Circulation: Heart Failure

ORIGINAL ARTICLE

Protein Biomarkers of New-Onset Heart Failure: Insights From the Heart Omics and Ageing Cohort, the Atherosclerosis Risk in Communities Study, and the Framingham Heart Study

Nicolas Girerd, MD, PhD*; Daniel Levy, MD*; Kevin Duarte, PhD; Joao Pedro Ferreira, MD, PhD; Christie Ballantyne, MD; Timothy Collier, MSc; Anne Pizard, PhD; Jens Björkman, PhD; Javed Butler, MD; Andrew Clark, MD, PhD; John G. Cleland, MD; Christian Delles, MD, PhD; Javier Diez, MD, PhD; Arantxa González, MD, PhD; Mark Hazebroek, MD, PhD; Jennifer Ho, MD; Anne-Cécile Huby, PhD; Shih-Jen Hwang, PhD; Roberto Latini, MD, PhD; Beatrice Mariottoni MD, PhD; Alexandre Mebazaa MD, PhD; Pierpaolo Pellicori MD, PhD; Naveed Sattar MBChB, PhD; Peter Sever MD, PhD; Jan A. Staessen, MD, PhD; Job Verdonschot, MD, PhD; Stephane Heymans, MD, PhD; Patrick Rossignol, MD, PhD; Faiez Zannad, MD, PhD

BACKGROUND: We sought to identify protein biomarkers of new-onset heart failure (HF) in 3 independent cohorts (HOMAGE cohort [Heart Omics and Ageing], ARIC study [Atherosclerosis Risk in Communities], and FHS [Framingham Heart Study]) and assess if and to what extent they improve HF risk prediction compared to clinical risk factors alone.

METHODS: A nested case-control design was used with cases (incident HF) and controls (without HF) matched on age and sex within each cohort. Plasma concentrations of 276 proteins were measured at baseline in ARIC (250 cases/250 controls), FHS (191/191), and HOMAGE cohort (562/871).

RESULTS: In single protein analysis, after adjusting for matching variables and clinical risk factors (and correcting for multiple testing), 62 proteins were associated with incident HF in ARIC, 16 in FHS, and 116 in HOMAGE cohort. Proteins associated with incident HF in all cohorts were BNP (brain natriuretic peptide), NT-proBNP (N-terminal pro-B-type natriuretic peptide), eukaryotic translation initiation factor 4E-BP1 (4E-binding protein 1), hepatocyte growth factor (HGF), Gal-9 (galectin-9), TGF-alpha (transforming growth factor alpha), THBS2 (thrombospondin-2), and U-PAR (urokinase plasminogen activator surface receptor). The increment in *C*-index for incident HF based on a multiprotein biomarker approach, in addition to clinical risk factors and NT-proBNP, was 11.1% (7.5%–14.7%) in ARIC, 5.9% (2.6%–9.2%) in FHS, and 7.5% (5.4%–9.5%) in HOMAGE cohort, all $P<0.001$), each of which was a larger increase than that for NT-proBNP on top of clinical risk factors. Complex network analysis revealed a number of overrepresented pathways related to inflammation (eg, tumor necrosis factor and interleukin) and remodeling (eg, extracellular matrix and apoptosis).

CONCLUSIONS: A multiprotein biomarker approach improves prediction of incident HF when added to natriuretic peptides and clinical risk factors.

Key Words: biomarkers ■ cardiovascular diseases ■ heart failure ■ protein interaction maps ■ proteomics

Correspondence to: Nicolas Girerd, MD, PhD, Centre d'Investigations Cliniques-INSERM CHU de Nancy, Institut Lorrain du Cœur et des Vaisseaux Louis Mathieu, 4 rue du Morvan, 54500 Vandoeuvre Lès Nancy, France. Email n.girerd@chru-nancy.fr

*N. Girerd and D. Levy contributed equally.

Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/CIRCHEARTFAILURE.122.009694.

For Sources of Funding and Disclosures, see page 440.

Circulation: Heart Failure is available at www.ahajournals.org/journal/circheartfailure

What Is New?

- A multiprotein approach improved the prediction of new-onset heart failure.
- The key proteins predicting heart are related to inflammatory (tumor necrosis factor and interleukins) and remodeling (extracellular matrix and apoptosis) pathways.
- These results arise from 3 large international cohorts, thus ensuring a good external validity.

What Are the Clinical Implications?

- The predictive value of a combined approach using clinical and numerous biological biomarkers is excellent.
- As predicting individuals at high risk for heart failure is the first step to a tailored intervention to prevent HF, this integrated approach opens up avenues for preventive strategies in a populational setting. These risk-stratification tools could aid in the design of preventive trials, seeking to decrease the incidence of heart failure with preserved ejection fraction.
- The pathways identified as key in the prediction of heart failure (mostly related to inflammation and remodeling) could also highlight promising therapeutic approaches in preventive cardiology.
- The pathways identified as key in the prediction of heart failure (mostly related to inflammation and remodeling) could also open therapeutic avenues in preventive cardiology.

Nonstandard Abbreviations and Acronyms

4E-BP1	4E-binding protein 1
ARIC	Atherosclerosis in Communities study
BNP	brain natriuretic peptide
CRP	C-reactive protein
CVD	cardiovascular disease
FHS	Framingham Heart Study
FS	follistatin
Gal-9	galectin-9
GO	gene ontology
HF	heart failure
HOMAGE	Heart Omics and Ageing cohort
LOD	limit of detection
MEPE	matrix extracellular phosphoglycoprotein
NT-proBNP	N-terminal pro-B-type natriuretic peptide
PON3	paraoxonase
SPON1	spondin-1
TGF-alpha	transforming growth factor alpha
THBS2	thrombospondin-2
TIM-1	T-cell immunoglobulin and mucin domain 1
TWEAK	tumor necrosis factor
U-PAR	urokinase plasminogen activator surface receptor

The incidence of heart failure (HF) is increasing due to ageing of the population and associated increases in the prevalence of and duration of exposure to several risk factor for HF including hypertension, diabetes, and obesity.[1–4] In the past several decades, important progress has been made in the treatment of patients with overt HF. Angiotensin-converting enzyme inhibitors, angiotensin II receptor blockers, beta blockers, and mineralocorticoid receptor antagonists have been shown to improve survival in patients with HF with reduced ejection fraction.[5,6] More recently, modest advances have also been made in the treatment of patients with HF with preserved ejection fraction. For example, spironolactone now has a grade II indication according to recent guidelines.[6] More recently, sodium/glucose cotransporter 2 inhibitors have been shown to improve prognosis regardless of HF type and have been proven to prevent HF onset in patients with diabetes.[7] The best effective personalized strategies to prevent HF, however, should be further explored to enable HF prevention at a population level.

Circulating proteins, most notably NT-proBNP (N-terminal pro-B-type natriuretic peptide), have been studied as predictors of incident HF.[8–11] The human plasma proteome is composed of proteins derived from multiple tissues and representing multiple biological pathways and can serve as a dynamic representation of the molecular states of diverse systems.[12] High throughput, automated, high-sensitivity protein assays can now be applied to large cohorts and clinical trials to identify protein biomarker profiles of HF. Multiprotein biomarker strategies may aid in the characterization of high-risk individuals in whom earlier treatment and ongoing surveillance may be beneficial and may provide information that is not captured by traditional HF risk factors.

A recent attempt to predict HF using 80 proteins in community-based cohorts resulted in a sizeable increase in the *C*-index (10.1%; *P*=0.006) on top of HF clinical risk factors.[11] The improvement in prediction from the circulating proteins biomarkers on top of natriuretic peptides, however, was a small and not significant (change in *C*-index 2.0%; *P*=0.40).[11] Ho et al[13] reported, based on findings in the FHS (Framingham Heart Study), that NT-proBNP, GDF-15 (growth/differentiation factor 15), CRP (C-reactive protein), and leptin (among 85 tested biomarkers) jointly predicted incident HF and were associated with a modest but statistically significant 3% improvement in the *C*-statistic on top of clinical risk factors (0.843–0.873; *P*<0.0001). We hypothesized, a priori, that a significantly larger panel of proteins, covering additional pathways and biological processes related to HF, would better identify individuals at high risk for

new-onset HF and highlight pathways contributing to its occurrence.

To that end, we undertook a study of new-onset heart failure in 3 independent cohorts (the HOMAGE cohort [Heart Omics and Ageing], the ARIC study [Atherosclerosis in Communities], and the FHS) with the aim to identify protein biomarkers for the prediction of HF.

METHODS

Data Availability

The data that support the findings of this study are available from N.G. and F.Z. upon reasonable request.

Study Populations

All studies were conducted in accordance with the Declaration of Helsinki and approved by each site ethics committees. All participants provided written informed consent.

ARIC Cohort

We used the data of the ARIC study, a longitudinal population-based cohort study of cardiovascular disease (CVD) and its risk factors ongoing since 1987. A detailed description of the ARIC study design and methods has been published elsewhere.[14] The analysis reported herein was conducted among individuals who participated in ARIC visit 5 (age 69–88 years), conducted between June 1, 2011 and August 30, 2013.

Participants with HF history at the 5th visit were not considered eligible for this analysis. We performed a classical nested matched case-control study (ie, controls were study participants who did not develop HF up until the end of follow-up). Controls were matched within each cohorts on follow-up time, age, and sex.

The FHS Cohort

The FHS is a community-based longitudinal study that began in 1948. Residents of Framingham, Massachusetts were invited to participate in a study of the prevalence, incidence, and natural history of CVDs. Participants attend on-site clinical examinations and provide information on cardiovascular health through face-to-face interviews, self-administered questionnaires, physician examinations, and provide blood samples for laboratory tests. For the current study, the 7th examination FHS Offspring cohort examination, conducted from 1998 to 2001, was used. Information on CVDs diagnosis was abstracted from the Framingham Heart Study 3-physician verified medical record data file using December 31, 2016 as the deadline of update.

A total of 43 (1.22%) participants had a diagnosis of HF prior to their seventh examination date who were excluded from eligibility for this study. The remaining 3231 exam attendance included 265 participants who had a HF diagnosis following their baseline examination. Of these cases, 261 had blood samples for the measurement of plasma proteins and were considered as candidate cases in the subsequent matching procedure. We applied 1-to-1 match strategy and identified 211 age-sex matched controls using a nested matched case-case design in which individuals who developed HF were considered to be at-risk, that is, eligible to be selected as controls up until the time they became a case. We also excluded all case/control pairs for which the case was not hospitalized for HF (n=20 pairs). This design can also be referred to as a case-cohort design.

HOMAGE Cohort

The HOMAGE consortium (NCT02556450) is a European Union funded program that aims to identify and validate omics biomarkers associated with incident HF in order to potentially develop new and personalized preventive strategies. The HOMAGE consortium included 20 completed and ongoing studies conducted in 8 European countries that enrolled healthy subjects, patients with HF, and patients at high risk of CVD, all of which were pooled in a common database.[15] From the HOMAGE population with >20 000 patients, we identified cohorts in whom individuals without a history of HF at baseline had been followed up until first hospitalization for HF. Participants from 2 suitable cohorts and 1 clinical trial population were identified: PREDICTOR[16] (patients recruited from 2007 to 2010), HEALTH-ABC[17] (recruited from 1997 to 1998), and PROSPER[15,18] (recruited from 1997 and 1999). As in the FHS cohort, we then employed a nested matched case-cohort design in which individuals who developed HF were considered to be at-risk, that is, eligible to be selected as controls up until the time they became a case.[19]

The initial purpose of the HOMAGE project was to identify underlying mechanistic pathways that may be associated with a primary incident diagnosis of HF.[20] The HOMAGE case-cohort study had 2 independent phases: discovery and replication. For the discovery phase, we selected 300 cases and 599 controls (1 case only had 1 match) randomly selected without replacement in a 1:2 proportion; because of 22 missing or poor-quality samples, the final match was 286 cases to 591 controls. For replication, we selected 315 cases and 315 controls randomly selected without replacement in a 1:1 proportion; because of 74 missing or poor-quality samples, the final match was 276 cases to 280 controls. Controls were matched within each cohorts on follow-up time, age, and sex. For the purpose of the current analysis, centered on risk prediction, we merged the datasets of the 2 phases (totalizing 562 cases and 871 controls).

Outcome

The outcome used in the HOMAGE, ARIC, and FHS studies was incident HF, which was defined as first hospitalization for HF as primary admission diagnosis (adjudicated by the investigators of the respective cohorts). NT-proBNP data were not required to ascertain the diagnosis, as most patients were included in an era during which natriuretic peptides were not yet standard of care.

Sample Handling

All sample shipments and sample data acquisition within the HOMAGE consortium (including the ARIC and FHS databases) are according to predefined standard operating procedures and material transfer agreements to maintain uniformity. Figure S1 shows the sample handling and storage per cohort and the sample flow until protein measurement at the TATAA-Biocenter

(Gothenburg, Sweden) and OLINK Proteomics (Sweden). The cases and controls were separately identified and selected by the study statistician. All participants' information was then removed and a randomly sorted list of patient/sample IDs for each cohort was sent to Maastricht University Medical Center. Aliquoting of the samples at Biobank Maastricht was performed using a multi-pipette in 1 run to reduce freeze/thaw cycles and batch effects. All samples were randomly aliquoted on each 96-well plate, independent of cohort allocation. The entire sampling handling/protein measurement was carried out fully blind to case-control status.

Assays and Studied Biomarkers

Baseline plasma samples were analyzed for protein biomarkers using the OLINK Proseek® Multiplex CVD II, CVD III, and inflammation panels for all cohorts. The HOMAGE analyses were carried out by the TATAA-Biocenter and by OLINK proteomics for the ARIC and FHS cohorts.

These panels were selected for the well-balanced inclusion of proteins with established associations with CV disease and HF (eg, BNP [brain natriuretic peptide], ST2 [suppression of tumorigenicity 2], and GDF-15 [growth and differentiation factor 15]) and others with less well-established associations (eg, TWEAK [tumor necrosis factor], PON3 [paraoxonase]; the full information on these panels is available at: https://www.olink.com/resources-support/document-download-center/). The assays use a proximity extension assay technology,[21] where 92 oligonucleotide-labeled antibody probe pairs per panel are allowed to bind to their respective targets in the sample in 96-well plate format. When binding to their correct targets, they give rise to new DNA amplicons with each ID-barcoding their respective antigens. The amplicons are subsequently quantified using a Fluidigm BioMark HD real-time PCR platform. The OLINK platform provides log2-normalized protein expression data. A detailed description of the Olink technology is depicted in the Supplemental Material. For 9 proteins measured in both the inflammation panel and CVD panels, the one from CVD panels was used for further data analyses (the results for these pairs of proteins were strongly correlated ≥0.9). In addition, 15 proteins that were below the limit of detection (LOD) were not included in the analysis. The Olink® quality control samples are considered as flagged if they deviate more than 0.3 log2-normalized protein expression from the median of all samples in 1 of 2 control assays for incubation and detection. The LOD is defined by the 3 negative controls run on each plate and set to 3 SDs above the measured background.

In the ARIC and FHS cohorts, biomarkers with 80% or more values below the LOD were excluded from candidate variables. When the proportion of below LOD value was 80% or less, the below limit of detection values were replaced by the LOD.

A total of 247 protein biomarkers was assessed in the baseline samples of the 3 cohorts. The abbreviations, full names, and respective Olink® multiplex panels of the studied proteins are described in Table S1.

The assays were performed blinded to case/control status with cases and controls randomly distributed across plates. The proteomic results were then merged with the baseline data, which included the case-control status, matching variables, and the clinical risk factors.

Statistical Considerations

For the baseline clinical characteristics, continuous variables are expressed as means±SD or as medians (interquartile range) if the distribution was skewed and categorical variables as frequencies and percentages. Participants' baseline characteristics were compared between cases and controls using Fisher exact test for categorical variables and *t* tests or nonparametric Wilcoxon tests for continuous variables, as appropriate.

Logistic regression models adjusting for the matching variables (age, sex, cohort, follow-up time), phase only for HOMAGE (discovery/replication), and relevant risk factor for HF were used to identify protein biomarkers associated with incident HF.[22] The association with HF risk was assessed for each biomarker individually (mono-marker models). Relevant risk factors for HF were prespecified clinical risk factors previously found to represent the best clinical prognostic model for incident HF in the HOMAGE population[23] (smoking, diabetes, history of coronary artery disease, serum creatinine, body mass index, systolic blood pressure, use of antihypertensive medication, heart rate) and FHS population (hypertension, history of atrial fibrillation, and ratio total/HDL cholesterol). In all 3 cohorts, we corrected for multiple testing using a false discovery rate of 5%, applying the Benjamini-Hochberg procedure.[24]

Because proteins expressions levels were log2-transformed, the odds ratio estimated the increase in odds of HF associated with a doubling of protein concentration.

To assess the added predictive value of biomarkers on top on routine HF predictors, we assessed the *C*-index of a baseline model including all the adjustment variables (matching variables, phase for HOMAGE, and the relevant risk factors described previously) in each cohort and assessed the increase in *C*-index when adding biomarkers individually (ie, 1 by 1) to the model.

In a subsequent analysis, several biomarkers were included in 1 unique model (multimarker model). Both BNP and NT-proBNP were measured but we kept only NT-proBNP in this multimarker model. To deal with the risk of multicollinearity, we used the variance inflation factor method within the logistic model. Biomarkers with the highest variance inflation factor were removed 1 by up until all variance inflation factor were 10 or less. One final model was fitted by cohort, and, as for the mono-marker models, the increase in C-index on top of routine HF predictors was assessed.

All statistical analysis were performed using R software (CRAN).

Network Analysis

The FHF-GKB resource, representing most available public knowledge about human protein-protein, protein-pathway and protein-Gene Ontology (GO) terms,[25] was queried to extract overrepresented pathways and GO terms associated with statistically relevant biomarkers. For each pathway/GO term linked to 1 selected biomarker, overrepresentation was computed by a Fisher exact test with the whole FHF-GKB as annotation background. As previously, values were corrected for multiple testing using false discovery rate. Overrepresented terms were displayed with their associated biomarkers in Cytoscape.[26] An additional step of clustering was performed for GO terms in order to obtain smaller and more understandable groups. This

was done using the Community Clustering (GLay) method[27,28] implemented in the Clustermaker2 Cytoscape app.[29]

RESULTS

Characteristics of Study Participants

The clinical characteristics of the 3 contributing cohorts are provided in Table 1. Almost half of the participants were women. In the ARIC cohort, 136 participants (27.2%) were Black. The median (IQR) follow-up time was 2.7 (1.5–3.7) years in ARIC, 4.1 (2.1–9.2) years in HOMAGE, and 8.1 (4.5–12.1) years in FHS. FHS participants were approximately 10 years younger than participants in the other 2 cohorts.

Association of Individual Protein Biomarkers With Incident HF

After adjusting for the matching variables and clinical HF risk factors, 142 proteins were individually associated with incident HF (correcting for multiple testing) in at least 1 cohort: 62 proteins in ARIC, 16 in FHS, and 116 in HOMAGE (Table 2; Table S2).

Proteins significantly associated with incident HF in all 3 cohorts were BNP, NT-proBNP, 4E-BP1 (4E-binding protein 1), HGF (hepatocyte growth factor), Gal-9 (galectin-9), TGF-alpha (transforming growth factor alpha), THBS2 (thrombospondin-2), and U-PAR (urokinase plasminogen activator surface receptor; Figure 1).

In a sensitivity analysis performed in the ARIC cohort by adding Black race to the clinical HF risk factors, 44 proteins were individually associated with incident HF: 41 proteins already associated with incident HF in the previous results and 3 new proteins (MEPE [matrix extracellular phosphoglycoprotein], brother of CDO, and MMP-2 [matrix metalloproteinase-2]; Table S3).

Predictive Value of Individual Protein Biomarkers for Incident HF on Top of Clinical HF Risk Factors

Sixty-eight proteins were found to improve HF prediction significantly (as assessed by a significant increase in the *C*-index) when considered individually on top of clinical HF risk factors in at least 1 cohort: 11 proteins in ARIC, 9 in FHS, and 56 in HOMAGE (Table 3; Table S4). Only BNP and NT-proBNP significantly improved prediction in all 3 cohorts (Figure 2).

Predictive Value of a Protein Multimarker Approach for Incident HF Risk Prediction

NT-proBNP significantly improved HF prediction on top of established clinical HF risk factors in all 3 cohorts (delta *C*-index=7.6% [95% CI, 3.6%–11.5%]) in ARIC,

Table 1. Characteristics of the 3 Cohorts

	ARIC			FHS			HOMAGE		
	Controls (n=250)	Incident HF (n=250)	*P* value*	Controls (n=191)	Incident HF (n=191)	*P* value*	Controls (n=871)	Incident HF (n=562)	*P* value*
Age, y	77.7±5.4	77.9±5.4	0.77	68.4±8.4	68.2±8.3	0.77	74.7±3.5	74.8±3.5	0.63
Men, n (%)	135 (54.0%)	136 (54.4%)	1.00	105 (55.0%)	105 (55.0%)	1.00	478 (54.9%)	311 (55.3%)	0.87
Follow-up time, y†	2.7 (1.5–3.7)	2.7 (1.5–3.7)	0.96	8.1 (4.4–12.1)	8.1 (4.4–12.1)	1.00	4.3 (2.1–9.5)	3.5 (2.0–9.0)	0.14
Body mass index, kg/m²	28.4±5.1	29.3±6.2	0.073	29.4±6.2	30.5±5.7	0.065	26.9±4.3	27.9±4.8	0.0001
Diabetes, n (%)	70 (29.0%)	95 (38.6%)	0.027	31 (16.2%)	56 (29.3%)	0.003	126 (14.5%)	127 (22.9%)	<0.0001
Systolic blood pressure, mmHg	133±19	134±21	0.44	134±18	136±20	0.23	143±23	145±23	0.046
Heart rate, bpm	645±11	67±12	0.042	64±10	68±12	0.0005	65±11	67±12	0.0001
Hypertension, n (%)	188 (76.4%)	209 (84.3%)	0.031	119 (62.3%)	140 (73.3%)	0.028	664 (76.3%)	478 (85.1%)	<0.0001
Hyperlipidemia treatment, n (%)	146 (58.9%)	155 (62.0%)	0.52	61 (31.9%)	71 (37.2%)	0.33	240 (27.6%)	155 (27.6%)	0.99
Current smoker, n (%)	12 (5.0%)	21 (8.6%)	0.14	12 (6.3%)	27 (14.1%)	0.017	118 (13.6%)	87 (15.5%)	0.31
History of coronary heart disease, n (%)	38 (15.4%)	66 (26.8%)	0.002	19 (9.9%)	38 (19.9%)	0.009	204 (23.6%)	217 (39.0%)	<0.0001
History of atrial fibrillation, n (%)	14 (5.6%)	40 (16.0%)	0.0002	10 (5.2%)	21 (11.0%)	0.059	NA	NA	NA
Creatinine, µmol/L	84 (72–103)	88 (72–110)	0.19	76 (68–88)	79 (65–95)	0.25	88 (80–106)	97 (80–115)	0.0003
Ratio total/HDL cholesterol	3.62±0.95	3.51±0.98	0.22	4.20±1.31	4.35±1.35	0.26	4.28±1.33	4.45±1.37	0.018

ARIC indicates Atherosclerosis Risk in Communities study; FHS, Framingham Heart Study; HDL, high-density lipoprotein; HF, heart failure; and HOMAGE, Heart Omics and Ageing cohort.

**P* value from *t* test or Wilcoxon test as appropriate for continuous variable, or Fisher exact test for categorical variable.

†Up until HF occurrence in the case group or time at which controls were considered for matching in the control group.

Table 2. ORs and 95% CIs for the Biomarkers Associated (*P*<0.05) With Incident Heart Failure on Mono-Biomarker Analysis After Adjusting for Matching and Clinical Variables and Correction for Multiple Comparisons in Each Cohort

ARIC			FHS			HOMAGE		
Biomarker	OR (CI 95%)	BH-adjusted *P* value	Biomarker	OR (CI 95%)	BH-adjusted *P* value	Biomarker	OR (CI 95%)	BH-adjusted *P* value
NT-proBNP*	1.95 (1.61–2.36)	<0.0001	NT-proBNP*	1.79 (1.41–2.26)	0.0003	BNP*	1.66 (1.50–1.85)	<0.0001
BNP*	1.65 (1.40–1.93)	<0.0001	GDF-15*	3.53 (2.05–6.09)	0.0006	NT-proBNP*	1.90 (1.65–2.18)	<0.0001
TRAIL-R2*	4.52 (2.44–8.40)	0.0001	BNP*	1.88 (1.42–2.48)	0.0009	TRAIL-R2*	2.43 (1.85–3.19)	<0.0001
ADM	3.59 (2.01–6.43)	0.0008	HGF	3.84 (2.00–7.40)	0.003	GDF-15*	2.00 (1.61–2.48)	<0.0001
CXCL5	0.74 (0.64–0.85)	0.0008	OSM	1.86 (1.36–2.55)	0.005	Gal-9	2.30 (1.69–3.14)	<0.0001
PDGF sub-unit B	0.68 (0.57–0.82)	0.002	TGF-alpha	4.84 (2.09–11.19)	0.009	FGF-23	1.53 (1.30–1.80)	<0.0001
CASP-3	0.66 (0.53–0.81)	0.002	Gal-4	2.46 (1.51–4.00)	0.010	TNFRS-F13B	1.75 (1.41–2.18)	<0.0001
THBS2	6.18 (2.52–15.17)	0.002	MMP-9	1.97 (1.36–2.84)	0.010	U-PAR	1.82 (1.44–2.31)	<0.0001
CXCL1	0.68 (0.55–0.82)	0.002	4E-BP1	1.86 (1.31–2.63)	0.013	TNFRS-F10A	1.94 (1.48–2.54)	<0.0001
PDGF sub-unit A	0.57 (0.43–0.76)	0.002	U-PAR	3.83 (1.78–8.24)	0.014	IGFBP-1	1.34 (1.19–1.52)	<0.0001
SRC	0.56 (0.42–0.76)	0.002	THBS2	6.09 (2.01–18.46)	0.028	SLAMF7	1.58 (1.30–1.92)	<0.0001
MMP12	1.79 (1.31–2.45)	0.004	ACE2	1.89 (1.28–2.80)	0.028	CCL16	1.47 (1.24–1.73)	0.0001
HB-EGF	0.51 (0.35–0.73)	0.004	Gal-9	4.40 (1.73–11.20)	0.035	TFF3	1.45 (1.23–1.71)	0.0001
HSP 27	0.27 (0.13–0.55)	0.004	IGFBP-7	3.25 (1.54–6.86)	0.035	IGFBP-2	1.51 (1.26–1.82)	0.0001
CD40-L	0.65 (0.52–0.82)	0.004	PLC	5.48 (1.83–16.42)	0.039	OPN	1.57 (1.28–1.92)	0.0001
ITGB1BP2	0.71 (0.59–0.86)	0.004	TNFSF14	2.31 (1.33–4.01)	0.043	FABP4	1.44 (1.22–1.70)	0.0001
PlgR	9.84 (2.80–34.55)	0.005				VEGFD	1.80 (1.38–2.34)	0.0001
SIRT2	0.71 (0.58–0.86)	0.006				CSTB	1.44 (1.22–1.69)	0.0002
ANG-1	0.69 (0.56–0.85)	0.006				CEACAM8	1.41 (1.20–1.66)	0.0003
AXIN1	0.67 (0.53–0.84)	0.006				TNFRS-F11A	1.57 (1.27–1.93)	0.0003
STK4	0.69 (0.56–0.85)	0.006				TGM2	1.34 (1.17–1.54)	0.0003
TFF3	2.29 (1.41–3.71)	0.008				SPON1	1.83 (1.38–2.43)	0.0003
DNER	0.27 (0.13–0.58)	0.008				TWEAK	0.60 (0.48–0.77)	0.0003
CCL17	0.72 (0.60–0.88)	0.009				ACE2	1.47 (1.22–1.76)	0.0003
PGF	2.99 (1.56–5.73)	0.009				HGF	1.61 (1.28–2.03)	0.0004
ST1A1	0.71 (0.58–0.88)	0.012				ADM	1.67 (1.31–2.14)	0.0004
Dkk-1	0.50 (0.33–0.77)	0.013				OPG	1.72 (1.32–2.25)	0.0005
CEACAM8	1.82 (1.26–2.64)	0.013				VEGFA	1.62 (1.28–2.06)	0.0006
SPON2	8.35 (2.23–31.25)	0.013				CXCL11	1.24 (1.11–1.38)	0.0007
LOX-1	2.19 (1.34–3.59)	0.014				CD4	1.65 (1.28–2.13)	0.0009
SELP	0.52 (0.34–0.78)	0.014				EN-RAGE	1.30 (1.14–1.49)	0.0009
JAM-A	0.52 (0.34–0.78)	0.014				RARRES2	1.92 (1.37–2.67)	0.0009
SPON1	3.19 (1.52–6.67)	0.016				REN	1.29 (1.13–1.47)	0.0009
TGF-alpha	3.28 (1.53–7.03)	0.016				LOX-1	1.39 (1.18–1.65)	0.0009
CCL19	1.44 (1.13–1.82)	0.018				TNF-R2	1.51 (1.22–1.87)	0.0009
IL-27	2.14 (1.29–3.53)	0.020				AGRP	1.56 (1.24–1.97)	0.001
IL6	1.42 (1.13–1.79)	0.020				IL2-RA	1.47 (1.20–1.80)	0.001
STAMPB	0.65 (0.49–0.87)	0.020				PSP-D	1.32 (1.14–1.53)	0.001
4E-BP1	0.66 (0.50–0.87)	0.020				TNF-R1	1.57 (1.24–1.99)	0.001
IL-4RA	2.31 (1.32–4.04)	0.020				Gal-4	1.45 (1.19–1.77)	0.001
NEMO	0.70 (0.56–0.89)	0.020				MMP-2	1.56 (1.23–1.97)	0.001
RAGE	2.01 (1.25–3.21)	0.021				SORT1	1.92 (1.35–2.72)	0.001

(*Continued*)

Table 2. Continued

ARIC			FHS			HOMAGE		
Biomarker	OR (CI 95%)	BH-adjusted *P* value	Biomarker	OR (CI 95%)	BH-adjusted *P* value	Biomarker	OR (CI 95%)	BH-adjusted *P* value
TR-AP	0.45 (0.26–0.77)	0.021				CASP-8	1.22 (1.10–1.36)	0.001
TWEAK	0.35 (0.17–0.71)	0.021				IL-27	1.63 (1.25–2.13)	0.001
CST5	0.59 (0.42–0.85)	0.023				SPON2	2.09 (1.39–3.13)	0.001
Gal-9	2.71 (1.36–5.38)	0.024				PGF	1.61 (1.24–2.09)	0.002
PAI	0.70 (0.55–0.90)	0.024				CXCL9	1.30 (1.12–1.50)	0.002
CD84	0.52 (0.33–0.82)	0.024				KIM1	1.28 (1.12–1.48)	0.002
U-PAR	2.29 (1.28–4.08)	0.025				PAR-1	1.54 (1.21–1.97)	0.002
VEGFD	2.02 (1.23–3.29)	0.025				MMP12	1.36 (1.14–1.61)	0.002
HGF	2.12 (1.25–3.61)	0.027				TNFRSF14	1.49 (1.19–1.87)	0.002
CD93	2.39 (1.29–4.43)	0.027				IL6	1.24 (1.10–1.40)	0.002
MCP-4	0.65 (0.48–0.88)	0.028				IGFBP-7	1.45 (1.17–1.79)	0.003
PTX3	1.66 (1.15–2.39)	0.028				PD-L2	1.66 (1.24–2.22)	0.003
FGF-23	1.58 (1.13–2.21)	0.033				TGF-alpha	1.41 (1.15–1.73)	0.003
PRELP	4.39 (1.43–13.50)	0.043				MCP-3	1.36 (1.13–1.64)	0.004
IL2-RA	1.74 (1.14–2.65)	0.046				RETN	1.38 (1.14–1.68)	0.004
IGFBP-2	1.53 (1.10–2.11)	0.046				PLC	1.61 (1.21–2.13)	0.004
EGFR	0.31 (0.13–0.77)	0.048				CD40	1.45 (1.16–1.81)	0.004
DECR1	0.72 (0.55–0.93)	0.048				IL16	1.36 (1.13–1.63)	0.005
CASP-8	0.64 (0.45–0.91)	0.049				ST2	1.36 (1.13–1.65)	0.005
MB	0.65 (0.46–0.91)	0.049				PGLYRP1	1.35 (1.12–1.63)	0.005
						IL-1ra	1.23 (1.08–1.41)	0.006
						HSP 27	1.44 (1.15–1.81)	0.006
						SLAMF1	1.31 (1.10–1.55)	0.007
						IL-4RA	1.40 (1.13–1.73)	0.007
						LTBR	1.45 (1.14–1.85)	0.008
						CCL11	1.40 (1.13–1.75)	0.008
						FAS	1.45 (1.13–1.85)	0.010
						IL-15RA	1.59 (1.17–2.16)	0.010
						PON3	0.78 (0.66–0.92)	0.010
						PRTN3	1.28 (1.09–1.52)	0.010
						TIMP4	1.34 (1.10–1.63)	0.010
						IL-6RA	1.42 (1.12–1.80)	0.011
						SOD2	1.89 (1.23–2.90)	0.011
						vWF	1.16 (1.05–1.28)	0.013
						CCL17	1.16 (1.05–1.29)	0.015
						MMP-3	1.26 (1.07–1.49)	0.016
						LAP TGF-beta-1	1.33 (1.09–1.63)	0.016
						TNFRSF9	1.27 (1.07–1.51)	0.017
						IL-18BP	1.36 (1.09–1.69)	0.017
						THBS2	1.62 (1.15–2.29)	0.017
						VSIG2	1.31 (1.08–1.60)	0.017
						MERTK	1.44 (1.11–1.86)	0.017
						CCL15	1.30 (1.08–1.57)	0.017
						IL18	1.20 (1.05–1.37)	0.018
						JAM-A	1.21 (1.05–1.38)	0.020

(*Continued*)

Table 2. Continued

ARIC			FHS			HOMAGE		
Biomarker	OR (CI 95%)	BH-adjusted *P* value	Biomarker	OR (CI 95%)	BH-adjusted *P* value	Biomarker	OR (CI 95%)	BH-adjusted *P* value
						ICAM-2	1.37 (1.09–1.73)	0.021
						4E-BP1	1.14 (1.04–1.26)	0.021
						PTX3	1.28 (1.07–1.53)	0.022
						TR	1.23 (1.05–1.43)	0.022
						GRN	1.49 (1.10–2.02)	0.025
						hOSCAR	1.63 (1.13–2.35)	0.025
						MMP-9	1.16 (1.04–1.31)	0.026
						PAPPA	1.27 (1.06–1.54)	0.029
						PARP-1	1.15 (1.03–1.28)	0.029
						EPHB4	1.39 (1.08–1.80)	0.030
						CXCL16	1.40 (1.07–1.82)	0.031
						CHI3L1	1.17 (1.03–1.32)	0.031
						CXCL10	1.18 (1.04–1.34)	0.031
						TNFSF14	1.21 (1.04–1.40)	0.032
						MPO	1.28 (1.05–1.55)	0.032
						OSM	1.17 (1.03–1.32)	0.036
						Gal-3	1.28 (1.05–1.57)	0.037
						PRELP	1.66 (1.10–2.51)	0.038
						PIgR	1.65 (1.09–2.49)	0.043
						CD93	1.42 (1.06–1.89)	0.043
						MCP-1	1.20 (1.03–1.40)	0.043
						FS	1.31 (1.04–1.65)	0.046
						HB-EGF	1.29 (1.04–1.59)	0.047
						DECR1	1.10 (1.01–1.19)	0.047
						CTSZ	1.31 (1.04–1.66)	0.047
						CCL20	1.12 (1.02–1.23)	0.047
						CCL23	1.29 (1.04–1.60)	0.049
						Dkk-1	1.21 (1.03–1.43)	0.049
						GT	1.19 (1.02–1.39)	0.049

Protein names are available in Table S1. ARIC indicates Atherosclerosis Risk in Communities study; BH, Benjamini-Hochberg; BNP, brain natriuretic peptide; FHS, Framingham Heart Study; Gal-9, galectin-9; HOMAGE, Heart Omics and Ageing cohort; NT-proBNP, N-terminal pro-B-type natriuretic peptide; OR, odds ratio; THBS2, thrombospondin-2; and U-PAR, urokinase plasminogen activator surface receptor.

*Four variables appeared among the top 10 variables in 2 or more datasets.

4.9% (1.3%–8.4%) in FHS, and 5.7% (3.7%–7.6%) in HOMAGE (all $P<0.01$; Table 3).

The added predictive value for incident HF of a multimarker approach, on top of clinical HF risk factors and NT-proBNP was significant in all 3 cohorts (delta *C*-index=11.1% [7.5%–14.7%] in ARIC, 5.9% [2.6%–9.2%] in FHS, and 7.5% [5.4%–9.5%] in HOMAGE; all $P<0.0001$) and of larger magnitude than that of NT-proBNP when considered alone on top of routine HF risk factors (Figure 3).

The total added predictive value of NT-proBNP and the multimarker approach was an increment in the *C*-index ranging from 10.8% (6.3%–15.3%) in FHS to 18.7% (14.0%–23.4%) in ARIC (Figure 3). Importantly, the achieved *C*-index (despite age and sex being part of the matching procedure) ranged from 79.8% in FHS to 88.1% in ARIC.

Network Analysis

When including the 142 proteins significantly associated with HF risk in complex network analysis (Figure 4), several overrepresented pathways were identified as relevant, mostly related to inflammation and remodeling. Involved inflammatory pathways were mainly related to TNF and interleukin. Remodeling pathways were mainly related to extracellular matrix and apoptosis.

When further narrowing the analysis to over-represented GO terms (Supplemental Material), we identified 9 mechanistic clusters: matrix metalloproteinase (MMP)/U-PAR, GDF-15/interleukin-4, apoptotic mechanisms, endocytosis/phagocytosis mechanisms, chemokines, regulation of inflammation and inflammation effectors, TNF/ nuclear factor-kappa B/TNF-related

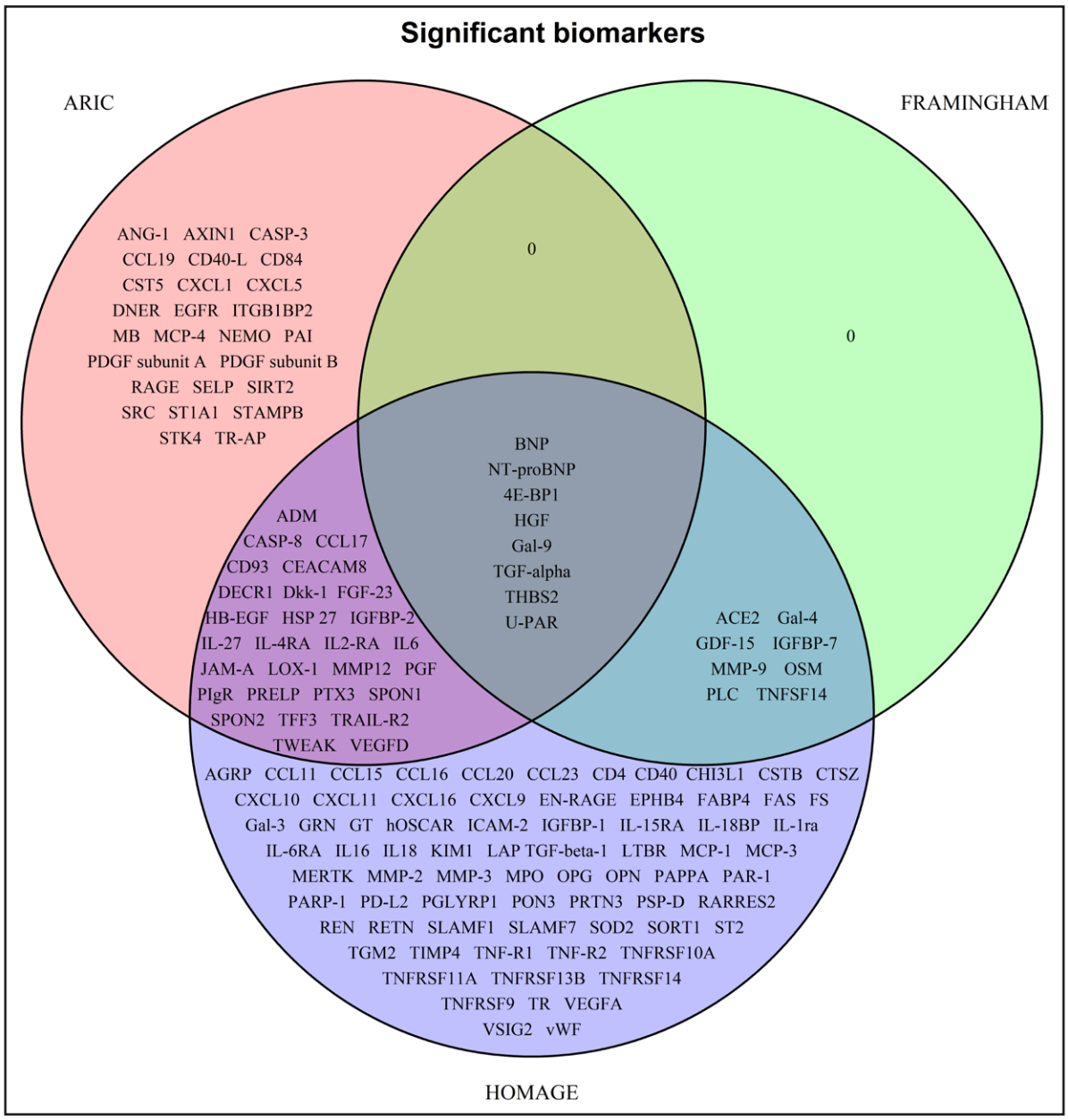

Figure 1. Venn diagram with the significant biomarkers identified in each cohort.
Protein names are available in Table S1.

apoptosis-inducing ligand, TNF/interleukin-12 and interleukin-related mechanisms.

DISCUSSION

We identified 142 protein biomarkers of incident HF in at least 1 of the 3 independent cohorts (ARIC, FHS, and HOMAGE). When we tested these proteomic biomarkers in a multimarker approach, we found that they significantly and substantially improved the prediction of incident HF on top of clinical HF risk factors in all 3 cohorts (*C*-index increases all >10%); the improvement in prediction for the protein multi-biomarkers was appreciable after accounting for HF risk factors and NT-proBNP (C-index increases all ≥5.9%). The proteins identified as significantly associated with HF varied across cohorts, but they were all related to inflammatory (eg, TNF and interleukins) and remodeling pathways (eg, extracellular matrix and apoptosis) in complex network analysis. In addition, we identified 9 mechanistic clusters related to these proteins, mostly related to inflammatory pathways.

Most previous biomarker studies of incident HF used single proteins or oligo-marker approaches. The natriuretic peptides have emerged as the most consistent protein predictors of risk,[5] and the results of our analyses confirm their importance in HF prediction. Some oligo-biomarkers strategies have been shown to improve HF risk prediction, such as the LIPID HF

Table 3. Biomarkers That Improved the *C*-Index When Added to the Clinical Model

ARIC			FHS			HOMAGE		
Biomarker	Δ *C*-index (CI 95%)	*P* value	Biomarker	Δ *C*-index (CI 95%)	*P* value	Biomarker	Δ *C*-index (CI 95%)	*P* value
NT-proBNP	7.6 (3.6 to 11.5)	0.0001	GDF-15	5.2 (1.8 to 8.6)	0.002	BNP	5.6 (3.7 to 7.5)	<0.0001
BNP	5.9 (2.4 to 9.5)	0.001	BNP	4.9 (1.7 to 8.2)	0.002	NT-proBNP	5.7 (3.7 to 7.6)	<0.0001
ADM	3.4 (0.6 to 6.2)	0.018	NT-proBNP	4.9 (1.3 to 8.4)	0.007	GDF-15	3.0 (1.5 to 4.5)	<0.0001
PDGF subunit B	3.0 (0.3 to 5.8)	0.030	HGF	4.0 (1.0 to 6.9)	0.009	TRAIL-R2	2.7 (1.2 to 4.3)	0.0004
TRAIL-R2	3.4 (0.3 to 6.6)	0.031	MMP-9	3.5 (0.6 to 6.3)	0.016	U-PAR	2.0 (0.8 to 3.2)	0.001
HB-EGF	2.7 (0.1 to 5.2)	0.038	OSM	3.4 (0.4 to 6.4)	0.024	IGFBP-1	1.9 (0.7 to 3.0)	0.001
CXCL5	3.0 (0.1 to 5.9)	0.041	TGF-alpha	3.1 (0.2 to 6.0)	0.037	FGF-23 (CVD II panel)	1.9 (0.7 to 3.1)	0.002
PDGF subunit A	2.7 (0.1 to 5.3)	0.041	THBS2	2.6 (0.1 to 5.1)	0.042	VEGFD	1.6 (0.5 to 2.7)	0.003
FGF-21	2.0 (0.0 to 3.9)	0.048	TWEAK	2.3 (0.0 to 4.6)	0.046	TFF3	1.6 (0.5 to 2.7)	0.003
DNER	2.4 (0.0 to 4.9)	0.048				SPON1	1.4 (0.4 to 2.4)	0.004
ANG-1	2.4 (0.0 to 4.9)	0.049				IGFBP-2	1.5 (0.4 to 2.5)	0.006
						OPG (CVD III panel)	1.4 (0.4 to 2.4)	0.006
						CCL16	1.5 (0.4 to 2.6)	0.006
						TNFRSF13B	1.7 (0.5 to 2.9)	0.007
						Gal-9	1.7 (0.4 to 3.0)	0.009
						IL6 (CVD II panel)	1.1 (0.3 to 2.0)	0.009
						FABP4	1.4 (0.3 to 2.5)	0.009
						CSTB	1.4 (0.3 to 2.4)	0.010
						MMP-2	1.2 (0.3 to 2.0)	0.011
						EN-RAGE	1.2 (0.3 to 2.1)	0.012
						TNFRSF10A	1.5 (0.3 to 2.7)	0.013
						TNF-R1	1.1 (0.2 to 2.1)	0.013
						HGF	1.2 (0.2 to 2.2)	0.014
						Gal-4	1.1 (0.2 to 2.0)	0.016
						OPN	1.3 (0.2 to 2.4)	0.017
						VEGFA	1.2 (0.2 to 2.1)	0.017
						IL2-RA	1.1 (0.2 to 2.0)	0.018
						PSP-D	1.1 (0.2 to 2.1)	0.018
						CXCL9	1.1 (0.2 to 2.0)	0.019
						TNF-R2	1.1 (0.2 to 2.0)	0.020
						MMP-3	0.9 (0.1 to 1.6)	0.021
						CEACAM8	1.2 (0.2 to 2.3)	0.021
						TNFRSF14	1.0 (0.1 to 1.9)	0.021
						RARRES2	1.0 (0.1 to 2.0)	0.026
						CXCL11	1.1 (0.1 to 2.1)	0.027
						TIMP4	0.8 (0.1 to 1.5)	0.028
						PGLYRP1	0.9 (0.1 to 1.7)	0.031
						AGRP	1.0 (0.1 to 2.0)	0.032
						TGF-alpha	0.9 (0.1 to 1.6)	0.032
						PLC	0.9 (0.1 to 1.7)	0.033
						SORT1	1.0 (0.1 to 1.9)	0.033
						TNFRSF11A	1.2 (0.1 to 2.2)	0.033
						SLAMF7	1.2 (0.1 to 2.3)	0.035
						LAP TGF-beta-1	0.7 (0.1 to 1.4)	0.035
						MCP-3	0.9 (0.1 to 1.7)	0.035
						IL-27	0.9 (0.1 to 1.8)	0.037

(*Continued*)

Table 3. Continued

ARIC			FHS			HOMAGE		
Biomarker	Δ *C*-index (CI 95%)	*P* value	Biomarker	Δ *C*-index (CI 95%)	*P* value	Biomarker	Δ *C*-index (CI 95%)	*P* value
						RETN	0.8 (0.0 to 1.6)	0.039
						IL-6RA	0.8 (0.0 to 1.5)	0.041
						LTBR	0.8 (0.0 to 1.5)	0.042
						IGFBP-7	0.8 (0.0 to 1.6)	0.043
						LOX-1	0.9 (0.0 to 1.9)	0.045
						vWF	0.7 (0.0 to 1.5)	0.046
						JAM-A	0.7 (0.0 to 1.3)	0.049
						ACE2	1.0 (0.0 to 2.0)	0.049
						TGM2	1.0 (0.0 to 2.1)	0.049
						PON3	0.7 (0.0 to 1.5)	0.049

Protein names are available in Table S1. ARIC indicates Atherosclerosis Risk in Communities study; BNP, brain natriuretic peptide; FHS, Framingham Heart Study; GDF-15, growth/differentiation factor 15; HGF, hepatocyte growth factor; HOMAGE, Heart Omics and Ageing cohort; MMP-9, matrix metalloproteinase-2; NT-proBNP, N-terminal pro-B-type natriuretic peptide; THBS2, thrombospondin-2; and TWEAK, tumor necrosis factor.

risk-prediction model (BNP, cystatin C, D-dimer, high-sensitivity CRP, and high-sensitivity troponin I), with an increment in *C*-index of 4% on top of clinical risk factor.[30] Stenemo et al[11] used data from 2 community-based cohorts from Uppsala, Sweden (mean age 70 and 78 years) and reported that the use of 18 proteins (out of 80 measured proteins) resulted in a sizeable increase in *C*-index (10.1%; *P*=0.006) on top of established HF risk factors. However, the incremental value of the protein biomarkers on top of clinical HF risk factors and natriuretic peptides was modest and not significant (change in *C*-index 2.0%; *P*=0.40).[11] Based on findings in the FHS, Ho et al[13] reported that NT-proBNP, GDF-15, CRP, and leptin (among 85 tested biomarkers) jointly predicted incident HF and were associated with a modest but statistically significant 3% improvement in the C-statistic on top of clinical risk factors (0.843–0.873; *P*<0.0001). This analysis, however, did not separately adjust for natriuretic peptides. We show that in all 3 cohorts there was a significant increase in *C*-index from the incorporation of multiple proteins on top of routine HF risk factors (whether excluding or including NT-proBNP in the models). In addition, we show that the increase in prediction arising from NT-proBNP on top of clinical HF risk factors was substantial (increment in *C*-index ranging from 4.9 to 7.6 in the 3 cohorts), but smaller than the increment arising from the further addition of multiple proteins on top of NT-proBNP (increment in *C*-index 5.9 to 11.1 in the 3 cohorts; Figure 3).

In the study by Stenemo et al,[11] 9 proteins (GDF-15, TIM-1 [T-cell immunoglobulin and mucin domain 1], TNF-related apoptosis-inducing ligand-R2, SPON1 [spondin-1], MMP-12, FS [follistatin], U-PAR, osteoprotegerin [OPG], and ST2) were associated with incident HF after adjusting for established HF risk factors. Among these 9 protein biomarkers, the only 1 not measured in our study was TIM-1 Of these previously reported proteins, only U-PAR was significantly associated with incident HF in each of our cohorts. GDF-15 was significantly associated with incident HF in FHS and HOMAGE, whereas TNF-related apoptosis-inducing ligand-R2, SPON1, and MMP-12 were significantly associated with incident HF in ARIC and HOMAGE. ST2 and FS were significantly associated with HF only in HOMAGE.

When including the HF-associated proteins in complex network analysis, all were related to inflammatory and remodeling pathways. This finding emphasizes the central importance of these pathways in HF pathogenesis, as emphasized previously in relation to HF with preserved ejection fraction.[31] In addition, when considering GO terms complex network analysis, a limited number of mechanistic clusters were highlighted, mostly related to inflammatory processes. These network results strongly emphasize the central importance of inflammation in HF.

The results presented herein expand the knowledge derived from a previous work from the HOMAGE project. For the purpose of the current analysis, we merged the datasets of the 2 phases of the previously reported HOMAGE study[20] (totalizing 562 cases and 871 controls) and conducted the same analysis in 3 large case-control studies: HOMAGE (merged), ARIC and FHS. This new analysis provides an important increase in the level of external validity achieve by our results, as they arise from both American and European settings. From this analysis that maximizes external validity, 8 proteins (BNP, NT-proBNP, 4E-BP1, HGF, Gal-9, TGF-alpha, THBS2, U-PAR) appeared as the most replicable biomarkers predicting incident HF in all cohorts. This top list of biomarkers, in light of the strength of the combined analysis we report herein, should be considered in future biomarker studies aiming to predict incident HF. Importantly, among this list of 8 proteins, 4E-BP1, HGF, TGF-alpha, THBS2, we not identified in our previous report as significant

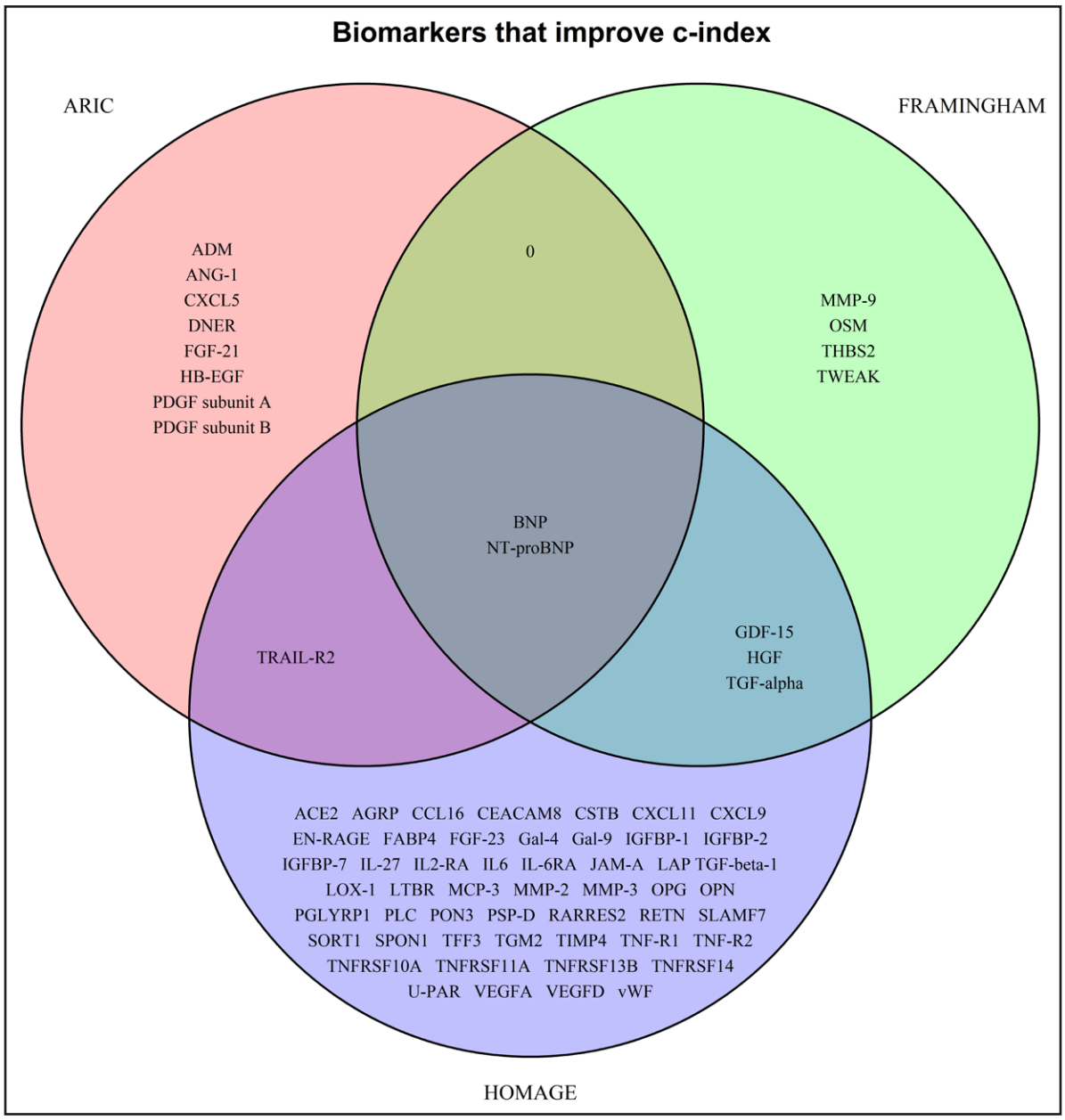

Figure 2. Venn diagram for the biomarkers in each cohort that improved the *C*-index when added to the clinical model.
Protein names are available in Table S1.

predictors of incident HF,[20] probably because of lower power compared with this new multi-cohort analysis. In addition, the current analysis also expands our mechanistic understanding of the biological pathways underlying the pathogenesis of HF. In our earlier analysis, we reported pathways related to inflammation, apoptosis, vascular function, matrix remodeling, blood pressure control, and metabolism.[20] With the additional data we report herein, we are able to further refine biological mechanisms an focus on a limited number of biological processes, namely inflammation (through tumor necrosis factor and interleukins) and remodeling (involving extracellular matrix and apoptosis). This may help guide future therapeutic development to target these key biological pathways in order to best prevent HF.

Perspectives

The total predictive value achieved using routine clinical and biological predictors is excellent (*C*-index ranging from 80% to 90%) despite cases and controls being matched on age and sex, that is, thus decreasing the *C*-index by removing these important variables from the model. This suggests an excellent total predictive value of an integrated approach based on numerous biological biomarkers and clinical variables. Adequately predicting individuals

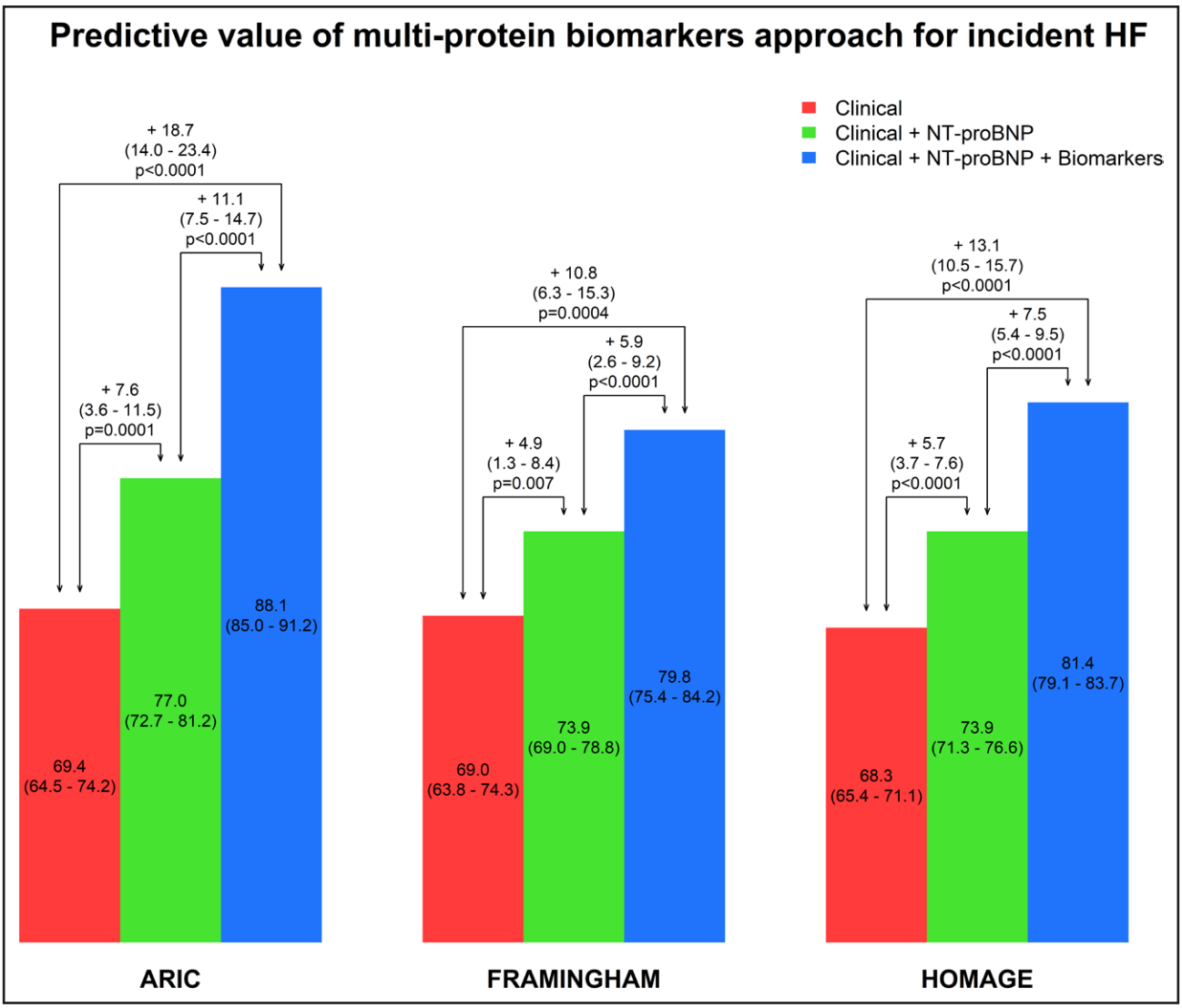

Figure 3. Predictive value of multiprotein biomarkers approach for incident heart failure (HF) in addition to routine clinical markers of incident heart failure in the 3 cohorts.

Biomarkers retained in multivariable model in each cohort (including NT-proBNP [N-terminal pro-B-type natriuretic peptide]) included the following: ARIC (Atherosclerosis Risk in Communities study; *P*=55): all significant BMs, except BNP (brain natriuretic peptide)+AXIN1, CASP-3, PDGF subunit A, PDGF subunit B, SRC, STAMPB; FHS (Framingham Heart Study; *P*=15): all significant BMs, except BNP; HOMAGE (Heart Omics and Ageing cohort; *P*=109): all significant BMs, except BNP, EPHB4, IL-18BP, OSM, TNF-R1, TNF-R2, TNFRSF14. The clinical model included the following variables: age, sex, smoking, diabetes, history of coronary artery disease, serum creatinine, body mass index, systolic blood pressure, use of antihypertensive medication, heart rate, hypertension, history of atrial fibrillation, and ratio total/high-density lipoprotein cholesterol.

at high risk for HF is the prerequisite of tailored interventions to prevent the onset of HF in a populational setting.

Identifying protein biomarkers of incident HF could have a major impact on our understanding of HF pathogenesis. Identifying pathway involved in the transition to HF with preserved ejection fraction in specific clinical phenotypes would considerably improve our understanding of this disease and could help tailor therapeutic interventions targeting specific biological processes relevant to specific clinical settings. In addition, improving the prediction of incident HF could pave the way for the design of preventive trials.

Preventive trials are of uttermost importance as the therapeutic management of HF remains imperfect, especially for HF with preserved ejection fraction. We believe that this new avenue in biomarker research is likely to greatly extend in the next few years and could have a major impact on preventive cardiology.

Limitations

Several limitations should be highlighted in the present study. First, this is an observational case-control study, hence causality cannot be ascertained.

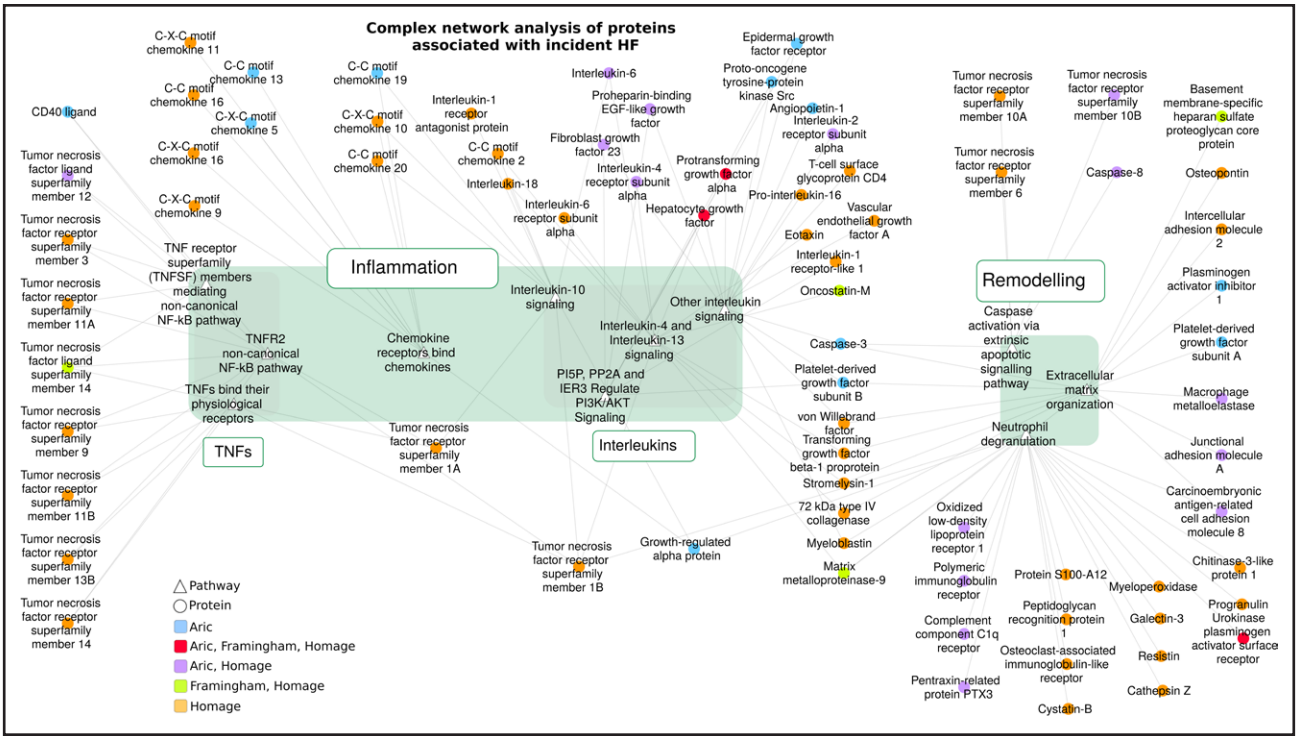

Figure 4. Complex network analysis of significant protein biomarkers.
Overrepresented pathways are symbolized by white triangles and biomarkers are ellipses. Biomarker color depends on the study where it was identified. Protein names are available in Table S1.

Natriuretic peptides were not routinely used to ascertain HF diagnosis in most patients considered in this analysis as it was not a standard of care in the era during which study participants were recruited. However, our definition required being hospitalized for HF, which may be sufficiently stringent to ensure patients did have HF despite not considering natriuretic peptides.

The mean age of patients included in this analysis ranged from 68 to 78 years according to the considered cohort. Our results consequently apply to the development of HF in fairly old population and do not provide insight on the development of specific causes of HF encountered usually/more frequently earlier in life such as peripartum, stress induced or tachycardia related cardiomyopathy.

We did not use large unbiased screens but rather preselected protein biomarkers based on mechanistic hypotheses and previous literature. The Olink panels used herein were mostly biomarkers related to cardiovascular and inflammatory diseases. We cannot consequently ascertain we did not miss important biomarkers not included in these panels.

Whether the biomarkers identified as associated with incident HF are causally involved in the development of HF and could serve as therapeutic targets is uncertain. It is possible that many biomarkers are actually bystander variables predicting HF incident because they are related to underlying causes of HF. The causality between these circulating biomarkers and HF genesis is to be further explored in animal/mechanistic models.

We did not have access to the left ventricular ejection fraction at the time of hospitalization; therefore, we cannot assess the potential value of these biomarkers in distinguishing progression to HF with reduced ejection fraction from HF with preserved ejection fraction and the HF cause.

The proteomics assay does not provide standard concentration units, making comparisons with clinically applied cutoffs difficult.

Most of the cohorts used in this analysis included a vast majority of Whites. Our results consequently do not improve our understanding of biological mechanisms involved in HF onset in other races.

Conclusions

In 3 large cohorts, a multiprotein approach improved HF prediction. The proteins identified as significantly associated with HF varied across cohorts but were all related to inflammatory (eg, TNF and interleukins) and remodeling (eg, extracellular matrix and apoptosis) pathways in complex network analysis. We identified 9 key mechanistic clusters related to these proteins, mostly related to inflammatory pathways, which further emphasizes the pivotal role of inflammation in the pathogenesis of HF.

ARTICLE INFORMATION

Received March 18, 2022; accepted March 3, 2023.

Affiliations

Université de Lorraine, Inserm, Centre d'Investigations Cliniques- Plurithématique 14-33, and Inserm U1116, CHRU, F-CRIN INI-CRCT (Cardiovascular and Renal Clinical Trialists), Nancy, France (N.G., K.D., J.P.F., A.P., A.-C.H., P.R., F.Z.). National Heart, Lung, and Blood Institute's and Boston University's Framingham Heart Study, Framingham, MA (D.L., J.H., S.-J.H). Population Sciences Branch, Division of Intramural Research, National Heart, Lung, and Blood Institute, Bethesda, MD (D.L., J.H., S.-J.H.). Baylor College of Medicine, Houston, TX (C.B.). London School of Hygiene and Tropical Medicine, United Kingdom (T.C.). Inserm 1024, Institut de Biologie de l'École Normale Supérieure (IBENS), PSL University of Paris, France (A.P.). TATAA Biocenter AB, Gothenburg, Sweden (J.B.). Department of Medicine, University of Mississippi School of Medicine, Jackson (J.B.). Hull York Medical School, Castle Hill Hospital, Cottingham, United Kingdom (A.C.). Robertson Centre for Biostatistics and Clinical Trials, Institute of Health and Wellbeing, University of Glasgow, United Kingdom (J.G.C., P.P.). National Heart and Lung Institute, Royal Brompton and Harefield Hospitals, Imperial College, London, United Kingdom (J.G.C.). Institute of Cardiovascular and Medical Sciences, University of Glasgow, United Kingdom (C.D.C.). Program of Cardiovascular Diseases, Centre for Applied Medical Research, University of Navarra, Pamplona, Spain (J.D.C., A.G.C.). CIBERCV, Carlos III Institute of Health, Madrid, Spain (J.D.C., A.G.C.). Instituto de Investigación Sanitaria de Navarra (IdiSNA), Spain (J.D.C., A.G.C.). Departments of Nephrology, and Cardiology and Cardiac Surgery, University of Navarra Clinic, Pamplona, Spain (A.G.C.). Department of Cardiology, Maastricht University Medical Centre, Center for Heart Failure Research, Cardiovascular Research Institute Maastricht (CARIM), University Hospital Maastricht, the Netherlands (M.H.C.). IRCCS-Istituto di Ricerche Farmacologiche Mario Negri, Milano, Italy (R.L.). Department of Cardiology, Cortona Hospital, Arezzo, Italy (B.M.). UMRS 942; University Paris Diderot; APHP, University Hospitals Saint Louis Lariboisière, France (A.M.). Institute of Cardiovascular and Medical Sciences, University of Glasgow, United Kingdom (N.S.). International Centre for Circulatory Health, National Heart and Lung Institute, Imperial College London, United Kingdom (P.S.). Studies Coordinating Centre, Research Unit Hypertension and Cardiovascular Epidemiology, KU Leuven Department of Cardiovascular Sciences, University of Leuven, Belgium (J.A.S.). Cardiovascular Research Institute Maastricht (CARIM), Maastricht University, the Netherlands (J.A.S., J.V.). Department of Cardiovascular Research, University of Leuven, UZ Herestraat, Belgium (S.H.). Netherlands Heart Institute (ICIN), Utrecht, the Netherlands (S.H.). Department of Clinical Genetics, Maastricht University Medical Center, the Netherlands (S.H.).

Sources of Funding

The research leading to these results has received funding from the European Union Commission's Seventh Framework Programme under grant agreement 305507 (HOMAGE [Heart Omics in Ageing consortium]). The authors acknowledge the support from the Netherlands Cardiovascular Research Initiative, an initiative with the support of the Dutch Heart Foundation CVON2016-Early HFPEF, and CVON 2017, ShePREDICTS.

Disclosures

Drs Ferreira, Girerd, Rossignol, and Zannad are supported by the French National Research Agency Fighting Heart Failure (ANR-15-RHU-0004), by the French PIA project Lorraine Université d'Excellence GEENAGE (ANR-15-IDEX-04-LUE) programs, and the Contrat de Plan Etat Région Lorraine and FEDER IT2MP. Dr Girerd reports honoraria from Novartis, AstraZeneka, Boehringer, and Vifor. Dr Rossignol received personal fees (consulting) from Idorsia and G3P, honoraria from AstraZeneca, Bayer, CVRx, Fresenius, Grunenthal, Novartis, Novo-Nordisk, Servier, StealthPeptides, Ablative Solutions, Corvidia, Relypsa, and Vifor Fresenius Medical Care Renal Pharma, outside the submitted work, and is the cofounder of CardioRenal. Dr Zannad reports steering committee personal fees from Applied Therapeutics, Bayer, Boehringer, Boston Scientific, Novartis, Janssen, and CVRx, advisory board personal fees from AstraZeneca, Vifor Fresenius, Cardior, Cereno Pharmaceutical, and Merck, stock options at G3Pharmaceutical, and being the founder of CardioRenal and CVCT. Dr Mebazaa received honoraria for lectures from Roche and Abbott, consultation fees from Sanofi and Servier, and research grants from Adrenomed and Sphyngotec. Dr Ballantyne is the coinventor on a provisional patent (patent 61721475) entitled Biomarkers to Improve Prediction of Heart Failure Risk filed by Roche and Baylor College of Medicine on their behalf. Dr Ballantyne has received grant support from Abbott Diagnostics and Roche Diagnostics and is a consultant for Roche Diagnostics and Abbott Diagnostics. The other authors report no conflicts.

Supplemental Material

Tables S1–S3
Figure S1

REFERENCES

1. McMurray JJ, Stewart S. Epidemiology, aetiology, and prognosis of heart failure. *Heart*. 2000;83:596–602. doi: 10.1136/heart.83.5.596
2. Petrie M, McMurray J. Changes in notions about heart failure. *Lancet*. 2001;358:432–434. doi: 10.1016/S0140-6736(01)05664-1
3. Conrad N, Judge A, Tran J, Mohseni H, Hedgecott D, Crespillo AP, Allison M, Hemingway H, Cleland JG, McMurray JJV, et al. Temporal trends and patterns in heart failure incidence: a population-based study of 4 million individuals. *Lancet*. 2018;391:572–580. doi: 10.1016/s0140-6736(17)32520-5
4. Manzano L, Babalis D, Roughton M, Shibata M, Anker SD, Ghio S, van Veldhuisen DJ, Cohen-Solal A, Coats AJ, Poole-Wilson PPA, et al. Predictors of clinical outcomes in elderly patients with heart failure. *Eur J Heart Fail*. 2011;13:528–536. doi: 10.1093/eurjhf/hfr030
5. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, Falk V, González-Juanatey JR, Harjola V-P, Jankowska EA, et al; Authors/Task Force Members. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur J Heart Fail*. 2016;18:891–975. doi: 10.1002/ejhf.592
6. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Colvin MM, Drazner MH, Filippatos GS, Fonarow GC, Givertz MM, et al. 2017 ACC/AHA/HFSA focused update of the 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines and the Heart Failure Society of America. *J Am Coll Cardiol*. 2017;70:776–803. doi: 10.1016/j.jacc.2017.04.025
7. Wiviott SD, Raz I, Bonaca MP, Mosenzon O, Kato ET, Cahn A, Silverman MG, Zelniker TA, Kuder JF, Murphy SA, et al; DECLARE–TIMI 58 Investigators. Dapagliflozin and cardiovascular outcomes in type 2 diabetes. *N Engl J Med*. 2019;380:347–357. doi: 10.1056/NEJMoa1812389
8. Wang TJ, Wollert KC, Larson MG, Coglianese E, McCabe EL, Cheng S, Ho JE, Fradley MG, Ghorbani A, Xanthakis V, et al. Prognostic utility of novel biomarkers of cardiovascular stress: the Framingham Heart Study. *Circulation*. 2012;126:1596–1604. doi: 10.1161/CIRCULATIONAHA.112.129437
9. Borne Y, Persson M, Melander O, Smith JG, Engstrom G. Increased plasma level of soluble urokinase plasminogen activator receptor is associated with incidence of heart failure but not atrial fibrillation. *Eur J Heart Fail*. 2014;16:377–383. doi: 10.1002/ejhf.49
10. Echouffo-Tcheugui JB, Greene SJ, Papadimitriou L, Zannad F, Yancy CW, Gheorghiade M, Butler J. Population risk prediction models for incident heart failure: a systematic review. *Circ Heart Fail*. 2015;8:438–447. doi: 10.1161/CIRCHEARTFAILURE.114.001896
11. Stenemo M, Nowak C, Byberg L, Sundström J, Giedraitis V, Lind L, Ingelsson E, Fall T, Ärnlöv J. Circulating proteins as predictors of incident heart failure in the elderly. *Eur J Heart Fail*. 2017;20:55–62. doi: 10.1002/ejhf.980
12. Smith JG, Gerszten RE. Emerging affinity-based proteomic technologies for large-scale plasma profiling in cardiovascular disease. *Circulation*. 2017;135:1651–1664. doi: 10.1161/CIRCULATIONAHA.116.025446
13. Ho JE, Lyass A, Courchesne P, Chen G, Liu C, Yin X, Hwang S-J, Massaro JM, Larson MG, Levy D. Protein biomarkers of cardiovascular disease and mortality in the community. *J Am Heart Assoc*. 2018;7:e008108. doi: 10.1161/JAHA.117.008108
14. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol*. 1989;129:687–702. doi: 10.1093/oxfordjournals.aje.a115184
15. Jacobs L, Thijs L, Jin Y, Zannad F, Mebazaa A, Rouet P, Pinet F, Bauters C, Pieske B, Tomaschitz A, et al. Heart "omics" in AGEing (HOMAGE): design, research objectives and characteristics of the common database. *J Biomed Res*. 2014;28:349–359. doi: 10.7555/JBR.28.20140045
16. Mureddu GF, Agabiti N, Rizzello V, Forastiere F, Latini R, Cesaroni G, Masson S, Cacciatore G, Colivicchi F, Uguccioni M, et al; PREDICTOR Study Group. Prevalence of preclinical and clinical heart failure in the elderly. A population-based study in Central Italy. *Eur J Heart Fail*. 2012;14:718–729. doi: 10.1093/eurjhf/hfs052

17. Beavers KM, Hsu FC, Houston DK, Beavers DP, Harris TB, Hue TF, Kim LJ, Koster A, Penninx BW, Simonsick EM, et al; Health ABC Study. The role of metabolic syndrome, adiposity, and inflammation in physical performance in the Health ABC Study. *J Gerontol A Biol Sci Med Sci*. 2013;68:617–623. doi: 10.1093/gerona/gls213
18. Shepherd J, Blauw GJ, Murphy MB, Cobbe SM, Bollen EL, Buckley BM, Ford I, Jukema JW, Hyland M, Gaw A, et al. The design of a prospective study of Pravastatin in the Elderly at Risk (PROSPER). PROSPER Study Group. PROspective Study of Pravastatin in the Elderly at Risk. *Am J Cardiol*. 1999;84:1192–1197. doi: 10.1016/s0002-9149(99)00533-0
19. Essebag V, Genest J Jr, Suissa S, Pilote L. The nested case-control study in cardiology. *Am Heart J*. 2003;146:581–590. doi: 10.1016/S0002-8703(03)00512-X
20. Ferreira JP, Verdonschot J, Collier T, Wang P, Pizard A, Bär C, Björkman J, Boccanelli A, Butler J, Clark A, et al. Proteomic bioprofiles and mechanistic pathways of progression to heart failure. *Circ Heart Fail*. 2019;12:e005897. doi: 10.1161/CIRCHEARTFAILURE.118.005897
21. Lundberg M, Eriksson A, Tran B, Assarsson E, Fredriksson S. Homogeneous antibody-based proximity extension assays provide sensitive and specific detection of low-abundant proteins in human blood. *Nucleic Acids Res*. 2011;39:e102. doi: 10.1093/nar/gkr424
22. Pearce N. Analysis of matched case-control studies. *BMJ*. 2016;352:i969. doi: 10.1136/bmj.i969
23. Jacobs L, Efremov L, Ferreira JP, Thijs L, Yang W-Y, Zhang Z-Y, Latini R, Masson S, Agabiti N, Sever P, et al; Heart "OMics" in AGEing (HOMAGE) investigators. Risk for incident heart failure: a subject-level meta-analysis from the Heart "OMics" in AGEing (HOMAGE) study. *J Am Heart Assoc*. 2017;6:e005231. doi: 10.1161/JAHA.116.005231
24. Green GH, Diggle PJ. On the operational characteristics of the Benjamini and Hochberg False Discovery Rate procedure. *Stat Appl Genet Mol Biol*. 2007;6:Article27. doi: 10.2202/1544-6115.1302
25. Girerd N, Bresso E, Devignes MD, Rossignol P. Insulin-like growth factor binding protein 2: a prognostic biomarker for heart failure hardly redundant with natriuretic peptides. *Int J Cardiol*. 2020;300:252–254. doi: 10.1016/j.ijcard.2019.11.100
26. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 2003;13:2498–2504. doi: 10.1101/gr.1239303
27. Newman ME, Girvan M. Finding and evaluating community structure in networks. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2004;69:026113. doi: 10.1103/PhysRevE.69.026113
28. Su G, Kuchinsky A, Morris JH, States DJ, Meng F. GLay: community structure analysis of biological networks. *Bioinformatics*. 2010;26:3135–3137. doi: 10.1093/bioinformatics/btq596
29. Morris JH, Apeltsin L, Newman AM, Baumbach J, Wittkop T, Su G, Bader GD, Ferrin TE. clusterMaker: a multi-algorithm clustering plugin for Cytoscape. *BMC Bioinf*. 2011;12:436. doi: 10.1186/1471-2105-12-436
30. Driscoll A, Barnes EH, Blankenberg S, Colquhoun DM, Hunt D, Nestel PJ, Stewart RA, West MJ, White HD, Simes J, et al. Predictors of incident heart failure in patients after an acute coronary syndrome: the LIPID heart failure risk-prediction model. *Int J Cardiol*. 2017;248:361–368. doi: 10.1016/j.ijcard.2017.06.098
31. Paulus WJ, Tschope C. A novel paradigm for heart failure with preserved ejection fraction: comorbidities drive myocardial dysfunction and remodeling through coronary microvascular endothelial inflammation. *J Am Coll Cardiol*. 2013;62:263–271. doi: 10.1016/j.jacc.2013.02.092