Heart failure subphenotypes based on repeated biomarker measurements are associated with clinical characteristics and adverse events (Bio-SHiFT study)

Iris de Lange a,1, Teun B. Petersen a,b,1, Marie de Bakker a,1, K. Martijn Akkerhuis a,1, Jasper J. Brugts a,1, Kadir Caliskan a,1, Olivier C. Manintveld a,1, Alina A. Constantinescu a,1, Tjeerd Germans c,1, Jan van Ramshorst c,1, Victor A.W.M. Umans c,1, Eric Boersma a,1, Dimitris Rizopoulos b,d,1, Isabella Kardys a,1,*

Department of Cardiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
Department of Biostatistics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
Department of Cardiology, Northwest Clinics, Alkmaar, the Netherlands
Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands

ARTICLE INFO

Keywords:
Heart failure
Phenotype
Cluster analysis
Biomarkers
Repeated measurements
Prognosis

ABSTRACT

Background: This study aimed to identify heart failure (HF) subphenotypes using 92 repeatedly measured circulating proteins in 250 patients with heart failure with reduced ejection fraction, and to investigate their clinical characteristics and prognosis.

Methods: Clinical data and blood samples were collected tri-monthly until the primary endpoint (PEP) or censoring occurred, with a maximum of 11 visits. The Olink Cardiovascular III panel was measured in baseline samples and the last two samples before the PEP (in 66 PEP cases), or the last sample before censoring (in 184 PEP-free patients). The PEP comprised cardiovascular death, heart transplantation, Left Ventricular Assist Device implantation, and hospitalization for HF. Cluster analysis was performed on individual biomarker trajectories to identify subphenotypes. Then biomarker profiles and clinical characteristics were investigated, and survival analysis was conducted.

Results: Clustering revealed three clinically diverse subphenotypes. Cluster 3 was older, with a longer duration of, and more advanced HF, and most comorbidities. Cluster 2 showed increasing levels over time of most biomarkers. Clusters 3 had elevated baseline levels and increasing levels over time of 16 remaining biomarkers. Median follow-up was 2.2 (1.4–2.5) years. Cluster 3 had a significantly poorer prognosis compared to cluster 1 (adjusted event-free survival time ratio 0.25 (95%CI:0.12–0.50), p < 0.001). Repeated measurements clusters showed incremental prognostic value compared to clusters using single measurements, or clinical characteristics only.

Conclusions: Clustering based on repeated biomarker measurements revealed three clinically diverse subphenotypes, of which one has a significantly worse prognosis, therefore contributing to improved (individualized) prognostication.

1. Introduction

The burden of heart failure (HF) is estimated to grow further in the coming years [1–3]. Despite advances in HF management, the 5-year mortality rate of heart failure with reduced ejection fraction (HFrEF) remains 67% [4]. Current HFrEF treatment guidelines do not sufficiently consider heterogeneity in clinical profiles and etiologies [5,6]. HFrEF involves many different pathophysiological mechanisms, including
International Journal of Cardiology 364 (2022) 77–84

78

International Journal of Cardiology 364 (2022) 77–84

2.1. Study population

This investigation was performed within the ‘Serial Biomarker Measurements and New Echocardiographic Techniques in Chronic Heart Failure Patients Result in Tailored Prediction of Prognosis’ (Bio-SHiFT) study. This is a prospective, multi-center cohort study of 398 patients with chronic HF, conducted in the Netherlands in Erasmus MC, Rotterdam, and Noordwest Ziekenhuis-group, Alkmaar. The medical ethics committee of the Erasmus MC approved the study, and it was conducted according to the ethical guidelines of the 1975 Declaration of Helsinki. The study is registered as NCT number 01851538 in ClinicalTrials.gov. Study design has been described in previous papers [18–22]. In brief, the study included outpatients with clinically stable HF (meaning no hospitalization in the past three months). All patients provided written informed consent. Inclusion criteria were ≥ 18 years old; capable of understanding and signing informed consent; and a diagnosis of chronic HF (with reduced or normal ejection fraction) according to the guidelines of the European Society of Cardiology (ESC) [23]. Exclusion criteria were HF secondary to circulatory high output conditions; a scheduled surgery or intervention for a coronary or non-coronary indication; severe renal failure requiring dialysis; pre-existing moderate or severe liver disease; Chronic Obstructive Pulmonary Disease (COPD) GOLD stage IV; congenital heart disease; life expectancy of ≤ 1 year; being unlikely to appear at all scheduled follow-up visits; and a linguistic barrier. The current investigation used data from 250 HFrEF patients enrolled in the first inclusion round between October 2011 and June 2013 and follow-up until November 2015.

2.2. Data-collection and follow-up

Patients received usual care from their physician at the outpatient clinic. Physicians were blinded to results of the biomarker measurements, because these were performed batch-wise after completion of follow-up. Parallel to usual care, study visits were scheduled at baseline and then every three months (± one month) until the end of follow-up or death. Maximum follow-up consisted of ten tri-monthly visits, consisting of a short medical assessment by a research physician and venipuncture. Clinical data were recorded in electronic case report forms. Collected data included demographics, medical history, medication use, and information on occurrence of adverse events.

2.3. Selection of blood samples

Blood samples were collected at the baseline and follow-up visits, with a maximum of 11 venipunctures per patient. In the first inclusion round, a total of 1984 samples was collected before occurrence of the primary endpoint (PEP) or censoring (median [25th–75th percentile]): 9 [5–10] samples per patient.

For reasons of efficiency, for the current investigation, a selection was made from these samples: all baseline samples, the last sample available in patients in whom the PEP did not occur during follow-up, and the two samples available closest in time before the PEP. Previous investigations in this cohort have demonstrated that several biomarker levels increase in the months preceding the clinical adverse event [20,22]. Thus, by selecting the last two samples before the endpoint, we aimed to capture this increase. Conversely, in event-free patients, previous investigations showed stable biomarker levels, in which case one additional sample suffices. Altogether, the selection resulted in 537 samples for the current analysis.

2.4. Biomarker measurements

Within two hours after collection, blood samples were processed and stored at ~80 °C. Biomarkers were measured batch-wise after completion of follow-up. Laboratory personnel were blinded to clinical patient information. The Olink Multiplex platform (Olink Proteomics AB, Uppsala, Sweden) Cardiovascular (CVD) III panel was used. The Olink assay is a Proximity Extension Assay (PEA) enabling high-throughput, multiplex immunooassay-qPCR that measures 92 proteins across 88 samples simultaneously [24,25]. The biomarkers are listed in supplemental table 1.

2.5. Primary endpoint

A clinical event committee adjudicated study endpoints based on hospital records and discharge letters, blinded to biomarker results. The PEP comprised cardiovascular death, heart transplantation (HTX), Left Ventricular Assist Device (LVAD) implantation, and hospitalization for the management of acute or worsened HF. Hospitalization for the management of acute or worsened HF was defined as hospitalization due to HF symptom exacerbation, combined with two of the following: BNP or NT-proBNP more than three times the upper limit of normal, signs of worsening HF, increased dosage or intravenous diuretics, or usage of positive inotropic agents [26].

2.6. Statistical methods

The full description of the statistical methods is included in the supplemental material, including power and sample size considerations. Continuous variables were described as mean (SD) when normally distributed, and as median (25th–75th percentile) when not normally distributed. Categorical variables were described as counts (percentage). The biomarker measurements provided by Olink Proteomics were already log-transformed.

To perform cluster analysis on repeated biomarker data, we first applied linear mixed modeling to compute individual intercepts and slopes of the temporal biomarker trajectories. Then winsorization [27] and standardization was performed.

The optimal number of clusters was determined using the NbClust R package [28]. Cluster stability was verified by resampling using several schemes (including bootstrapping and noise replacement) [29,30]. Then

neurohormonal, inflammatory and cardio-renal pathways [7,8]. Circulating plasma proteins representing these mechanisms may be used as biomarkers to provide more detailed insights into underlying biological processes in individual HF patients.

Accordingly, previous studies have identified HF subphenotypes based on these proteins and clinical characteristics [5,9–13]. More specific phenotyping may enable treatment strategies directed towards the different HF subphenotypes’ mechanisms. Moreover, improved phenotyping may provide opportunities for more accurate prognostication by associating subphenotypes with the occurrence of clinical endpoints [14–16]. These prognostic insights may ultimately contribute to personalized monitoring and timing of treatment strategies.

In this context, subphenotypes of heart failure with preserved ejection fraction (HFrEF) have been investigated in several studies [9–11,13,17]. However, studies investigating HFrEF subphenotypes are scarce [5,12]. The latter have based subphenotyping on patient characteristics, medical history, quality of life, exercise capability, and baseline biomarker measurements. They emphasize the heterogeneity within the HFrEF population and the necessity of improved phenotyping. In addition, the syndrome of HF tends to be progressive and includes episodes of decompensation. Thus, using serial biomarker measurements to distinguish subphenotypes, as opposed to single measurements, may enable us to capture this dynamic nature [18–22]. Therefore, in the current study, we aimed to identify HFrEF subphenotypes based on a set of 92 proteins measured in repeatedly collected blood samples of 250 HFrEF patients, and to investigate the association of these subphenotypes with clinical characteristics and adverse events.

Patients received usual care from their physician at the outpatient clinic. Physicians were blinded to results of the biomarker measurements, because these were performed batch-wise after completion of follow-up. Parallel to usual care, study visits were scheduled at baseline and then every three months (± one month) until the end of follow-up or death. Maximum follow-up consisted of ten tri-monthly visits, consisting of a short medical assessment by a research physician and venipuncture. Clinical data were recorded in electronic case report forms. Collected data included demographics, medical history, medication use, and information on occurrence of adverse events.

Blood samples were collected at the baseline and follow-up visits, with a maximum of 11 venipunctures per patient. In the first inclusion round, a total of 1984 samples was collected before occurrence of the primary endpoint (PEP) or censoring (median [25th–75th percentile]): 9 [5–10] samples per patient.

For reasons of efficiency, for the current investigation, a selection was made from these samples: all baseline samples, the last sample available in patients in whom the PEP did not occur during follow-up, and the two samples available closest in time before the PEP. Previous investigations in this cohort have demonstrated that several biomarker levels increase in the months preceding the clinical adverse event [20,22]. Thus, by selecting the last two samples before the endpoint, we aimed to capture this increase. Conversely, in event-free patients, previous investigations showed stable biomarker levels, in which case one additional sample suffices. Altogether, the selection resulted in 537 samples for the current analysis.

Within two hours after collection, blood samples were processed and stored at ~80 °C. Biomarkers were measured batch-wise after completion of follow-up. Laboratory personnel were blinded to clinical patient information. The Olink Multiplex platform (Olink Proteomics AB, Uppsala, Sweden) Cardiovascular (CVD) III panel was used. The Olink assay is a Proximity Extension Assay (PEA) enabling high-throughput, multiplex immunooassay-qPCR that measures 92 proteins across 88 samples simultaneously [24,25]. The biomarkers are listed in supplemental table 1.

A clinical event committee adjudicated study endpoints based on hospital records and discharge letters, blinded to biomarker results. The PEP comprised cardiovascular death, heart transplantation (HTX), Left Ventricular Assist Device (LVAD) implantation, and hospitalization for the management of acute or worsened HF. Hospitalization for the management of acute or worsened HF was defined as hospitalization due to HF symptom exacerbation, combined with two of the following: BNP or NT-proBNP more than three times the upper limit of normal, signs of worsening HF, increased dosage or intravenous diuretics, or usage of positive inotropic agents [26].

The full description of the statistical methods is included in the supplemental material, including power and sample size considerations. Continuous variables were described as mean (SD) when normally distributed, and as median (25th–75th percentile) when not normally distributed. Categorical variables were described as counts (percentage). The biomarker measurements provided by Olink Proteomics were already log-transformed.

To perform cluster analysis on repeated biomarker data, we first applied linear mixed modeling to compute individual intercepts and slopes of the temporal biomarker trajectories. Then winsorization [27] and standardization was performed.

The optimal number of clusters was determined using the NbClust R package [28]. Cluster stability was verified by resampling using several schemes (including bootstrapping and noise replacement) [29,30]. Then
unsupervised K-means cluster analysis partitioned the patients into this optimal number of clusters. [31].

We repeated clustering and the subsequent analysis described below using 1) solely baseline biomarker measurements, and 2) solely individual intercepts from linear mixed models containing random intercepts only.

Once the clusters were constructed, the biomarker profiles of the clusters were displayed using a heatmap [32]. Individual biomarkers were compared across clusters using one-way analysis of variance (ANOVA) or t-tests, depending on the number of clusters, and were corrected for multiple testing with a Bonferroni correction.

Clinical characteristics were described according to cluster, including age, gender, medical history, risk factors, underlying etiology of left ventricular dysfunction, medication use, symptoms of HF, and echocardiographic measurements. We compared characteristics across clusters using t-tests, Mann-Whitney tests, ANOVA, Kruskal-Wallis tests, χ²-tests, and Fisher’s exact tests, depending on variable distributions and the number of clusters. Pairwise comparisons were made, including a Bonferroni correction.

The association between cluster membership and the PEP was investigated with Kaplan-Meier curves and Tarone Ware tests. We also fitted accelerated failure time (AFT) models, which produce a time ratio. This time ratio is the exponentiated regression coefficient for cluster membership from the log-linear form of the AFT model, which shows the mathematical relation between the log of time and a set of covariates. Thus the time ratio can be interpreted as the average event-free survival time ratio.

### Table 1
Baseline characteristics according to clusters based on repeated measurements.

<table>
<thead>
<tr>
<th>Total cohort (N = 250)</th>
<th>Cluster 1 (N = 93)</th>
<th>Cluster 2 (N = 78)</th>
<th>Cluster 3 (N = 79)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>67.6 (58.3–75.8)</td>
<td>63.7 (53.6–69.8)</td>
<td>65.5 (58.4–74.6)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Male</td>
<td>184 (73.6)</td>
<td>72 (77.4)</td>
<td>55 (70.5)</td>
<td>0.558</td>
</tr>
<tr>
<td>Clinical characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>27.4 (4.7)</td>
<td>27.6 (4.4)</td>
<td>27.6 (4.9)</td>
<td>0.505</td>
</tr>
<tr>
<td>Systolic BP (mmHg)</td>
<td>121.6 (20.6)</td>
<td>123.1 (18.5)</td>
<td>120.6 (19.1)</td>
<td>0.567</td>
</tr>
<tr>
<td>Diastolic BP (mmHg)</td>
<td>72.2 (10.9)</td>
<td>72.7 (10.3)</td>
<td>72.8 (10.6)</td>
<td>0.317</td>
</tr>
<tr>
<td>Heart rate (bpm)</td>
<td>67.1 (11.6)</td>
<td>65.8 (10.4)</td>
<td>67.6 (11.4)</td>
<td>0.217</td>
</tr>
<tr>
<td>Ejection fraction (%)</td>
<td>47.1 (18.7–9.7)</td>
<td>5.1 (20.9–6.6)</td>
<td>2.6 (0.7–7.1)</td>
<td>0.004</td>
</tr>
<tr>
<td>NYHA class</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>75 (30.0)</td>
<td>36 (38.7)</td>
<td>25 (32.1)</td>
<td>14.177</td>
</tr>
<tr>
<td>II</td>
<td>113 (45.2)</td>
<td>41 (44.1)</td>
<td>37 (47.4)</td>
<td>35.443</td>
</tr>
<tr>
<td>III</td>
<td>62 (24.8)</td>
<td>16 (17.2)</td>
<td>16 (20.5)</td>
<td>0.380</td>
</tr>
<tr>
<td>IV</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Ejection fraction (%)</td>
<td>29.9 (9.7)</td>
<td>30.2 (10.9)</td>
<td>31.1 (8.7)</td>
<td>0.281</td>
</tr>
<tr>
<td>Etiology of HF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ischemic heart disease</td>
<td>116 (46.4)</td>
<td>41 (44.1)</td>
<td>39 (50.0)</td>
<td>0.730</td>
</tr>
<tr>
<td>Hypertension</td>
<td>31 (12.4)</td>
<td>6 (6.5)</td>
<td>10 (12.8)</td>
<td>0.045</td>
</tr>
<tr>
<td>Secundary to valvular heart disease</td>
<td>10 (4.0)</td>
<td>3 (3.2)</td>
<td>2 (2.6)</td>
<td>0.513</td>
</tr>
<tr>
<td>Cardiomyopathy</td>
<td>63 (25.2)</td>
<td>34 (36.6)</td>
<td>14 (17.9)</td>
<td>0.006</td>
</tr>
<tr>
<td>Dilated</td>
<td>49 (19.6)</td>
<td>29 (31.2)</td>
<td>13 (16.7)</td>
<td>7.893</td>
</tr>
<tr>
<td>Hypertrophic</td>
<td>7 (2.8)</td>
<td>1 (1.1)</td>
<td>0 (0)</td>
<td>0.766</td>
</tr>
<tr>
<td>Non contraction</td>
<td>4 (1.6)</td>
<td>3 (3.2)</td>
<td>0 (0)</td>
<td>1.13</td>
</tr>
<tr>
<td>Unclassified</td>
<td>3 (1.2)</td>
<td>1 (1.1)</td>
<td>1 (1.3)</td>
<td>1.300</td>
</tr>
<tr>
<td>Other</td>
<td>12 (4.8)</td>
<td>4 (4.3)</td>
<td>3 (3.8)</td>
<td>0.810</td>
</tr>
<tr>
<td>Unknown</td>
<td>18 (7.2)</td>
<td>5 (5.4)</td>
<td>10 (12.8)</td>
<td>3.068</td>
</tr>
<tr>
<td>Risk factors &amp; Medical history</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>77 (30.8)</td>
<td>21 (22.6)</td>
<td>20 (25.6)</td>
<td>0.001</td>
</tr>
<tr>
<td>Hypertension</td>
<td>113 (45.2)</td>
<td>30 (32.3)</td>
<td>35 (44.9)</td>
<td>0.001</td>
</tr>
<tr>
<td>Hypercholesterolemia</td>
<td>94 (37.6)</td>
<td>29 (31.2)</td>
<td>28 (35.9)</td>
<td>0.100</td>
</tr>
<tr>
<td>MI, PCI and/or CABG</td>
<td>123 (49.2)</td>
<td>44 (47.3)</td>
<td>40 (51.3)</td>
<td>0.874</td>
</tr>
<tr>
<td>Arrial fibrillation</td>
<td>97 (38.8)</td>
<td>23 (24.7)</td>
<td>27 (34.6)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Other arrhythmia</td>
<td>82 (32.3)</td>
<td>33 (35.5)</td>
<td>24 (31.6)</td>
<td>0.850</td>
</tr>
<tr>
<td>Pacemaker, ICD and/or CRT</td>
<td>158 (63.2)</td>
<td>66 (71.0)</td>
<td>42 (53.8)</td>
<td>0.064</td>
</tr>
<tr>
<td>Valvular heart disease</td>
<td>127 (50.8)</td>
<td>46 (49.5)</td>
<td>33 (42.3)</td>
<td>0.047</td>
</tr>
<tr>
<td>Stroke (CVA/TIA)</td>
<td>39 (15.6)</td>
<td>10 (10.8)</td>
<td>12 (15.4)</td>
<td>0.152</td>
</tr>
<tr>
<td>COPD</td>
<td>31 (12.4)</td>
<td>7 (7.5)</td>
<td>14 (17.9)</td>
<td>0.120</td>
</tr>
<tr>
<td>Sleep apnea</td>
<td>23 (9.3)</td>
<td>8 (8.9)</td>
<td>7 (9.3)</td>
<td>0.962</td>
</tr>
<tr>
<td>Medication</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACE inhibitors</td>
<td>169 (67.6)</td>
<td>63 (67.7)</td>
<td>58 (74.4)</td>
<td>0.191</td>
</tr>
<tr>
<td>ARBs</td>
<td>72 (28.8)</td>
<td>30 (32.3)</td>
<td>20 (25.6)</td>
<td>0.620</td>
</tr>
<tr>
<td>Beta blockers</td>
<td>225 (90.0)</td>
<td>81 (87.1)</td>
<td>73 (93.6)</td>
<td>0.370</td>
</tr>
<tr>
<td>Aldosterone antagonists</td>
<td>174 (69.6)</td>
<td>63 (67.7)</td>
<td>59 (75.6)</td>
<td>0.362</td>
</tr>
<tr>
<td>Diuretics</td>
<td>227 (90.8)</td>
<td>82 (88.2)</td>
<td>70 (89.7)</td>
<td>0.274</td>
</tr>
</tbody>
</table>

Values are mean (standard deviation) or median (25th–75th percentile) for continuous variables and number of patients (%) for categorical variables.

BMI = body mass index, BP = blood pressure, HF = heart failure, NYHA = New York Heart Association, LVEF = left ventricular ejection fraction, eGFR = estimated glomerular filtration rate, MI = myocardial infarction, PCI = Percutaneous Coronary Intervention, CABG = coronary artery bypass graft surgery, ICD = implantable cardioverter-defibrillator, CRT = Cardiac Resynchronization Therapy, CVA = cerebrovascular accident, TIA = transient ischemic attack, COPD = chronic obstructive pulmonary disease, ACE = angiotensin converting enzyme, ARBs = angiotensin receptor blockers.

† P < 0.05 compared to clusters 1 and 2.
‡ P < 0.05 compared to clusters 2 and 3.
§ P < 0.05 compared to cluster 2.
‡‡ P < 0.05 compared to cluster 1.

No pairwise significant differences.

LVEF was measured at baseline in a subset of 191 patients.

Myocarditis, cytotoxic agents, alcohol and peripartum cardiomyopathy.
time of one cluster compared to the average event-free survival time of the reference cluster. First, we fitted a univariable model (model 1). Then we investigated whether the association was independent of potential confounders by fitting multivariable AFT models. Covariates were chosen based on existing literature, as well as significant differences in clinical characteristics between the clusters: model 2 was corrected for age and sex; model 3 was corrected for age, sex, New York Heart Association (NYHA) Functional Classification, duration of HF and HF etiology cardiomyopathy; and model 4 was corrected for age, sex, diabetes mellitus, estimated Glomerular Filtration Rate (eGFR), body mass index (BMI), systolic blood pressure, and atrial fibrillation. Lastly, model performance was compared to that of a model containing only clinical characteristics: age, NYHA class, diabetes mellitus, eGFR, BMI, systolic blood pressure, and atrial fibrillation. Missing values in baseline covariates were handled using single imputation. AFT model performance was assessed using the Akaike Information Criterion (AIC), and a difference of two AIC units was considered significant. All analyses were performed with R Statistical Software version 4.1.1.

3. Results

3.1. Baseline characteristics

The cohort consisted of 184 (73.6%) males, and the median age was 67.6 (58.3–75.8) years. The median time since diagnosis of HF was 4.7 (1.8–9.7) years and mean left ventricular ejection fraction (LVEF) was 29.9% (SD 9.7). There were 75 (30.0%) patients in NYHA class I, 113 (45.2%) patients in class II and 62 (24.8%) patients in class III. The most frequent HF etiologies were ischemic heart disease in 116 (46.4%) patients, cardiomyopathy in 63 (25.2%) patients, and hypertension in 31 (12.4%) patients. All characteristics are presented in Table 1. The baseline biomarker levels are shown in supplemental table 2.

3.2. Clustering based on repeated biomarker measurements

The optimal number of clusters was three, and the K-means clustering algorithm resulted in stable clusters of 93, 78 and 79 patients. The biomarker profiles of the three clusters are displayed in Fig. 1. The ANOVAs showed that 70 out of 92 (76.1%) biomarker intercepts and 56 out of 92 (60.9%) biomarker slopes differed significantly between the clusters. Supplemental table 3 depicts the statistical significance of the differences per biomarker, ranked by p-value. The biomarkers that were in the top 10% both in terms of intercepts and slopes included tumor necrosis factor receptor 1, growth differentiation factor 15, lymphotoxin-beta receptor, trefoil factor 3, and interleukin-18 binding protein. Overall, cluster 3 showed higher intercepts than clusters 1 and 2, signifying higher modeled baseline biomarker levels. Cluster 2 showed larger slopes for most of the biomarkers compared to clusters 1 and 3, which signifies a larger change of levels over time. Additionally, 16 biomarkers showed larger slopes in cluster 3 compared to cluster 2. We identified these biomarkers by eyeballing the heatmap patterns (supplemental fig. 1). Several of these proteins, including NT-proBNP, ST2, GDF-15 and galectin-3, have previously been shown to be important in cardiovascular disease and HF. Based on previous literature, the biological processes in which these 16 proteins are involved include proteolysis, catabolic processes, inflammatory responses, and cardiac remodeling.

3.3. Clinical characteristics per cluster

In Table 1, baseline characteristics are presented. Additionally, an overview of the subphenotypes’ characteristics can be found in Fig. 2. Patients in cluster 1 were significantly younger than patients in clusters 2 and 3 (63.7 (53.6–69.8) vs. 65.5 (58.4–74.6) and 72.8 (66.6–80.0) years old respectively, p < 0.001). Additionally, NYHA classification differed significantly between clusters 1 and 3 (pairwise p-value = 0.004): patients in cluster 1 were significantly more often classified as NYHA class I than patients in cluster 3 (36 (38.7%) vs. 14 (17.7%)). Patients in cluster 1 also had significantly more frequently dilated cardiomyopathy as their HF etiology compared to patients in cluster 3 (29 (31.2%) vs. 7 (8.9%), pairwise p-value = 0.002). In comparison to cluster 2, cardiomyopathy in general occurred significantly more often as the etiology of HF in cluster 1 (34 (36.6%) vs. 15 (19.0%), pairwise p-
The average duration of HF differed significantly between clusters 2 and 3. Specifically, patients in cluster 2 had a significantly shorter average duration of HF than patients in in cluster 3 (2.6 (0.7–7.1) vs. 6.6 (3.5–11.5) years, pairwise p-value = 0.035).

Cluster 3 showed most significant differences in clinical characteristics compared to the other 2 clusters. When comparing cluster 3 to clusters 1 and 2, we saw that the average age was significantly higher (p < 0.001), and the average eGFR was significantly lower (50.9 (SD 21.4) vs. 67.1 (SD 21.0) and 61.4 (SD 21.3) mL/min/1.73m² respectively, p-value <0.001). Also, when comparing cluster 3 to clusters 1 and 2, the proportion of patients with a medical history of diabetes mellitus was significantly larger (36 (45.6%) vs. 21 (22.6%) and 20 (25.6%), p-value = 0.002), as well as that of atrial fibrillation (47 (59.5%) vs. 23 (24.7%) and 27 (34.6%), p-value <0.001). Moreover, the average duration of HF was significantly longer in cluster 3 compared to cluster 2 (6.6 (3.5–11.5) vs. 2.6 (0.7–7.1) years, pairwise p-value = 0.035).

Additionally, NYHA classification differed significantly between clusters 3 and 1 (pairwise p-value = 0.004): a significantly larger proportion of patients in cluster 3 were classified as NYHA class III (30 (38.0%) vs. 16 (17.2%)). Lastly, compared to cluster 1, patients in cluster 3 significantly more often had a history of hypertension (48 (60.8%) vs. 30 (32.3%), pairwise p-value = 0.001).

### 3.4. Survival analysis

Median follow-up was 2.2 (1.4–2.5) years, and a total of 66 (26.4%) patients experienced the PEP. At median follow-up, the event-free survival probability was 0.85, 0.76 and 0.53 in clusters 1, 2, and 3 respectively. Fig. 2 shows the Kaplan-Meier plot. Cluster 3 showed the worst prognosis (p < 0.001).

The results of the AFT models can be found in Table 2. According to univariable AFT model 1, the time ratio for repeated measurements cluster 3 was 0.21 (95% CI 0.11–0.40, p-value <0.001). Thus, patients in
Table 2

<table>
<thead>
<tr>
<th>Clustering based on baseline or repeated biomarker measurements, and the primary endpoint.</th>
<th>Time ratio* (95% CI)</th>
<th>P-value</th>
<th>Time ratio* (95% CI)</th>
<th>P-value</th>
<th>Time ratio* (95% CI)</th>
<th>P-value</th>
<th>AIC**</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Baseline measurements</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>Ref</td>
<td>Ref</td>
<td>0.42 (0.25-0.69)</td>
<td>0.001</td>
<td>–</td>
<td>–</td>
<td>1159.82</td>
</tr>
<tr>
<td>Model 2</td>
<td>Ref</td>
<td>Ref</td>
<td>0.45 (0.27-0.76)</td>
<td>0.003</td>
<td>–</td>
<td>–</td>
<td>1162.05</td>
</tr>
<tr>
<td>Model 3</td>
<td>Ref</td>
<td>Ref</td>
<td>0.49 (0.29-0.82)</td>
<td>0.007</td>
<td>–</td>
<td>–</td>
<td>1164.33</td>
</tr>
<tr>
<td>Model 4</td>
<td>Ref</td>
<td>Ref</td>
<td>0.55 (0.32-0.95)</td>
<td>0.033</td>
<td>–</td>
<td>–</td>
<td>1150.00</td>
</tr>
<tr>
<td><strong>Repeated measurements</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>Ref</td>
<td>Ref</td>
<td>0.60 (0.28-1.26)</td>
<td>0.180</td>
<td>0.21 (0.11-0.40)</td>
<td>&lt;0.001</td>
<td>1147.34</td>
</tr>
<tr>
<td>Model 2</td>
<td>Ref</td>
<td>Ref</td>
<td>0.58 (0.27-1.24)</td>
<td>0.160</td>
<td>0.21 (0.11-0.42)</td>
<td>&lt;0.001</td>
<td>1149.60</td>
</tr>
<tr>
<td>Model 3</td>
<td>Ref</td>
<td>Ref</td>
<td>0.58 (0.27-1.23)</td>
<td>0.157</td>
<td>0.25 (0.13-0.48)</td>
<td>&lt;0.001</td>
<td>1138.86</td>
</tr>
<tr>
<td>Model 4</td>
<td>Ref</td>
<td>Ref</td>
<td>0.61 (0.29-1.31)</td>
<td>0.204</td>
<td>0.25 (0.12-0.50)</td>
<td>&lt;0.001</td>
<td>1152.95</td>
</tr>
</tbody>
</table>

Clinical characteristics only: age, sex, NYHA class (1–2 vs 3–4), diabetes mellitus, eGFR, BMI, systolic blood pressure and atrial fibrillation. Cluster 1B and 2B are based on baseline biomarker measurements, and Cluster 1IS, Cluster 2IS and Cluster 3IS are based on intercepts and slopes from linear mixed models (using repeated measurements).

Model 1: univariable.

Model 2: corrected for age and sex.

Model 3: corrected for age, sex, NYHA class (1–2 vs 3–4), duration of HF and HF etiology cardiomyopathy.

Model 4: corrected for age, sex, diabetes mellitus, eGFR, BMI, systolic blood pressure and atrial fibrillation.

The clinical characteristics only model and model 4 should be interpreted with caution as inclusion of this multitude of variables could have resulted in overfitting.

CI = confidence interval, AIC = Akaike Information Criterion, NYHA = New York Heart Association, HF = heart failure, eGFR = estimated glomerular filtration rate, BMI = body mass index.

* Time ratio can be interpreted as the ratio of the average event-free survival time of cluster 2 or 3 compared to the average event-free survival time of cluster 1.

** Lower AIC indicates better model fit.

cluster 3 had a statistically significantly poorer prognosis than patients in cluster 1. This difference persisted when correcting for potential confounders in models 2 to 4, with model 4 being the most elaborate. According to model 4, the adjusted time ratio for cluster 3 compared to cluster 1 was 0.25 (95% CI 0.12–0.50, p-value <0.001). The time ratios for cluster 2 compared to cluster 1 were not statistically significant in any AFT model. All models including cluster membership (models 1–4) showed a better model fit than the clinical characteristics only model.

3.5. Comparison to single biomarker measurements

For convenience, the clusters based on intercepts and slopes from the repeated measurements mixed models described so far, will be referred to as cluster 1IS, cluster 2IS, and cluster 3IS here. When only using baseline biomarker measurements, 2 clusters of sizes 141 and 109 were found: cluster 1B and cluster 2B. Cluster 2B had higher, and cluster 1B lower baseline biomarker levels (supplemental fig. 2). Cluster 1B mostly overlapped with cluster 1IS and 2IS, while cluster 2B mostly overlapped with cluster 3IS (supplemental fig. 3). Clinical characteristics of clusters 1B and 2B showed similar patterns to those of the repeated measurements clusters, cluster 2B most resembling cluster 3IS (supplemental table 4). Cluster 2B had a significantly poorer prognosis compared to cluster 1B (supplemental fig. 4), also when correcting for potential confounders. Model 3 (corrected for age, sex, NYHA class, duration of HF and HF etiology cardiomyopathy) was the only model that, according to the AIC values, showed a better model fit than the clinical characteristics only model. Notably, the repeated measurements AFT models all had better model fit than the baseline AFT models (Table 2).

Clustering based on linear mixed models with random intercepts and slopes from repeated measurements for clustering, which accounts for the dynamic nature of HF and therefore may uncover additional properties of HFrEF subphenotypes. The clusters showed significant differences in clinical characteristics and survival. Moreover, survival models including clusters based on repeated biomarker measurements showed better model fit than models including clusters based on baseline biomarker measurements only, as well as a model containing clinical characteristics only. Altogether, our results underscore that identification of subphenotypes based on repeated measurements of a large set of cardiovascular-related biomarkers carries potential to improve prognostication of HFrEF patients.

4. Discussion

In this investigation, we used a set of 92 repeatedly measured blood biomarkers in a cohort of 250 HFrEF patients. Based on their temporal biomarker patterns, we identified three HFrEF clusters, or subphenotypes. The clusters showed significant differences in clinical characteristics and survival. Moreover, survival models including clusters based on repeated biomarker measurements showed better model fit than models including clusters based on baseline biomarker measurements only, as well as a model containing clinical characteristics only. Altogether, our results underscore that identification of subphenotypes based on repeated measurements of a large set of cardiovascular-related biomarkers carries potential to improve prognostication of HFrEF patients.

Proteomics phenotyping is upcoming in patients with heart failure with preserved ejection fraction (HFpEF) [9–11,13]. However, studies on HFrEF subphenotypes are still scarce. Ahmad et al. [5] clustered 1619 HFrEF patients based on 45 baseline clinical variables and found 4 clusters. Tromp et al. [12] found 6 clusters based on 92 single ‘baseline’ cardiovascular biomarker measurements in 1802 HFrEF patients. Similarly to our investigation, these studies found an association between cluster membership and prognosis. These studies based their clustering on patient characteristics, medical history, quality of life scores, exercise capability, and baseline biomarker measurements. The current investigation extends their findings by using repeated biomarker measurements for clustering, which accounts for the dynamic nature of HF and therefore may uncover additional properties of HFrEF subphenotypes.

Cluster 3 showed the highest average age, a longer duration of HF, more advanced HF, lower eGFR levels, and more pre-existing diabetes mellitus, hypertension, and atrial fibrillation. As well as the highest baseline biomarker levels and the worst prognosis. Comparing our results to previous studies, a ‘phenotype’ similar to our cluster 3 was found by Ahmad et al. [5]. Patients in their cluster 1 were older, had pre-existing hypertension, diabetes mellitus and atrial fibrillation, and had more advanced HF. Ahmad et al. also found higher levels of some blood biomarkers (NT-proBNP, galectin-3 and ST2) and the worst prognosis in this cluster. The clusters found by Tromp et al. [12] had less similar characteristics, with ‘endotype’ 4 most resembling cluster 3 of the current study. Namely, endotype 4 more often showed pre-existing atrial fibrillation, had higher NT-proBNP levels and had the worst prognosis.
To our knowledge, this is the first study that associates HFrEF subphenotypes based on temporal evolutions of circulating proteins with prognosis. Circulating protein levels fluctuate over time and repeated protein measurements may account for these fluctuations and thus enhance the prognostic value of the ensuing subphenotypes. Indeed, our prognostic models containing clusters that were based on full temporal trajectories of biomarkers, performed better compared to those containing clusters based on biomarkers measured at one single time-point. This demonstrates the advantage of using temporal biomarker trajectories in subphenotyping of HFrEF.

In addition, we found interesting differences between the clusters’ dynamic biomarker profiles. Cluster 2 showed relatively low modeled baseline biomarker levels, with increases in levels of the majority of the biomarkers over time, while cluster 3 showed higher modeled baseline biomarker levels, and increasing levels over time of the 16 remaining biomarkers. Considering the lower average age, shorter duration of HF, and less advanced HF in cluster 2 compared to cluster 3, a tentative hypothesis may be that these phenotypes represent chronologically and less advanced HF in cluster 2 compared to cluster 3, a tentative proportion of women is limited.

Phenotype 3 is reached, whereafter the remaining biomarker levels rise over time in patients with phenotype 2, until the high ‘baseline phenotype 3 (more advanced HF); i.e., the biomarkers might increase successive states. Phenotype 2 (milder HF) could transition into other strategies in subphenotyping of HFrEF.

In conclusion, we identified three HFrEF subphenotypes based on a significant differences in clinical characteristics and are significantly associated with adverse events. Thus, phenotyping based on temporal biomarker patterns carries potential to contribute to improved prognostication and personalized risk assessment.

Funding
This work was supported by the Jaap Schouten Foundation and the Noordwest Academie. The funding sources had no role in the study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Declaration of Competing Interest
The authors report no relationships that could be construed as a conflict of interest.

Acknowledgements
None.

Appendix A. Supplementary data
Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijcard.2022.06.020.

References


