ESC European Society of Cardiology

European Heart Journal - Digital Health (2023) **4**, 444–454
https://doi.org/10.1093/ehjdh/ztad056

ORIGINAL ARTICLE

Machine learning–based biomarker profile derived from 4210 serially measured proteins predicts clinical outcome of patients with heart failure

Marie de Bakker[1], Teun B. Petersen[1,2], Anja J. Rueten-Budde[2], K. Martijn Akkerhuis[1], Victor A. Umans[3], Jasper J. Brugts[1], Tjeerd Germans[3], Marcel J.T. Reinders[4], Peter D. Katsikis[5], Peter J. van der Spek[6], Rachel Ostroff[7], Ruicong She[8], David Lanfear[9,10], Folkert W. Asselbergs[11,12], Eric Boersma[1], Dimitris Rizopoulos[2,13], and Isabella Kardys[1,*]

[1]Department of Cardiology, Erasmus MC, University Medical Center Rotterdam, Dr. Molenwaterplein 40, 3015GD, Rotterdam, The Netherlands; [2]Department of Biostatistics, Erasmus MC, University Medical Center Rotterdam, Dr. Molenwaterplein 40, 3015GD, Rotterdam, The Netherlands; [3]Department of Cardiology, Northwest Clinics, Wilhelminalaan 12, 1815 JD, Alkmaar, The Netherlands; [4]Delft Bioinformatics Lab, Delft University of Technology, Van Mourik Broekmanweg 6, 2628 XE, Delft, The Netherlands; [5]Department of Immunology, Erasmus MC, University Medical Center Rotterdam, Dr. Molenwaterplein 40, 3015GD, Rotterdam, The Netherlands; [6]Department of Pathology, Erasmus MC, University Medical Center Rotterdam, Dr. Molenwaterplein 40, 3015GD, Rotterdam, The Netherlands; [7]SomaLogic, Inc., 2945 Wilderness Pl., Boulder, CO 80301, USA; [8]Department of Public Health Sciences, Henry Ford Health System, 1 Ford Pl, Detroit, MI 48202, USA; [9]Center for Individualized and Genomic Medicine Research (CIGMA), Henry Ford Hospital, 2799 W. Grand Boulevard, Detroit MI, 48202, USA; [10]Heart and Vascular Institute, Henry Ford Hospital, 2799 W. Grand Boulevard, Detroit, MI 48202, USA; [11]Amsterdam University Medical Centers, Department of Cardiology, University of Amsterdam, Meibergdreef 9, 1105 AZ, Amsterdam, The Netherlands; [12]Health Data Research UK and Institute of Health Informatics, University College London, Gower St, London, WC1E 6BT, UK; and [13]Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Dr. Molenwaterplein 40, 3015GD, Rotterdam, The Netherlands

Received 21 June 2023; revised 6 September 2023; accepted 3 October 2023; online publish-ahead-of-print 4 October 2023

Aims

Risk assessment tools are needed for timely identification of patients with heart failure (HF) with reduced ejection fraction (HFrEF) who are at high risk of adverse events. In this study, we aim to derive a small set out of 4210 repeatedly measured proteins, which, along with clinical characteristics and established biomarkers, carry optimal prognostic capacity for adverse events, in patients with HFrEF.

Methods and results

In 382 patients, we performed repeated blood sampling (median follow-up: 2.1 years) and applied an aptamer-based multiplex proteomic approach. We used machine learning to select the optimal set of predictors for the primary endpoint (PEP: composite of cardiovascular death, heart transplantation, left ventricular assist device implantation, and HF hospitalization). The association between repeated measures of selected proteins and PEP was investigated by multivariable joint models. Internal validation (cross-validated *c*-index) and external validation (Henry Ford HF PharmacoGenomic Registry cohort) were performed. Nine proteins were selected in addition to the MAGGIC risk score, N-terminal pro-hormone B-type natriuretic peptide, and troponin T: suppression of tumourigenicity 2, tryptophanyl-tRNA synthetase cytoplasmic, histone H2A Type 3, angiotensinogen, deltex-1, thrombospondin-4, ADAMTS-like protein 2, anthrax toxin receptor 1, and cathepsin D. N-terminal pro-hormone B-type natriuretic peptide and angiotensinogen showed the strongest associations [hazard ratio (95% confidence interval): 1.96 (1.17–3.40) and 0.66 (0.49–0.88), respectively]. The multivariable model yielded a *c*-index of 0.85 upon internal validation and *c*-indices up to 0.80 upon external validation. The *c*-index was higher than that of a model containing established risk factors ($P = 0.021$).

Conclusion

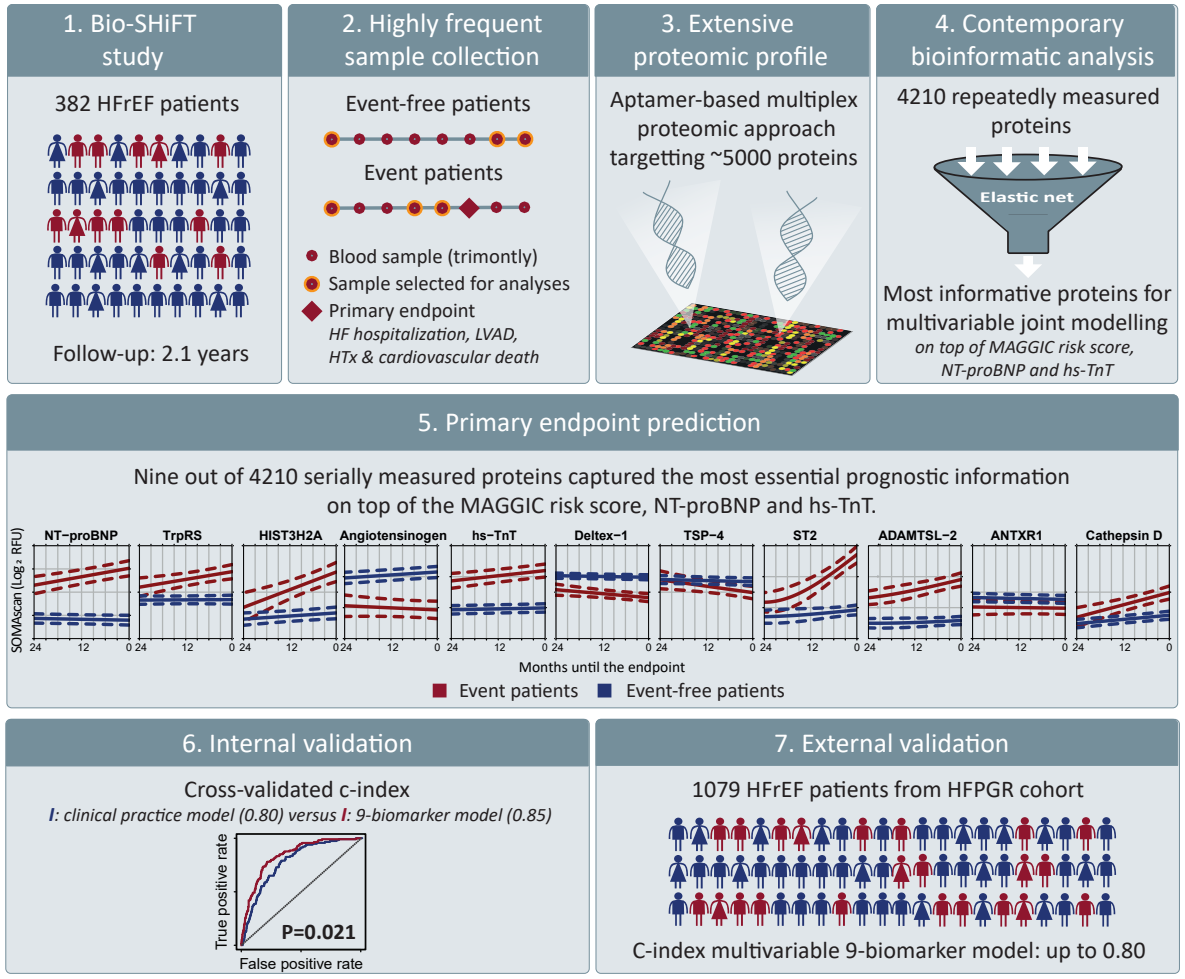
Nine serially measured proteins captured the most essential prognostic information for the occurrence of adverse events in patients with HFrEF, and provided incremental value for HF prognostication beyond established risk factors. These proteins

* Corresponding author. Tel: +31 6 50032051, Email: i.kardys@erasmusmc.nl

could be used for dynamic, individual risk assessment in a prospective setting. These findings also illustrate the potential value of relatively 'novel' biomarkers for prognostication.

Clinical Trial Registration https://clinicaltrials.gov/ct2/show/NCT01851538?term=nCT01851538&draw=2&rank=1 24

Graphical Abstract



Keywords Heart failure • Proteomics • Repeated measurements • Prediction • Elastic net • NT-proBNP

Introduction

Despite recent advances in the management of patients with heart failure (HF) with reduced ejection fraction (HFrEF), mortality and re-hospitalization rates remain high.[1,2] Blood biomarkers may contribute to risk assessment, and consequently, improve the timing of treatment.[3] Accordingly, current guidelines recommend the use of several biomarkers for prognostication and risk stratification.[1,2]

Nevertheless, the poor prognosis of HFrEF illustrates that there is still room for improvement in personalized risk assessment. Given the complex pathophysiology of HFrEF, deep multiple-marker testing panels carry potential for such improved risk assessment. Recently, robust affinity-based methods have been developed to systematically assay elaborate sets (>1000) of circulating proteins that represent various biological processes.[4] Deriving a reduced number of proteins that capture essential prognostic information from such a large set could provide an efficient yet comprehensive approach to prognostication. In the context of HF, so far there are only a few studies that have investigated the value of comprehensive proteomic assays for prognostication. Cuvelliez *et al.*[5] examined 1310 circulating proteins and identified 6 that were associated with cardiovascular death in patients with systolic HF. A recent study showed that a baseline multiprotein score, consisting of 8 circulating proteins selected out of nearly 5000, improved risk stratification for all-cause mortality in

patients with HFrEF.[6] A baseline 27-protein model improved 4-year risk prediction of major cardiovascular outcomes, including HF hospitalization, and all-cause mortality in high-risk populations.[7] Further studies are warranted.

Identifying patient-specific temporal biomarker evolutions may provide another opportunity to improve individualized risk assessment. Previous studies on the prognostic value of circulating proteins in HF have generally performed measurements at study baseline and related them to adverse events occurring over many years thereafter or have left several years in between (usually two) repeated measurements.[8] However, given the dynamic nature of HF, distinguishing patients at different levels of risk of adverse events based on a single biomarker measurement, or based on a simplified representation of temporal biomarker evolution, is challenging.

Therefore, we have performed serial measurements of an elaborate set of 4210 circulating proteins in 382 patients with stable HFrEF in order to develop a comprehensive, dynamic prediction model that adds to conventional risk estimation. We aimed to derive a smaller set of up to 10 out of the 4210 proteins, with optimal prognostic capacity for adverse clinical events, which along with clinical characteristics and established biomarkers [i.e. N-terminal pro-hormone B-type natriuretic peptide (NT-proBNP), and high-sensitivity troponin T (hs-TnT)],[9] could be used for dynamic, individual risk assessment in clinical practice. The final set containing the most informative proteins was externally validated in a cohort of 1079 patients.

Methods

Study population

The Serial Biomarker Measurements and New Echocardiographic Techniques in Chronic Heart Failure Patients Result in Tailored Prediction of Prognosis (Bio-SHiFT) study is a prospective cohort study of stable patients with chronic heart failure (CHF), conducted at Erasmus MC, Rotterdam, and Northwest Clinics, Alkmaar, The Netherlands. The study design has been described in more detail previously.[10] In brief, patients were recruited during their regular outpatient visits and were included if they were aged ≥18 years, capable of understanding and signing informed consent, diagnosed with CHF ≥3 months before inclusion according to the European Society of Cardiology (ESC) guidelines,[11,12] and if they had not been hospitalized for HF in the past 3 months. Study follow-up visits were predefined and scheduled every 3 months (±1 month). At baseline and at each follow-up visit, a short medical evaluation was performed by a research physician or research nurse, blood samples were collected, and the occurrence of adverse cardiovascular events since the last visit was recorded. According to the ESC guidelines, the routine outpatient follow-up and treatment by the treating physician continued in parallel with the study visits. The medical ethics committee of the Erasmus Medical Center in Rotterdam approved the study protocol, and all patients provided written informed consent. The study was conducted in accordance with the Declaration of Helsinki and registered in ClinicalTrial.gov (NCT01851538). Between August 2011 and January 2018, a total of 398 patients with CHF were enrolled. In the current investigation, 382 patients with HFrEF were evaluated.

Baseline assessment

All patients were evaluated by research physicians, who collected information on HF-related symptoms and New York Heart Association (NYHA) class and performed a physical examination. Information on HF aetiology, left ventricular ejection fraction (LVEF), cardiovascular risk factors, medical history, and treatment was retrieved primarily from hospital records and was checked in case of ambiguities.

Sample collection and processing

Blood samples were collected at baseline and at each study follow-up visit. Within 2 h after collection, blood samples were processed, and ethylenediaminetetraacetic acid (EDTA) plasma was stored at −80°C. Accordingly, at the time of the outpatient visits, results of the proteomic analysis were not available to treating physicians. Laboratory personnel were blinded to clinical data and patient outcomes. For the current investigation, all baseline blood samples were selected. Additionally, the last two samples drawn before the occurrence of the primary endpoint (PEP), or the last two samples that were available before censoring for patients who remained endpoint free, were selected (visualized in Supplementary material online, *S1*). In total, 1070 samples were available for the current study and 86% of the patients with HFrEF had three samples available for analysis. As per study design, 'missing' samples (i.e. availability of less than three samples per patient) could occur only when the PEP or censoring occurred before the selected second or third, prescheduled study visit. In total, 327 (86%) patients had 3 available samples, 30 (8%) patients had 2 available samples, and 25 (7%) patients had 1 available sample. Previous investigations using all available samples in our patient cohort have demonstrated that the concentration of several plasma and urine biomarker candidates changes in the months preceding the occurrence of an adverse event.[10,13] By selecting the last two samples prior to the incident study endpoint, we aimed to capture these changes while improving efficiency.

Proteomic analysis

Plasma protein concentrations were measured in a single batch using the aptamer-based proteomic SOMAscan platform (Somalogic, Boulder, CO, USA), as previously described.[14] SOMAscan utilizes single-stranded DNA-based protein affinity reagents called Slow Off-rate Modified Aptamers (SOMAmers). The SOMAmers bind proteins with high specificity and affinity, and slow dissociation rates, minimizing non-specific binding interactions. The readout of the SOMAscan assay is in normalized relative fluorescent units (RFUs). These intensities are directly proportional to the amount of target protein in the initial sample. Previous studies reported high assay reproducibility and low technical variability of SOMAscan.[15,16]

Somalogic's previously described standard processes for normalization, calibration, and quality control were followed (see Supplementary material online, *S2*). The SOMAmers with non-human and/or not validated targets were excluded. In addition, when multiple SOMAmer versions were present, those with the highest binding affinity were used. Thus, out of the total 5284 modified aptamers, aptamers against 4210 proteins were included in the current analyses. Individual sample quality was judged by comparing normalized median signal relative to the external reference standard. Data from 1066 samples passed quality-control criteria.

Clinical study endpoints

A clinical event committee, blinded by the proteomic results, reviewed hospital records and discharge letters and adjudicated the study endpoints. The PEP comprised the composite of cardiovascular death, heart transplantation (HTx), left ventricular assist device (LVAD) implantation, and hospitalization to manage acute or worsened HF. In patients who reached multiple endpoints, only the first was used for analysis. Hospitalization for acute or worsened HF was defined as hospitalization for an exacerbation of HF symptoms, in combination with two of the following: brain natriuretic peptide or NT-proBNP > 3× normal upper limit, signs of worsening HF, such as pulmonary rales, raised jugular venous pressure or peripheral oedema, increased dose or intravenous administration of diuretics, or administration of positive inotropic agents.[11]

External validation cohort

For external validation, data from the Henry Ford HF PharmacoGenomic Registry (HFPGR) were used. The HFPGR is a prospective observational registry of patients in the Henry Ford Health System in Detroit, MI, USA. The study was approved by the Henry Ford Hospital Institutional Review Board, and all participants provided written informed consent. The study design has previously been described in detail.[17] Briefly, patients were included if they were aged ≥18 years, insured, and met the definition of HF as defined by the Framingham Heart Study. For the current investigation, only patients with HFrEF and SOMAscan measurements were included ($n = 1079$). A composite of cardiovascular death and hospitalization with a primary discharge diagnosis of HF was examined.

Statistical analysis

A full description of the statistical analyses is given in Supplementary material online, *S3*. First, differences in estimated baseline concentration and evolution during follow-up between patients who reached the PEP and those who remained endpoint free were evaluated using linear mixed-effect (LMEs) modelling adjusted for the MAGGIC risk score.[18] By using LME modelling, we effectively account for the inherent correlation among repeated measurements within each patient.

Next, protein concentrations estimated by unadjusted LME models were standardized and used in the regularized time-varying Cox regression model. We used a regularized time-varying Cox regression model with an elastic net penalty to select a lean subset of circulating proteins with strong prognostic value on top of clinical predictors and established biomarkers. In particular, the Cox model included the MAGGIC risk score, repeated measures of clinical NT-proBNP, and hs-TnT (all unpenalized) as well as standardized values of all 4210 repeatedly measured SOMAscan proteins (except for NT-proBNP and TnT; penalized). For the selection of the optimal shrinkage penalty (λ), 10-fold cross-validation of the model was performed.

The discriminative ability of the final prediction model (i.e. the model containing the MAGGIC risk score, NT-proBNP, hs-TnT, and circulating proteins selected in the previous step) was determined using internal and external validation. For internal validation, a 10-fold cross-validated concordance index (*c*-index) was used. We compared the discriminative ability to that of a model containing clinical characteristics (i.e. the MAGGIC risk score), a clinical practice–based biomarker model (i.e. the MAGGIC risk score, NT-proBNP, and hs-TnT) and a literature-based biomarker model [i.e. the MAGGIC risk score, NT-proBNP, hs-TnT, C-reactive protein (CRP), growth/differentiation factor 15 (GDF-15), interleukin 1 receptor-like 1 (ST2), and galectin-3 (Gal-3)]. For external validation, we calculated the *c*-index in HFPGR to determine the discriminative ability of the final prediction model as derived from Bio-SHiFT and the models described above.

Because the regularized Cox model optimizes predictive ability at the expense of biasedly estimated regression coefficients, to estimate the unbiased associations between the repeatedly measured variables selected in the previous step, and the PEP, we applied joint modelling (JM). Both univariable and multivariable JMs were used. Results are given as hazard ratios (HRs) and 95% confidence intervals (CIs) per 1 standard deviation (SD) difference of the absolute $\log_2$-tranformed protein intensities at any point in time during follow-up.

All analyses were performed in R version 4.0.3. A two-sided *P*-value <0.05 or false discovery rate (FDR) < 0.05 was considered statistically significant, depending on the context.

Results

Baseline characteristics

In total, 382 patients with stable HFrEF were included in the current study. The mean (±SD) age was 63.3 (±13.1) years and 72.8% were males (*Table 1*). Patients who experienced the PEP during follow-up had significantly lower systolic blood pressure, higher NYHA class, lower LVEF, longer duration of HF at baseline, and were more frequently on diuretics and anticoagulants compared with patients who remained endpoint free. Moreover, the prevalence of comorbidities was also higher in those patients, as were the baseline levels of 'established' biomarkers.

Follow-up and study endpoints

During a median (25th–75th percentile) follow-up of 25 (13–31) months, a total of 114 (29.8%) patients reached the PEP. Specifically, 90 patients were re-hospitalized for acute or worsened HF, 13 patients underwent LVAD placement, 17 patients underwent HTx, and 33 patients died of cardiovascular causes. Only the first event was used for analysis in patients who reached multiple endpoints during follow-up (39 out of 114). Hence, the PEP consisted of 90 patients with re-hospitalization for HF, 6 patients with LVAD placement, 10 patients with HTx, and 8 who died of cardiovascular causes.

Longitudinal evolution of protein concentrations

A statistically significant difference in estimated protein concentration at baseline between patients who reached the PEP and patients who remained endpoint free was shown for 1214 out of the 4210 proteins after adjustment for the MAGGIC risk score and correction for multiple testing. The top 100 proteins with the largest difference at baseline are presented in Supplementary material online, *S4*. Moreover, 356 proteins showed a significantly different evolution during follow-up between patients with and without the PEP after correction for multiple testing. Proteins that showed at least a 10% relative difference in trajectory per year are depicted in *Figure 1* and Supplementary material online, *S5*. The majority of these proteins showing a significantly different evolution during follow-up were involved in biological mechanisms that have been implicated in HF, for example, cardiac stress (e.g. NT-proBNP), cardiac remodelling (e.g. ST2 and GDF-15), inflammation (e.g. CRP and serum amyloid A), iron homeostasis (e.g. transferrin receptor), oxidative stress (e.g. peroxidasin and erythropoietin), and cholesterol metabolism (e.g. apolipoprotein F).

Prognostic value of serially measured protein expression

Penalized multivariable time-dependent Cox regression using repeated measurements of 4210 circulating proteins resulted in a selection of 9 proteins on top of the MAGGIC risk score, NT-proBNP and hs-TnT: ST2, tryptophanyl-tRNA synthetase cytoplasmic (TrpRS), histone H2A Type 3 (HIST3H2A), angiotensinogen, deltex-1, thrombospondin-4 (TSP-4), ADAMTS-like protein 2 (ADAMTSL-2), anthrax toxin receptor 1 (ANTXR1), and cathepsin D. These results imply that for optimal prognostic performance, a model containing the MAGGIC risk score, serially measured NT-proBNP and hs-TnT, and these nine repeatedly measured circulating proteins suffices. *Figure 2* shows the average temporal patterns of NT-proBNP and hs-TnT as well as these nine proteins in patients with and without the PEP during follow-up. Twenty-four months before the occurrence of the PEP, levels of NT-proBNP, TrpRS, hs-TnT, and ADAMTSL-2 were already higher in patients who ultimately reached the PEP compared with patients who remained event free. Furthermore, NT-proBNP significantly increased as the endpoint approached but remained stable in endpoint-free patients. HIST3H2A, ST2, and cathepsin D showed similar patterns, although sometimes less pronounced. In contrast, deltex-1 and TSP-4 decreased as the endpoint approached but did remain stable in endpoint-free patients. During follow-up, levels of angiotensinogen and ANTXR1 remained lower in patients who ultimately reached the PEP compared with patients who remained event free.

The multivariable regression model, including the MAGGIC risk score and serial measurements of NT-proBNP, hs-TnT, and the nine circulating proteins, showed high discriminative ability, with a cross-validated *c*-index of 0.85 (*Table 2*, Model 7). The *c*-index was significantly higher than that of a model containing baseline clinical predictors and repeated measurements of established biomarkers NT-proBNP and hs-TnT (*c*-index: 0.80, $P = 0.021$; *Table 2*, Model 5). In contrast, the discriminative ability of a model containing clinical predictors and biomarkers previously shown to be involved in HF (NT-proBNP, hs-TnT, CRP, GDF-15, ST2, and Gal-3; *Table 2*, Model 6; *c*-index: 0.83) did not significantly improve performance compared with this clinical practice–based model ($P = 0.087$). Moreover, the *c*-indices of all multivariable models including repeated measurements exceeded the *c*-indices of models consisting of only baseline measurements of the same respective sets of biomarkers (*Table 2*, Models 2–4).

Table 1 Baseline characteristics of the study population

	Total population ($n = 382$)	No endpoint ($n = 268$)	Endpoint ($n = 114$)	*P*-value
Demographics				
Age [mean (SD)]	63.3 (13.1)	63.0 (12.7)	64.0 (14.0)	0.486
Sex (% male)	278 (72.8)	187 (69.8)	91 (79.8)	0.058
Caucasian ethnicity (%)	351 (92.6)	246 (92.8)	105 (92.1)	0.973
Clinical characteristics				
Body mass index, kg/m^2 [mean (SD)]	27.2 (4.5)	27.4 (4.6)	26.7 (4.4)	0.211
Systolic blood pressure, mmHg [mean (SD)]	115.3 (21.3)	118.2 (21.6)	108.5 (19.1)	**<0.001**
Diastolic blood pressure, mmHg [mean (SD)]	70.0 (10.5)	71.0 (10.9)	67.5 (9.2)	**0.003**
Established biomarker levels				
NT-proBNP (pmol/L)[a]	145.0 (54.7–289.0)	95.3 (34.0–211.1)	297.4 (179.9–524.6)	**<0.001**
Hs-TnT (ng/L)[a]	18.0 (10.3–34.0)	14.5 (8.9–25.0)	32.0 (20.0–49.5)	**<0.001**
CRP (mg/L)[a]	2.0 (0.9–4.7)	1.6 (0.8–3.8)	2.8 (1.4–5.4)	**0.004**
Features of heart failure				
Duration of HF, years[a]	4.2 (1.6–9.5)	3.7 (1.2–7.8)	5.7 (2.6–13.0)	**<0.001**
NYHA class (%)				**<0.001**
NYHA Classes I and II	276 (72.6)	214 (80.1)	62 (54.9)	
NYHA Classes III and IV	104 (27.4)	53 (19.9)	51 (45.1)	
LVEF [mean (SD)][b]	29.8 (10.3)	31.2 (9.9)	25.6 (10.1)	**<0.001**
Heart failure aetiology				
Ischaemic heart disease (% yes)	166 (43.5)	113 (42.2)	53 (46.5)	0.504
Hypertension (% yes)	33 (8.6)	25 (9.3)	8 (7.0)	0.592
Secondary to valvular heart disease (% yes)	12 (3.1)	6 (2.2)	6 (5.3)	0.219
Cardiomyopathy (% yes)	122 (31.9)	82 (30.6)	40 (35.1)	0.458
Unknown aetiology (% yes)	27 (7.1)	24 (9.0)	3 (2.6)	**0.047**
Other aetiology (% yes)	26 (6.8)	21 (7.8)	5 (4.4)	0.316
Medical history				
Myocardial infarction (% yes)	145 (38.5)	96 (36.5)	49 (43.0)	0.283
PCI (% yes)	126 (33.0)	88 (32.8)	38 (33.3)	1.000
CABG (% yes)	54 (14.1)	35 (13.1)	19 (16.7)	0.444
Atrial fibrillation (% yes)	137 (36.3)	80 (30.3)	57 (50.4)	**<0.001**
CRT (% yes)	113 (29.7)	74 (27.6)	39 (34.5)	0.221
Pacemaker (% yes)	85 (23.0)	54 (20.7)	31 (28.7)	0.127
Chronic renal failure (% yes)	181 (47.6)	112 (41.9)	69 (61.1)	**0.001**
Diabetes mellitus (% yes)	98 (25.7)	61 (22.8)	37 (32.5)	0.063
Known hypercholesterolaemia (% yes)	160 (42.9)	107 (41.2)	53 (46.9)	0.359
COPD (% yes)	50 (13.4)	29 (11.1)	21 (18.8)	0.063
Intoxications				
Smoking (%)				0.330
Never	109 (28.7)	81 (30.3)	28 (24.8)	
Current	37 (9.7)	28 (10.5)	9 (8.0)	
Former (>30 days)	234 (61.6)	158 (59.2)	76 (67.3)	
Medication use				
Beta-blockers (% yes)	350 (91.9)	249 (93.3)	101 (88.6)	0.187
ACE-I (% yes)	258 (67.7)	187 (70.0)	71 (62.3)	0.173
ARB (% yes)	107 (28.0)	75 (28.0)	32 (28.1)	1.000
Aldosterone antagonist (% yes)	293 (76.7)	199 (74.3)	94 (82.5)	0.109
Loop diuretics (% yes)	353 (92.4)	241 (89.9)	112 (98.2)	**0.009**
Thiazide diuretics (% yes)	12 (3.1)	4 (1.5)	8 (7.0)	**0.012**
Aspirin (% yes)	77 (20.2)	59 (22.0)	18 (15.9)	0.226

Continued

Table 1 ***Continued***

	Total population ($n = 382$)	No endpoint ($n = 268$)	Endpoint ($n = 114$)	*P*-value
Anticoagulants (% yes)	279 (73.0)	185 (69.0)	94 (82.5)	**0.010**
MAGGIC risk score	20.3 (7.2)	18.6 (6.8)	24.4 (6.4)	**<0.001**

A *P*-value <0.05 is considered statistically significant and presented in bold typeface.
ACE, angiotensin-converting enzyme; ARB, angiotensin receptor blocker; CABG, coronary artery bypass graft; COPD, chronic obstructive pulmonary disease; CRP, C-reactive protein; CRT, cardiac resynchronization therapy; hs-TnT, high-sensitivity troponin T; LVEF, left ventricular ejection fraction; NT-proBNP, N-terminal pro-hormone B-type natriuretic peptide; NYHA, New York Heart Association; PCI, percutaneous coronary intervention; SD, standard deviation.
[a]All biomarker levels and duration of heart failure are presented as median (25th–75th percentile).
[b]Missing for 81 patients.

Although from a statistical point of view, the *P*-value is not the deciding criterion in elastic net regression, all nine circulating proteins were significantly associated with the occurrence of the PEP in the univariable joint models (*Table 3*). In the multivariable joint model, NT-proBNP [HR (95% CI): 1.96 (1.17–3.40), $P = 0.006$] and angiotensinogen [0.66 (0.49–0.88), $P = 0.002$] were statistically significantly associated with the PEP. N-terminal pro-hormone B-type natriuretic peptide showed the association that was numerically the strongest, and implied that if a patient has a 1 SD higher NT-proBNP level compared with another patient at any point in time, the HR for that patient for the PEP is 1.96, independent of all other proteins.

External validation

Patients in the HFPGR were on average 67.5 years old, and 65.1% were males (see Supplementary material online, *S6*). During an average follow-up duration of 3.8 years, 320 (29.7%) patients died.

The discriminative ability of our nine-biomarker prediction model was high in the HFPGR, with a *c*-index of 0.79 in the total cohort and 0.80 in Caucasian patients only for the composite of cardiovascular death and HF hospitalization (*Table 2* and Supplementary material online, *S7*). Moreover, the *c*-index of our nine-biomarker model was significantly higher than that of a model containing clinical predictors and measurements of established biomarkers NT-proBNP and TnT in the total cohort as well as in Caucasian patients only ($P < 0.001$ and $P = 0.012$, respectively). The discriminative ability of a model containing clinical predictors and biomarkers previously shown to be involved in HF did not significantly improve compared with the clinical practice–based model in the total cohort as well as in Caucasian patients only ($P = 0.299$ and $P = 0.155$, respectively).

Discussion

We performed a prospective study with repeated measurements of 4210 proteins in almost 400 patients with HFrEF. Nine repeatedly measured proteins derived by machine-learning methods showed to carry the most essential prognostic information for adverse cardiovascular events on top of conventional risk factors, including clinical characteristics and established biomarkers NT-proBNP and hs-TnT. Discriminative ability of this set was high, with a cross-validated *c*-index of 0.85 upon internal validation and *c*-indices up to 0.80 upon external validation. The set consisted of established as well as 'novel' markers in the context of HF, namely ST2, TrpRS, HIST3H2A, angiotensinogen, deltex-1, TSP-4, ADAMTSL-2, ANTXR1, and cathepsin D.

Large-scale proteomic approaches, such as aptamer-based multiplex platforms, are becoming increasingly important in the discovery of biomarkers relevant for disease, and individual proteomic signatures may contribute to improved personalized risk assessment. To the best of our knowledge, this is the first study that uses repeatedly applied affinity-based methods, which target nearly 5000 proteins, to construct a prognostic model in patients with HFrEF. By performing repeated blood sampling at fixed 3-month intervals over the full course of the follow-up, our study extends current knowledge while addressing previous limitations. So far, studies on circulating proteins in HF have generally examined limited numbers of proteins at a time, have traditionally performed measurements at study baseline only, or have left several years between (usually at most two) repeated measurements. In contrast, our study design enabled us to account for the dynamic nature of HF and to select blood samples taken on average several weeks before the occurrence of an adverse event, hence describing the temporal protein evolution in a detailed and appropriate way. Importantly, the model we applied also enables subsequent use of the derived risk estimates for dynamic, individual risk assessment in a prospective setting. Specifically, based on individual temporal protein trajectories, personalized screening intervals (i.e. the optimal timing of the patient's next outpatient visit) can be calculated using the joint model as we have described earlier.[19] Such a dynamic risk prediction tool provides an individualized approach to the timing of treatment adaptations for patients with HFrEF, in order to improve outcomes and provide additional public health benefits.

Our data were both elaborate (4210 proteins) and correlated (repeated measurements). By using sophisticated statistical techniques, we were able to derive the most informative proteins from the comprehensive, repeatedly measured panel, and to subsequently use the obtained model for dynamic, individualized prognostication. We determined the nine-biomarker model based on the elastic net penalization method with cross-validation, ensuring that its prognostic capacity was maximized, and internally validated the model by cross-validation. The discriminative ability of our nine-biomarker model was significantly higher than that of a model containing baseline clinical predictors and repeated measurements of established biomarkers NT-proBNP and hs-TnT (i.e. a clinical practice–based model), whereas the discriminative ability of a model containing clinical predictors and biomarkers previously shown to be involved in HF (NT-proBNP, hs-TnT, CRP, GDF-15, ST2, and Gal-3) did not significantly improve compared with this clinical practice–based model. Despite our nine-biomarker model significantly improving *c*-indices compared with a clinical practice–based model, improvements in absolute terms were modest. It has previously been demonstrated that, although frequently used in clinical practice, *c*-indices are very conservative and thus rather insensitive to improvements in predictive ability.[20] The discriminative ability of our model being higher than that of a set of biomarkers previously shown to be involved in HF, illustrates the potential value of relatively 'novel' biomarkers for prognostication. It should be noted that during internal validation, the discriminative ability of our model was determined using the same data set that the proteins were derived from. Nonetheless,

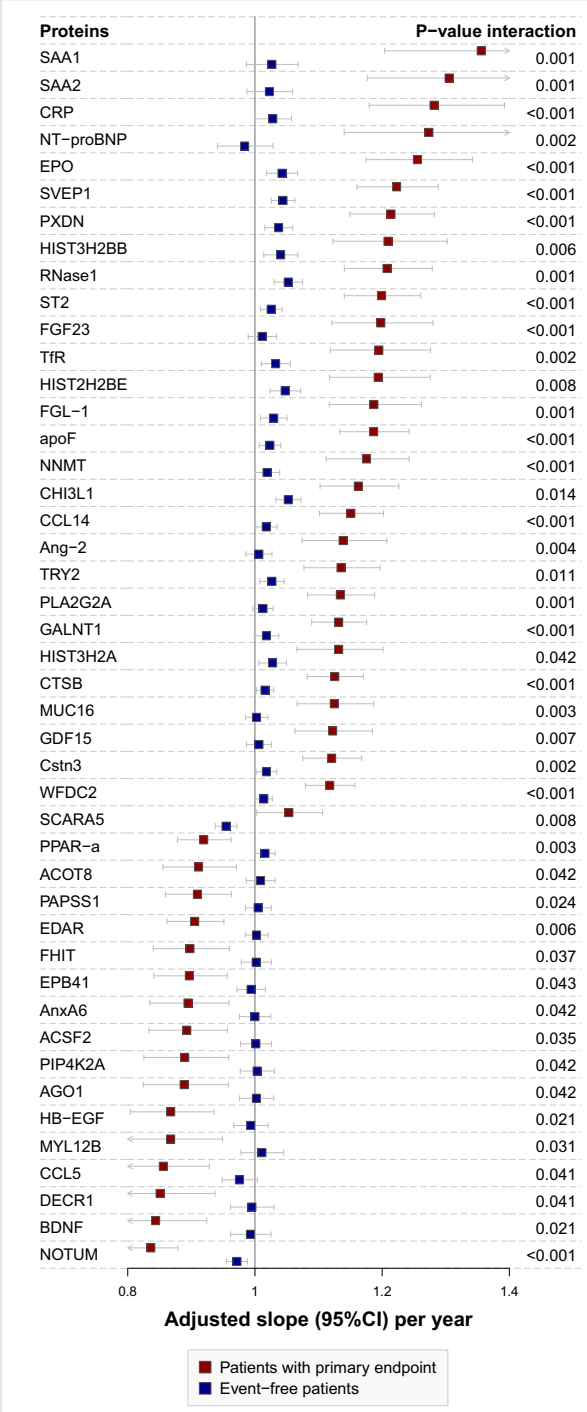


Figure 1 Proteins with a significantly different average evolution between patients with and without the primary endpoint. The estimated average evolution (per year) of circulating proteins with a significantly different trajectory (FDR < 0.05 and |relative difference in slope| > 10%) is depicted separately in patients with the primary endpoint and those who remained endpoint free. Average evolutions are estimated using linear mixed-effect regression models and are adjusted for the MAGGIC risk score. The red box depicts the average evolution of proteins in patients who reached the study endpoint, and the blue box depicts the average evolutions in patients who remained endpoint free.

the discriminative ability of our nine-biomarker model was also high in the external validation cohort, with *c*-indices up to 0.80.

Our results carry important clinical implications. Our statistical approach is able to capture markers that together represent essential information from the proteome, may improve individualized risk assessment, and could ultimately aid in designing more effective biomarker-guided therapies. For instance, patients who are at higher risk of adverse clinical outcomes, such as those with an unfavourable, high-risk biomarker profile, may benefit from more intensive monitoring and aggressive treatment strategies. Our joint models enable the calculation of personalized screening intervals,[19] and in these high-risk cases, the model will indicate that for optimal timing of the patient's next outpatient visit, the visit should be preponed. Likewise, it is important to identify patients with a favourable, low-risk biomarker profile. Here, the model will indicate that the outpatient visit may be postponed safely. These patients can be reassured and potentially followed up less intensively, considering their lower risk of adverse clinical outcome. For such individualized monitoring, ideally, cardiovascular biomarkers should be used as continuous, longitudinal measures in a dynamic cardiovascular risk prediction tool that also incorporates clinical features, and which may be updated as further information, including additional biomarker measurements or alterations in patients' comorbidities, become available. Such a dynamic measure of risk could guide clinicians to intensify monitoring and escalate therapy where risk remains elevated, and encourage patients to follow recommended treatment. By incorporating both favourable and unfavourable biomarker profiles into clinical practice, we can allocate resources more effectively and provide targeted care to individuals who may require different levels of monitoring and optimization of management.

Among the repeatedly measured proteins included in our final model, ST2 is the most elaborately investigated protein so far in the context of HF and its prognostic value has been substantiated extensively.[21,22] Our results confirm the importance of this protein for prognostication and underline the validity of our study. Other proteins that have previously been reported to carry prognostic value in patients with HF (e.g. GDF-15) showed a statistically significant different evolution during follow-up between patients with and without the PEP. However, in a multivariable prediction model, not all of these proteins did provide incremental information for risk assessment, possibly in part due to correlations between proteins. Nonetheless, here we show that the prognostic value of serially measured NT-proBNP, hs-TnT, and ST2 as reported previously[10,23] may be further increased when these 'established' markers are measured together with eight other proteins (TrpRS, HIST3H2A, angiotensinogen, deltex-1, TSP-4, ADAMTSL-2, ANTXR1, and cathepsin D) related to processes that have been implicated in HF or are dysregulated in cardiovascular disease (e.g. cardiac remodelling, atherosclerosis, the renin-angiotensin-aldosterone system). Interestingly, serial measurements of NT-proBNP and angiotensinogen showed the strongest associations with adverse outcomes in the current study. These proteins have been previously related to cardiac remodelling and the renin-angiotensin-aldosterone system, respectively, which are both central features in the pathophysiology of heart failure and which may thus explain the prominent role of these proteins. For some of these circulating proteins (cathepsin D and angiotensinogen), the prognostic value of a single measurement or gene polymorphism in HF has been reported in previous studies. For example, cathepsin D, recognized as a marker for oxidative stress, is associated with HF severity and predicts 2-year death or HF hospitalization in patients with CHF.[24] A polymorphism associated with higher plasma levels of angiotensinogen has been associated with increased risk of HF in a Caucasian population.[25,26] Nonetheless, reduced levels of angiotensinogen have been reported in patients with severe or end-stage HF and were associated with occurrence of the PEP in the current study.[27–29] The most likely explanation for this phenomenon is the

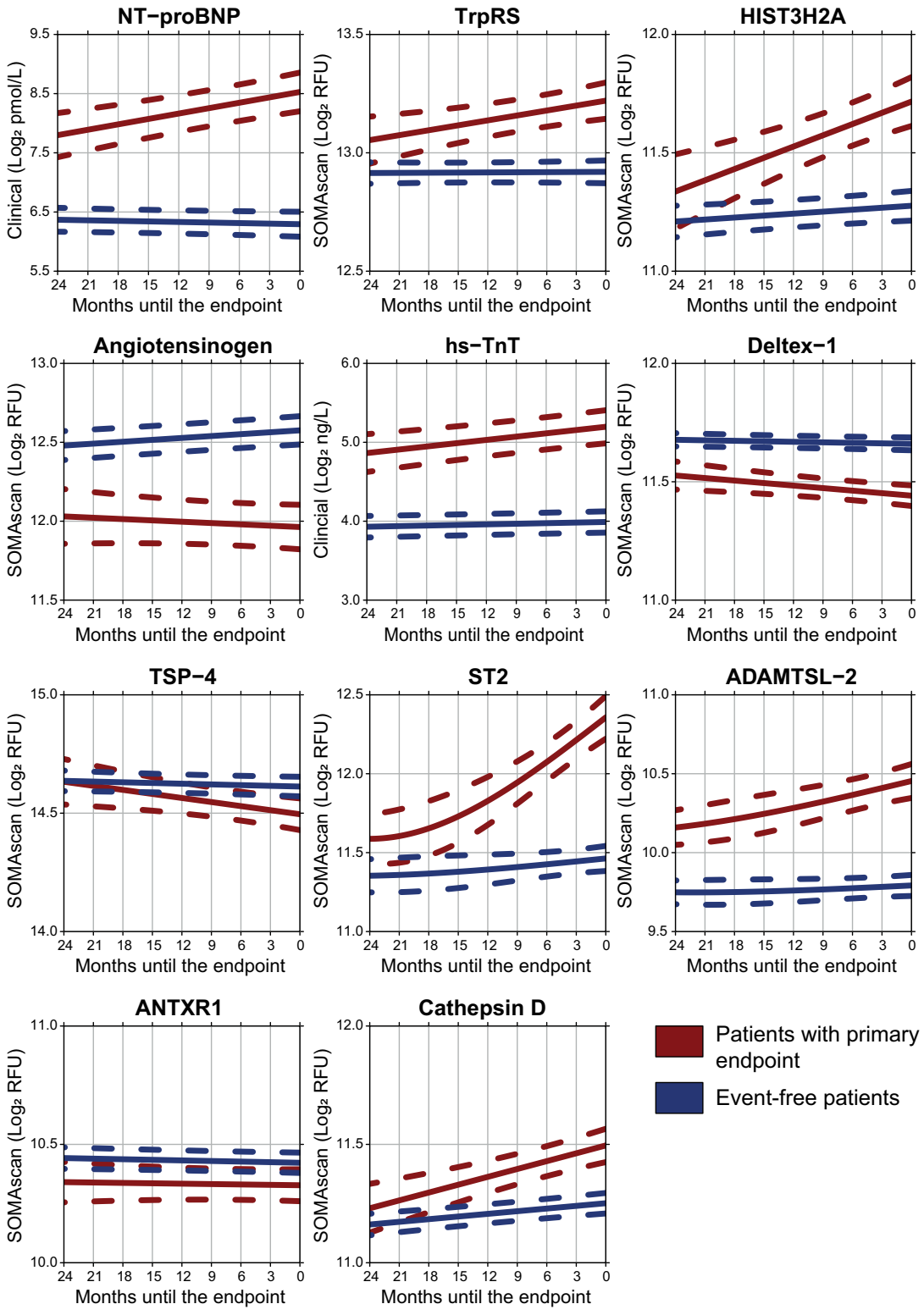


Figure 2 Average evolutions of the 10 proteins selected based on penalized regression. The average evolution of circulating proteins is depicted during the 2 years preceding a primary endpoint in patients with chronic heart failure who reached the study endpoint and last sample moment in patients who remained endpoint free. 'Time zero' is defined as the occurrence of the endpoint or censoring and is depicted on the right side of the *x*-axis; inherently to this representation, baseline sampling preceded this 'time zero'. The solid red line depicts the average evolution of proteins in patients who reached the study endpoint, and the solid blue line depicts the average evolution in patients who remained endpoint free. The dashed lines represent the 95% confidence intervals.

Table 2 Discriminative ability of the models

	c-index (95% CI)	*P*-value	Cross-validated *c*-index (95% CI)	*P*-value	Externally validated *c*-index (95% CI)	*P*-value
Clinical characteristics						
Model 1: MAGGIC risk score	0.70 (0.65–0.74)	—	0.70 (0.66–0.73)	—	0.71 (0.68–0.74)	—
Baseline measurements						
Model 2: MAGGIC risk score + NT-proBNP + hs-TnT (biomarker selection based on clinical practice)	0.77 (0.73–0.82)	Reference	0.77 (0.74–0.80)	Reference	0.77 (0.74–0.79)	Reference
Model 3: MAGGIC risk score + NT-proBNP + hs-TnT + biomarker selection based on literature	0.80 (0.76–0.84)	**0.005**	0.78 (0.75–0.81)	0.138	0.77 (0.75–0.80)	0.299
Model 4: MAGGIC risk score + NT-proBNP + hs-TnT + biomarker selection based on penalized regression	0.84 (0.80–0.88)	**<0.001**	0.82 (0.79–0.86)	**0.018**	0.79 (0.76–0.81)	**<0.001**
Repeated measurements						
Model 5: MAGGIC risk score + NT-proBNP + hs-TnT (biomarker selection based on clinical practice)	0.80 (0.76–0.85)	Reference	0.80 (0.78–0.83)	Reference		
Model 6: MAGGIC risk score + NT-proBNP + hs-TnT + biomarker selection based on literature	0.85 (0.81–0.88)	**<0.001**	0.83 (0.80–0.86)	0.087		
Model 7: MAGGIC risk score + NT-proBNP + hs-TnT + biomarker selection based on penalized regression	0.86 (0.83–0.90)	**<0.001**	0.85 (0.83–0.88)	**0.021**		

Model 1: MAGGIC risk score. Model 2: MAGGIC risk score and baseline measurements of biomarkers selected based on clinical practice (NT-proBNP and high-sensitivity TnT). Model 3: MAGGIC risk score, baseline measurements of NT-proBNP and high-sensitivity TnT, and baseline measurements of biomarkers selected from proteomic panel based on literature (CRP, GDF-15, ST2, and Gal-3). Model 4: MAGGIC risk score, baseline measurements of NT-proBNP and high-sensitivity TnT, and baseline measurements of biomarkers selected from proteomic panel based on penalized regression (elastic net; TrpRS, HIST3H2A, angiotensinogen, deltex 1, TSP-4, ADAMTSL-2, ANTXR1, and cathepsin D). Model 5: MAGGIC risk score and serially measured biomarkers selected based on clinical practice (NT-proBNP and high-sensitivity TnT). Model 6: MAGGIC risk score, serially measured NT-proBNP and high-sensitivity TnT, and serially measured biomarkers selected from proteomic panel based on literature (CRP, GDF-15, ST2, and Gal-3). Model 7: MAGGIC risk score, serially measured NT-proBNP and high-sensitivity TnT, and serially measured biomarkers selected from proteomic panel based on penalized regression (elastic net; TrpRS, HIST3H2A, angiotensinogen, deltex 1, TSP-4, ADAMTSL-2, ANTXR1, and cathepsin D).

A *P*-value <0.05 is considered statistically significant and presented in bold typeface. ADAMTSL2, a disintegrin and metalloproteinase with thrombospondin repeats like protein 2; ANTXR1, anthrax toxin receptor 1; CI, confidence interval; *c*-index, concordance index; CRP, C-reactive protein; Gal-3, galectin-3; GDF-15, growth/differentiation factor 15; HIST3H2A, histone H2A type 3; NT-proBNP, N-terminal pro-hormone B-type natriuretic peptide; TnT, troponin T; TSP-4, thrombospondin-4; ST2, interleukin 1 receptor-like 1; TrpRS, tryptophanyl-tRNA synthetase 1.

Table 3 Joint models predicting adverse clinical outcome in the study population

Biomarkers selected based on penalized regression	Univariable joint models HR (95% CI)	*P*-value	Multivariable joint model HR (95% CI)	*P*-value
NT-proBNP	3.83 (2.68–5.67)	**<0.001**	1.96 (1.17–3.40)	**0.006**
TrpRS	2.95 (1.05–16.3)	**0.028**	1.89 (0.86–4.12)	0.111
Histone H2A Type 3	2.93 (1.74–4.87)	**<0.001**	1.52 (0.76–3.20)	0.242
Angiotensinogen	0.54 (0.45–0.65)	**<0.001**	0.66 (0.49–0.88)	**0.002**
High-sensitivity troponin T	2.07 (1.60–2.62)	**<0.001**	1.50 (0.98–2.36)	0.060
Deltex 1	0.37 (0.28–0.49)	**<0.001**	0.67 (0.40–1.09)	0.111
Thrombospondin-4	0.50 (0.38–0.67)	**<0.001**	0.71 (0.47–1.05)	0.086
Interleukin-1 receptor-like 1	2.46 (1.91–3.20)	**<0.001**	1.29 (0.86–1.96)	0.233
ADAMTS-like protein 2	2.04 (1.68–2.47)	**<0.001**	0.81 (0.46–1.37)	0.426
Anthrax toxin receptor 1	0.60 (0.46–0.78)	**<0.001**	0.85 (0.62–1.15)	0.323
Cathepsin D	1.70 (1.31–2.19)	**<0.001**	1.13 (0.68–1.95)	0.641

A *P*-value <0.05 is considered statistically significant and presented in bold typeface. ADAMTS, a disintegrin and metalloproteinase with thrombospondin repeats; CI, confidence interval; HIST3H2A, histone H2A Type 3; HR, hazard ratio; NT-proBNP, N-terminal pro-hormone B-type natriuretic peptide; TrpRS, tryptophanyl-tRNA synthetase 1.

increase in renin concentration and renin activity with HF progression,[30] resulting in a depletion of its substrate angiotensinogen in patients with end-stage disease.

TrpRS, TSP-4, ADAMTSL-2, and ANTXR1 have been previously indicated to play a role in cardiac ischaemia and cardiac remodelling. An isoform of TrpRS affects the myocardial infarction area in an animal model of myocardial infarction, inhibits angiogenesis, and prevents transcription of shear stress-responsive genes, suggesting the role of TrpRS in atherosclerosis, vascular remodelling, and blood pressure regulation.[31–33] Thrombospondin-4 is involved in regulating fibrosis and remodelling of the myocardium in response to pressure overload in rodents.[34,35] Increased activity of TSP-4 is associated with recurrent coronary risk in post-infarction patients.[36] Upregulation of ADAMTSL-2 was found in the hearts of mice and patients with fibrosis and heart failure.[37] Fibroblasts isolated from mice deficient in ANTXR1 show an increased expression of collagen and fibronectin, consequently leading to fibrosis.[38]

For HIST3H2A and deltex-1 limited evidence of cardiovascular involvement is available. Modifications of histone proteins, such as HIST3H2A, have emerged as pivotal players in the development of heart failure.[39] HIST3H2A has been previously associated with manifest heart failure and HTx dynamics.[40] Nonetheless, the exact role of HIST3H2A and modifications thereof in the context of heart failure requires further research. The deltex-1 protein has ubiquitin E3 ligase activity. Cardiac E3 ubiquitin ligases regulate processes involved in heart failure and ischaemic heart disease, such as dysregulation of cardiac protein turnover and myocardial apoptosis.[41]

Some aspects of this study warrant consideration. First, SOMAmer reagents are selected against proteins in their native folded conformations. Hence, unfolded and denatured proteins are not detected. Moreover, the SOMAscan assay does not provide absolute concentrations, but RFUs. While these values can be used for comparing patients and changes over time within a patient, for clinical applications, absolute concentrations based on validated assays (e.g. enzyme-linked immunosorbent assay [ELISA]) are recommended. In the current study, immunoassay-based measurements for five biomarkers previously associated with heart failure (i.e. NT-proBNP, troponin T, CRP, cystatin C, and neutrophil gelatinase–associated lipocalin) showed high correlations with their SOMAscan measurements (Pearson correlation coefficient ranging from 0.74 to 0.94, Supplementary material online, *S8*). Protein features, specificity conformation, and precision of aptamer target binding of the nine proteins included in our final model are reported in the Supplementary material online, *S9* and *S10*. Lastly, the Bio-SHiFT study comprises a mostly white population, and thus, generalizing our findings to other ethnic groups should be done with caution.

Conclusions

Nine proteins, related to cardiac remodelling and atherosclerosis, and derived from 4210 serially measured circulating proteins, provided the optimal multivariable, dynamic model for the occurrence of adverse clinical events in patients with HFrEF, along with the MAGGIC risk score, NT-proBNP, and hs-TnT: ST2, TrpRS, HIST3H2A, angiotensinogen, deltex-1, TSP-4, ADAMTSL-2, ANTXR1, and cathepsin D. Two proteins showed the strongest associations (NT-proBNP and angiotensinogen). Altogether, our study shows that proteomic profiling could provide information for risk assessment beyond established risk factors, and underlines that repeated measurements of multiple circulating proteins may convey incremental prognostic value over clinical characteristics and repeatedly measured established biomarkers.

Supplementary material

Supplementary material is available at *European Heart Journal – Digital Health*.

Acknowledgements

F.W.A. is supported by the UCL Hospitals NIHR Biomedical Research Centre.

Funding

This work was supported by the EU/EFPIA Innovative Medicines Initiative 2 Joint Undertaking BigData@Heart grant no. 116074, the Jaap Schouten Foundation, and Noordwest Academie.

Conflict of interest: R.O. is employed by Somalogic Inc. D.L. reports non-financial support from Somalogic Inc., during the conduct of the study; grants and personal fees from Janssen, personal fees from Ortho Diagnostics, personal fees from DCRI (Novartis), grants from Bayer, grants from Astra Zeneca, grants from Critical Diagnostics, non-financial support from Somalogic, grants from Lilly, personal fees from ACI (Abbott Laboratories), personal fees from Martin Pharmaceuticals, personal fees from Illumina, personal fees from Vicardia, other from Hridaya, grants and personal fees from Amgen, personal fees from Cytokinetics, outside the submitted work. In addition, D.L. has a patent genomic predictors of BB response issued. The other authors have no disclosures to report.

Data availability

Anonymized data that support the findings of this study will be made available to other researchers for purposes of reproducing the results upon reasonable request and in accordance with a data-sharing agreement.

Author contributions

M.d.B.: conception and design of the work, the acquisition, analysis and interpretation of data, drafting and revision of the manuscript. T.B.P., A.J.R-B., and R.S.: analysis and interpretation of data, revising the manuscript. K.M.A. and V.A.U.: conception and design of the work, the acquisition and interpretation of data, revising the manuscript. J.J.B. and T.G.: the acquisition and interpretation of data, revising the manuscript. M.J.T.R., P.D.K., and P.J.v.d.S.: conception and design of the work, interpretation of data, revising the manuscript. R.O.: interpretation of data, revising the manuscript. D.L.: design of the work, interpretation of data, revising the manuscript. F.W.A.: conception and design of the work, interpretation of data, revising the manuscript, handling funding. E.B.: conception and design of the work, interpretation of data, revising the manuscript, handling funding and supervision. D.R.: conception and design of the work, analysis and interpretation of data, revising the manuscript, supervision. I.K.: conception and design of the work, the acquisition, analysis and interpretation of data, revising the manuscript, handling funding and supervision.

References

1. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, *et al.* 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: developed by the task force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J* 2021;**42**:3599–3726.
2. Heidenreich PA, Bozkurt B, Aguilar D, Allen LA, Byun JJ, Colvin MM, *et al.* 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *J Am Coll Cardiol* 2022;**79**:e263–e421.
3. Chow SL, Maisel AS, Anand I, Bozkurt B, de Boer RA, Felker GM, *et al.* Role of biomarkers for the prevention, assessment, and management of heart failure: a scientific statement from the American Heart Association. *Circulation* 2017;**135**:e1054–e1091.
4. Smith JG, Gerszten RE. Emerging affinity-based proteomic technologies for large-scale plasma profiling in cardiovascular disease. *Circulation* 2017;**135**:1651–1664.

5. Cuvelliez M, Vandewalle V, Brunin M, Beseme O, Hulot A, de Groote P, *et al.* Circulating proteomic signature of early death in heart failure patients with reduced ejection fraction. *Sci Rep* 2019;**9**:19202.
6. Gui H, She R, Luzum J, Li J, Bryson TD, Pinto Y, *et al.* Plasma proteomic profile predicts survival in heart failure with reduced ejection fraction. *Circ Genom Precis Med* 2021;**14**: e003140.
7. Williams SA, Ostroff R, Hinterberg MA, Coresh J, Ballantyne CM, Matsushita K, *et al.* A proteomic surrogate for cardiovascular outcomes that is sensitive to multiple mechanisms of change in risk. *Sci Transl Med* 2022;**14**:eabj9625.
8. Cresci S, Pereira NL, Ahmad F, Byku M, de las Fuentes L, Lanfear DE, *et al.* Heart failure in the era of precision medicine: a scientific statement from the American Heart Association. *Circ Genom Precis Med* 2019;**12**:458–485.
9. Pocock SJ, Ferreira JP, Packer M, Zannad F, Filippatos G, Kondo T, *et al.* Biomarker-driven prognostic models in chronic heart failure with preserved ejection fraction: the EMPEROR–preserved trial. *Eur J Heart Fail* 2022;**24**:1869–1878.
10. van Boven N, Battes LC, Akkerhuis KM, Rizopoulos D, Caliskan K, Anroedh SS, *et al.* Toward personalized risk assessment in patients with chronic heart failure: detailed temporal patterns of NT-proBNP, troponin T, and CRP in the Bio-SHiFT study. *Am Heart J* 2018;**196**:36–48.
11. McMurray JJ, Adamopoulos S, Anker SD, Auricchio A, Böhm M, Dickstein K, *et al.* ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: the task force for the diagnosis and treatment of acute and chronic heart failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. *Eur Heart J* 2012;**33**:1787–1847.
12. Paulus WJ, Tschöpe C, Sanderson JE, Rusconi C, Flachskampf FA, Rademakers FE, *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**:2539–2550.
13. Brankovic M, Akkerhuis KM, van Boven N, Anroedh S, Constantinescu A, Caliskan K, *et al.* Patient-specific evolution of renal function in chronic heart failure patients dynamically predicts clinical outcome in the Bio-SHiFT study. *Kidney Int* 2018;**93**:952–960.
14. Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, *et al.* Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One* 2010;**5**:e15004.
15. Kim CH, Tworoger SS, Stampfer MJ, Dillon ST, Gu X, Sawyer SJ, *et al.* Stability and reproducibility of proteomic profiles measured with an aptamer-based platform. *Sci Rep* 2018;**8**:8382.
16. Candia J, Cheung F, Kotliarov Y, Fantoni G, Sellers B, Griesman T, *et al.* Assessment of variability in the SOMAscan assay. *Sci Rep* 2017;**7**:14248.
17. Luzum JA, Peterson E, Li J, She R, Gui H, Liu B, *et al.* Race and beta-blocker survival benefit in patients with heart failure: an investigation of self-reported race and proportion of African genetic ancestry. *J Am Heart Assoc* 2018;**7**:e007956.
18. Pocock SJ, Ariti CA, McMurray JJV, Maggioni A, Køber L, Squire IB, *et al.* Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies. *Eur Heart J* 2013;**34**:1404–1413.
19. Schuurman AS, Tomer A, Akkerhuis KM, Brugts JJ, Constantinescu AA, Ramshorst JV, *et al.* Personalized screening intervals for measurement of N-terminal pro-B-type natriuretic peptide improve efficiency of prognostication in patients with chronic heart failure. *Eur J Prev Cardiol* 2021:**28**:e11–e14.
20. Pepe MS, Kerr KF, Longton G, Wang Z. Testing for improvement in prediction model performance. *Stat Med* 2013;**32**:1467–1482.
21. Savarese G, Musella F, D'Amore C, Vassallo E, Losco T, Gambardella F, *et al.* Changes of natriuretic peptides predict hospital admissions in patients with chronic heart failure: a meta-analysis. *JACC Heart Fail* 2014;**2**:148–158.
22. Emdin M, Aimo A, Vergaro G, Bayes-Genis A, Lupón J, Latini R, *et al.* sST2 predicts outcome in chronic heart failure beyond NT-proBNP and high-sensitivity troponin T. *J Am Coll Cardiol* 2018;**72**:2309–2320.
23. Bouwens E, Brankovic M, Mouthaan H, Baart S, Rizopoulos D, van Boven N, *et al.* Temporal patterns of 14 blood biomarker candidates of cardiac remodeling in relation to prognosis of patients with chronic heart failure-the Bio-SHiFT study. *J Am Heart Assoc* 2019;**8**:e009555.
24. Hoes MF, Tromp J, Ouwerkerk W, Bomer N, Oberdorf-Maass SU, Samani NJ, *et al.* The role of cathepsin D in the pathophysiology of heart failure and its potentially beneficial properties: a translational approach. *Eur J Heart Fail* 2020;**22**:2102–2111.
25. Chen S, Zhang L, Wang H-W, Wang X-Y, Li X-Q, Zhang L-L. The M235T polymorphism in the angiotensinogen gene and heart failure: a meta-analysis. *J Renin Angiotensin Aldosterone Syst* 2014;**15**:190–195.
26. Zhai C, Cong H, Zhang H, Hou K, Zhang Y, Zhang Y. M235t polymorphism in the angiotensinogen gene and cardiovascular disease: an updated meta-analysis of 39 case-control comparisons. *Anatol J Cardiol* 2019;**21**:222–232.
27. Danser AHJ, van Kesteren CAM, Bax WA, Tavenier M, Derkx FHM, Saxena PR, *et al.* Prorenin, renin, angiotensinogen, and angiotensin-converting enzyme in normal and failing human hearts. *Circulation* 1997;**96**:220–226.
28. Klotz S, Burkhoff D, Garrelds IM, Boomsma F, Danser AHJ. The impact of left ventricular assist device-induced left ventricular unloading on the myocardial renin–angiotensin–aldosterone system: therapeutic consequences? *Eur Heart J* 2009;**30**:805–812.
29. Arnal JF, Cudek P, Plouin P-F, Guyenne T-T, Michel J-B, Corvol P. Low angiotensinogen levels are related to the severity and liver dysfunction of congestive heart failure: implications for renin measurements. *Am J Med* 1991;**90**:17–22.
30. Sullivan RD, Mehta RM, Tripathi R, Reed GL, Gladysheva IP. Renin activity in heart failure with reduced systolic function-new insights. *Int J Mol Sci* 2019;**20**:3182.
31. Zeng R, Chen YC, Zeng Z, Liu WQ, Jiang XF, Liu, *et al.* Effect of mini-tyrosyl-tRNA synthetase/mini-tryptophanyl-tRNA synthetase on ischemic angiogenesis in rats: proliferation and migration of endothelial cells. *Heart Vessels* 2011;**26**:69–80.
32. Zou Y, Yang Y, Fu X, He X, Liu M, Zong T, *et al.* The regulatory roles of aminoacyl-tRNA synthetase in cardiovascular disease. *Mol Ther Nucleic Acids* 2021;**25**:372–387.
33. Tzima E, Reader JS, Irani-Tehrani M, Ewalt KL, Schwartz MA, Schimmel P. Biologically active fragment of a human tRNA synthetase inhibits fluid shear stress-activated responses of endothelial cells. *Proc Natl Acad Sci U S A* 2003;**100**:14903–14907.
34. Frolova EG, Sopko N, Blech L, Popović ZB, Li J, Vasanji A, *et al.* Thrombospondin-4 regulates fibrosis and remodeling of the myocardium in response to pressure overload. *FASEB J* 2012;**26**:2363–2373.
35. Chistiakov DA, Melnichenko AA, Myasoedova VA, Grechko AV, Orekhov AN. Thrombospondins: a role in cardiovascular disease. *Int J Mol Sci* 2017;**18**:1540.
36. Corsetti JP, Ryan D, Moss A, McCarthy J, Goldenberg I, Zareba W, *et al.* Thrombospondin-4 polymorphism (A387P) predicts cardiovascular risk in postinfarction patients with high HDL cholesterol and C-reactive protein levels. *Thromb Haemost* 2011;**106**:1170–1178.
37. Rypdal KB, Erusappan PM, Melleby AO, Seifert DE, Palmero S, Strand ME, *et al.* The extracellular matrix glycoprotein ADAMTSL2 is increased in heart failure and inhibits TGFβ signalling in cardiac fibroblasts. *Sci Rep* 2021;**11**:19757.
38. Besschetnova TY, Ichimura T, Katebi N, St Croix B, Bonventre JV, Olsen BR, *et al.* Regulatory mechanisms of anthrax toxin receptor 1-dependent vascular and connective tissue homeostasis. *Matrix Biol* 2015;**42**:56–73.
39. Yerra VG, Advani A. Histones and heart failure in diabetes. *Cell Mol Life Sci* 2018;**75**: 3193–3213.
40. Egerstedt A, Berntsson J, Smith ML, Gidlöf O, Nilsson R, Benson M, *et al.* Profiling of the plasma proteome across different stages of human heart failure. *Nat Commun* 2019;**10**: 5830.
41. Willis MS, Bevilacqua A, Pulinilkunnil T, Kienesberger P, Tannu M, Patterson C. The role of ubiquitin ligases in cardiac disease. *J Mol Cell Cardiol* 2014;**71**:43–53.