved ejection fraction (HFpEF) phenogroups (false discovery rate, FDR<0.05). The adjusted protein levels across the phenogroups are shown in online supplementary figure S5

| Protein	Gene	Description	P value
ESM-1	ESM1	Endothelial cell-specific molecule 1 (Endocan)	2.53E-04
sST2	IL1RL1	ST2 protein	6.61E-04
FGF-23	FGF23	Fibroblast growth factor 23	2.63E-03
NEMO	IKBKG	NF-κB essential modulator	4.78E-03
TIE2	TEK/TIE2	Angiopoietin-1 receptor	6.76E-03
TRAIL-R2	TNFRSF10B	TNF-related apoptosis-inducing ligand receptor 2	7.90E-03
HB-EGF	HBEGF	Heparin-binding EGF-like growth factor	0.011
PlGF	PGF	Placenta growth factor	0.024
IL16	IL16	Pro-interleukin-16	0.024
BNP	NPPB	Natriuretic peptides B	0.026
U-PAR	PLAUR	Urokinase plasminogen activator surface receptor	0.027
REN	REN	Renin	0.029
CCL20	CCL20	C-C motif chemokine 20	0.033
MMP-1	MMP1	Matrix metalloproteinase-1	0.038
AGRP	AGRP	Agouti-related protein	0.039
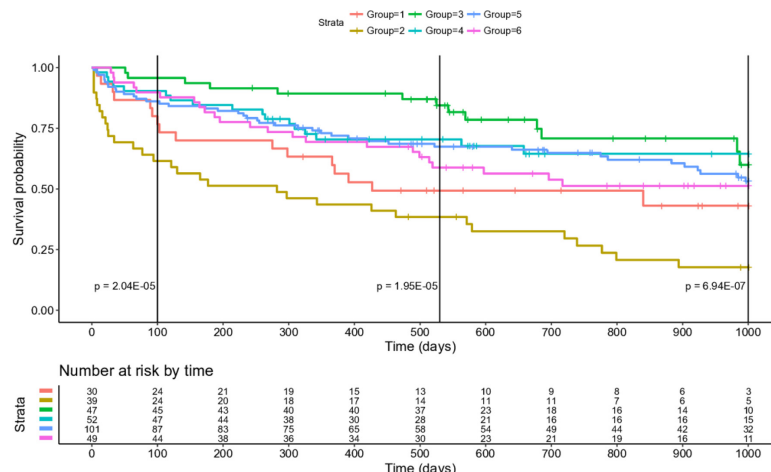
Figure 2 Kaplan-Meier curves of composite end points during 1000 days of follow-up from stable condition for phenogroups. The log rank p values for the composite end point at an early time point (100 days) and a mid-range time point (18 months) are also shown.

variables as used in the clustering. To avoid overfitting, while optimising the models, we used a nested cross-validation procedure. The best performance was observed for the Elastic Net model, with an average multiclass AUC of 0.79 (online supplementary figure S7, online supplementary table S11), indicating reasonable prediction performance across phenogroups. In the final model on all individuals (multinomial Elastic Net) we found a number of important predictors for falling into each of the phenogroups (online supplementary figure S8, online supplementary table S12). Variables with the largest coefficients in the overall model included E/A ratio, NT-proBNP, sodium,

creatinine and haemoglobin, diastolic blood pressure, heart rate and body surface area. Finally, we validated the performance of the complete procedure, that is, clustering patients to derive phenogroups, training an Elastic Net model for assigning new patients to phenogroups, and assessing the difference in the composite outcome between the newly assigned phenogroups (online supplementary methods). Averaged over all trial splits (50:50 patients), we could detect phenogroups, reliably assign new patients to these phenogroups and show a differential composite outcome significantly different from random (online supplementary figure S9, supplementary results).
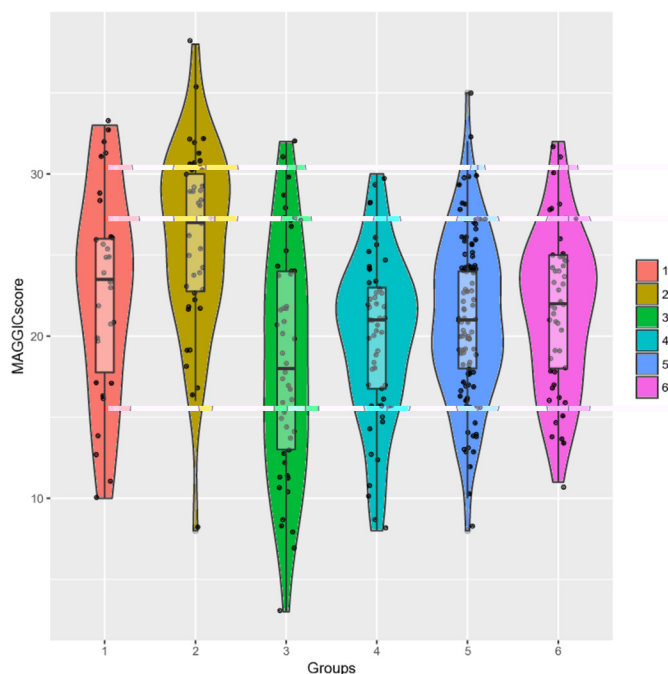
DISCUSSION

In this study we identify phenotypically distinct groups of patients with HFpEF, with characteristics indicating differential burden of comorbidities, severity of disease and outcomes. We also demonstrate the association of phenogroups with differential levels of 15 inflammatory and cardiovascular proteins. Furthermore, we optimise a model (using the same phenotypical variables as in clustering) for categorisation of patients into these phenogroups, which may also be used to assign new patients into the stratified groups, given the same phenotype variables. Using this model we also validate the complete clustering procedure, using random resampling of patients, showing good reproducibility. Our work confirms the feasibility of phenotyping of patients with HFpEF via statistical machine learning[15 16] and extends the findings of published literature by exploring this in a deeply phenotyped European population as well as correlating phenogroups to inflammatory and cardiovascular proteomics.

HFpEF phenogroups

A wide range of characteristics were significantly different between phenogroups, and we note that, for example, AF, kidney dysfunction, anaemia, age, sex, cardiac structural and functional alterations seem to be important in substratifying patients with HFpEF. The same comorbidities and demographics also showed differences across subgroups in previous studies, identifying three groups (applying model-based clustering to 46 continuous variables[15]) and six groups (applying latent class analysis to 11 categorical variables[16]), respectively. However, in our study of European patients with HFpEF we did not observe a phenogroup of obese patients (as previously found in an US setting[15]). This most likely reflects

Figure 3 Violin plots of Meta-analysis Global Group in Chronic Heart Failure (MAGGIC) Scores of patients in phenogroups. Significant differences in MAGGIC Score, an established mortality risk score for patients with heart failure (HF) predicting 3-year mortality, were observed between phenogroups (overall p=3.1×10–08), with patients in phenogroups with the most events having higher MAGGIC Score. A mean MAGGIC Score of 26, as observed for phenogroup 2, corresponds to an approximate 50% risk of 3-year mortality (Sartipy *et al* [24]).

the difference in prevalence of obesity and morbid obesity between these populations.[25] While we did not specifically seek to separate patients on outcome, our analysis indicated that clustering based on baseline characteristics only identified clusters with distinctly different rates of the composite end points for the phenogroups, suggesting that clustering can be used prospectively to estimate event rates, for example, in trial design. The two phenogroups that fared the worst during follow-up had different comorbidities and clinical phenotypical profile: (phenogroup 1) all with a history of hypertension, with the highest prevalence of any groups of CAD, kidney dysfunction, anaemia and diabetes, consistent with primary myocardial dysfunction and at least in part forward failure, and more resemblance to HFrEF; and (phenogroup 2) most with AF, high prevalence of COPD, old age, kidney dysfunction, anaemia, increased evidence of diastolic dysfunction and right ventricular dysfunction, consistent with a more distinct HFpEF and backward failure phenotype. We have previously demonstrated many of these phenotypes (including AF and flutter, age and anaemia) to be predictors of the composite end point in the KaRen Study population.[7] On the other hand, phenogroup 3 with the least severe HF, were younger, obese, with high blood pressure but with low prevalence of anaemia and advanced kidney dysfunction, and less cardiac structural and functional alterations. Although study population and analytical strategy differed, phenogroups 1 and 2 shared characteristics with the subgroups that fared the worst in a previous larger study:[16] phenogroup 1 (diabetes, anaemia and kidney dysfunction, but not obesity; subgroup C) and phenogroup 2 (old age, AF, kidney dysfunction, anaemia; subgroup F). Furthermore, characteristics such as RV remodelling and dysfunction, old age, AF and CKD were observed for the group with the highest event rates here and in a previous similarly conducted study in an US population.[15] This suggests that the groups derived using clustering may in part be reproducible across methods and populations.

Protein correlates of phenogroups

In addition to clinical and laboratory differences of phenogroups we also found 15 plasma proteins at differential levels across groups, including sST2 and BNP, markers of HF severity or prognosis.[26 27] Soluble ST2 had the highest mean level in phenogroup 2, consistent with the poor outcome of these patients, and lowest in phenogroup 3, with the least severe HF. Nine of the 15 proteins with differential levels across phenogroups have previously been associated with HF and comorbidities in independent data sets using the same proximity eThese trials have relied only on signsxtension assay proteomics chip (online supplementary table S7), and an additional two proteins in previous work on partly the same patients with HFpEF as in this study,[18] lending support to the importance of these proteins in HF disease processes. Of these FGF-23, PlGF, NEMO, sST2, TRAIL-R2, U-PAR have previously been associated with HF incidence and HFpEF,[28 29] and all, except NEMO and sST2, have also been associated with CKD or kidney dysfunction.[30] FGF-23 and U-PAR have also been associated with incident AF.[31] Further evidence indicates the importance of some of the associated proteins in HF. NEMO (NF-κ-B essential modulator), found at highest levels in phenogroups 3 and 5, has been suggested to be important in cardiac function, as disruption in mice leads to a HF phenotype.[32] Levels of ESM-1, with the strongest association to phenogroups of any of the proteins,

has previously been negatively associated with BMI and insulin resistance.[33 34] Further evidence supports a role for this protein as a marker of endothelial dysfunction (upregulated in acute MI, CKD and hypertension[35]), which is also consistent with a purported model of global inflammation driving microvascular endothelial dysfunction in HFpEF.[10 11] Furthermore, lower levels of both FGF-23 and TRAIL-R2 were recently found to associate with resilience against new cardiovascular events in a cohort of patients with stable CAD.[36] These two protein markers were found at the highest levels in the two highest-risk groups in our study.

Assigning patients into phenogroups

We developed a model for assigning patients to phenogroups, which may also be used to assign new patients into the stratified groups, given the same variables. Results of this analysis also indicated the relative importance of the phenotypes (predictors) in stratifying the patients with HFpEF. Using the method of assigning new patients into phenogroups we also validated the complete clustering procedure showing that using only half the patients for detecting phenogroups and assessing outcome differences we could reliably assign new patients into phenogroups of differential outcome increasing our confidence in the robustness of the complete procedure of deriving phenogroups. This finding increases our confidence that the detected phenogroups and outcome differences will hold up in populations similar to the KaRen cohort studied here.

Potential clinical implications

To date there is no evidence-based therapy for HFpEF. HFpEF is a heterogeneous syndrome, thought to be driven by a range of comorbidities, and the diversity of patients may partly explain the failure to demonstrate clinical efficacy in randomised control trials of ACE inhibitors, angiotensin receptor blockers and spironolactone. These trials have relied only on signs and symptoms of HF, which are non-specific for HF and poorly correlated with outcomes. Newly updated stricter criteria for diagnosis of HFpEF may lead to a more well defined HFpEF phenotype,[1] mitigating some of these problems, however, further categorisation of patients into well-defined phenogroups with diverse characteristics and event rates may be useful in study design and when investigating novel intervention strategies in HFpEF.

Strengths and limitations

The main strengths of this study include the well-phenotyped patient population from two countries with long-term follow-up of disease outcome with the inclusion of proteomics data on a subset of individuals that help us to draw conclusions regarding molecular correlates of the phenogroups. This study also has some limitations. First, we acknowledge that the most important limitations are the small number of patients (n=320) and the lack of external validation of the phenogroups, making generalisability to the general HFpEF population uncertain. We provide a prediction model, which, assuming the study has the same clinical phenotypes, may be used to assign patients into the same phenogroups of patients with HFpEF. We aim to validate this in further studies, and in addition, this prediction model may be used to validate the results in this study by other researchers. Second, proteomics data were only available for a subset of patients, all Swedish. We assume, but cannot confirm, that the conclusions would also apply to other patients within the same phenogroup.

In conclusion, in this study we identify distinct subgroupings of patients with HFpEF in a deeply phenotyped European study using statistical machine learning, extending previous knowledge on HFpEF subtypes. Adding to previous work, we also integrate our results with protein biomarkers, providing evidence of differential levels of cardiovascular biomarkers relating to HF, AF and kidney function between phenogroups. Categorisation of patients into well-defined phenogroups with diverse characteristics and event rates may be useful in study design and when investigating novel intervention strategies in HFpEF.

Key questions

What is already known on this subject?
► HFpEF is a heterogeneous syndrome, however previous studies indicate that clustering of patients with HFpEF based on clinical data and biochemical data can reveal subgroups of patients at greater risk of clinical end points.

What might this study add?
► In this study we apply this approach to a well-defined European population and in addition use plasma proteomics on these patients to define molecular patterns of the newly defined HFpEF phenogroups.

How might this impact on clinical practice?
► Categorisation of patients with HFpEF into well-defined phenogroups with diverse characteristics and event rates may be useful in future interventional trial design and when investigating novel intervention strategies in HFpEF.

Author affiliations
[1]Cardiovascular Medicine Unit, Department of Medicine Solna, Karolinska Institute, Stockholm, Sweden
[2]Pfizer Global Research and Development, Stockholm, Sweden
[3]Department of Medicine Solna, Unit of Cardiology, Karolinska Institute, Stockholm, Sweden
[4]Heart and Vascular Theme, Karolinska University Hospital, Stockholm, Sweden
[5]Pfizer Global Research and Development, Boston, Massachusetts, USA
[6]Department of Cardiology and Department of Molecular and Clinical Medicine, Institute of Medicine, Sahlgrenska Academy, Goteborgs Universitet, Goteborg, Sweden
[7]Early Clinical Development, Early CVRM BioPharmaceuticals R&D, AstraZeneca FoU Goteborg, Goteborg, Sweden
[8]Division of Cardiology, Department of Medicine, Feinberg School of Medicine, Northwestern University, Evanston, Illinois, USA
[9]Cardiology and CIC-IT1414, CHU de Rennes LTSI, Universite Rennes, Rennes, France
[10]Cardiology, University Hospital Rennes, Rennes, Bretagne, France
[11]Pfizer Global Research and Development, Berlin, Germany

ORCID iDs
Åsa K Hedman http://orcid.org/0000-0001-5413-204X
Erwan Donal http://orcid.org/0000-0002-9083-1582
Anders Mälarstig http://orcid.org/0000-0003-2608-1358

REFERENCES

1 Pieske B, Tschöpe C, de Boer RA, *et al*. How to diagnose heart failure with preserved ejection fraction: the HFA-PEFF diagnostic algorithm: a consensus recommendation from the heart failure association (HFA) of the European Society of cardiology (ESC). *Eur Heart J* 2019;40:3297–317.
2 Owan TE, Hodge DO, Herges RM, *et al*. Trends in prevalence and outcome of heart failure with preserved ejection fraction. *N Engl J Med Overseas Ed* 2006;355:251–9.
3 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.
4 Patel RB, Shah SJ, Fonarow GC, *et al*. Designing future clinical trials in heart failure with preserved ejection fraction: lessons from TOPCAT. *Curr Heart Fail Rep* 2017;14:217–22.
5 Kelly JP, Mentz RJ, Mebazaa A, *et al*. Patient selection in heart failure with preserved ejection fraction clinical trials. *J Am Coll Cardiol* 2015;65:1668–82.
6 Butler J, Fonarow GC, Zile MR, *et al*. Developing therapies for heart failure with preserved ejection fraction: current state and future directions. *JACC Heart Fail* 2014;2:97–112.
7 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.
8 Koh AS, Tay WT, Teng THK, *et al*. A comprehensive population-based characterization of heart failure with mid-range ejection fraction. *Eur J Heart Fail* 2017;19:1624–34.
9 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.
10 Paulus WJ, Tschöpe 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.
11 Lam CSP, Lund LH. Microvascular endothelial dysfunction in heart failure with preserved ejection fraction. *Heart* 2016;102:257–9.
12 Sørlie T, Perou CM, Tibshirani R, *et al*. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98:10869–74.
13 Parker JS, Mullins M, Cheang MCU, *et al*. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009;27:1160–7.
14 Ahlqvist E, Storm P, Käräjämäki A, *et al*. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol* 2018;6:361–9.
15 Shah SJ, Katz DH, Selvaraj S, *et al*. Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation* 2015;131:269–79.
16 Kao DP, Lewsey JD, Anand IS, *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.
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 Hage C, Michaëlsson E, Linde C, *et al*. Inflammatory biomarkers predict heart failure severity and prognosis in patients with heart failure with preserved ejection fraction: a holistic proteomic approach. *Circ Cardiovasc Genet* 2017;10:e001633.
19 Donal E, Lund LH, Oger E, *et al*. New echocardiographic predictors of clinical outcome in patients presenting with heart failure and a preserved left ventricular ejection fraction: a subanalysis of the KA (Karolinska) ren (Rennes) study. *Eur J Heart Fail* 2015;17:680–8.
20 Pocock SJ, Ariti CA, McMurray JJV, *et al*. Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies. *Eur Heart J* 2013;34:1404–13.
21 Fraley C, Raftery AE. Model-Based clustering, discriminant analysis, and density estimation. *J Am Stat Assoc* 2002;97:611–31.
22 Fisher RA. *Statistical methods for research workers*. Edinburgh, UK: Oliver & Boyd, 1934.
23 Kuhn M. Building Predictive Models in *R* Using the caret Package. *J Stat Softw* 2008;28:1–26.
24 Sartipy U, Dahlström U, Edner M, *et al*. Predicting survival in heart failure: validation of the MAGGIC heart failure risk score in 51 043 patients from the Swedish heart failure registry. *Eur J Heart Fail* 2014;16:173–9.

25 World Health Organization. *Global status report on noncommunicable diseases 2014*. Geneva: World Health Organization, 2015.

26 Weinberg EO, Shimpo M, Hurwitz S, *et al*. Identification of serum soluble ST2 receptor as a novel heart failure biomarker. *Circulation* 2003;107:721–6.

27 Oremus M, Don-Wauchope A, McKelvie R, *et al*. BNP and NT-proBNP as prognostic markers in persons with chronic stable heart failure. *Heart Fail Rev* 2014;19:471–505.

28 Stenemo M, Nowak C, Byberg L, *et al*. Circulating proteins as predictors of incident heart failure in the elderly. *Eur J Heart Fail* 2018;20:55–62.

29 Drum CL, Tan WKY, Chan Siew-Pang, *et al*. Thymosin Beta-4 is elevated in women with heart failure with preserved ejection fraction. *J Am Heart Assoc* 2017;6:pii:e005586.

30 Carlsson AC, Ingelsson E, Sundström J, *et al*. Use of proteomics to investigate kidney function decline over 5 years. *Clin J Am Soc Nephrol* 2017;12:1226–35.

31 Lind L, Sundström J, Stenemo M, *et al*. Discovery of new biomarkers for atrial fibrillation using a custom-made proteomics CHIP. *Heart* 2017;103:377–82.

32 Kratsios P, Huth M, Temmerman L, *et al*. Antioxidant amelioration of dilated cardiomyopathy caused by conditional deletion of NEMO/IKKgamma in cardiomyocytes. *Circ Res* 2010;106:133–44.

33 Westergren HU, Svedlund S, Momo RA, *et al*. Insulin resistance, endothelial function, angiogenic factors and clinical outcome in non-diabetic patients with chest pain without myocardial perfusion defects. *Cardiovasc Diabetol* 2016;15:36.

34 Harlid S, Myte R, Van Guelpen B. The metabolic syndrome, inflammation, and colorectal cancer risk: an evaluation of large panels of plasma protein markers using repeated, prediagnostic samples. *Mediators Inflamm* 2017;2017:4803156

35 Qiu C-R, Fu Q, Sui J, *et al*. Serum endothelial cell-specific molecule 1 (Endocan) levels in patients with acute myocardial infarction and its clinical significance. *Angiology* 2017;68:354–9.

36 Wallentin L, Eriksson N, Siegbahn A. Screening for protein biomarkers associated with resilence against cardiovascular and cancer events in elderly patients with stable coronary artery disease using the proximity extension assay. In: *Advancing precision medicine: current and future proteogenomic strategies for biomarker discovery and development*. Washington, DC: Science/AAAS, 2017: 18–21.