Circulation

ORIGINAL RESEARCH ARTICLE

Therapeutic Targets for Heart Failure Identified Using Proteomics and Mendelian Randomization

Albert Henry, MD, MSc; María Gordillo-Marañón, MSc; Chris Finan, PhD; Amand F. Schmidt, MSc, PhD; João Pedro Ferreira, MD, PhD; Ravi Karra, MD, MHS; Johan Sundström, MD, PhD; Lars Lind, MD, PhD; Johan Ärnlöv, MD, PhD; Faiez Zannad, MD, PhD; Anders Mälarstig, MSc, PhD*; Aroon D. Hingorani, MBBS, PhD*; R. Thomas Lumbers, MB BChir, PhD*; HERMES and SCALLOP Consortia

BACKGROUND: Heart failure (HF) is a highly prevalent disorder for which disease mechanisms are incompletely understood. The discovery of disease-associated proteins with causal genetic evidence provides an opportunity to identify new therapeutic targets.

METHODS: We investigated the observational and causal associations of 90 cardiovascular proteins, which were measured using affinity-based proteomic assays. First, we estimated the associations of 90 cardiovascular proteins with incident heart failure by means of a fixed-effect meta-analysis of 4 population-based studies, composed of a total of 3019 participants with 732 HF events. The causal effects of HF-associated proteins were then investigated by Mendelian randomization, using *cis*-protein quantitative loci genetic instruments identified from genomewide association studies in more than 30 000 individuals. To improve the precision of causal estimates, we implemented an Mendelian randomization model that accounted for linkage disequilibrium between instruments and tested the robustness of causal estimates through a multiverse sensitivity analysis that included up to 120 combinations of instrument selection parameters and Mendelian randomization models per protein. The druggability of candidate proteins was surveyed, and mechanism of action and potential on-target side effects were explored with cross-trait Mendelian randomization analysis.

RESULTS: Forty-four of ninety proteins were positively associated with risk of incident HF ($P<6.0\times10^{-4}$). Among these, 8 proteins had evidence of a causal association with HF that was robust to multiverse sensitivity analysis: higher CSF-1 (macrophage colony-stimulating factor 1), Gal-3 (galectin-3) and KIM-1 (kidney injury molecule 1) were positively associated with risk of HF, whereas higher ADM (adrenomedullin), CHI3L1 (chitinase-3-like protein 1), CTSL1 (cathepsin L1), FGF-23 (fibroblast growth factor 23), and MMP-12 (matrix metalloproteinase-12) were protective. Therapeutics targeting ADM and Gal-3 are currently under evaluation in clinical trials, and all the remaining proteins were considered druggable, except KIM-1.

CONCLUSIONS: We identified 44 circulating proteins that were associated with incident HF, of which 8 showed evidence of a causal relationship and 7 were druggable, including adrenomedullin, which represents a particularly promising drug target. Our approach demonstrates a tractable roadmap for the triangulation of population genomic and proteomic data for the prioritization of therapeutic targets for complex human diseases.

Key Words: drug target prediction ■ heart failure ■ Mendelian randomization analysis ■ proteomics

Circulation is available at www.ahajournals.org/journal/circ

Clinical Perspective

What Is New?

- Among 90 proteins investigated for their association with heart failure onset, 44 were observationally associated, and 8 were causally associated, 2 of which are the target of drugs in early clinical trials for heart failure.
- Targeting adrenomedullin was estimated to protect against new-onset heart failure consistent with the agonist effect of adrenomedullin drug antibodies, which are under evaluation in clinical trials.

What Are the Clinical Implications?

- Findings provide confirmatory evidence for the development and evaluation of therapeutics targeting galectin-3 and adrenomedullin, which are currently being pursued in clinical trials for heart failure.
- Integrating population-scale genomic and proteomic data through triangulation of observational and Mendelian randomization analyses facilitates prioritization of drug targets and provides insights into molecular mechanisms of a complex clinical syndrome.

Nonstandard Abbreviations and Acronyms

BNP	B-type natriuretic peptide
CAD	coronary artery disease
***cis*-MR**	Mendelian randomization using *cis*-acting protein quantitative trait loci instruments
GWAS	genome-wide association study
Health ABC	Health Aging and Body Composition
HERMES	Heart Failure Molecular Epidemiology for Therapeutic Targets Consortium
HF	heart failure
HOMAGE	Heart Omics in Ageing
IVW	Mendelian randomization with inverse-variance weighted estimator
LD	linkage disequilibrium
MR	Mendelian randomization
NT-proBNP	N-terminal pro-BNP
PIVUS	Prospective Investigation of the Vasculature in Uppsala Seniors
PREDICTOR	Valutazione della Prevalenza di Disfunzione Cardiaca Asintomatica e di Scompenso Cardiaco
PROSPER	Prospective Study of Pravastatin in the Elderly at Risk
RR	risk ratio
SCALLOP	Systematic and Combined Analysis of Olink Proteins Consortium
ULSAM	Uppsala Longitudinal Study of Adult Men

Heart failure (HF) is a clinical syndrome arising from disease processes that either injure or overload the heart muscle leading to inadequate function at normal filling pressures.[1] Despite primary prevention through treatment of known antecedent risk factors, the prevalence is rising, and the burden of associated morbidity and mortality remains high.[2] The challenge of recapitulating a complex age-associated disease entity such as HF in model systems is reflected in a history of late-stage failures of new therapeutics in clinical trials.[3–5] More robust approaches to drug target identification and validation for HF are therefore required.[5]

Proteins are frequently the principal regulators of molecular pathways and the target of the majority of drugs.[6] The circulating proteome is composed of proteins derived from almost all cells and tissues, which are either actively or passively secreted into the circulation or released during cell damage or turnover.[7] Studies of the human circulating proteome measured using affinity or aptamer-based multiplexed assays have identified a large number of circulating proteins associated with HF onset, progression, and recovery.[8–10] However, the causal relevance of associations from these nonrandomized, observational studies (referred to as observational associations in the present article) remains largely undetermined; they may arise because of confounding factors, reverse causation, or inclusion of undetected or asymptomatic prevalent cases at the time of protein measurement.

Mendelian randomization (MR) can be used to estimate the causal effect of protein levels on disease outcomes,[11] on the condition that 3 core assumptions are met: that genetic instrumental variables are associated with the exposure (relevance assumption); that they are not associated with confounding factors (independence assumption); and that they affect the outcome only through their effects on the exposure of interest (exclusion restriction assumption).[12,13] In addition to the biological relevance of proteins, the use of genetic variants associated with protein level (protein quantitative trait loci) as instrumental variables in MR has desirable properties in relation to these assumptions[14] (Figure 1, Table S1). Protein quantitative trait loci variants are frequently derived from genomewide association studies (GWAS) using population-based genetic and circulating protein level data,[15,16] fulfilling the relevance assumption by definition. The selection of genetic instruments mapping to the vicinity of the transcriptional gene unit (*cis*-acting variants), as opposed to those located more remotely (*trans*-acting variants), limits the scope for violating the exclusion restriction assumption, because protein quantitative trait loci variant effects on the outcome are likely mediated through expressions of the protein under consideration (no horizontal pleiotropy).[17] Last, on the basis of the central dogma of molecular biology, it is implausible that *cis* variant instruments for protein exposures
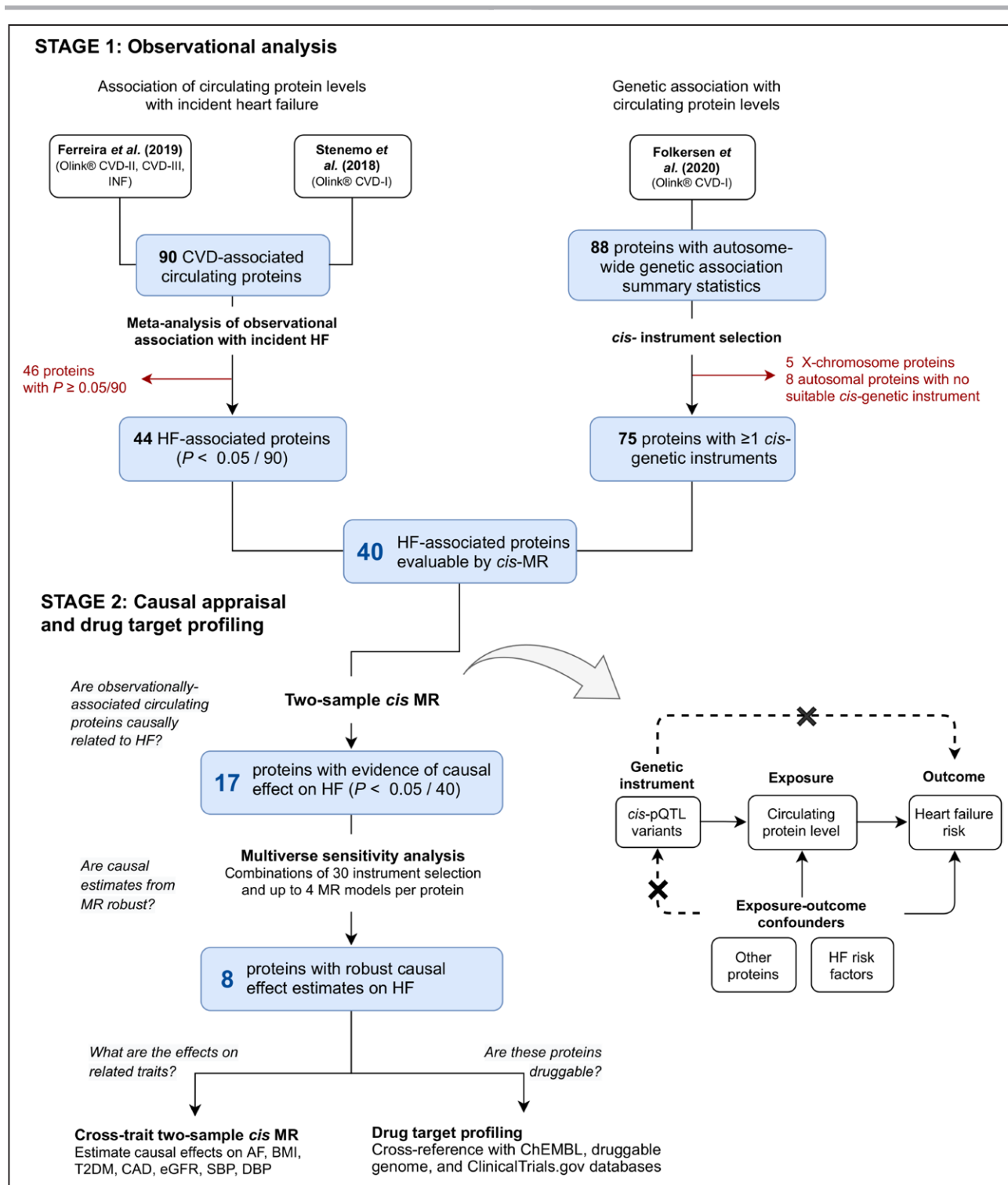
Figure 1. A flow chart of the study design and a schematic illustration of *cis*-MR.
AF indicates atrial fibrillation; BMI, body mass index; CAD, coronary artery disease; *cis*-MR, Mendelian randomization using cis-acting protein quantitative trait loci instruments; DBP, diastolic blood pressure; eGFR, glomerular filtration rate; GWAS, genome-wide association study; HF, heart failure; LD, linkage disequilibrium; MR, Mendelian randomization; pQTL, protein quantitative trait loci; SBP, systolic blood pressure; and T2DM, type 2 diabetes mellitus.

are conditional on the disease outcome, and therefore, it is reasonable to assume that protein traits are upstream of the disease outcome in any causal model. Throughout

the article, we refer to this technique as *cis*-MR, which has been demonstrated to be able to predict efficacy of known drug targets for coronary heart disease.[14]

Here, we report an integrated observational and *cis*-MR analyses of circulating protein levels for therapeutic target identification and prioritization in HF, focusing on up to 90 cardiovascular disease–related circulating proteins measured with the Olink Cardiovascular I circulating protein biomarker panel (Olink CVD-1) multiplexed affinity-based proximity extension assay[18] (Figure 1). We perform meta-analysis of observational associations between circulating protein levels with incident HF[8,9] estimated from 4 independent samples. We estimate the causality of these associations with *cis*-MR analysis by leveraging summary-level data from large GWASs of circulating levels of proteins under study[15] and HF risk.[19] We identify several likely causal proteins, report the anticipated effects on HF-related traits estimated through cross-trait *cis*-MR analysis, and characterize the druggability properties of these proteins as potential therapeutic targets for HF.

METHODS

Data and Code Availability

For purposes of reproducing the results or replicating the procedure, the data and analysis code used in the main analysis have been made available to other researchers at https://github.com/alhenry/cvd1-hf. Other supporting data are available in the article, supplemental files, and referenced public datasets.

Circulating Protein Level Measurement

Circulating protein levels were assessed using Olink Proseek Multiplex proximity extension assay[7,18] technology and were quantified in a normalized protein expression unit, where 1 U difference represents a doubling of protein concentration.[20] The present study focused on cardiovascular-disease related proteins available on the Olink CVD-1 panel, for which both observational associations with HF and genetic association estimates for *cis*-MR analysis were uniquely available at the time of the study. Observational association estimates with incident HF were available for 90 proteins reported in Ferreira et al[9] and Stenemo et al,[8] of which 88 had autosomewide genetic association results reported in Folkersen et al.[15] In the observational studies, protein measures were taken at baseline. A detailed description of the methods used for protein quantification and the proteins measured by each of the included studies is provided in the Supplemental Methods, Table S2, and Figure 1.

Study Population for Observational Analysis

We meta-analyzed observational association estimates between circulating protein level and incident HF from 4 independent samples reported in Ferreira et al[9] and Stenemo et al[8]: HOMAGE (Heart Omics in Ageing)[21] discovery, HOMAGE validation, PIVUS (Prospective Investigation of the Vasculature in Uppsala Seniors),[22] and ULSAM (Uppsala Longitudinal Study of Adult Men).[23] The HOMAGE discovery and validation samples were derived from 2 population cohorts and 1 clinical trial population: Health ABC (Health Aging and Body Composition),[24]

PREDICTOR (Valutazione della Prevalenza di Disfunzione Cardiaca Asintomatica e di Scompenso Cardiaco),[25,26] and PROSPER (Prospective Study of Pravastatin in the Elderly at Risk).[27–29] Individuals with prevalent HF at enrollment were excluded from the analysis. Incident HF was defined as the first diagnosis of HF, ascertained on the basis of hospital record review by trained physicians. The combined sample was composed of 3019 individuals (median age ranged from 70 to 78 years), among whom 732 incident HF events were observed during follow-up (median follow-up time ranged from 1.8 to 10 years). The studies were not able to differentiate between HF with reduced and preserved ejection fraction because of a lack of data on left ventricular ejection fraction. Characteristics of included studies are provided in Table 1 and the Supplemental Methods and in previous reports.[8,9]

Statistical Analysis

Meta-Analysis of Observational Associations

We performed a fixed-effect meta-analysis using effect estimates from (1) HOMAGE discovery, (2) HOMAGE replication, (3) PIVUS, and (4) ULSAM. Effect estimates for HOMAGE discovery and HOMAGE replication were extracted from odds ratios calculated using multivariable logistic regression adjusting for age, sex, cohort, and follow-up time—which were used as matching variables in a matched, nested case-control design.[9] For PIVUS and ULSAM, effect estimates were taken from hazard ratios calculated using Cox proportional hazard regression adjusting for age and sex.[8] Hazard ratios and odds ratios were assumed to approximate to an equivalent risk ratio (RR), given that the outcome is rare.[30] To make results comparable across studies and proteins, study-level circulating protein measures in the normalized protein expression unit are standardized by setting the mean to 0 and SD to 1 before running regression models, with an assumption that the SDs of circulating protein levels are similar across studies. To account for multiple testing, we implemented a Bonferroni-corrected allowable type I error rate (α) of 0.05/90 (number of proteins under study).

MR Analysis

We assessed the causality of associations for proteins that survived multiple testing correction in the observational analysis by performing 2-sample *cis*-MR using estimates of genetic association with circulating protein levels under study and with HF. Genetic associations with circulating protein levels were extracted from a GWAS meta-analysis of 14 cohorts composed of 30931 subjects of European ancestry included in the SCALLOP consortium (Systematic and Combined Analysis of Olink Proteins).[15] Genetic associations with HF were extracted from a GWAS meta-analysis of 47309 all-cause HF cases from 26 studies of European ancestries included in the HERMES consortium (Heart Failure Molecular Epidemiology for Therapeutic Targets).[19] Details of participating studies in each GWAS meta-analysis are provided in Tables S3 and S4.

Genetic instruments for proteins were selected from all biallelic single-nucleotide polymorphism available in both protein and outcome GWAS summary statistics with minor allele frequency >0.01 and located within 200 kbp upstream or downstream of the cognate protein-encoding transcription start and stop sites. Given that a gene *cis*- region constitutes only

Table 1. Summary of Study Characteristics Included in the Observational Meta-Analysis

Reporting study	Study design	Cohort	Median age at baseline (years)	Median follow-up (years) [range]	Heart failure events/N total sample
Ferreira et al (2019)[9]	Nested-matched case-control (logistic regression)	HOMAGE[21] discovery			
		Health ABC[24]	74	8.3 [0.0–14.4]	215/648
		PREDICTOR[25,26]	77	2.5 [0.2–3.8]	15/44
		PROSPER[27–29]	77	2.0 [0.2–3.9]	56/185
		HOMAGE[21] validation			
		Health ABC[24]	73	9.0 [0.1–14.4]	109/208
		PREDICTOR[25,26]	76	2.2 [0.1–4.5]	29/58
		PROSPER[27–29]	76	1.8 [0.1–3.8]	138/290
Stenemo et al (2018)[8]	Time-to-event analysis (Cox proportional hazards regression)	PIVUS[22]	70.2	10.0 [0.1–10.9]	80/901
		ULSAM[23]	77.8	8.0 [0.2–10.9]	90/685

Health ABC indicates Health Aging and Body Composition; HOMAGE, Heart Omics in Ageing; PIVUS, Prospective Investigation of the Vasculature in Uppsala Seniors; PREDICTOR, Valutazione della Prevalenza di Disfunzione Cardiaca Asintomatica e di Scompenso Cardiaco; PROSPER, Prospective Study of Pravastatin in the Elderly at Risk; and ULSAM, Uppsala Longitudinal Study of Adult Men.

a small proportion of the genome, we relaxed the conventional genomewide significance P value threshold for instrument selection to $P<1\times10^{-4}$. To allow for an increased statistical power to detect an association, we implemented a relaxed linkage disequilibrium (LD) r^2 threshold 0.4 and used MR models accounting for residual correlation.[31] This threshold was based on a simulation study finding that unstable estimates caused by multicollinearity started to occur at a threshold correlation of around r^2=0.36.[32] Using these thresholds, we performed variant clumping implemented in PLINK 1.9[33] to select *cis*- genetic instruments for each protein, with an LD model derived from individual-level genotype data imputed against the Haplotype Reference Consortium[34] reference panel from a random sample of 10,000 UK Biobank[35] participants.

MR estimates were calculated using the Wald ratio estimator for proteins with a single instrument selected, or the inverse-variance weighted (IVW) estimator for proteins with 2 or more instruments. The Wald ratio estimates are calculated as the regression coefficient for genetic association with the outcome divided by the regression coefficient for genetic association with circulating protein levels. The IVW estimates are calculated as the average of instrument ratio coefficients weighted by the inverse variance. Both estimates from observational association and MR analyses approximate a RR of HF per 1 SD increase in normalized protein expression unit (equivalent to per SD per doubling circulating protein concentration).

Multiverse Sensitivity Analysis for MR

To test the robustness of estimates from the primary MR analysis, proteins with MR estimates surviving multiple testing correction (P value <0.05/numbers of observationally associated proteins with at least 1 instrument) were taken forward to undergo an in-depth, multiverse sensitivity analysis[36] in which the stability of the effect estimates was evaluated under a wide combinations of instrument selection parameters and MR models. Thresholds for instrument selection (P value and r^2) and alternative MR models were prioritized more than other possible parameters, such as LD reference population and genomic distance, because these parameters were observed to have the

greatest influence on estimate stability in a previous systematic evaluation of methods for drug target MR.[14] For each MR model, we computed causal estimates for all possible combinations of 5 LD r^2 thresholds (0.05, 0.1, 0.2, 0.4, and 0.6) and 6 P value thresholds (5×10^{-8}, 1×10^{-5}, 1×10^{-4}, 1×10^{-3}, 1×10^{-2}, and 1/no threshold). These combinations included the parameters used in the primary MR analysis above and stringent parameters commonly used in conventional MR analysis of complex trait exposures.[37] For proteins with a single *cis* instrument, the Wald ratio was the only model that could be tested; where 2 or more instruments were available, estimates were calculated with the IVW estimator and MR models using principal components[32] with 90% variance and 99% variance explained; and where there were 3 or more instruments, we in addition calculated estimates using MR with Egger regression estimator[12] (Figure S1). MR with principal components is an alternative model to account for correlation between instruments,[32] and MR with Egger regression estimator provides estimates accounting for residual horizontal pleiotropy.[12] To reduce spurious associations that may arise because of excess multicollinearity or bias toward the null because of weak instruments in 2-sample MR,[14] outlier point estimates with a value outside 1.5 times the interquartile range above the upper quartile and below the lower quartile were removed. An association was declared as robust if all point estimates from the multiverse sensitivity analysis were directionally concordant with estimates from the primary MR analysis, including those on the basis of strict instrument selection parameters and a standard IVW model.

The IVW and MR with Egger regression estimates were calculated using the *MendelianRandomization* package in R,[38] with a fixed-effect model for 3 or fewer genetic instruments, or a multiplicative random-effects model otherwise. To minimize erroneously low P value caused by a multicollinearity issue, correlation between instruments was accounted for by incorporating the instrument pairwise LD correlation matrix in the IVW and MR with Egger regression estimator models.[14,31] The MR method with principal components was implemented using sample codes from the original publication.[32] Genomic coordinates for all relevant analyses were based on Ensembl *GRCh37* reference.[39]

Cross-Trait MR Analysis

To investigate the potential mechanisms through which candidate target proteins may influence HF risk, we performed an exploratory cross-trait MR to estimate the causal association of genetically predicted circulating protein levels with common risk factors and comorbidities of HF: coronary artery disease (CAD), atrial fibrillation, estimated glomerular filtration rate (eGFR), systolic blood pressure, diastolic blood pressure, type 2 diabetes, and body mass index. MR analysis was performed with the primary instrument selection strategy and MR model described in the MR Analysis section using publicly available GWAS statistics for the relevant traits (Table S5).[40–45] To allow comparison across protein-trait pairs, effect estimates were converted to Z scores, calculated as log odd ratios divided by their SEs. The protein-trait MR association was considered potentially causal if the P value from the MR analysis was less than a conservative Bonferroni adjusted threshold of 0.05 divided by the number of protein-trait pairs.

Evaluation of Druggability and Clinical Development Activity

We extracted the druggability profile of candidate target proteins from an updated list of druggable genes.[6] To evaluate clinical development activity of candidate drugs targeting the candidate proteins, we queried the ChEMBL[46] (release 27) database to get information on drug molecule types, approved indications, and target outcomes in clinical trials. We complemented this query by performing a manual search through the https://www.ClinicalTrials.gov website for each candidate target.

Ethical Statement

All included studies were ethically approved by local institutional review boards, and all participants provided written informed consent. The analysis was conducted in accordance with guidelines for study procedures provided by the University College London Research Ethics Committee.

RESULTS

Meta-Analysis of Observational Studies Reveals 44 Circulating Proteins Associated With Incident HF

Through a meta-analysis of observational associations from 4 independent samples, composing up to 732 incident HF events in 3019 subjects, we found 44 out of the 90 proteins were associated with incident HF after multiple testing adjustment at $P<6.0\times10^{-4}$ ($\alpha=0.05/90$ proteins), including 22 associations that were not reported in the individual participating studies.[8,9] Increasing circulating levels of all the 44 observationally associated proteins showed a risk-increasing effect on incident HF, with a median RR of 1.33 (interquartile range, 1.26–1.46). The largest effect sizes were observed in BNP (B-type natriuretic peptide; RR, 1.92 [95% CI, 1.70–2.18]) and NT-proBNP (N-terminal pro-BNP; RR, 1.85 [95%

CI, 1.63–2.10]), 2 biomarkers that have been routinely used in the clinic to diagnose HF. We found no evidence of heterogeneity of the effect estimates after adjustment for multiple testing ($P_{heterogeneity}<0.05/44$). Full study-level and meta-analysis estimates are provided in Table S6.

Causal Effect Estimation With cis-MR

Of the 90 proteins being studied, cis region genetic association summary statistics were available for 83 proteins encoded by autosomal genes (Table S2). Cis region sizes varied according to gene length from 401 to 705 kbp and contained a mean of 1181 variants (SD, 498). Using the primary instrument selection parameter with LD r^2 threshold of 0.4 and P value threshold of 10^{-4}, we identified 75 proteins with 1 to 125 (median, 23) cis- genetic instrument, including 40 of the 44 observationally associated proteins. For comparison, conventional instrument selection parameters (LD $r^2<0.05$, $P<5\times10^{-8}$) identified 70 proteins with 1 to 28 (median, 5) cis- genetic instruments. Instrument-specific estimates are provided in the data and code at https://github.com/alhenry/cvd1-hf/tree/main/resources.

The primary MR analysis suggested causal relationships for 17 of the 40 (43%) observationally associated proteins ($P<0.05/40$). The direction of effects for 16 of 17 proteins were consistent with those calculated using conventional MR parameters; however, only CHI3L1 survived the multiple testing correction (Figure 2). We also investigated the remaining 35 proteins that did not have an observational association with HF and with at least 1 cis- genetic instrument. Of these, we found an additional 9 proteins (26%) with evidence suggestive of a causal association with HF in MR ($P<0.05/35$). Full MR results are provided in Table S7.

Multiverse Sensitivity Analysis Demonstrates Robust Causal Estimates for 8 HF-Associated Proteins

Noting that MR estimates are highly sensitive to choice of parameters for instrument and model selection,[17,47] we tested the stability of the association estimates for each of the 17 HF-associated proteins for which the primary MR analysis suggested underlying causal effects, using a multiverse sensitivity analysis. We tested up to 120 combinations of commonly used parameters for instrument selection and MR models per protein, focusing on parameters that explain the largest variability in MR estimates on the basis of previous simulation and empirical studies,[14,32] resulting in a total of 1850 individual effect estimates. We evaluated the distribution of the point estimates generated and compared these with the primary cis-MR analysis estimates and with estimates from conventional instrument selection parameters (Figure 2b, Table S8). For all 17 proteins under analysis, estimates from the primary cis-MR analysis were directionally
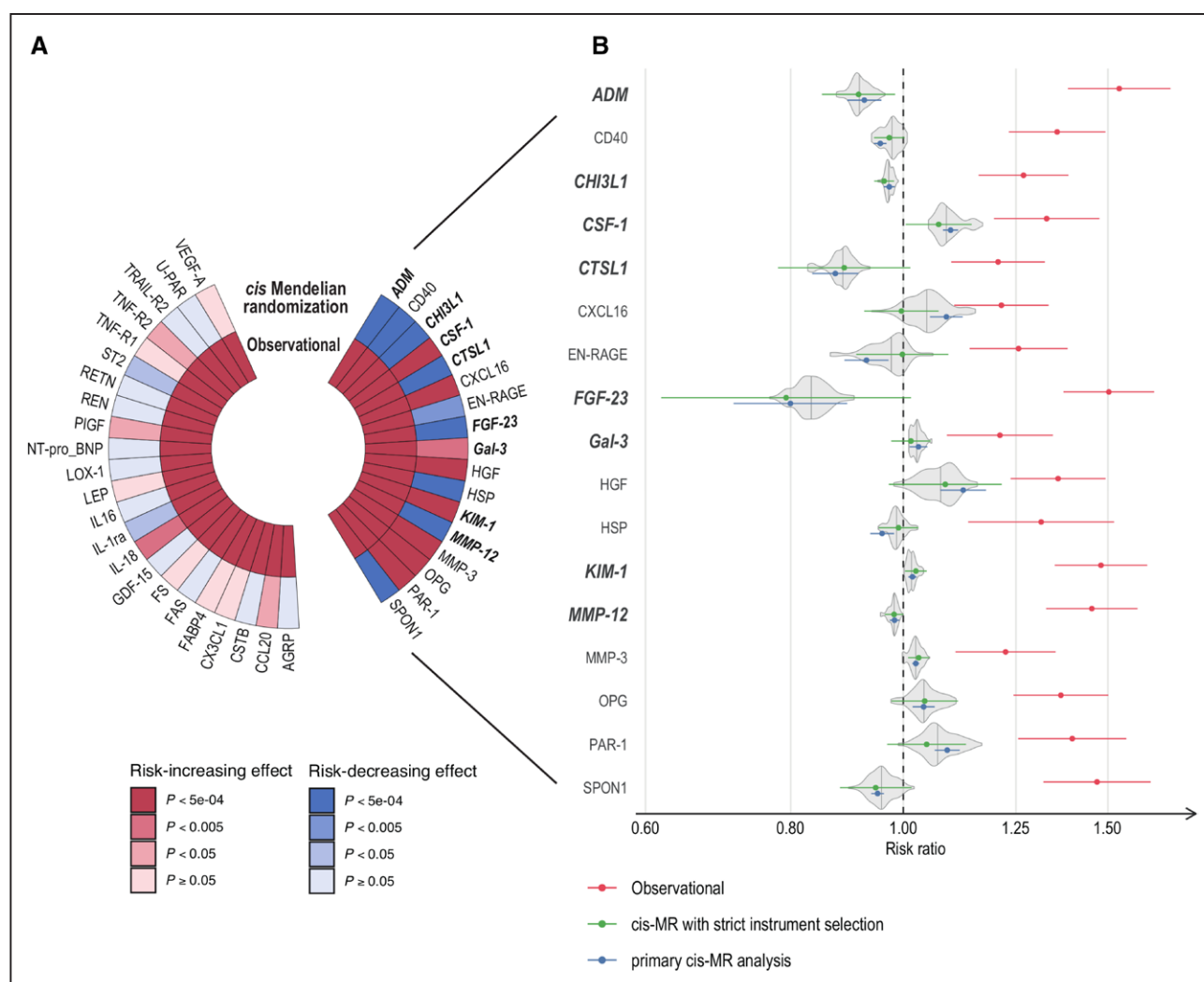
Figure 2. Observational and MR estimates of protein–heart failure association.
A, Circular heatmap of association from 40 proteins associated with incident heart failure in observational studies (*P*<0.05/83=0.0006). The 2 circular lanes refer to results from 2 analyses: (1) observational analysis and (2) *cis*-MR with partially correlated instruments. Color represents direction of effect and strength of association with heart failure measured by *P* value. **B**, Forest plot of risk ratio (hazard ratio from observational analysis and odds ratio from MR analysis) from 17 proteins associated with heart failure in MR analysis (*P*<0.05/40=0.001). Colored dots and error bars indicate the point estimate and 95% CIs. The gray violin plots around the MR estimates illustrate the distributions of odds ratio point estimates estimated from combinations of up to 30 instrument selection parameters and 4 MR models in multiverse sensitivity analysis, with medians of distribution shown as vertical lines within the violin plot. Proteins with consistent direction of effect as indicated by multiverse sensitivity analysis are highlighted in bold and italic font. Full protein names are provided in Table S2. *cis*-MR indicates Mendelian randomization using *cis*-acting protein quantitative trait loci instruments; and MR, Mendelian randomization.

concordant with median values of the multiverse analysis point estimate distributions and showed overlapping 95% CIs with estimates from *cis*-MR using conventional strict instrument selection parameters.

Furthermore, we identified robust evidence of a causal association with HF as indicated by sign concordance of all MR point estimates from the multiverse sensitivity analysis for 8 proteins: ADM (adrenomedullin), CHI3L1 (chitinase-3-like protein 1), CSF-1 (macrophage colony-stimulating factor 1), CTSL1 (cathepsin L1), FGF-23 (fibroblast growth factor 23), Gal-3 (galectin-3), MMP-12 (matrix metalloproteinase-12), and KIM-1 (kidney injury molecule 1). Increasing circulating levels of all 8

proteins were positively associated with risk of incident HF in the observational analysis. In the MR analysis, however, only 3 proteins (CSF-1, Gal-3, and KIM-1) showed positive associations with risk of HF, whereas the remaining 5 (ADM, CHI3L1, CTSL1, FGF-23, and MMP-12) showed negative associations, suggesting causal protective effects (Figure 3).

Cross-Trait MR Analysis for Candidate Therapeutic Targets for HF

We took forward the 8 proteins robustly associated with HF and explored their causal effects on 7 HF-related
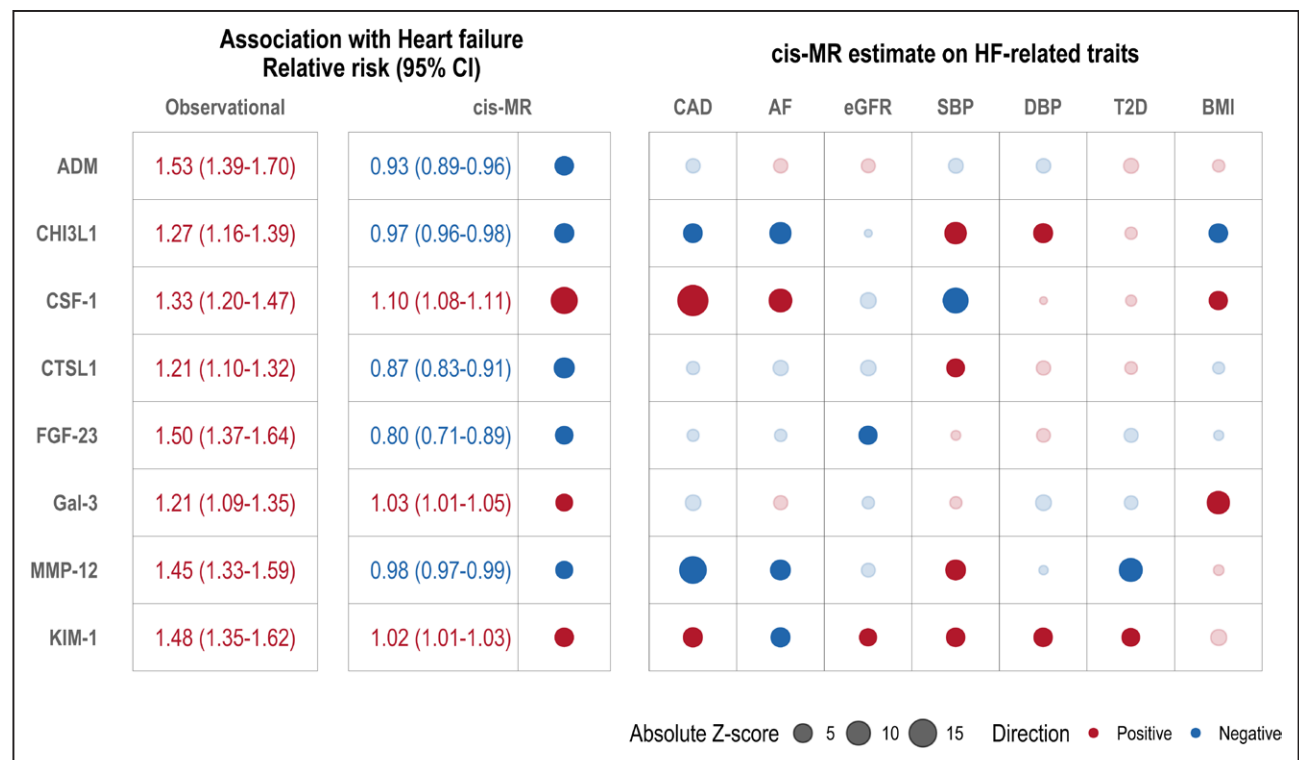
Figure 3. Estimated effect of prioritized circulating protein levels with HF and related traits.
Left, Approximate relative risks of HF per doubling circulating protein levels as estimated with meta-analysis of observational data and *cis*-MR.
Right, A matrix of estimated causal effect size of prioritized circulating protein levels (rows) on HF and related traits (columns) from cis-MR analysis as represented by bullet points. The size of the bullet represents the magnitude of estimated causal effect measured in absolute *Z* score. Bullet points with a darker shade indicate associations that survived multiple testing at *P* value <0.0009 (α=0.05/[8 proteins*7 traits, excluding HF]). Red indicates a risk/trait-increasing effect, and blue indicates a risk/trait-decreasing effect. ADM indicates adrenomedullin; AF, atrial fibrillation; BMI, body mass index; CAD, coronary artery disease; CHI3L1, chitinase 3-like 1; CSF-1, colony stimulating factor 1; CTSL1, cathepsin L; DBP, diastolic blood pressure; eGFR, estimated glomerular filtration rate; FGF-23, fibroblast growth factor 23; Gal-3, galectin-3; HF, heart failure; KIM-1, kidney injury molecule 1; MMP-12, matrix metallopeptidase 12; MR, Mendelian randomization; SBP, systolic blood pressure; and T2D, type 2 diabetes.

traits (CAD, atrial fibrillation, estimated glomerular filtration rate, systolic blood pressure, diastolic blood pressure, type 2 diabetes, and body mass index), using the primary *cis*-MR analysis method (Figure 3). Of the 8 candidate proteins, 1 (ADM) was not associated with any trait other than HF, whereas the remaining 7 were associated with at least 1 other trait after multiple testing correction (*P*<0.05/8 proteins/7 traits excluding HF). Consistent with evidence from overexpression perturbation studies in animal models, Gal-3[48] and CSF-1[49] were positively associated with body mass index, a biomarker of adiposity and a known risk factor for HF.[50] CHI3L1 and CTSL1 were protective for CAD, consistent with reports of cardioprotective effects in animal models of cardiac ischemia.[51,52] A higher circulating CSF-1 level was associated with an increased risk of CAD,[53] whereas MMP-12 showed a protective effect, consistent with previous reports.[16] A higher level of FGF-23 was associated with a lower estimated glomerular filtration rate, consistent with findings from preclinical models in which FGF-23 deficiency was associated with worsening renal failure and cardiac hypertrophy.[54]

Appraisal of Druggability and Existing Approved or Clinical-Phase Drug Candidates for Candidate Protein Targets

To evaluate the druggability and drug development activities of candidate targets, we searched through a list of druggable genes,[6] the ChEMBL (release 27) drug discovery database, and a clinical trial registry (https://www.clinicaltrials.gov, accessed on December 1, 2020). We grouped candidate targets into 3 categories corresponding to the highest status in the drug development pipeline: *approved* (targeted by drugs already approved for 1 or more conditions), *in development* (currently being investigated in clinical trials), and *druggable* (listed as druggable targets; Table 2). A candidate drug targeting adrenomedullin, adrecizumab (a humanized, monoclonal, nonneutralizing antibody against the N terminus of ADM[55]), is entering phase II trials for septic shock (URL: https://www.ClinicalTrials.gov; Unique identifier: NCT03085758), cardiogenic shock (Unique identifier: NCT03989531), and acute HF (Unique identifier: NCT04252937). A modified citrus pectin Gal-3 inhibitor

Table 2. Summary of Druggability and Clinical Development Activity for Heart Failure Associated With Causal Associations on Mendelian Randomization Analysis

Target	Status	Compound name	Molecule type	Action type	Clinical development activities
ADM (adrenomedullin)	In development	Adrecizumab†	Antibody	Redistributing interstitial to plasma ADM	Phase I/II trials: septic shock (NCT03085758)† Cardiogenic shock (NCT03989531)† Acute heart failure (NCT04252937)†
CHI3L1 (chitinase 3-like 1)	Druggable*	–	Antibody*	–	–
CSF-1 (colony stimulating factor 1)	In development	MCS-110‡ (CHEMBL2109512)	Antibody	Antagonist	Phase I/II trials: melanoma, stomach neoplasms, breast neoplasms
		PD-360324‡ (CHEMBL2109513)	Antibody	Antagonist	Phase I/II trials: arthritis, rheumatoid arthritis, lupus erythematosus, cutaneous sarcoidosis
CTSL1 (cathepsin L)	Druggable*	–	Small molecule*	–	–
FGF-23 (fibroblast growth factor 23)	Approved	Burosumab‡ (CHEMBL3707326)	Antibody	Antagonist	Approved for X-linked hypophosphatemia, hypophosphatemic rickets
Gal-3 (galectin-3)	In development	GB1211†	Small molecule*	Antagonist	Phase I/II trials: nonalcoholic steatohepatitis (NCT04607655)†
	In development	Modified citrus pectin (MCP)	Carbohydrate oral supplement	Antagonist	Phase I/II trials: hypertension (NCT01960946)†
MMP-12 (matrix metallopeptidase 12)	In development	FP-025†	Small molecule	Antagonist	Phase I/II trials: asthma (NCT03858686)†
		MARIMASTAT‡ (CHEMBL279785)	Small molecule	Antagonist	Phase III trials‡: lung neoplasms, breast neoplasms
KIM-1 (kidney injury molecule 1)	Not currently listed as druggable	–	–	–	–

ADM indicates adrenomedullin; CHI3L1, chitinase 3-like 1; CSF-1, colony stimulating factor 1; CTSL1, cathepsin L; FGF-23, fibroblast growth factor 23; Gal-3, galectin-3; KIM-1, kidney injury molecule 1; and MMP-12, matrix metallopeptidase 12.

*Data from druggable gene list.[6]

†Data from https://www.ClinicalTrials.gov (clinical trial ID in brackets).

‡Data from ChEMBL release 27[46] (compound ID in brackets).

has been evaluated for effects markers of collagen metabolism in patients with hypertension in a proof-of-concept clinical trial for cardiac fibrosis.[56] CSF-1 and MMP-12 inhibitors are currently being evaluated in clinical trials for non-HF conditions. Burosumab, a monoclonal antibody FGF-23 inhibitor, has already been approved for treating X-linked hypophosphatemia and hypophosphatemic rickets. Although we found no ongoing trials specific for CHI3L1 or CTSL1, inhibition of CTSL1 is proposed as potential treatment for SARS-CoV-2 infection, and several approved agents show inhibitory activity against CTSL1.[57] With the exception of KIM-1, all 7 other proteins are predicted to be secreted in at least 1 tissue according to the Human Protein Atlas database.[58] KIM-1 is also not currently listed as a potential drug target according to the druggable gene list, ChEMBL release 27, and https://www.ClinicalTrials.gov databases.

DISCUSSION

Principal Findings

We investigated 90 circulating proteins for their association with incident HF in a population of 3019 individuals with 732 events. A total of 44 proteins had positive associations with risk of incident HF, 22 of which were not reported in the participating studies. These included associations with incident HF reported elsewhere such as Gal-3, HGF, and Resistin,[59–61] proteins such as CXCL16 with reported associations with prognosis in HF,[62] and with cardiac fibrosis on cardiac magnetic resonance imaging in HF including MMP3.[63] Among the novel associations to highlight, CTSL1 is a potent endoprotease linked to the development of dilated cardiomyopathy and HF in mouse models.[64,65] We used *cis*-MR to estimate whether the observational protein-HF associations reflected an underlying causal relationship. Of the 40 proteins for which *cis* genetic instruments were available, 17 showed evidence suggestive of causal effects, of which 8 were robust to multiverse sensitivity analysis. Among these 8 HF-associated proteins, 3 were positively associated with risk of HF (CSF-1, Gal-3, and KIM-1), and 5 were negatively associated, consistent with causally protective effects (ADM, CHI3L1, CTSL1, FGF-23, and MMP-12). Seven are known or predicted to be druggable by conventional therapeutic modalities, and therapeutic agents targeting 2 of the identified proteins are currently under evaluation in phase II clinical trials: adrecizumab, an ADM agonist, for acute HF and cardiogenic shock,[55] and modified citrus pectin, a Gal-3 antagonist, for cardiac fibrosis.[56] We note that CTSL1 inhibition has been proposed as a potential treatment for COVID-19[66]; our results signal HF as a potential safety liability of this therapeutic approach. Our findings provide evidence supporting the therapeutic

Concordance of Observational and Causal Associations for Identified Proteins

One of the key strengths of study is the triangulation of evidence between observational and MR analyses for a consistently measured set of cardiovascular proteins. For all the protein-HF associations that were identified in our meta-analysis, there was a positive association, ie, a higher protein concentration was associated with an increased risk of incident HF. This is consistent with previously reported biomarker association studies with incident HF; for example, a study of incident HF in the Framingham population identified 18 associated circulating biomarkers, of which 17 were positive associations.[67] When we estimated the causal association of the observationally associated HF proteins, however, we found that the observational and causal association estimates were frequently discordant with opposing direction of effects. For example, 5 proteins with an estimated causally protective effect were found to have a positive association with incident HF, including MMP-12 and ADM. In the case of MMP-12, our findings are consistent with previous reports on the associations between MMP-12 and CAD.[16,68] These discordant findings may be explained by subclinical or predisease leading to higher levels of these proteins that precedes the clinical diagnosis of HF, potentially as an adaptive feedback response to mitigate the disease process. The median baseline age in the included studies ranged from 70 to 78 years, and it is likely that subclinical alterations in cardiac structure and function occurred before incident HF, which was defined as the first HF hospitalization. Concordant observational and causal associations (CSF-1, Gal-3, and KIM-1) may be explained either by upstream processes driving risk or by reverse causation where a positive feedback loop exists between the HF and expression of the protein. For several proteins, including established clinical biomarkers NT-proBNP and ST2, we found positive observational associations but were unable to detect causal effects by MR analysis. In these, the observational associations may be interpreted as noncausal, arising from reverse causation. We cannot, however, exclude a type 2 error caused by imprecision of the MR estimates.

Comparison With Other Studies

To our knowledge, our study represents the first large-scale analysis of incident HF that combines observational associations of circulating proteins with a systematic appraisal of causal effects using MR. Our results were consistent with previously reported findings from MR studies of NT-proBNP and GDF-15, which did not report evidence of a strong causal relationship between these proteins and risk of HF.[69,70] Our approach of triangulating evidence from observational association and MR represents a pragmatic approach to screen and prioritize targets for therapeutic development, according to the relative strength of evidence from analysis of the data available.[71] In our study, we used a method for *cis*-MR that incorporates the LD correlation structure within the causal model and provides estimates with higher precision.[31] We combined this primary approach with a new technique to evaluate the robustness of the identified protein-HF associations that involved systematically testing multiple combinations of model parameter selection in a multiverse sensitivity analysis, enabling us to deprioritize proteins with unstable estimates. Using this framework, we found evidence supporting a causal relationship for 8 of the 40 HF-associated proteins tested, compared with a single association for CHI3L1 that was identified using conventional approaches. For example, the estimates for CTSL1 and FGF-23 generated with this approach more clearly suggest a causal effect compared with those on the basis of more stringent instrument selection (Figure 2b, Table S7).

Implications for Therapeutics Targeting HF

All 8 proteins with estimated causal effects, except ADM, were associated with HF-related traits in an exploratory *cis*-MR cross-trait analysis, including upstream HF risk factors. Distinct pathobiological pathways and proteomic signatures are described for subgroups of patients with HF, such as those defined by left ventricular ejection fraction[72]; however, we were unable to perform a stratified analysis because of the limited phenotype data available at the time of HF diagnosis. To leverage the full potential of proteomics and genomics in understanding HF and identifying drug targets, there is a need to decompose HF into phenotypic components, including those of cardiac dysfunction and fluid congestion, which characterize this condition. ADM and CTSL1 are notable among our findings because their protective effect against the risk of HF was not explained by association with upstream risk factor traits. ADM is a circulating peptide hormone synthesized by endothelial and vascular smooth muscle cells, the biologically active form of which has been proposed as a marker and inhibitor of tissue fluid congestion, a hallmark feature of HF.[55] Consistent with our results, it has been hypothesized that ADM may play a protective role in HF development and progression by maintaining vascular integrity, inducing vasodilatation, and inhibiting the renin-angiotensin-aldosterone system.[55]

Limitations

Although the clinical ascertainment of HF was consistent across the studies included in the observational analysis

and in HF GWAS, the interpretation of our findings is limited by the lack of detailed phenotyping by pathogenesis and phenotypes of cardiac structure and function. Our MR framework, including the prioritization of parameters for the multiverse analysis, was based on previous studies of gene transcript exposures which demonstrated robust and reproducible MR estimates[73]; however the scope of our multiverse analysis was limited by the computational burden inherent in the approach. There is a lack of consensus about the optimal approach to *cis*-MR, and we were unable to empirically replicate our findings in an independent sample because none were available at the time of the study. It is possible for proteins with an important causal contribution to HF risk to have a null observational association in this study because of negative confounding or imprecision of the estimates. Given that circulating protein concentrations are measured in a relative normalized protein expression unit,[20] the derived effect estimates are rarely representative of the absolute magnitude of effect on HF and are not directly comparable across proteins. The expected causal direction of effects, however, can inform potential efficacy and on-target side effects, which can be formally investigated further in clinical trials. Further studies are needed to corroborate and extend our findings, to include a larger number of protein biomarkers, and to explore the relationship of the identified proteins with disease subtypes. These studies will be enabled by the rapidly increasing availability of proteomic and genomic information in large populations from large health care–linked biobanks.

CONCLUSIONS

In conclusion, we evaluated 90 cardiovascular-related proteins through observational and MR analysis using population-based proteomic data and identified 7 candidate drug targets for HF. Of these, 2 proteins (ADM and Gal-3) are currently under evaluation in clinical trials for HF, and 5 (CHI3L1, CSF-1, CTSL1, FGF-23, and MMP-12) represent novel putative therapeutic targets for HF. This study provides an example of the opportunities for human target prioritization that are enabled by emerging population-based genomic and proteomic data resources. Proteomewide studies incorporating both direct association with target outcomes and genetic-based inference through MR are likely to provide important new tools for therapeutic target discovery and prioritization.

Affiliations

Institute of Cardiovascular Science (A.H., M.G.-M., C.F., A.F.S., A.D.H.), British Heart Foundation Research Accelerator (A.H., M.G.-M., C.F., A.F.S., A.D.H., R.T.L.), Institute of Health Informatics (A.H., R.T.L.), Health Data Research UK London (R.T.L.), University College London, United Kingdom. Department of Cardiology, Division of Heart and Lungs, University Medical Center Utrecht, The Netherlands (C.F., A.F.S.). Unidade de Investigação e Desenvolvimento Cardiovascular, Rede de Investigação em Saúde, Department of Surgery and Physiology, Faculty of Medicine of the University of Porto, Portugal (J.P.F.). Université de Lorraine, Inserm, Centre d'Investigations Cliniques - Plurithématique 14-33, and Inserm U1116, Centre Hospitalier Régional Universitaire, French Clinical Research Infrastructure Network, Investigation Network Initiative - Cardiovascular and Renal Clinical Trialists, Nancy, France (J.P.F., F.Z.). Division of Cardiology, Department of Medicine (R.K.), Department of Pathology (R.K.), Duke University Medical Center, Durham, NC. Department of Medical Sciences, Uppsala University, Sweden (J.S., L.L.). The George Institute for Global Health, University of New South Wales, Sydney, Australia (J.S.). School of Health and Social Studies, Dalarna University, Falun, Sweden (J.Ä.). Division of Family Medicine and Primary Care, Department of Neurobiology, Care Science and Society, Karolinska Institutet, Huddinge, Sweden (J.Ä.). Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Solna, Sweden (A.M.). Emerging Science and Innovation, Pfizer Worldwide Research, Development and Medical, Cambridge, MA (A.M.).

REFERENCES

1. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, Falk V, González-Juanatey JR, Harjola VP, Jankowska EA, et al; ESC Scientific Document Group. 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 Heart J*. 2016;37:2129–2200. doi: 10.1093/eurheartj/ehw128

2. Roger VL, Weston SA, Redfield MM, Hellermann-Homan JP, Killian J, Yawn BP, Jacobsen SJ. Trends in heart failure incidence and survival in a community-based population. *JAMA*. 2004;292:344–350. doi: 10.1001/jama.292.3.344

3. Redfield MM, Anstrom KJ, Levine JA, Koepp GA, Borlaug BA, Chen HH, LeWinter MM, Joseph SM, Shah SJ, Semigran MJ, et al; NHLBI Heart Failure Clinical Research Network. Isosorbide mononitrate in heart failure with preserved ejection fraction. *N Engl J Med*. 2015;373:2314–2324. doi: 10.1056/NEJMoa1510774

4. Gheorghiade M, Greene SJ, Butler J, Filippatos G, Lam CS, Maggioni AP, Ponikowski P, Shah SJ, Solomon SD, Kraigher-Krainer E, et al; SOCRATES-REDUCED Investigators and Coordinators. Effect of vericiguat, a soluble guanylate cyclase stimulator, on natriuretic peptide levels in patients with worsening chronic heart failure and reduced ejection fraction: the SOCRATES-REDUCED Randomized Trial. *JAMA*. 2015;314:2251–2262. doi: 10.1001/jama.2015.15734

5. Fordyce CB, Roe MT, Ahmad T, Libby P, Borer JS, Hiatt WR, Bristow MR, Packer M, Wasserman SM, Braunstein N, et al. Cardiovascular drug development: is it dead or just hibernating? *J Am Coll Cardiol*. 2015;65:1567–1582. doi: 10.1016/j.jacc.2015.03.016

6. Finan C, Gaulton A, Kruger FA, Lumbers RT, Shah T, Engmann J, Galver L, Kelley R, Karlsson A, Santos R, et al. The druggable genome and support for target identification and validation in drug development. *Sci Transl Med*. 2017;9:eaag1166. doi: 10.1126/scitranslmed.aag1166

7. 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

8. 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*. 2018;20:55–62. doi: 10.1002/ejhf.980

9. 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

10. Egerstedt A, Berntsson J, Smith ML, Gidlöf O, Nilsson R, Benson M, Wells QS, Celik S, Lejonberg C, Farrell L, et al. Profiling of the plasma proteome across different stages of human heart failure. *Nat Commun*. 2019;10:5830. doi: 10.1038/s41467-019-13306-y

11. Suhre K, McCarthy MI, Schwenk JM. Genetics meets proteomics: perspectives for large population-based studies. *Nat Rev Genet*. 2021;22:19–37. doi: 10.1038/s41576-020-0268-2

12. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44:512–525. doi: 10.1093/ije/dyv080

13. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ*. 2018;362:k601. doi: 10.1136/bmj.k601

14. Schmidt AF, Finan C, Gordillo-Marañón M, Asselbergs FW, Freitag DF, Patel RS, Tyl B, Chopade S, Faraway R, Zwierzyna M, et al. Genetic drug target validation using Mendelian randomisation. *Nat Commun*. 2020;11:3255. doi: 10.1038/s41467-020-16969-0

15. Folkersen L, Gustafsson S, Wang Q, Hansen DH, Hedman ÅK, Schork A, Page K, Zhernakova DV, Wu Y, Peters J, et al. Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. *Nat Metab*. 2020;2:1135–1148. doi: 10.1038/s42255-020-00287-2

16. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E, Surendran P, et al. Genomic atlas of the human plasma proteome. *Nature*. 2018;558:73–79. doi: 10.1038/s41586-018-0175-2

17. Swerdlow DI, Kuchenbaecker KB, Shah S, Sofat R, Holmes MV, White J, Mindell JS, Kivimaki M, Brunner EJ, Whittaker JC, et al. Selecting instruments for Mendelian randomization in the wake of genome-wide association studies. *Int J Epidemiol*. 2016;45:1600–1616. doi: 10.1093/ije/dyw088

18. Assarsson E, Lundberg M, Holmquist G, Björkesten J, Thorsen SB, Ekman D, Eriksson A, Rennel Dickens E, Ohlsson S, Edfeldt G, et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS One*. 2014;9:e95192. doi: 10.1371/journal.pone.0095192

19. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, Hedman ÅK, Wilk JB, Morley MP, Chaffin MD, et al; Regeneron Genetics Center. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. *Nat Commun*. 2020;11:163. doi: 10.1038/s41467-019-13690-5

20. What is NPX? Olink. https://www.olink.com/question/what-is-npx/. Accessed November 24, 2020.

21. 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

22. Lind L, Fors N, Hall J, Marttala K, Stenborg A. A comparison of three different methods to evaluate endothelium-dependent vasodilation in the elderly: the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study. *Arterioscler Thromb Vasc Biol*. 2005;25:2368–2375. doi: 10.1161/01.ATV.0000184769.22061.da

23. Hedstrand H. A study of middle-aged men with particular reference to risk factors for cardiovascular disease. *Ups J Med Sci Suppl*. 1975;19:1–61.

24. Kalogeropoulos A, Georgiopoulou V, Kritchevsky SB, Psaty BM, Smith NL, Newman AB, Rodondi N, Satterfield S, Bauer DC, Bibbins-Domingo K, et al. Epidemiology of incident heart failure in a contemporary elderly cohort: the health, aging, and body composition study. *Arch Intern Med*. 2009;169:708–715. doi: 10.1001/archinternmed.2009.40

25. 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

26. Mureddu GF, Tarantini L, Agabiti N, Faggiano P, Masson S, Latini R, Cesaroni G, Miceli M, Forastiere F, Scardovi AB, et al. Evaluation of different strategies for identifying asymptomatic left ventricular dysfunction and pre-clinical (stage B) heart failure in the elderly. Results from 'PREDICTOR', a population based-study in central Italy. *Eur J Heart Fail*. 2013;15:1102–1112. doi: 10.1093/eurjhf/hft098

27. 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

28. Shepherd J, Blauw GJ, Murphy MB, Bollen EL, Buckley BM, Cobbe SM, Ford I, Gaw A, Hyland M, Jukema JW, et al; PROSPER study group. PROspective Study of Pravastatin in the Elderly at Risk. Pravastatin in elderly individuals at risk of vascular disease (PROSPER): a randomised controlled trial. *Lancet*. 2002;360:1623–1630. doi: 10.1016/s0140-6736(02)11600-x

29. Nanchen D, Stott DJ, Gussekloo J, Mooijaart SP, Westendorp RG, Jukema JW, Macfarlane PW, Cornuz J, Rodondi N, Buckley BM, et al; PROSPER Group. Resting heart rate and incident heart failure and cardiovascular mortality in older adults: role of inflammation and endothelial dysfunction: the PROSPER study. *Eur J Heart Fail*. 2013;15:581–588. doi: 10.1093/eurjhf/hfs195

30. VanderWeele TJ. Optimal approximate conversions of odds ratios and hazard ratios to risk ratios. *Biometrics*. 2020;76:746–752. doi: 10.1111/biom.13197

31. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in Mendelian randomization: comparison of allele score and summarized data methods. *Stat Med*. 2016;35:1880–1906. doi: 10.1002/sim.6835

32. Burgess S, Zuber V, Valdes-Marquez E, Sun BB, Hopewell JC. Mendelian randomization with fine-mapped genetic data: choosing from large numbers of correlated instrumental variables. *Genet Epidemiol*. 2017;41:714–725. doi: 10.1002/gepi.22077

33. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7. doi: 10.1186/s13742-015-0047-8

34. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, Kang HM, Fuchsberger C, Danecek P, Sharp K, et al; Haplotype Reference Consortium. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*. 2016;48:1279–1283. doi: 10.1038/ng.3643

35. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562:203–209. doi: 10.1038/s41586-018-0579-z

36. Steegen S, Tuerlinckx F, Gelman A, Vanpaemel W. Increasing transparency through a multiverse analysis. *Perspect Psychol Sci*. 2016;11:702–712. doi: 10.1177/1745691616658637

37. Burgess S, Davey Smith G, Davies NM, Dudbridge F, Gill D, Glymour MM, Hartwig FP, Holmes MV, Minelli C, Relton CL, et al. Guidelines for performing Mendelian randomization investigations. *Wellcome Open Res*. 2019;4:186. doi: 10.12688/wellcomeopenres.15555.2

38. Yavorska OO, Burgess S. MendelianRandomization: an R package for performing Mendelian randomization analyses using summarized data. *Int J Epidemiol*. 2017;46:1734–1739. doi: 10.1093/ije/dyx034

39. Yates AD, Achuthan P, Akanni W, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, et al. Ensembl 2020. *Nucleic Acids Res*. 2020;48:D682–D688. doi: 10.1093/nar/gkz966

40. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, Hopewell JC, et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47:1121–1130. doi: 10.1038/ng.3396

41. Roselli C, Chaffin MD, Weng LC, Aeschbacher S, Ahlberg G, Albert CM, Almgren P, Alonso A, Anderson CD, Aragam KG, et al. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet*. 2018;50:1225–1233. doi: 10.1038/s41588-018-0133-9

42. Wuttke M, Li Y, Li M, Sieber KB, Feitosa MF, Gorski M, Tin A, Wang L, Chu AY, Hoppmann A, et al; Lifelines Cohort Study; V. A. Million Veteran Program. A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat Genet*. 2019;51:957–972. doi: 10.1038/s41588-019-0407-x

43. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, Ntritsos G, Dimou N, Cabrera CP, Karaman I, et al; Million Veteran Program. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat Genet*. 2018;50:1412–1425. doi: 10.1038/s41588-018-0205-x

44. Mahajan A, Taliun D, Thurner M, Robertson NR, Torres JM, Rayner NW, Payne AJ, Steinthorsdottir V, Scott RA, Grarup N, et al. Fine-mapping type 2 diabetes loci to single-variant resolution using high-density imputation and islet-specific epigenome maps. *Nat Genet*. 2018;50:1505–1513. doi: 10.1038/s41588-018-0241-6

45. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, Yengo L, Ferreira T, Marouli E, Ji Y, et al; GIANT Consortium. Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum Mol Genet*. 2019;28:166–174. doi: 10.1093/hmg/ddy327

46. Mendez D, Gaulton A, Bento AP, Chambers J, De Veij M, Félix E, Magariños MP, Mosquera JF, Mutowo P, Nowotka M, et al. ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res*. 2019;47:D930–D940. doi: 10.1093/nar/gky1075

47. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG. Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple Genetic variants. *Epidemiology*. 2017;28:30–42. doi: 10.1097/EDE.0000000000000559

48. Pejnovic NN, Pantic JM, Jovanovic IP, Radosavljevic GD, Milovanovic MZ, Nikolic IG, Zdravkovic NS, Djukic AL, Arsenijevic NN, Lukic ML. Galectin-3 deficiency accelerates high-fat diet-induced obesity and amplifies inflammation in adipose tissue and pancreatic islets. *Diabetes*. 2013;62:1932–1944. doi: 10.2337/db12-0222

49. Levine JA, Jensen MD, Eberhardt NL, O'Brien T. Adipocyte macrophage colony-stimulating factor is a mediator of adipose tissue growth. *J Clin Invest*. 1998;101:1557–1564. doi: 10.1172/JCI2293

50. Lumbers T, Katsoulis M, Henry A, Mordi I, Lang C, Hemingway H, Langenberg C, Holmes M, Sattar N. Body mass index and heart failure risk: a cohort study in 1.5 million individuals and Mendelian randomisation analysis [published online September 25, 2020]. *medRxiv*. https://doi.org/10.1101/2020.09.23.20200360. doi: 10.1101/2020.09.23.20200360

51. Harutyunyan M, Christiansen M, Johansen JS, Køber L, Torp-Petersen C, Kastrup J. The inflammatory biomarker YKL-40 as a new prognostic marker for all-cause mortality in patients with heart failure. *Immunobiology*. 2012;217:652–656. doi: 10.1016/j.imbio.2011.11.003

52. Liu CL, Guo J, Zhang X, Sukhova GK, Libby P, Shi GP. Cysteine protease cathepsins in cardiovascular disease: from basic research to clinical trials. *Nat Rev Cardiol*. 2018;15:351–370. doi: 10.1038/s41569-018-0002-3

53. Sjaarda J, Gerstein H, Chong M, Yusuf S, Meyre D, Anand SS, Hess S, Paré G. Blood CSF1 and CXCL12 as causal mediators of coronary artery disease. *J Am Coll Cardiol*. 2018;72:300–310. doi: 10.1016/j.jacc.2018.04.067

54. Clinkenbeard EL, Noonan ML, Thomas JC, Ni P, Hum JM, Aref M, Swallow EA, Moe SM, Allen MR, White KE. Increased FGF23 protects against det-

55. Voors AA, Kremer D, Geven C, Ter Maaten JM, Struck J, Bergmann A, Pickkers P, Metra M, Mebazaa A, Düngen HD, et al. Adrenomedullin in heart failure: pathophysiology and therapeutic application. *Eur J Heart Fail*. 2019;21:163–171. doi: 10.1002/ejhf.1366

56. Lau ES, Liu E, Paniagua SM, Sarma AA, Zampierollo G, López B, Díez J, Wang TJ, Ho JE. Galectin-3 inhibition with modified citrus pectin in hypertension. *JACC Basic Transl Sci*. 2021;6:12–21. doi: 10.1016/j.jacbts.2020.10.006

57. Liu T, Luo S, Libby P, Shi GP. Cathepsin L-selective inhibitors: a potentially promising treatment for COVID-19 patients. *Pharmacol Ther*. 2020;213:107587. doi: 10.1016/j.pharmthera.2020.107587

58. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, et al. Proteomics. Tissue-based map of the human proteome. *Science*. 2015;347:1260419. doi: 10.1126/science.1260419

59. Ho JE, Liu C, Lyass A, Courchesne P, Pencina MJ, Vasan RS, Larson MG, Levy D. Galectin-3, a marker of cardiac fibrosis, predicts incident heart failure in the community. *J Am Coll Cardiol*. 2012;60:1249–1256. doi: 10.1016/j.jacc.2012.04.053

60. Ferraro RA, Ogunmoroti O, Zhao D, Ndumele CE, Rao V, Pandey A, Larson NB, Bielinski SJ, Michos ED. Hepatocyte growth factor and incident heart failure subtypes: the Multi-Ethnic Study of Atherosclerosis (MESA). *J Card Fail*. 2021;27:981–990. doi: 10.1016/j.cardfail.2021.04.022

61. Frankel DS, Vasan RS, D'Agostino RB Sr, Benjamin EJ, Levy D, Wang TJ, Meigs JB. Resistin, adiponectin, and risk of heart failure the Framingham offspring study. *J Am Coll Cardiol*. 2009;53:754–762. doi: 10.1016/j.jacc.2008.07.073

62. Borst O, Schaub M, Walker B, Sauter M, Muenzer P, Gramlich M, Mueller K, Geisler T, Lang F, Klingel K, et al. CXCL16 is a novel diagnostic marker and predictor of mortality in inflammatory cardiomyopathy and heart failure. *Int J Cardiol*. 2014;176:896–903. doi: 10.1016/j.ijcard.2014.08.033

63. Kanagala P, Arnold JR, Singh A, Chan DCS, Cheng ASH, Khan JN, Gulsin GS, Yang J, Zhao L, Gupta P, et al. Characterizing heart failure with preserved and reduced ejection fraction: an imaging and plasma biomarker approach. *PLoS One*. 2020;15:e0232280. doi: 10.1371/journal.pone.0232280

64. Petermann I, Mayer C, Stypmann J, Biniossek ML, Tobin DJ, Engelen MA, Dandekar T, Grune T, Schild L, Peters C, et al. Lysosomal, cytoskeletal, and metabolic alterations in cardiomyopathy of cathepsin L knockout mice. *FASEB J*. 2006;20:1266–1268. doi: 10.1096/fj.05-5517fje

65. Sun M, Chen M, Liu Y, Fukuoka M, Zhou K, Li G, Dawood F, Gramolini A, Liu PP. Cathepsin-L contributes to cardiac repair and remodelling post-infarction. *Cardiovasc Res*. 2011;89:374–383. doi: 10.1093/cvr/cvq328

66. Zhao MM, Yang WL, Yang FY, Zhang L, Huang WJ, Hou W, Fan CF, Jin RH, Feng YM, Wang YC, et al. Cathepsin L plays a key role in SARS-CoV-2 infection in humans and humanized mice and is a promising target for new drug development. *Signal Transduct Target Ther*. 2021;6:134. doi: 10.1038/s41392-021-00558-8

67. Ho JE, Lyass A, Courchesne P, Chen G, Liu C, Yin X, Hwang SJ, 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

68. Ganz P, Heidecker B, Hveem K, Jonasson C, Kato S, Segal MR, Sterling DG, Williams SA. Development and validation of a protein-based risk score for cardiovascular outcomes among patients with stable coronary heart disease. *JAMA*. 2016;315:2532–2541. doi: 10.1001/jama.2016.5951

69. Geelhoed B, Börschel CS, Niiranen T, Palosaari T, Havulinna AS, Fouodo CJK, Scheinhardt MO, Blankenberg S, Jousilahti P, Kuulasmaa K, et al. Assessment of causality of natriuretic peptides and atrial fibrillation and heart failure: a Mendelian randomization study in the FINRISK cohort. *Europace*. 2020;22:1463–1469. doi: 10.1093/europace/euaa158

70. Wang Z, Yang F, Ma M, Bao Q, Shen J, Ye F, Xie X. The impact of growth differentiation factor 15 on the risk of cardiovascular diseases: two-sample Mendelian randomization study. *BMC Cardiovasc Disord*. 2020;20:462. doi: 10.1186/s12872-020-01744-2

71. Lawlor DA, Tilling K, Davey Smith G. Triangulation in aetiological epidemiology. *Int J Epidemiol*. 2016;45:1866–1886. doi: 10.1093/ije/dyw314

72. Adamo L, Yu J, Rocha-Resende C, Javaheri A, Head RD, Mann DL. Proteomic signatures of heart failure in relation to left ventricular ejection fraction. *J Am Coll Cardiol*. 2020;76:1982–1994. doi: 10.1016/j.jacc.2020.08.061

73. Ndungu A, Payne A, Torres JM, van de Bunt M, McCarthy MI. A multi-tissue transcriptome analysis of human metabolites guides interpretability of associations based on multi-SNP models for gene expression. *Am J Hum Genet*. 2020;106:188–201. doi: 10.1016/j.ajhg.2020.01.003