ORIGINAL RESEARCH

Large-Scale Plasma Protein Profiling of Incident Myocardial Infarction, Ischemic Stroke, and Heart Failure

Lars Lind [iD], MD, PhD; Daniela Zanetti [iD], PhD; Martin Ingelsson [iD], MD, PhD; Stefan Gustafsson, PhD; Johan Ärnlöv [iD], MD, PhD; Themistocles L. Assimes [iD], MD, PhD

BACKGROUND: We recently reported a link between plasma levels of 2 of 84 cardiovascular disease (CVD)–related proteins and the 3 major CVDs, myocardial infarction, ischemic stroke, and heart failure. The present study investigated whether measurement of almost 10 times the number of proteins could lead to discovery of additional risk markers for CVD.

METHODS AND RESULTS: We measured 742 proteins using the proximity extension assay in 826 male participants of ULSAM (Uppsala Longitudinal Study of Adult Men) who were free from CVD at the age of 70 years. Cox proportional hazards models were adjusted for age only, as well as all traditional risk factors. During a 12.5-year median follow-up (maximal, 22.0 years), 283 incident CVDs occurred. Forty-one proteins were significantly (false discovery rate <0.05) related to the combined end point of incident CVD, with N-terminal pro–brain natriuretic peptide as the top finding, while 53 proteins were related to incident myocardial infarction. A total of 13 and 16 proteins were significantly related to incident ischemic stroke and heart failure, respectively. Growth differentiation factor 15, 4-disulfide core domain protein 2, and kidney injury molecule were related to all of the 3 major CVD outcomes. A lasso selection of 11 proteins improved discrimination of incident CVD by 5.0% (P=0.0038).

CONCLUSIONS: Large-scale proteomics seem useful for the discovery of new risk markers for CVD and to improve risk prediction in an elderly population of men. Further studies are needed to replicate the findings in independent samples of both men and women of different ages.

Key Words: cardiovascular disease ■ heart failure ■ myocardial infarction ■ proteomics ■ stroke

Genomics, proteomics, and metabolomics are increasingly being applied to the study of a variety of cardiovascular diseases (CVDs) in order to better understand the pathophysiology of these diseases[1–10] and improve our ability to predict their future occurrence. Some CVDs share common features, such as atherosclerosis for both myocardial infarction (MI) and ischemic stroke, or poor systolic function in MI and nonischemic causes of heart failure (HF). One can, therefore, anticipate that some aspects of the genomic, proteomic, and metabolomic profiles of these diseases are shared, while other aspects are likely unique to a specific disease and/or feature.[11] Determining the overlap as well as the nonoverlap in the genomic,

proteomic, and metabolomic profiles of CVDs of interest is expected to facilitate the identification of shared and nonshared pathophysiological pathways.

We recently conducted an analysis of shared genetic loci and plasma proteomic signatures of MI, ischemic stroke, and HF. Using existing results from genome-wide association studies, we found 4 genetic loci in common between MI and HF and 5 loci overlapping between MI and ischemic stroke, while no locus was identified for all 3 of these CVDs.[11] In that study, a nontargeted approach for selection of genetic loci was used. We also investigated the overlap between 92 preselected CVD-related proteins measured in the plasma by the proximity extension technique (PEA)

CLINICAL PERSPECTIVE

What Is New

- We measured 742 proteins with the proximity extension high-throughput assay in 826 men aged 70 years and determined which subset of proteins predicted the development of a first cardiovascular disease (CVD) event.
- A subset of 11 proteins selected through a specialized lasso regression analysis was able to improve discrimination of first CVD events by 5% over established clinical risk factors.
- Some proteins tracking with disease are well known from prior studies, while others offer potential novel insight into the pathophysiology of CVD.

What Are the Clinical Implications?

- Plasma proteins associated with CVD may provide new clues on the pathophysiology of CVD and potential new drug targets.
- Levels of multiple proteins in plasma may one day be routinely used to improve our ability to predict who will develop CVD.
- This improved prediction may lead to improved outcomes by allowing us to more efficiently implement established primary prevention strategies.

Nonstandard Abbreviations and Acronyms

CARDIoGRAMplusC4D	Coronary Artery Disease Genome Wide Replication and Meta-Analysis (CARDIoGRAM) plus the Coronary Artery Disease (C4D) Genetics
FDR	false discovery rate
GDF-15	growth differentiation factor 15
GPNMB	transmembrane glycoprotein NMB
HE4	human epididymis protein 4
KIM-1	kidney injury molecule
LOD	limit of detection
MMP-12	macrophage metalloelastase 12
MR	Mendelian randomization
PEA	proximity extension assay
PIVUS	Prospective Investigation of Vasculature in Uppsala Seniors
ULSAM	Uppsala Longitudinal Study of Adult Men
WFDC2	WAP 4-disulfide core domain protein 2
WISP-1	WNT1-inducible signaling pathway protein 1

and the same 3 major CVDs in the ULSAM (Uppsala Longitudinal Study of Adult Men) cohort. We found GDF-15 (growth differentiation factor 15) and tumor necrosis factor–related apoptosis-inducing ligand receptor 2 to be related to all 3 CVDs, 3 proteins shared between MI and congestive HF (CHF), and 5 proteins shared between ischemic stroke and CHF.[12]

A major disadvantage with our proteomic analysis to date is restricting the analyzed proteins to a limited number already known or suspected to be linked to CVD. This design limitation restricted our ability to identify new pathophysiological pathways or targets of CVD. To partially overcome this problem, we applied the same PEA technology to study almost 10 times the number of proteins including many not previously linked to CVD. The primary aim was to identify novel proteins linked to ≥1 of the CVDs of interest. A secondary aim was to see whether protein determinations could improve discrimination for CVD beyond that possible with traditional risk factors. To achieve these aims, we once again analyzed the prospectively followed ULSAM cohort, which documented incident CVD outcomes for almost 2 decades after blood collected at the age 70 clinic visit in the early 1990s.

METHODS

The data that support the findings of this study are available from the corresponding author upon reasonable request. ULSAM is a population-based cohort study initiated in 1970–1974 when all men in Uppsala City, Sweden, aged 50 years were invited to a health survey aiming to evaluate risk factors for cardiometabolic disease. A total of 2322 men participated (82% of all men invited).[13] This cohort has since been reexamined at ages 60, 70, 77, 82, and 88 years. The present study sample derives from the reexamination at the age of 70 years, which included 1221 of the original participants. However, only 954 could be evaluated regarding plasma proteins, since we excluded 9 attributable to

technical problems of the measurements with the majority of the protein panels and another 12 with multiple missing protein data, possibly because of a shortage of plasma. Another 107 patients with prevalent CVD at age 70 years were also excluded, resulting in a sample of 826 individuals in the final analyses.

Traditional Risk Factors

Blood was drawn into EDTA plasma tubes in the morning after an overnight fast. The tubes were immediately spun with a cooled centrifuge and within 1 hour the plasma was placed in vials and frozen at −80 °C. The vials were kept frozen at that temperature until protein analysis except for 1 thaw/freeze cycle 5 years before the protein analysis. Serum levels of cholesterol, triglycerides, and high-density lipoprotein were assayed by enzymatic techniques on fresh samples. Friedewald formula was used to calculate low-density lipoprotein cholesterol. Blood glucose was measured using an oxidase method on fresh samples. Supine systolic and diastolic blood pressures were measured twice in the right arm after 10 minutes of rest, and the means of these 2 measures were calculated. Data on smoking status and medications at baseline were derived from a questionnaire.

CVD Diagnosis

Data on causes of death and hospitalizations were retrieved from the Swedish Cause of Death Register and the Swedish Hospital Discharge Register, respectively. The 3 major CVDs were defined as: MI (*International Classification of Diseases, Eighth Revision* [*ICD-8*] code 410; *International Classification of Diseases, Ninth Revision* [*ICD-9*] code 410; or *International Classification of Diseases, Tenth Revision* [*ICD-10*] code I20); ischemic stroke (*ICD-8* codes 431, 433–436; *ICD-9* code 431, 433–436; *ICD-10* code I63-I66); and CHF (*ICD-8* codes 427.00, 427.10, 428.99; *ICD-9* code 428; and *ICD-10* codes I50 and I11.0). The accuracy of these diagnoses in the Swedish registers have been deemed to be high,[14] but since CHF diagnosis is less precise, we performed additional chart review–based validation of HF events, as previously described.[15]

Protein Analysis

We used 9 Olink Proseek Multiplex 96×96 panels (Cardiometabolic, Cardiovascular II, Cardiovascular III, Development, Immune Response, Inflammation, Metabolism, Oncology, and Organ Damage) to simultaneously measure proteins in plasma by real-time polymerase chain reaction using the Fluidigm BioMark HD (Olink).[16]

Olink provides a relative protein measure even when it falls below the limit of detection (LOD) for that protein.

While these below LOD measures can be used for analyses, they are substantially less reliable than measures above the LOD. Thus, association with proteins carrying a high fraction of measures below the LOD are much less likely to be reproducible even in the same sample set. We thus chose to exclude any protein with >25% of the values below the LOD. Such a cutoff has been used in previously published studies using PEA technology, as well as in studies using measurements of other compounds with other techniques.[4,12,17] This exclusion led to the following proteins being removed from our analyses: Cardiometabolic panel: prolylcarboxypeptidase, regenerating family member 3 alpha, superoxide dismutase 1, integrin subunit alpha M, and defensin alpha 1; Cardiovascular II panel: brain natriuretic peptide; Cardiovascular III panel: peptidase inhibitor 3 and kallikrein-6; Development panel: cytosolic phospholipase A2, bone marrow stromal cell antigen 1, myocilin, tyrosine-protein phosphatase non-receptor type 6, stress-induced-phosphoprotein 1, and serine protease inhibitor Kazal-type 5; Immune Response panel: zinc finger and BTB domain-containing protein 16, interleukin 1 receptor-associated kinase 4, diacylglycerol kinase zeta, thioredoxin-dependent peroxide reductase, fibroblast growth factor 2, corneodesmosin, FXYD domain-containing ion transport regulator 5, Egl nine homolog 1, eukaryotic translation initiation factor 5A-1, baculoviral IAP repeat-containing protein 2, merlin, SH2B adapter protein 3, transcription factor AP-1, protein kinase C theta type, aryl hydrocarbon receptor nuclear translocator, islet cell autoantigen 1, protein-arginine deiminase type-2, TRAF family member-associated nuclear factor-kappa B activator, importin subunit alpha-5, and interleukin 5; Inflammation panel: interleukin 2 receptor subunit beta, interleukin 1 alpha, interleukin 2, thymic stromal lymphopoietin, interleukin 22 receptor subunit alpha-1, beta-nerve growth factor, interleukin 24, interleukin 13, artemin, interleukin 20, interleukin 33, interleukin 4, leukemia inhibitory factor, neurturin, and interleukin 5; Metabolism panel: protein S100-P, glutaredoxin-1, diablo homolog, annexin A4, annexin A11, catechol O-methyltransferase, hepatoma-derived growth factor, N-ribosyldihydronicotinamide:quinone reductase 2, arginase-1, peptidyl-prolyl cis-trans isomerase FKBP4, and disabled homolog 2; Oncology panel: no protein; and Organ Damage panel: claspin, mothers against decapentaplegic homolog 1, casein kinase I isoform delta, anterior gradient protein 2 homolog, troponin I, NAD-dependent protein deacylase sirtuin-5, tyrosine-protein kinase Fes/Fps, killer cell immunoglobulin like receptor, three Ig domains and long cytoplasmic tail 1, REST corepressor 1, protein phosphatase 1B, pyruvate dehydrogenase [acetyl-transferring]-phosphatase 1, tubulinyl-tyr carboxypeptidase 1, erbin, Ras GTPase-activating protein 1, probetacellulin, melanoma-associated antigen D1,

protein max, integrin beta-1-binding protein 1, NEDD8 ultimate buster 1, methionine aminopeptidase 1, proteasome subunit alpha type-1, syntaxin-binding protein 3, peptidyl-prolyl cis-trans isomerase FKBP1B, Wiskott-Aldrich syndrome protein, receptor-type tyrosine-protein phosphatase eta, and ribonucleoside-diphosphate reductase subunit M2 B. Further details regarding LODs, reproducibility, and validations are available at https://www.olink.com/resources-support/document-download-center/.

Statistical Analysis

All proteins were ranked inverse normalized to obtain a normal distribution and a similar scale for easy comparison between the proteins. To first obtain an overview of protein correlations, we calculated a pairwise correlation heat map using the lattice package in R version 3.6 (R Foundation for Statistical Computing).

Cox proportional hazard analysis was then applied to determine the presence of an association between each of the proteins and the composite end point of CVD (MI, ischemic stroke, or HF). In the first set of models, adjustment was performed for age only. In the second set of models, adjustment was further made for traditional cardiovascular risk factors; systolic blood pressure, diabetes, smoking, low-density lipoprotein and high-density lipoprotein cholesterol, and body mass index.

Thereafter, similar analyses were performed for the 3 CVD outcomes in isolation: MI, ischemic stroke, or HF. We regarded a false discovery rate (FDR) <0.05 according to Benjamini-Hochberg for the age-adjusted analyses and nominal $P<0.05$ for the cardiovascular risk factor–adjusted analyses to be a significant finding for a protein versus an outcome. We used FDR <0.05 instead of the more stringent Bonferroni adjustment for multiple testing because we anticipated that many of the protein levels would be correlated making a Bonferroni-adjustment too conservative.

Next, we evaluated whether a plasma protein "signature" could improve discrimination for the composite end point of CVD (MI or ischemic stroke or HF) over and above that afforded by traditional risk factors. We first used least absolute shrinkage and selection operator (lasso) for logistic regression with the split sample technique (split 50/50 with cross-validation) with all 752 proteins (forcing age into the model) to select a subset of the most independently predictive proteins. Thereafter, we compared the discrimination (C statistic, expressed as area under the curve) for a logistic regression model with the traditional risk factors (systolic blood pressure, diabetes, high- and low-density lipoprotein cholesterol, body mass index, and smoking) with a model that further included the proteins identified using lasso. STATA16 (StataCorp LLC) was used for these calculations.

Last, pathway enrichment analysis was conducted using Reactome online software (https://reactome.

Table 1. Basic Characteristics in the Study Sample Free From Prevalent CVD (n=826)

Variable	Mean (SD) or proportion
Age, y	71.2 (0.6)
Systolic blood pressure, mm Hg	147 (19)
HDL-C, mmol/L	1.3 (0.4)
LDL-C, mmol/L	3.9 (0.9)
Current smoking	21
BMI, kg/m^2	26.3 (3.4)
Diabetes prevalence	11

BMI indicates body mass index; CVD, cardiovascular disease; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein direct.

org).[18] The hypergeometric test was used for enrichment analyses. FDR <0.05 was regarded as a significant pathway.

The study was approved by the ethics committee of Uppsala University on August 9, 2017 (Dnr 2017/286) and patients gave informed consent.

RESULTS

The risk factor profile of the cohort is given in Table 1. Mean intra-assay and interassay variation of all proteins measured was 8% and 12%, respectively. A correlation matrix of all proteins is provided in Figure 1 showing that many pairs of proteins are correlated, with many of the correlations in the range of 0.20 to 0.40.

A total of 283 failures occurred during a median follow-up of 12.5 years (maximal follow-up 22.0 years, and 9869 person-years at risk) for the combined end point CVD (MI, stroke, or HF).

We found a total of 38 proteins to be significantly related to incident CVD using FDR <0.05 for the age-adjusted analyses and $P<0.05$ for the cardiovascular risk factors–adjusted analyses (Table 2). The top 5 proteins were NT-proBNP (N-terminal pro–brain natriuretic peptide), GDF-15, WFDC2 (WAP 4-disulfide core domain protein 2), MMP-12 (macrophage metalloelastase 12), and KIM-1 (kidney injury molecule).

A total of 125 incident cases of MI occurred during follow-up, and 53 proteins were significantly related to this specific outcome using FDR <0.05 (Table 3). The corresponding number of incident events of ischemic stroke and proteins associated with ischemic stroke was 135 and 13, respectively (Table 4). Lastly, we found 16 proteins to be significantly associated with 155 incident HF outcomes (Table 5).

We found 3 proteins (GDF-15, KIM-1, and WFDC2) to be related to all 3 major CVD outcomes. NT-proBNP was related to both stroke and HF, metallopeptidase inhibitor 1 and WISP-1 (WNT1-inducible signaling pathway protein 1) were related to both MI and stroke, and MMP-12 and GPNMB (transmembrane glycoprotein
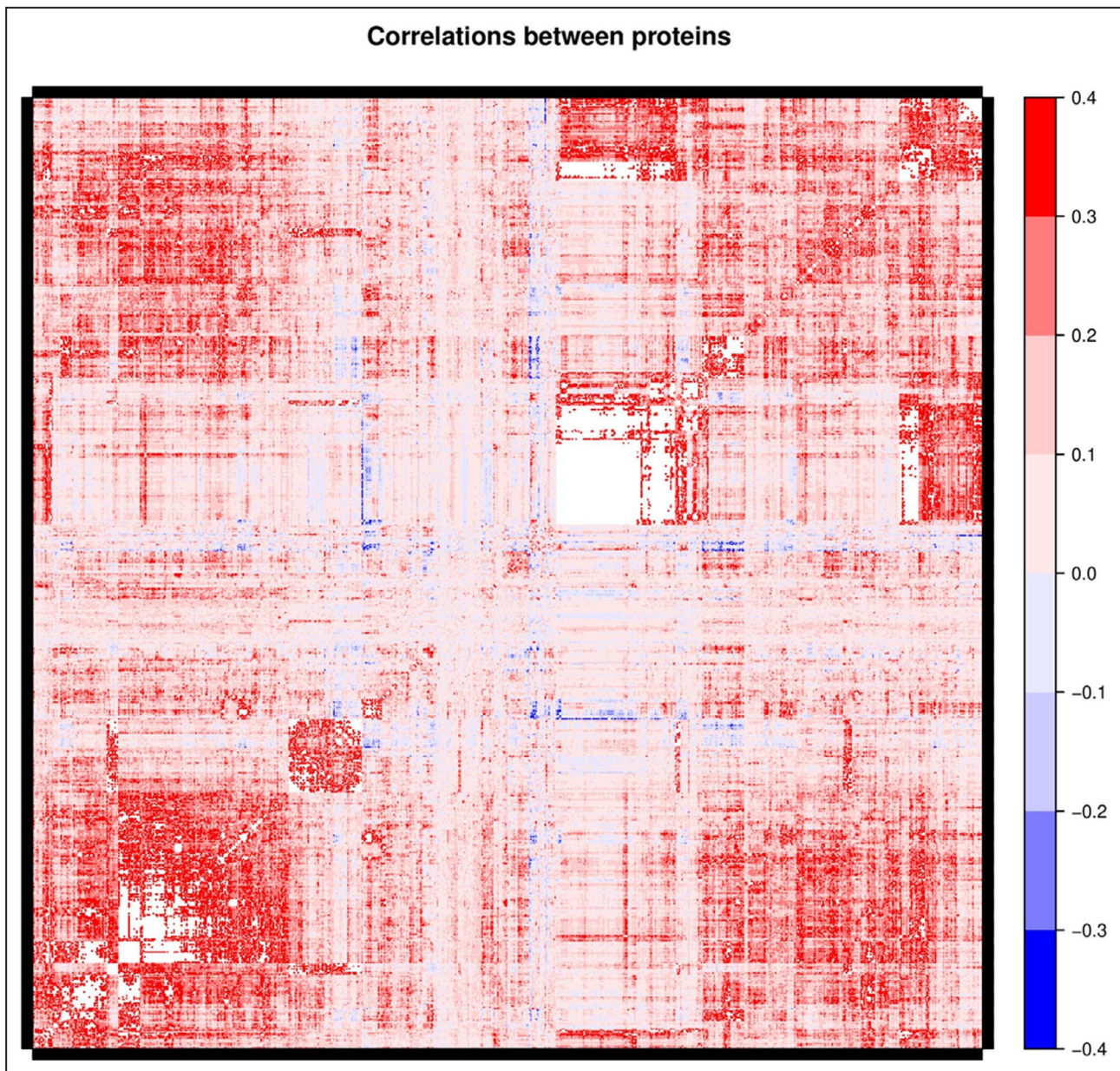
Figure 1.   Pairwise correlations between the 742 proteins used in the analyses.
The correlation coefficients are given by the color-coding at the right. The names of the individual proteins cannot be given in the figure because of limitation in space, therefore the figure gives only an overview of the correlation matrix of the measured proteins.

NMB) were related to both MI and CHF. An overview of the overlap between the protein associations is given in Figure 2.

Lasso logistic regression disclosed 11 proteins to be linked to the combined CVD end point with a lambda of 0.049: NT-proBNP, WFDC2, macrophage metalloelastase 2, KIM-1, interleukin 1 receptor antagonist, lipoprotein lipase (inverse relationship), macrophage receptor with collagenous structure, Ras association domain-containing protein 2 (inverse), stromal cell-derived factor 1, Fc receptor-like protein 6, and

Thymosin beta-10 (inverse). With these 11 proteins, we observed an improvement in the area under the curve of 5.0% ($P=0.0038$, from 0.649 [95% CI, 0.609–0.689] to 0.699 [95% CI, 0.661–0.738]) compared with a model restricted to traditional risk factors (systolic blood pressure, diabetes, smoking, low- and high-density lipoprotein cholesterol, and body mass index). NT-proBNP was by far the most predictive protein in this respect, but a model testing only the remaining 10 selected proteins improved the area under the curve by 2.8% ($P=0.039$).

Table 2. Relationships Between Multiple Proteins at Age 70 Years and Incident CVD in ULSAM

Variable	Gene name	Age-adjusted				Adjusted for traditional cardiovascular risk factors			
		HR	95%CI lower	95%CI higher	P value	HR	95%CI lower	95%CI higher	P value
N-terminal prohormone brain natriuretic peptide	NTproBNP	1.53	1.33	1.75	6.22E-10	1.54	1.34	1.78	3.00E-09
Growth differentiation factor 15	GDF15	1.44	1.27	1.63	1.27E-08	1.31	1.15	1.49	5.05E-05
WAP 4-disulfide core domain protein 2	WFDC2	1.41	1.24	1.59	1.15E-07	1.36	1.19	1.56	7.64E-06
Macrophage metalloelastase	MMP12	1.36	1.2	1.54	9.06E-07	1.24	1.08	1.41	1.76E-03
Hepatitis A virus cellular receptor 1	KIM1	1.34	1.18	1.52	5.34E-06	1.2	1.05	1.37	7.26E-03
Tumor necrosis factor receptor superfamily member 10B	TRAILR2	1.32	1.17	1.49	1.08E-05	1.26	1.1	1.43	5.24E-04
CCN family member 4	WISP1	1.28	1.13	1.44	5.36E-05	1.18	1.04	1.34	9.55E-03
Thrombospondin-2	THBS2	1.28	1.14	1.45	5.81E-05	1.18	1.03	1.34	1.60E-02
Trefoil factor 3	TFF3	1.27	1.13	1.44	1.18E-04	1.26	1.11	1.44	3.55E-04
Interleukin 1 receptor antagonist protein	IL1ra	1.25	1.12	1.41	1.47E-04	1.18	1.03	1.36	1.69E-02
Fatty acid-binding protein, adipocyte	FABP4	1.26	1.12	1.42	1.51E-04	1.21	1.05	1.39	8.09E-03
Interleukin 6	IL6	1.24	1.1	1.39	3.27E-04	1.15	1.01	1.31	3.97E-02
Metalloproteinase inhibitor 1	TIMP1	1.24	1.1	1.4	3.46E-04	1.14	1.01	1.29	3.93E-02
Galectin-9	Gal9	1.24	1.1	1.39	4.38E-04	1.16	1.02	1.33	2.34E-02
Tumor necrosis factor receptor superfamily member 19	TNFRSF19	1.24	1.1	1.41	4.98E-04	1.28	1.12	1.46	2.25E-04
Keratin, type I cytoskeletal 19	KRT19	1.24	1.1	1.39	5.43E-04	1.17	1.03	1.33	1.64E-02
Urokinase plasminogen activator surface receptor	UPAR	1.23	1.09	1.38	5.92E-04	1.15	1.02	1.31	2.65E-02
Transmembrane glycoprotein NMB	GPNMB	1.22	1.09	1.38	6.78E-04	1.21	1.07	1.37	1.86E-03
V-set and immunoglobulin domain-containing protein 4	VSIG4	1.23	1.09	1.39	8.19E-04	1.22	1.07	1.38	2.74E-03
Corticotropin-releasing factor-binding protein	CRHBP	1.23	1.09	1.38	8.26E-04	1.15	1.02	1.31	2.66E-02
Transferrin receptor protein 1	TR	1.23	1.09	1.39	8.39E-04	1.23	1.08	1.4	1.70E-03
Uromodulin	UMOD	0.82	0.73	0.92	9.60E-04	0.86	0.76	0.97	1.43E-02
Protein S100-A12	ENRAGE	1.22	1.08	1.37	1.26E-03	1.15	1.01	1.31	4.13E-02
Macrophage receptor MARCO	MARCO	1.22	1.08	1.37	1.33E-03	1.2	1.06	1.36	3.34E-03
Tumor necrosis factor receptor superfamily member 19L	RELT	1.23	1.08	1.39	1.35E-03	1.18	1.03	1.34	1.38E-02
Amphiregulin	AREG	1.22	1.08	1.38	1.36E-03	1.19	1.05	1.35	7.81E-03
Tumor necrosis factor ligand superfamily member 12	TWEAK	0.83	0.74	0.93	1.84E-03	0.87	0.77	0.98	2.62E-02
Metalloproteinase inhibitor 4	TIMP4	1.21	1.07	1.37	1.96E-03	1.23	1.08	1.4	1.57E-03
Inactive tyrosine-protein kinase 7	PTK7	1.21	1.07	1.37	1.96E-03	1.19	1.05	1.35	5.11E-03

(Continued)

Table 2. Continued

Variable	Gene name	Age-adjusted				Adjusted for traditional cardiovascular risk factors			
		HR	95%CI lower	95%CI higher	P value	HR	95%CI lower	95%CI higher	P value
Follistatin-related protein 3	FSTL3	1.22	1.07	1.38	2.05E-03	1.17	1.03	1.33	1.45E-02
Tissue factor pathway inhibitor 2	TFPI2	1.21	1.07	1.36	2.26E-03	1.18	1.04	1.34	1.12E-02
Leukocyte immunoglobulin-like receptor subfamily B member 4	LILRB4	1.21	1.07	1.36	2.72E-03	1.19	1.05	1.35	6.50E-03
Collectin-12	COLEC12	1.21	1.07	1.36	2.76E-03	1.21	1.07	1.38	2.52E-03
Insulin-like growth factor-binding protein 7	IGFBP7	1.21	1.07	1.37	2.80E-03	1.14	1	1.29	4.86E-02
Tumor necrosis factor receptor superfamily member 10A	TNFRSF10A	1.2	1.06	1.35	3.37E-03	1.14	1.01	1.29	3.73E-02
Triggering receptor expressed on myeloid cells 1	TREM1	1.2	1.06	1.35	3.45E-03	1.18	1.04	1.34	1.13E-02
Carcinoembryonic antigen-related cell adhesion molecule 5	CEACAM5	1.2	1.06	1.35	3.59E-03	1.17	1.03	1.34	1.72E-02
Leukocyte-associated immunoglobulin-like receptor 1	LAIR1	1.19	1.06	1.35	4.11E-03	1.16	1.03	1.32	1.85E-02

Only proteins with a false discovery rate <0.05 for age-adjusted P value and multiadjusted P value <0.05 are shown (n=38). CVD indicates cardiovascular disease; HR, hazard ratio; MARCO, macrophage receptor with collagenous structure; and ULSAM, Uppsala Longitudinal Study of Adult Men.

Pathway enrichment analysis disclosed 13 pathways to be enriched for MI, 2 for ischemic stroke, and none for HF (Table 6).

DISCUSSION

The main finding of our study was the identification of 3 proteins measured in the plasma using the proximity extension assay to be related to the prediction of all 3 major CVD outcomes: GDF-15, KIM-1, and WFDC2. These 3 proteins, together with 8 others selected by lasso, improved discrimination of incident CVD by 5% versus traditional risk factors.

GDF-15 and KIM-1 are among the proteins linked to multiple CVDs in our previously published study of 84 proteins.[12] However, the association between WFDC2 and all 3 major CVDs in the present study is novel. WFDC2 is a broad range protease inhibitor[19] expressed in a number of tissues, such as the cortex of the kidney, multiple urogenital tissues, salivary gland, and the thyroid, but not specifically tissues related to CVD (https://www.gtexportal.org/home/gene/WFDC2). The most well-known application of WFDC2 is as a biomarker of ovarian carcinoma.[20] A search of this gene on Phenoscanner (http://www.phenoscanner.medschl.cam.ac.uk/) showed some genetic associations in the UK Biobank but no convincing evidence for a link with CVD. The WFDC2 protein is, however, also known as HE4 (human epididymis protein 4), and some studies have reported an increased risk of HF in patients with chronic kidney disease in those with high levels of HE4,[21] as well as recurrence of atrial fibrillation after catheter ablation[22] and incident CVD in patients with chronic obstructive pulmonary disease,[23] supporting a potential role of the WFDC2/HE4 protein in CVD.

Mendelian randomization (MR) studies may provide vital insights on whether plasma protein levels are causally associated with phenotypes of interest. Unfortunately, no cis-protein quantitative trait locus has been published for WFDC2 that could be used in an MR study to evaluate whether WFDC2 is causally linked to CVD. In a search for genetic loci in and around the WFDC2 gene (location Chr20:44098346-44110172 using hg19) in the larger genome-wide association study for coronary artery disease including MI, ischemic stroke, and HF, the lowest P value for any single nucleotide polymorphism in that region was 0.03 in CARDIoGRAMplusC4D (Coronary Artery Disease Genome Wide Replication and Meta-Analysis [CARDIoGRAM] plus the Coronary Artery Disease [C4D] Genetics) consortium and the UK Biobank meta-analysis, 0.02 in MEGASTROKE,[2] and 0.04 in Heart Failure Molecular Epidemiology for Therapeutic Targets (HERMES)[3] on HF. Thus, formal MR studies for this protein with these 3 CVDs is unlikely to generate

Table 3. Relationships Between Multiple Proteins at Age 70 Years and Incident MI in ULSAM

Variable	Gene name	Age-adjusted				Adjusted for traditional cardiovascular risk factors			
		HR	95%CI lower	95%CI higher	P value	HR	95% CI lower	95% CI higher	P value
WAP 4-disulfide core domain protein 2	WFDC2	1.66	1.37	2	1.67E-07	1.6	1.3	1.96	6.65E-06
Growth differentiation factor 15	GDF15	1.63	1.35	1.97	3.33E-07	1.51	1.23	1.84	6.12E-05
Tumor necrosis factor receptor superfamily member 10B	TRAILR2	1.55	1.28	1.86	4.42E-06	1.51	1.24	1.84	4.45E-05
Retinoic acid receptor responder protein 2	RARRES2	1.52	1.27	1.83	6.45E-06	1.51	1.24	1.85	5.84E-05
Tumor necrosis factor receptor superfamily member 1A	TNFR1	1.52	1.26	1.83	1.35E-05	1.51	1.24	1.84	4.21E-05
Macrophage metalloelastase	MMP12	1.48	1.23	1.78	2.54E-05	1.36	1.11	1.67	3.00E-03
Tumor necrosis factor receptor superfamily member 19	TNFRSF19	1.47	1.23	1.77	3.21E-05	1.52	1.25	1.85	2.43E-05
Trefoil factor 3	TFF3	1.46	1.22	1.75	3.64E-05	1.46	1.2	1.77	1.50E-04
Inactive tyrosine-protein kinase 7	PTK7	1.44	1.2	1.73	7.78E-05	1.42	1.18	1.71	2.60E-04
Macrophage receptor MARCO	MARCO	1.43	1.2	1.71	9.85E-05	1.39	1.15	1.68	6.06E-04
CD27 antigen	CD27	1.43	1.19	1.73	1.72E-04	1.42	1.16	1.72	5.56E-04
Urokinase plasminogen activator surface receptor	UPAR	1.4	1.17	1.67	1.72E-04	1.32	1.09	1.59	3.90E-03
Alpha-1-microglobulin	AMBP	1.41	1.17	1.69	2.24E-04	1.38	1.14	1.68	1.12E-03
Hepatitis A virus cellular receptor 1	KIM1	1.42	1.18	1.71	2.59E-04	1.27	1.03	1.55	2.31E-02
Tumor necrosis factor receptor superfamily member 4	TNFRSF4	1.41	1.17	1.7	3.12E-04	1.39	1.15	1.69	8.88E-04
Inhibin beta C chain	INHBC	1.39	1.16	1.66	3.17E-04	1.35	1.11	1.64	2.30E-03
Tumor necrosis factor receptor superfamily member 1B	TNFR2	1.4	1.16	1.68	3.31E-04	1.41	1.16	1.71	4.76E-04
Tumor necrosis factor receptor superfamily member 19L	RELT	1.4	1.16	1.68	3.65E-04	1.38	1.14	1.67	1.11E-03
CCN family member 4	WISP1	1.37	1.15	1.64	4.60E-04	1.31	1.09	1.58	4.86E-03
C-X-C motif chemokine 16	CXCL16	1.38	1.15	1.65	4.83E-04	1.3	1.08	1.57	5.66E-03
Corticotropin-releasing factor-binding protein	CRHBP	1.37	1.15	1.64	5.03E-04	1.3	1.08	1.57	6.70E-03
Cathepsin Z	CTSZ	1.37	1.15	1.63	5.19E-04	1.38	1.14	1.68	8.88E-04
Receptor tyrosine-protein kinase erbB-3	ERBB3	1.36	1.14	1.62	5.68E-04	1.28	1.06	1.55	9.96E-03
Tumor necrosis factor receptor superfamily member 10A	TNFRSF10A	1.38	1.15	1.66	5.73E-04	1.38	1.14	1.66	1.07E-03
C-type lectin domain family 5 member A	CLEC5A	1.37	1.15	1.65	5.95E-04	1.31	1.07	1.6	8.00E-03
Leukocyte-associated immunoglobulin-like receptor 1	LAIR1	1.37	1.14	1.64	7.32E-04	1.4	1.16	1.7	5.31E-04

(Continued)

Table 3. Continued

Variable	Gene name	Age-adjusted				Adjusted for traditional cardiovascular risk factors			
		HR	95%CI lower	95%CI higher	P value	HR	95% CI lower	95% CI higher	P value
Protein S100-A11	S100A11	1.35	1.13	1.61	7.44E-04	1.26	1.04	1.52	1.57E-02
Transmembrane glycoprotein NMB	GPNMB	1.35	1.13	1.61	8.84E-04	1.27	1.06	1.53	9.58E-03
Collectin-12	COLEC12	1.36	1.13	1.63	9.13E-04	1.4	1.16	1.69	5.06E-04
Metalloproteinase inhibitor 4	TIMP4	1.36	1.13	1.62	9.90E-04	1.43	1.17	1.74	3.47E-04
Triggering receptor expressed on myeloid cells 1	TREM1	1.36	1.13	1.63	1.17E-03	1.33	1.09	1.61	4.08E-03
Ephrin type-A receptor 2	EPHA2	1.37	1.13	1.65	1.18E-03	1.42	1.16	1.74	5.87E-04
Cystatin-B	CSTB	1.33	1.12	1.59	1.27E-03	1.32	1.09	1.6	3.97E-03
Metalloproteinase inhibitor 1	TIMP1	1.34	1.12	1.6	1.32E-03	1.22	1.01	1.47	3.79E-02
Tumor necrosis factor receptor superfamily member 11B	OPG	1.34	1.12	1.6	1.35E-03	1.26	1.04	1.53	1.78E-02
C-C motif chemokine 7	MCP3	1.32	1.11	1.57	1.48E-03	1.23	1.01	1.48	3.63E-02
Tumor necrosis factor ligand superfamily member 14	TNFSF14	1.33	1.12	1.6	1.60E-03	1.33	1.1	1.6	3.64E-03
Tumor necrosis factor receptor superfamily member 14	TNFRSF14	1.33	1.12	1.6	1.60E-03	1.33	1.1	1.6	3.64E-03
C-X-C motif chemokine 17	CXL17	1.34	1.12	1.6	1.65E-03	1.25	1.02	1.54	3.22E-02
CUB domain-containing protein 1	CDCP1	1.32	1.11	1.58	1.85E-03	1.27	1.06	1.53	1.01E-02
Nectin-2	NECTIN2	1.34	1.11	1.61	1.96E-03	1.33	1.09	1.62	5.16E-03
Carbonic anhydrase 6	CA6	0.76	0.64	0.9	2.06E-03	0.76	0.63	0.92	4.67E-03
Osteoclast-associated immunoglobulin-like receptor	hOSCAR	1.33	1.11	1.59	2.12E-03	1.32	1.09	1.59	4.67E-03
Tumor necrosis factor receptor superfamily member 6	FAS	1.33	1.11	1.59	2.20E-03	1.26	1.04	1.53	1.84E-02
Proprotein convertase subtilisin/kexin type 9	PCSK9	1.32	1.1	1.57	2.27E-03	1.22	1.01	1.48	3.47E-02
CD99 antigen-like protein 2	CD99L2	1.32	1.1	1.57	2.29E-03	1.32	1.1	1.59	3.38E-03
Spondin-2	SPON2	1.33	1.11	1.61	2.31E-03	1.31	1.08	1.59	6.84E-03
Follistatin-related protein 3	FSTL3	1.33	1.11	1.61	2.38E-03	1.34	1.1	1.63	3.97E-03
Glycoprotein Xg	XG	1.32	1.1	1.58	2.41E-03	1.37	1.13	1.66	1.16E-03
Hepatitis A virus cellular receptor 2	HAVCR2	1.33	1.11	1.6	2.61E-03	1.33	1.1	1.62	4.04E-03
Desmocollin-2	DSC2	1.32	1.1	1.58	2.62E-03	1.37	1.13	1.66	1.15E-03
Cystatin-C	CST3	1.32	1.1	1.59	2.93E-03	1.24	1.02	1.5	2.77E-02
Nectin-4	PVRL4	1.31	1.09	1.57	3.57E-03	1.33	1.09	1.61	4.11E-03

Only proteins with a false discovery rate <0.05 for age-adjusted P value and multiadjusted P value <0.05 are shown. HR indicates hazard ratio; MARCO, macrophage receptor with collagenous structure; MI, myocardial infarction; and ULSAM, Uppsala Longitudinal Study of Adult Men.

Table 4. Relationships Between Multiple Proteins at Age 70 Years and Incident Ischemic Stroke in ULSAM

Variable	Gene name	Age-adjusted				Adjusted for traditional cardiovascular risk factors			
		HR	95%CI lower	95%CI higher	P value	HR	95%CI lower	95%CI higher	P value
N-terminal prohormone brain natriuretic peptide	NTproBNP	1.62	1.35	1.95	1.61E-07	1.63	1.35	1.97	4.96E-07
Growth differentiation factor 15	GDF15	1.5	1.26	1.79	7.14E-06	1.42	1.18	1.7	1.52E-04
Syndecan-1	SYND1	1.45	1.21	1.73	3.52E-05	1.46	1.21	1.75	6.97E-05
Thrombospondin-2	THBS2	1.43	1.2	1.7	5.45E-05	1.39	1.15	1.68	7.09E-04
Metalloproteinase inhibitor 1	TIMP1	1.42	1.19	1.69	9.36E-05	1.35	1.13	1.62	9.44E-04
Adhesion G protein–coupled receptor G1	ADGRG1	1.38	1.16	1.63	1.97E-04	1.31	1.07	1.61	9.86E-03
WAP 4-disulfide core domain protein 2	WFDC2	1.4	1.17	1.68	2.67E-04	1.41	1.17	1.71	4.07E-04
Tenascin	TNC	1.36	1.15	1.62	3.45E-04	1.32	1.1	1.58	2.23E-03
Leukocyte immunoglobulin-like receptor subfamily B member 4	LILRB4	1.38	1.16	1.64	3.67E-04	1.39	1.16	1.67	4.27E-04
Hepatitis A virus cellular receptor 1	KIM1	1.37	1.14	1.64	5.82E-04	1.25	1.04	1.51	1.68E-02
Polypeptide N-acetylgalactosaminyltransferase 10	GALNT10	1.35	1.14	1.6	6.74E-04	1.27	1.06	1.52	9.57E-03
Insulin-like growth factor-binding protein 7	IGFBP7	1.35	1.13	1.6	8.46E-04	1.3	1.09	1.55	2.96E-03
CCN family member 4	WISP1	1.34	1.13	1.6	9.25E-04	1.28	1.07	1.53	6.84E-03

Only proteins with a false discovery rate <0.05 for age-adjusted P value and multiadjusted P value <0.05 are shown. HR indicates hazard ratio; and ULSAM, Uppsala Longitudinal Study of Adult Men.

a causal estimate for the WFDC2 protein at the present time. The loci identified in CARDIoGRAMplusC4D and the UK Biobank meta-analysis being closest to the WFDC2 gene was located about 200 kb downstream the WFDC2 gene (position Chr 20:44586023_C_T), but the data from the present study do not suggest that this

Table 5. Relationships Between Multiple Proteins at Age 70 Years and Incident HF in ULSAM

Variable	Gene name	Age-adjusted				Adjusted for traditional cardiovascular risk factors			
		HR	95%CI lower	95%CI higher	P value	HR	95%CI lower	95%CI higher	P value
N-terminal prohormone brain natriuretic peptide	NTproBNP	1.88	1.58	2.24	1.20E-12	1.98	1.64	2.39	9.80E-13
Leukocyte immunoglobulin-like receptor subfamily A member 5	LILRA5	1.42	1.2	1.67	3.48E-05	1.25	1.04	1.49	1.57E-02
Macrophage metalloelastase	MMP12	1.42	1.2	1.68	3.74E-05	1.32	1.1	1.59	2.56E-03
WAP 4-disulfide core domain protein 2	WFDC2	1.41	1.19	1.67	9.51E-05	1.41	1.18	1.7	2.36E-04
Amphiregulin	AREG	1.39	1.17	1.64	1.18E-04	1.31	1.11	1.56	1.77E-03
Protein fosB	FOSB	1.37	1.17	1.6	1.19E-04	1.36	1.16	1.61	2.40E-04
Interleukin 6	IL6	1.37	1.17	1.61	1.36E-04	1.25	1.05	1.5	1.33E-02
Tumor necrosis factor ligand superfamily member 12	TWEAK	0.75	0.64	0.87	2.70E-04	0.8	0.68	0.95	9.91E-03
Heat shock protein beta-1	HSP27	1.29	1.11	1.5	1.14E-03	1.29	1.1	1.51	2.10E-03
Fibroblast growth factor 23	FGF23	1.32	1.12	1.57	1.14E-03	1.23	1.03	1.47	1.96E-02
Uromodulin	UMOD	0.77	0.66	0.9	1.15E-03	0.82	0.69	0.97	2.13E-02
Contactin-1	CNTN1	0.77	0.65	0.9	1.15E-03	0.77	0.65	0.91	2.39E-03
C-X-C motif chemokine 6	CXCL6	0.77	0.66	0.9	1.23E-03	0.75	0.64	0.88	4.67E-04
Growth differentiation factor 15	GDF15	1.31	1.11	1.55	1.41E-03	1.19	1	1.42	4.71E-02
Beta-1,4-glucuronyltransferase 1	B4GAT1	0.78	0.66	.91	1.51E-03	0.83	0.7	0.98	2.72E-02
Transmembrane glycoprotein NMB	GPNMB	1.29	1.1	1.52	1.59E-03	1.3	1.1	1.53	2.06E-03

Only proteins with a false discovery rate <0.05 for age-adjusted P value and multiadjusted P value <0.05 are shown. HF indicates heart failure; HR, hazard ratio; and ULSAM, Uppsala longitudinal study of adult men.
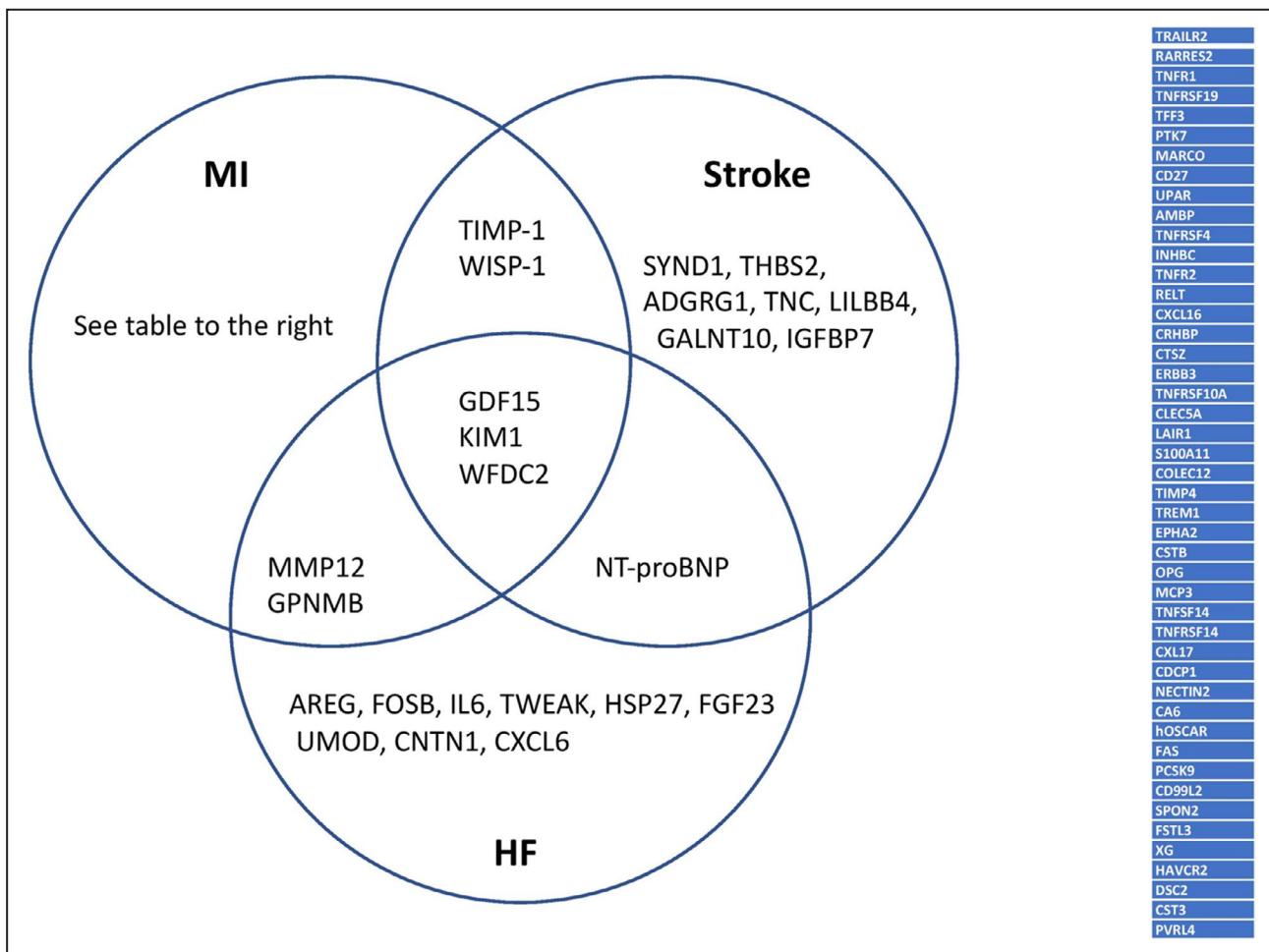
Figure 2. Overview of relationships between protein associations and the 3 different cardiovascular diseases.
The table to the right gives the proteins related to myocardial infarction (MI) only. HF indicates heart failure.

locus is a quantitative trait locus for WFDC2. Although both of the other top findings, GDF-15 and KIM-1, have previously been related to CVD in observational studies,[12] causality versus CVDs was not demonstrated for these 2 proteins in a previously published MR study including multiple CVD-related proteins.[24]

We found several novel proteins to be associated with 2 of our 3 CVD outcomes in this study. First, we found NT-proBNP to be associated with ischemic stroke and CHF. Based on what is already known about the pathophysiology of elevated NT-proBNP, elevated levels of this natriuretic peptide are most likely to be a consequence of certain CVDs rather than the cause. Previous MR studies have also shown that genetically determined high NT-proBNP levels are causally related to reduced blood pressure and a reduced risk of stroke and cardiovascular mortality.[25,26] Second, we found WISP-1, also known as CCN4, to be associated with both MI and stroke. WISP-1 is a downstream regulator in the Wnt/Frizzled signaling pathway, being

associated with cell survival. The protein attenuates p53-mediated apoptosis in response to DNA damage through activation of protein kinase B. Genetic loci in the WISP-1 gene have been linked to hypertension in the UK Biobank in an online-only report (http://www.nealelab.is/uk-biobank). Last, we found GPNMB to be associated with MI and CHF. GPNMB is a type I transmembrane glycoprotein, which shows homology to the Pmel17 precursor, a melanocyte-specific protein. The protein has previously been linked to melanomas, and polymorphisms in *GPNMB* have been associated with HF mortality and treatment for hypertension in the online-only UK Biobank analyses.

Our lasso selection of 11 proteins improved the discrimination statistic from 0.65 to 0.70 versus traditional risk factors when validated in the ULSAM cohort of patients at age 70 years. Since these 11 proteins were not tested in an independent sample, the data are at risk of overfitting. In comparison, in our previous study, the addition of 7 proteins selected by lasso in the PIVUS

Table 6. Pathway Enrichment Analysis for the Proteins Found in Table 3 to Table 5 for the 3 CVDs

Pathway name	Entities found	Entities total	P value	FDR
Myocardial infarction				
TNFs bind their physiological receptors	9	30	1.98E-14	3.36E-12
TP53 regulates transcription of death receptors and ligands	6	18	2.81E-10	2.39E-08
TNFR2 noncanonical nuclear factor-kB pathway	9	104	1.10E-09	6.18E-08
TP53 regulates transcription of cell death genes	6	83	2.16E-06	8.98E-05
Interleukin 10 signaling	6	86	2.64E-06	8.98E-05
Nectin/Necl trans heterodimerization	3	7	4.78E-06	1.34E-04
Immune system	27	2681	1.37E-05	3.29E-04
Regulation by c-FLIP	3	11	1.83E-05	3.30E-04
Dimerization of procaspase-8	3	11	1.83E-05	3.30E-04
CASP8 activity is inhibited	3	14	3.74E-05	6.36E-04
Caspase activation via death receptors in the presence of ligand	3	20	1.07E-04	1.61E-03
Neutrophil degranulation	9	480	2.82E-04	3.72E-03
Cytokine signaling in immune system	14	1092	2.86E-04	3.72E-03
Ischemic stroke				
Post-translational protein phosphorylation	3	109	2.66e-04	1.20E-02
Regulation of IGF transport and uptake by IGFBPs	3	127	4.15e-04	1.20E-02
HF				
None				

Only pathways with a false discovery rate (FDR) <0.05 are shown. c-FLIP indicates cellular FLICE-like inhibitory protein; CVD, cardiovascular disease; HF, heart failure; IGF, insulin-like growth factor; IGFBP, insulin-like growth factor–binding protein; and TNF, tumor necrosis factor.

(Prospective Investigation of Vasculature in Uppsala Seniors) cohort improved discrimination from 0.64 to 0.72 versus traditional risk factors when the calculation of discrimination was performed in the ULSAM cohort at age 77 years.[12] Thus, we observed comparable improvement in discrimination in the 2 studies despite having access to many more proteins in this study. While these results might be caused by differences in the risk factor profile of the 2 cohorts, we note that NT-proBNP was the major protein contributing to the improvement in discrimination in both studies, while other proteins played a minor role in relation to NT-proBNP. A similar conclusion was drawn in an analysis of the ULSAM cohort performed before the -omics era in which 4 biomarkers (NT-proBNP, C-reactive protein, cystatin C, and troponin T) were measured by ELISA.[27] The present results suggest that NT-proBNP as measured by the PEA technique is a powerful predictor of CVD that may be useful in the risk stratification of the general population in addition to the care and identification of patients with HF.

The pathway enrichment analyses revealed significant findings only for the proteins associated with MI, highlighting a potential role of the immune system and cytokines, as well as cell death and apoptosis. The role of the immune system in the initiation of atherosclerosis, as well as rupture of the vulnerable coronary plaque, is well studied, especially in the experimental setting.[28] Also, apoptosis, autophagy,

and cell death seem to be an important feature of atherosclerosis.[29]

A major strength of this study is the measurement of multiple proteins in a cohort with a long follow-up of CVD events. Furthermore, a minority of measured proteins examined have known links to CVD, making our approach more nontargeted than our previous efforts. We acknowledge, however, that a completely unbiased approach can only truly be implemented with antibody-based techniques, such as PEA, sometime in the future when almost all proteins in plasma may be included in the assay. Significant limitations of our study include lack of an adequately powered replication cohort with the same proteins measured and the fact that we studied men only of European descent. Our findings, therefore, still need to be replicated in other cohorts that include women, as well as other age, geographic, and ethnic groups.

Another potential limitation of our study was the prolonged storage time of the plasma used for the protein analyses. During the almost 25 years of storage, some proteins might have degraded. However, we suspect that such degradation would only serve to drive associations observed toward the null hypothesis and would not produce false-positive findings. We are further reassured that some of our top findings overlap with findings from one of our previous studies on protein levels and CVD involving a different cohort with substantially shorter freezer storage time of plasma samples.[12] The

fact that the present plasma samples were subjected to 1 thaw/freeze cycle is probably of less importance, since it has been shown that up to 8 such cycles have little effect on the PEA measurements.[30]

In conclusion, we demonstrate the potential utility of a high-throughput plasma proteomic study to discover new proteins linked to the causes or consequences of CVD. In addition to providing insights on pathophysiology, associated proteins may one day allow for clinically meaningful improvement in risk prediction. However, more definitive conclusions in this respect await replication of our findings in independent samples with the same or even more proteins.

ARTICLE INFORMATION

Affiliations

Department of Medical Sciences, Uppsala University, Sweden (L.L., S.G.); Division of Cardiovascular Medicine, Department of Medicine, Stanford University School of Medicine, Stanford, CA (D.Z., T.L.A.); Department of Public Health and Caring Sciences/Geriatrics, Uppsala University, Sweden (M.I.); Division of Family Medicine and Primary Care, Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, Huddinge, Sweden (J.Ä.); School of Health and Social Sciences, Dalarna University, Falun, Sweden (J.Ä.); and Palo Alto VA Healthcare System, Palo Alto, CA (T.L.A.).

Disclosures

Dr Ingelsson serves as a consultant for Bioarctic AB. Dr Gustafsson is employed by Sence Research AB. Dr Ärnlöv is a consultant for Boehringer Ingelheim and has financial relationships with Novartis and AstraZeneca. The remaining authors have no disclosures to report.

REFERENCES

1. Van der Harst P, Verweij N. Identification of 64 novel genetic loci provides an expanded view on the genetic architecture of coronary artery disease. *Circ Res*. 2018;122:433–443. doi: 10.1161/CIRCRESAHA.117.312086

2. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, Rutten-Jacobs L, Giese AK, van der Laan SW, Gretarsdottir S, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet*. 2018;50:524–537. doi: 10.1038/s41588-018-0058-3

3. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, Hedman ÅK, Wilk JB, Morley MP, Chaffin MD, et al. 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

4. Lind L, Gigante B, Borné Y, Feldreich T, Leppert J, Hedberg P, Östgren CJ, Nyström FH, Sundström J, Ärnlöv J, et al. Plasma protein profile of carotid artery atherosclerosis and atherosclerotic outcomes: meta-analyses and mendelian randomization analyses. *Arterioscler Thromb Vasc Biol*. 2021;41:1777–1788. doi: 10.1161/ATVBAHA.120.315597

5. Zheng J, Haberland V, Baird D, Walker V, Haycock PC, Hurle MR, Gutteridge A, Erola P, Liu Y, Luo S, et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat Genet*. 2020;52:1122–1131. doi: 10.1038/s41588-020-0682-6

6. Chong M, Sjaarda J, Pigeyre M, Mohammadi-Shemirani P, Lali R, Shoamanesh A, Gerstein HC, Pare G. Novel drug targets for ischemic stroke identified through Mendelian randomization analysis of the blood proteome. *Circulation*. 2019;140:819–830. doi: 10.1161/CIRCULATIONAHA.119.040180

7. 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

8. Lind L, Salihovic S, Ganna A, Sundström J, Broeckling CD, Magnusson PK, Pedersen NL, Siegbahn A, Prenni J, Fall T, et al. A multi-cohort metabolomics analysis discloses sphingomyelin (32:1) levels to be inversely related to incident ischemic stroke. *J Stroke Cerebrovasc Dis*. 2020;29:104476. doi: 10.1016/j.jstrokecerebrovasdis.2019.104476

9. Stenemo M, Ganna A, Salihovic S, Nowak C, Sundström J, Giedraitis V, Broeckling CD, Prenni JE, Svensson P, Magnusson PK, et al. The metabolites urobilin and sphingomyelin (30:1) are associated with incident heart failure in the general population. *ESC Heart Fail*. 2019;6:764–773. doi: 10.1002/ehf2.12453

10. Ganna A, Salihovic S, Sundström J, Broeckling CD, Hedman ÅK, Magnusson PK, Pedersen NL, Larsson A, Siegbahn A, Zilmer M, et al. Large-scale metabolomic profiling identifies novel biomarkers for incident coronary heart disease. *PLoS Genet*. 2014;10:e1004801. doi: 10.1371/journal.pgen.1004801

11. Lind L, Sundstrom J, Arnlov J, Ingelsson M, Henry A, Lumbers RT, Lampa E. Life-time covariation of major cardiovascular diseases: a 40-year longitudinal study and genetic studies. *Circ Genom Precis Med*. 2021;14:e002963. doi: 10.1161/CIRCGEN.120.002963

12. Lind L, Arnlov J, Sundstrom J. Plasma protein profile of incident myocardial infarction, ischemic stroke, and heart failure in 2 cohorts. *J Am Heart Assoc*. 2021;10:e017900. doi: 10.1161/JAHA.120.017900

13. Arnlov J, Ingelsson E, Sundstrom J, Lind L. Impact of body mass index and the metabolic syndrome on the risk of cardiovascular disease and death in middle-aged men. *Circulation*. 2010;121:230–236. doi: 10.1161/CIRCULATIONAHA.109.887521

14. Merlo J, Lindblad U, Pessah-Rasmussen H, Hedblad B, Rastam J, Isacsson SO, Janzon L, Rastam L. Comparison of different procedures to identify probable cases of myocardial infarction and stroke in two Swedish prospective cohort studies using local and national routine registers. *Eur J Epidemiol*. 2000;16:235–243. doi: 10.1023/a:1007634722658

15. Ingelsson E, Arnlov J, Sundstrom J, Lind L. The validity of a diagnosis of heart failure in a hospital discharge register. *Eur J Heart Fail*. 2005;7:787–791. doi: 10.1016/j.ejheart.2004.12.007

16. Assarsson E, Lundberg M, Holmquist G, Björkesten J, Bucht Thorsen S, 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

17. Salihovic S, Lind L, Larsson A, Lind PM. Plasma perfluoroalkyls are associated with decreased levels of proteomic inflammatory markers in a cross-sectional study of an elderly population. *Environ Int*. 2020;145:106099. doi: 10.1016/j.envint.2020.106099

18. Jassal B, Matthews L, Viteri G, Gong C, Lorente P, Fabregat A, Sidiropoulos K, Cook J, Gillespie M, Haw R, et al. The reactome pathway knowledgebase. *Nucleic Acids Res*. 2020;48:D498–D503. doi: 10.1093/nar/gkz1031

19. Chari PR. Hydatid disease of scapula and upper third of humerus treated by en bloc excision and fibular bone grafting. *Indian J Orthop*. 2007;41:241–243. doi: 10.4103/0019-5413.33691

20. Molina R, Escudero JM, Auge JM, Filella X, Foj L, Torne A, Lejarcegui J, Pahisa J. HE4 a novel tumour marker for ovarian cancer: comparison with CA 125 and ROMA algorithm in patients with gynaecological diseases. *Tumour Biol*. 2011;32:1087–1095. doi: 10.1007/s13277-011-0204-3

21. Huang Y, Jiang H, Zhu L. Human epididymis protein 4 as an indicator of acute heart failure in patients with chronic kidney disease. *Lab Med*. 2020;51:169–175. doi: 10.1093/labmed/lmz041

22. Mao ZJ, Lin Y, Hu XK, Chen YH, Huang WJ, Huang ZQ. Serum human epididymis protein 4 levels in the prediction of the recurrence of atrial fibrillation after catheter ablation. *Heart Vessels*. 2021;36:686–692. doi: 10.1007/s00380-020-01744-4

23. Lin H, Xiao J, Su X, Song B. The association between serum human epididymis protein 4 level and cardiovascular events in patients with

chronic obstructive pulmonary disease. *Lab Med*. 2021;52:260–266. doi: 10.1093/labmed/lmaa076

24. 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

25. Seidelmann SB, Vardeny O, Claggett B, Yu B, Shah AM, Ballantyne CM, Selvin E, MacRae CA, Boerwinkle E, Solomon SD. An NPPB promoter polymorphism associated with elevated N-terminal pro-B-type natriuretic peptide and lower blood pressure, hypertension, and mortality. *J Am Heart Assoc*. 2017;6:e005257. doi: 10.1161/JAHA.116.005257

26. Theriault S, Sjaarda J, Chong M, Hess S, Gerstein H, Pare G. Identification of circulating proteins associated with blood pressure using mendelian randomization. *Circ Genom Precis Med*. 2020;13:e002605. doi: 10.1161/CIRCGEN.119.002605

27. Zethelius B, Berglund L, Sundstrom J, Ingelsson E, Basu S, Larsson A, Venge P, Arnlov J. Use of multiple biomarkers to improve the prediction of death from cardiovascular causes. *New Engl J Med*. 2008;358:2107–2116. doi: 10.1056/NEJMoa0707064

28. Libby P, Buring JE, Badimon L, Hansson GK, Deanfield J, Bittencourt MS, Tokgozoglu L, Lewis EF. Atherosclerosis. *Nat Rev Dis Primers*. 2019;5:56. doi: 10.1038/s41572-019-0106-z

29. Shan R, Liu N, Yan Y, Liu B. Apoptosis, autophagy and atherosclerosis: relationships and the role of Hsp27. *Pharmacol Res*. 2021;166:105169. doi: 10.1016/j.phrs.2020.105169

30. Shen Q, Bjorkesten J, Galli J, Ekman D, Broberg J, Nordberg N, Tillander A, Kamali-Moghaddam M, Tybring G, Landegren U. Strong impact on plasma protein profiles by precentrifugation delay but not by repeated freeze-thaw cycles, as analyzed using multiplex proximity extension assays. *Clin Chem Lab Med*. 2018;56:582–594. doi: 10.1515/cclm-2017-0648