## Supplementary Table 1: Studies of Clustering Analysis of Patients with HFrEF

<table>
<thead>
<tr>
<th>Authors</th>
<th>Data source</th>
<th>Study outcomes</th>
<th>Phenotypic approach</th>
<th>Number of features included</th>
<th>Number of clusters</th>
<th>Characteristics of each phenotype (proportion of patients out of studied patients, %)</th>
</tr>
</thead>
</table>
| Shah et al.              | 397 patients from Northwestern University HFrEF program | CV and non-CV hospitalisation                      | Hierarchical clustering              | 67                        | 3                  | 1. **Younger**, lower BNP, least myocardial remodeling and dysfunction (32%)  
2. Highest prevalence of obesity, DM, obstructive sleep apnoea, highest PCWP and PVR (30%)  
3. **Oldest**, high prevalence of AF, CKD, high BNP, high MAGGIC score, longest QRS duration, highest LVMI, worst RV function, worst outcome (38%)  
*Cluster 3 had the highest HR of 1.7 (95% CI 0.6-4.9) versus Group 1 for all-cause death*  
HF duration was similar across phenotypes |
| Kao et al.               | 4,113 patients from I-PRESERVE                   | All-cause mortality or CV hospitalisation          | Latent class analysis                | 11                        | 6                  | 1. 100% **men**, younger, more alcohol use, low rates of AF, renal disease, valvular disease (15%)  
2. Predominantly women, **younger**, more anaemia, low rates of AF, renal disease, valvular disease (17%)  
3. High rates of obesity, DM, hyperlipidaemia, CAD, anaemia, **worse renal function** (8%)  
4. 100% women, average rates of DM, hyperlipidaemia, obesity, renal insufficiency (30%)  
5. 100% **men**, lower BMI, excess AF and CAD (18%)  
6. Predominantly **women advanced age**, lower BMI, higher rates of AF, valvular disease, renal dysfunction, anaemia, worst outcome (12%)  
*Cluster 6 had the highest HR of 3.1 (95% CI 2.4-4.0) versus Group 1 for all-cause death or CV hospitalisation* |
| Segar et al.             | 654 patients from TOPCAT Outcome;                | Composite of aborted cardiac arrest, HF hospitalisation or CV death | Penalized finite mixture model-based clustering analysis | 61                        | 3                  | 1. Higher BMI, severe HF symptoms, higher NP, highest prevalence of DM, dyslipidaemia, and atherosclerotic CV disease, worst abnormalities in LV and LA structure and function, worst outcome (21%)  
2. **Lowest DM**, mild HF symptoms, lower CV disease, higher burden of diastolic dysfunction (17%)  
3. lowest BMI, intermediate burden of atherosclerotic vascular disease, DM, lowest NP level, most favourable diastolic function profile (63%)  
*Cluster 1 had the highest HR of 1.9 (95% CI 1.4-2.7) versus Group 3 for the composite of aborted cardiac arrest, HF hospitalisation or CV death* |
| Hedman, et al.           | 320 patients from KaRen study                    | Composite of all-cause mortality and                | Model-based clustering               | 32 echo variables         | 6                  | 1. **Younger**, risk factors such as HTN, CAD, DM, larger LV volumes and CKD (9%)  
2. **Older**, less hypertrophy but **worse LV and RV systolic function** and more severe HF, and notably high prevalence of COPD, worst outcome (13%) |
<table>
<thead>
<tr>
<th>Study</th>
<th>Number of Patients</th>
<th>Composite Event</th>
<th>Methodology</th>
<th>Cluster Characteristics</th>
</tr>
</thead>
</table>
| Cohen et al. | 3442 | Composite of CV death, HF hospitalisation, or aborted cardiac arrest | Latent class analysis | 1. Younger, smoking, preserved functional class, highest smoking rate, preserved renal function, low DM, least evidence of LV hypertrophy and arterial stiffness, lowest NT-proBNP (35%)<br>2. Older, high prevalence of AF and CKD, low DM and obesity, small concentric LV with lowest LV mass, largest LA enlargement, large-artery stiffening, biomarkers of innate immunity and vascular calcification (IL-8, pentraxin-3, soluble intercellular adhesion molecule-1) (39%)<br>3. Very high prevalence of obesity and DM, high prevalence of CKD and depression, impaired functional class, concentric LV hypertrophy, high TNF-alpha, liver fibrosis, tissue remodeling (26%)<br>*Cluster 2 had the highest HR of 5.73 (95% CI 2.57 to 12.77) versus Group 3 for Composite of all-cause mortality or HFH (13%)
| Schrub et al. | 356 | Composite of all-cause mortality or first hospitalisation for HF | Hierarchical clustering | 1. Younger male, highest rate of HTN, DM, obesity and renal insufficiency (36%)<br>2. Female dominant, lowest rates of DM, less sinus rhythm, preserved renal function, subnormal LV function (38%)<br>3. Oldest, female prevalent, highest rate of rhythm disorder, highest incidence of MR and atrial remodeling (26%)<br>*No statistical difference in all-cause death or HF hospitalisation
| Woolley et al. | 429 | Occurrence of death or HF hospitalisation | Hierarchical clustering | 1. Highest prevalence of DM, and renal disease, activation of inflammatory pathways (14%)<br>2. Oldest and frequent age-related comorbidities (38%)<br>3. Youngest, highest prevalence of obesity, least symptoms, lowest NT-proBNP (39%)<br>4. Highest prevalence of COPD, ischaemic aetiology, smoking, more symptom, highest NT-proBNP and troponin (10%)<br>*Occurrence of death or HFH was highest in Group 1 (62.1%) and 4 (62.8%) and lowest in Group 3 (25.6%)
| Uijl et al. | 6909 | Composite of CV mortality and HF hospitalisation | Latent class analysis | 1. Young, low comorbidity burden, highest implantable device (10%)<br>2. AF, HTN, without DM (30%)<br>3. Oldest, high prevalence of AF (25%)<br>4. High prevalence of obesity, DM and HTN (15%)
5. Older, high prevalence of AF, HTN, low eGFR, frequent diuretic prescription (20%)
*Cluster 5 had the highest HR of 4.7 (95%CI 3.4-6.4) versus Group 1 for composite of all-cause death or HFH

*Features used to further classify into five common phenotypes (*Figure 1*) are bolded.

Abbreviations: AF = atrial fibrillation; BIOSTAT-CHF = Biology Study to Tailored Treatment in Chronic Heart Failure; BMI = body mass index; BNP = brain natriuretic peptide; CAD = coronary artery disease; CI = confidence interval; CKD = chronic kidney disease; CV = cardiovascular; DM = diabetes mellitus; HFH = heart failure hospitalisation; eGFR = estimated glomerular filtration rate; HFpEF = heart failure with preserved ejection fraction; HTN = hypertension; HR = hazard ratio; IHD = ischaemic heart disease; IL = interleukin; I-PRESERVE = Irbesartan in Heart Failure with Preserved Ejection Fraction Study; KaRen, Karolinska-Rennes; LA = left atrial; LV = left ventricular; LVMI = left ventricle mass index; MAGGIC = Meta-Analysis Global Group in Chronic Heart Failure; MR = mitral regurgitation; NT-proBNP = N-terminal pro brain natriuretic peptide; NP = natriuretic peptide; NYHA = New York Heart Association classification; PCWP = pulmonary capillary wedge pressure; PMI = pacemaker implantation; PVR = pulmonary vascular resistance; RV = right ventricle; SBP = systolic blood pressure; SwedeHF = Swedish Heart Failure Registry; TNF = tumour necrosis factor; TOPCAT = Treatment of Preserved Cardiac Function Heart Failure with an Aldosterone Antagonist.