CARDIOVASCULAR DISEASE

A proteomic surrogate for cardiovascular outcomes that is sensitive to multiple mechanisms of change in risk

Stephen A. Williams[1]*†, Rachel Ostroff[1], Michael A. Hinterberg[1], Josef Coresh[2], Christie M. Ballantyne[3], Kunihiro Matsushita[2], Christian E. Mueller[4], Joan Walter[4,5], Christian Jonasson[6], Rury R. Holman[7], Svati H. Shah[8], Naveed Sattar[9], Roy Taylor[10], Michael E. Lean[11], Shintaro Kato[12], Hiroaki Shimokawa[13,14], Yasuhiko Sakata[13], Kotaro Nochioka[13], Chirag R. Parikh[2], Steven G. Coca[15], Torbjørn Omland[16], Jessica Chadwick[1], David Astling[1], Yolanda Hagar[1], Natasha Kureshi[1], Kelsey Loupy[1], Clare Paterson[1], Jeremy Primus[1], Missy Simpson[1], Nelson P. Trujillo[17], Peter Ganz[18]†

A reliable, individualized, and dynamic surrogate of cardiovascular risk, synoptic for key biologic mechanisms, could shorten the path for drug development, enhance drug cost-effectiveness and improve patient outcomes. We used highly multiplexed proteomics to address these objectives, measuring about 5000 proteins in each of 32,130 archived plasma samples from 22,849 participants in nine clinical studies. We used machine learning to derive a 27-protein model predicting 4-year likelihood of myocardial infarction, stroke, heart failure, or death. The 27 proteins encompassed 10 biologic systems, and 12 were associated with relevant causal genetic traits. We independently validated results in 11,609 participants. Compared to a clinical model, the ratio of observed events in quintile 5 to quintile 1 was 6.7 for proteins versus 2.9 for the clinical model, AUCs (95% CI) were 0.73 (0.72 to 0.74) versus 0.64 (0.62 to 0.65), *c*-statistics were 0.71 (0.69 to 0.72) versus 0.62 (0.60 to 0.63), and the net reclassification index was +0.43. Adding the clinical model to the proteins only improved discrimination metrics by 0.01 to 0.02. Event rates in four predefined protein risk categories were 5.6, 11.2, 20.0, and 43.4% within 4 years; median time to event was 1.71 years. Protein predictions were directionally concordant with changed outcomes. Adverse risks were predicted for aging, approaching an event, anthracycline chemotherapy, diabetes, smoking, rheumatoid arthritis, cancer history, cardiovascular disease, high systolic blood pressure, and lipids. Reduced risks were predicted for weight loss and exenatide. The 27-protein model has potential as a "universal" surrogate end point for cardiovascular risk.

INTRODUCTION

In its seminal report in 2004, "Innovation or Stagnation; Challenge or Opportunity on the Critical Path to New Medical Products" (*1*), the Food and Drug Administration (FDA) recognized that a key cause of the increasing time and cost of drug development was that clinical efficacy and safety claims depended on empirical outcomes testing, using "20th century tools in the 21st century." For cardiovascular disease, they specifically recommended developing and qualifying biomarkers "to improve innovation in a field affecting millions of Americans" and "in clinical practice to evaluate patient risk and to assist physicians and patients in developing treatment strategies." Since that report 17 years ago, although therapeutic options for reducing cardiovascular events in high-risk populations have expanded (*2*–*8*), all clinical trials for safety and all cardiovascular efficacy trials other than cholesterol and blood pressure lowering remain in the same empirical state as they were 20 or more years ago, dependent on counting events, hospitalizations, and deaths. It is hard to avoid the conclusion that dependency on outcome measures as lone pivotal end points led to the too-late finding of the adverse cardiovascular effect of torcetrapib that caused excess deaths and events during its phase 3 trial (*9*), created years of uncertainty around rosiglitazone, leading to the FDA requiring cardiovascular outcome safety studies for all new antidiabetic agents (*10*), and delayed the translation of unexpected cardiovascular and mortality benefits of sodium-glucose cotransporter-2 inhibitors (SGLT2i), which manifested only in late development, into approved claims.

Medical practice faces a different risk assessment problem: individualizing residual cardiovascular risk assessment to enable precision allocation and monitoring of the benefit of cardioprotective therapies (*11*). However, traditional risk assessment tools are only modestly prognostic in individuals most eligible for new drug classes, including

[1]SomaLogic Inc., Boulder, CO 80301, USA. [2]Johns Hopkins University, Baltimore, MD 21218, USA. [3]Baylor College of Medicine, Houston, TX 77030, USA. [4]Cardiovascular Research Institute, University of Basel, Basel 4001, Switzerland. [5]Institute of Diagnostic and Interventional Radiology, University Hospital Zürich, University of Zürich, Zürich 7491, Switzerland. [6]Jebsen Centre for Genetic Epidemiology, Department of Public Health and Nursing, Norwegian University of Science and Technology, Trondheim 7491, Norway. [7]Diabetes Trials Unit, Radcliffe Department of Medicine, University of Oxford, Oxford OX3 9DU, UK. [8]Division of Cardiology, Duke Department of Medicine, and Duke Molecular Physiology Institute, Duke University, Durham, NC 27710, USA. [9]Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow G12 8QQ, UK. [10]Newcastle Magnetic Resonance Centre, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, UK. [11]School of Medicine, Nursing and Dentistry, University of Glasgow, Glasgow G12 8QQ, UK. [12]NEC Solution Innovators Ltd., Tokyo 136-0082, Japan. [13]Tohoku University Graduate School of Medicine, Sendai 980-8576, Japan. [14]Graduate School, International University of Health and Welfare, Narita 286-8686, Japan. [15]Mt Sinai Clinical and Translational Science Research Unit, Icahn School of Medicine at Mount Sinai, New York, NY 11766, USA. [16]Department of Cardiology, Akershus University Hospital and University of Oslo, Oslo 1478, Norway. [17]Boulder Community Hospital, Boulder, CO 80301, USA. [18]Zuckerberg San Francisco General Hospital, University of California, San Francisco, San Francisco, CA 94110, USA.
*Corresponding author. Email: swilliams@somalogic.com
†These authors contributed equally to this work.

people with known cardiovascular disease [who may remain at risk despite adequately controlled low-density lipoprotein (LDL)–cholesterol and blood pressure] (*12*), multimorbid groups (*13*), or the elderly without confirmed cardiovascular disease (*14*).

Last, the insensitivity to improvements of many traditional risk factors (age, sex, race, diabetes status, and hypertension history) and imaging measures (coronary artery calcium score and carotid and coronary imaging) is a problem for both clinical trials and medical practice. This is important because recently developed agents, such as SGLT2i, glucagon-like peptide 1 receptor agonists (GLP-1 RAs), or anti-inflammatory drugs such as canakinumab, reduce cardiovascular risk independently of changes in these factors (*3*, *4*, *6*).

These unmet needs led an expert commentary for the American College of Cardiology to conclude that "new cardiovascular risk models incorporating novel risk markers are needed" (*15*), echoing the FDA's comments made 15 years earlier. The idealized requirements of precise, sensitive prognostics that respond agnostically and reliably to all changes in outcomes regardless of intervention mechanism are key features of a surrogate end point (*16*)—a biomarker that is intended to substitute for a clinical end point. Fortunately, one issue that has been addressed as part of the FDA's Critical Path initiative is the development of an evidentiary framework for biomarker qualification (*16*). Because previous studies suggested that blood protein patterns may have some of the desired characteristics of a surrogate end point (*17*–*20*), we proposed the hypothetical design of the proteomic program described herein at the "Framework for Defining Evidentiary Criteria; Surrogate Endpoint Qualification Workshop" in 2018, cosponsored by the FDA and Foundation for the National Institutes of Health.

The aims of this research program were therefore to derive and validate a proteomic prognostic test that predicted all major cardiovascular outcomes and deaths in a time scale consistent with clinical outcome trials, encompassed all detectable biologic mechanisms relating to changes in risk, and was reliably directionally sensitive to all adverse and beneficial changes in outcome. The proposed two contexts of use for a validated test would be as a candidate surrogate end point for phase 2 studies in cardiovascular disease and diabetes and subsequently as a reasonably likely surrogate end point for accelerated drug approval of breakthrough products. In medical practice, it ideally would also be used as a test for individualized and cost-effective cardioprotective drug allocation and for monitoring of responses.

RESULTS

Development of the proteomic prognostic model

Using machine learning applied to the 5000 proteins measured in each plasma sample, a fully parametric accelerated failure time (AFT) Weibull prognostic model was developed in secondary disease subsets from The Trøndelag Health Study (HUNT) and visit 5 of the Atherosclerosis Risk in Communities (ARIC) study, hypothesized to be enriched for the biologic mechanisms of risk and with well-adjudicated outcomes, as shown in the top half of Fig. 1. The final model was subsequently applied to a number of independent cohorts of multimorbid populations: Biochemical and Electrocardiographic Signatures in the Detection of Exercise-induced Myocardial Ischemia (BASEL VIII), Chronic Heart Failure Analysis and Registry in the Tohoku District 2 Trial (CHART-2), Exenatide Study of Cardiovascular Event Lowering (EXSCEL), the Action to Control Cardiovascular Risk in Diabetes (ACCORD), the Diabetes Remission Clinical Trial (DiRECT), Prevention of Cardiac Dysfunction During Adjuvant Breast Cancer Therapy (PRADA), and the fraction of the ARIC visit 5 cohort over 65 without known coronary disease. Longitudinal change was assessed in paired samples from EXSCEL, ACCORD, PRADA, and DiRECT. Cross-sectional relation to predictions with known elevated cardiovascular risks was evaluated in the entire ARIC visit 3 cohort. Table 1 shows the breakdown of the 13,167 participants in the nine studies or study fractions used for training and validation with known 4-year outcomes; further details of the parent studies are described in Materials and Methods and tables S1 to S6. Table 1 also shows how the six independent validation study fractions with 4-year outcomes were combined into a single validation metacohort of 11,609 participants, with an overall 4-year event rate of 21.9%. There were 2540 events: 972 deaths (38%), 622 hospitalizations for heart failure (24.5%), 601 myocardial infarctions (23.6%), and 345 strokes (13.6%).

The model consisted of 27 proteins and predicted the absolute likelihood within 4 years of any component of the composite end point after the blood sample: myocardial infarction, stroke, hospitalization for heart failure, and all-cause death. The proteins included in the model are listed in Table 2, with their proportionate contribution to the model output, along with the coefficients of variation for each analyte. Fourteen of the proteins were positively correlated with risk, and 13 were negatively correlated. Although the population average contribution of each protein ranged from 1 to 23%, within individuals, this varied.

Biologic relevance of the 27 proteins and inferences of causality

A formal statistical biologic pathway enrichment analysis would not be valid because of intentional biases in machine learning, which limit the inclusion of correlated features. Nevertheless, we approximated thematic biologic groupings across at least 10 biological processes: blood volume and natriuresis [natriuretic peptides B (NTproBNP) and atrial natriuretic factor (ANP)], vesicle biogenesis [adenosine 5′-diphosphate (ADP)–ribosylation factor-like protein 11 (ARL11)], matrix/tissue modeling, growth, angiogenesis or adhesion [anthrax toxin receptor 2 (ANTR2), cartilage intermediate layer protein 2 (CILP2), mucin-16 (CA125*), Golgi membrane protein 1 (GOLM1), spondin-1*, Sushi von Willebrand factor type A (SVEP1*), receptor-type tyrosine-protein phosphatase eta (PTRPJ), inter-alpha-trypsin inhibitor heavy-chain H2 (ITI heavy-chain 2*), protein kinase C–binding protein NELL1 (NELL1), and growth/differentiation factor 11/8 shared epitope (GDF11/8*)], cellular immunity [macrophage metalloelastase (MMP12*), receptor tyrosine-protein kinase erbB-3 (ERBB3), and neural cell adhesion molecule 1, 120-kDa isoform (NCAM-120*)], calcium channel modulation [voltage-dependent calcium channel subunit alpha-2/delta-3 (CA2D3*)], glomerular filtration rate [trefoil factor 3 (TFF3)], immunoglobulins/receptors [immunoglobulin superfamily DCC subclass member 4 (IGDC4), junctional adhesion molecule B (JAM-B), and triggering receptor expressed on myeloid cells (sTREM1*)], metabolism and lipids [nicotinamide adenine dinucleotide (NAD)–dependent protein deacetylase sirtuin-2 (SIRT2), protein phosphatase 1 regulatory subunit 1A (PPR1A), and LDL receptor–regulated protein 11 (LRP11*)], inflammation [urokinase plasminogen activator surface receptor (suPAR*) and bifunctional heparan sulfate *N*-deacetylase/*N*-sulfotransferase 1 (NDST1)], and coagulation [A disintegrin and metalloproteinase with thrombospondin motifs 13 (ATS13*)]. The potential causal relationship between these proteins and cardiovascular disease and its risk factors was explored through Mendelian randomization analysis

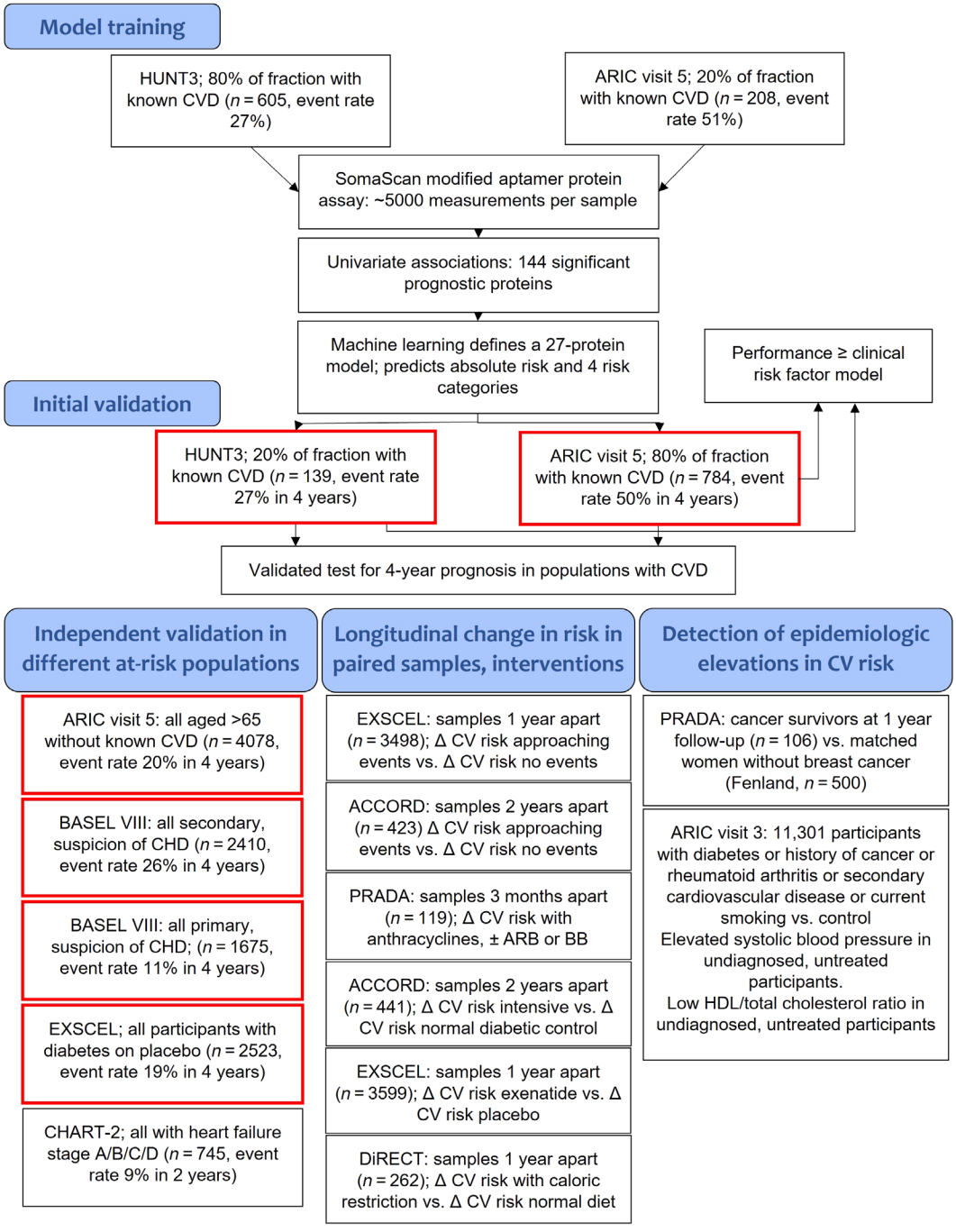

Fig. 1. Contributing clinical study fractions and their roles in the surrogate discovery and validation program. Studies and study fractions outlined in red represent validation studies with 4-year outcomes. These were analyzed individually and also merged into a single validation meta-cohort. CV, cardiovascular; CVD, cardiovascular disease; CHD, coronary heart disease; HUNT, The Trøndelag Health Study; ARIC, Atherosclerosis Risk in Communities; BASEL VIII, Biochemical and Electrocardiographic Signatures in the Detection of Exercise-induced Myocardial Ischemia; CHART-2, Chronic Heart Failure Analysis and Registry in the Tohoku District 2 Trial; EXSCEL, Exenatide Study of Cardiovascular Event Lowering; ACCORD, the Action to Control Cardiovascular Risk in Diabetes; DiRECT, Diabetes Remission Clinical Trial; PRADA, Prevention of Cardiac Dysfunction During Adjuvant Breast Cancer Therapy; ARB, angiotensin receptor blocker; BB, beta blocker.

available for 989 proteins in the PheWAS database (*21*). Sixteen of the 27 model proteins were included in the database, 12 of which were associated with at least one cardiovascular disease–related trait, denoted by the asterisks in the list above and detailed in Table 2 and table S7. Table S7 also shows their relation to cardiovascular disease in the literature, where known.

The equation that described the relationship of the 27 proteins listed in Table 2 to generate the likelihood output was generated as

Table 1. Baseline characteristics of study fractions used for prognostic model training and validation. CHART-2 was not included in the meta-cohort because the follow-up period was only 2 years. Race/ethnicity was not recorded for HUNT3 or BASEL VIII. HUNT, The Trøndelag Health Study; ARIC, Atherosclerosis Risk in Communities; BASEL VIII, Biochemical and Electrocardiographic Signatures in the Detection of Exercise-induced Myocardial Ischemia; CHART-2, Chronic Heart Failure Analysis and Registry in the Tohoku District 2 Trial; EXSCEL, Exenatide Study of Cardiovascular Event Lowering. CHF, congestive heart failure; MI, myocardial infarction; n/a, not available.

Covariate	Measure	HUNT3 secondary	ARIC secondary	CHART-2 heart failure	HUNT3 secondary	ARIC secondary	ARIC primary elderly	BASEL VIII secondary	BASEL VIII primary	EXSCEL placebo baseline	Validation meta-cohort with 4-year outcomes
Population	Purpose/ duration	Model training, 4-year outcomes		2-year outcomes	Validation studies or study fractions with 4-year outcomes included in validation meta-cohort						All 4-year studies merged dataset
	Study fraction and morbidity	80% of total secondary population	20% of total visit 5 secondary population	100% prevalent CHF stages A/B/C/D	20% of total secondary population	80% of total visit 5 secondary population	100% of visit 5 primary population aged >65	100% secondary with CAD symptoms	100% primary with CAD symptoms	100% of placebo; type 2 diabetes	Multimorbidity
	Sample size	605	208	745	139	784	4078	2410	1675	2523	11,609
Composite CV event	Event (%)	166 (27.4%)	106 (51.0%)	68 (9.1%)	37 (26.6%)	390 (49.7%)	817 (20.0%)	627 (26.0%)	181 (10.8%)	488 (19.3%)	2,540 (21.9%)
		Death = 33	Death = 25	Death = 26	Death = 14	Death = 72	Death = 350	Death = 236	Death = 124	Death = 176	Death = 972
		CHF = 42	CHF = 48	CHF = 27	CHF = 10	CHF = 207	CHF = 253	CHF = 77	CHF = 18	CHF = 57	CHF = 622
		MI = 57	MI = 23	MI = 3	MI = 10	MI = 80	MI = 109	MI = 218	MI = 11	MI = 173	MI = 601
		Stroke = 34	Stroke = 10	Stroke = 12	Stroke = 3	Stroke = 31	Stroke = 105	Stroke = 96	Stroke = 28	Stroke = 82	Stroke = 345
	No event (%)	439 (72.6%)	102 (49.0%)	677 (90.9%)	102 (73.4%)	394 (50.3%)	3261 (80.0%)	1783 (74.0%)	1494 (89.2%)	2035 (80.7%)	9,069 (78.1%)
Follow-up time (days)	Mean events (SD), range	806 (514), 5–1897	846 (582), 27–2027	294 (205), 10–714	638 (432), 4–1675	910 (616), 7–2232	1136 (580), 4–2375	776 (647), 1–2958	863 (679), 7–2822	725 (477), 1–1950	907 (614), 1–2,958
	Mean no events (SD), range	1687 (176), 1407–2030	2022 (269), 484–2399	567 (152), 46–1092	1721 (175), 1407–2035	2019 (257), 238–2389	2042 (235), 134–2404	1421 (669), 361–3024	1392 (681), 267–3054	1423 (346), 1–2112	1,674 (558), 1–3,054
Age (years)	Mean (SD), range	69.4 (10.4), 32–99	76.8 (5.2), 67–89	68.5 (13.1), 29–94	69.5 (10.8), 44–92	77.2 (5.3), 67–90	75.3 (5.1), 66–90	68.1 (10.6), 34–93	65.9 (11.8), 26–95	62.8 (9.4), 29–88	69.8 (10.1), 26–95
Sex	Male (%)	438 (72.4%)	124 (59.6%)	494 (66.0%)	108 (77.7%)	498 (63.5%)	1581 (38.8%)	1891 (78.5%)	855 (51.0%)	1520 (60.2%)	6,453 (55.6%)
	Female (%)	167 (27.6%)	84 (40.4%)	251 (34.0%)	31 (22.3%)	286 (36.5%)	2497 (61.2%)	519 (21.5%)	820 (49.0%)	1003 (39.8%)	5,156 (44.4%)
Race or ethnicity	White (%)	n/a	171 (82.2%)	0	n/a	652 (83.2%)	3322 (81.5%)	n/a	n/a	2032 (80.5%)	10,230 (88.2%)
	Black (%)	n/a	37 (17.8%)	0	n/a	132 (16.8%)	756 (18.5%)	n/a	n/a	44 (1.7%)	932 (8%)
	Asian (%)	n/a	0	745 (100%)	n/a	0	0	n/a	n/a	176 (7.0%)	176 (1.5%)
	Hispanic (%)	n/a	0	0	n/a	0	0	n/a	n/a	260 (10.3%)	260 (2.2%)
	Other (%)	n/a	0	0	n/a	0	0	n/a	n/a	11 (0.4%)	11 (0.10%)

Table 2. The 27 proteins selected by machine learning for inclusion in the model. Their contributions to the risk prediction total in the population ("population effect") are shown for the HUNT3 training dataset; a negative sign indicates that increases in the analyte were negatively correlated with risk. Magnitudes correspond to linear predictor beta coefficients in the model, with each analyte centered and scaled to its respective distribution in the training data. FDR, false discovery rate, *P* values corrected for 5000 measurements relating to the univariate prognostic performance of each analyte in a Cox model. CV, coefficients of variation, derived from 6300 production runs of quality control samples included in each 96-well plate of the SomaScan assay.

Proteins included in the model	Abbreviation	Population effect	Univariate *P* value FDR corrected	Analyte CV (%)
Natriuretic peptides B	NTproBNP	0.23	2.60×10^{-10}	7.47
ADP-ribosylation factor-like protein 11	ARL11	−0.22	2.94×10^{-3}	4.19
Anthrax toxin receptor 2	ANTR2	−0.14	5.27×10^{-7}	3.85
Macrophage metalloelastase*	MMP-12	0.14	1.16×10^{-5}	8.84
Cartilage intermediate layer protein 2	CILP2	−0.13	2.75×10^{-6}	5.95
Mucin-16*	CA125	0.11	5.55×10^{-4}	7.43
Receptor tyrosine-protein kinase erbB-3	ERBB3	−0.11	9.46×10^{-4}	6.1
Voltage-dependent calcium channel subunit alpha-2/delta-3*	CA2D3	−0.1	6.36×10^{-4}	5.7
Golgi membrane protein 1	GOLM1	0.1	1.38×10^{-3}	5.89
Neural cell adhesion molecule 1, 120-kDa isoform*	NCAM-120	−0.1	4.96×10^{-3}	4.8
Spondin-1*	Spondin-1	0.1	1.09×10^{-5}	5.15
Immunoglobulin superfamily DCC subclass member 4	IGDC4	−0.09	1.44×10^{-11}	8.4
Trefoil factor 3	TFF3	0.09	3.00×10^{-7}	7.81
Protein phosphatase 1 regulatory subunit 1A	PPR1A	0.08	3.55×10^{-4}	6.31
NAD-dependent protein deacetylase sirtuin-2	SIRT2	−0.08	2.11×10^{-3}	6.5
Atrial natriuretic factor	ANP	0.07	1.24×10^{-4}	5.5
Junctional adhesion molecule B	JAM-B	0.07	2.29×10^{-4}	6.24
Low-density lipoprotein receptor-related protein 11*	LRP11	0.07	1.06×10^{-4}	7.08
Sushi, von Willebrand factor type A, EGF, and pentraxin domain-containing protein 1*	SVEP1	0.05	1.22×10^{-7}	5.01
Receptor-type tyrosine-protein phosphatase eta	PTPRJ	−0.04	3.54×10^{-5}	4.98
Inter-alpha-trypsin inhibitor heavy-chain H2*	ITI heavy-chain H2	−0.03	2.56×10^{-6}	6.44
Protein kinase C–binding protein NELL1	NELL1	0.03	4.94×10^{-7}	5.97
Urokinase plasminogen activator surface receptor*	suPAR	0.03	7.27×10^{-4}	6.52
A disintegrin and metalloproteinase with thrombospondin motifs 13*	ATS13	−0.02	3.00×10^{-7}	4.08
Bifunctional heparan sulfate *N*-deacetylase/*N*-sulfotransferase 1	NDST1	−0.02	8.11×10^{-6}	2.7
Triggering receptor expressed on myeloid cells 1*	sTREM-1	0.02	1.31×10^{-6}	7.66

continued on next page

Proteins included in the model	Abbreviation	Population effect	Univariate *P* value FDR corrected	Analyte CV (%)
Growth/differentiation factor 11/8 (shared epitope)*	GDF-11/8	−0.01	3.00×10^{-7}	3.97

*Potentially causal model proteins significantly associated with at least one cardiovascular disease–related trait.

follows: The survival function for a Weibull AFT model could be expressed as

$$S(t \mid \theta) = \exp\{-[\![(\theta * t)]\!] \hat{} a\}$$

where a is the shape parameter of the Weibull distribution and is estimated with the other parameters during model development. For the 27-protein test, $a\hat{} = 1.1$, and the time t is set to 4 years to generate the probability of an event within 4 years.

The relationship of the 27 proteins with full names listed in Table 2 was a linear combination of protein measurements, as follows

$$\theta\hat{} = \exp\{-(2.83 + -0.09 * \text{TFF3} + -0.23 * \text{BNP} + -0.05 * \text{SVEP1} + 0.01 * \text{"GDF} - 11/8\text{"} + -0.02 * \text{“sTREM} - 1\text{”} + 0.09 * \text{IGDC4} + -0.03 * \text{NELL1} + -0.14 * \text{"MMP} - 12\text{"} + 0.02 * \text{ATS13} + -0.03 * \text{suPAR} + 0.13 * \text{CILP2} + 0.02 * \text{NDST1} + -0.1 * \text{"Spondin} - 1\text{"} + 0.14 * \text{ANTR2} + 0.04 * \text{PTPRJ} + -0.07 * \text{LRP11} + -0.07 * \text{ANP} + -0.07 * \text{"JAM} - \text{B"} + 0.08 * \text{SIRT2} + -0.11 * \text{CA125} + 0.1 * \text{CA2D3} + 0.03 * \text{“ITI heavy} - \text{chain H2”} + 0.11 * \text{ERBB3} + -0.1 * \text{GOLM1} + -0.08 * \text{PPR1A} + 0.22 * \text{ARL11} + 0.1 * \text{"NCAM} - 120\text{"})\}$$

Dynamic range of stratification and reclassification of risk

The dynamic range of prognostic stratification was assessed using the ratio of observed 4-year event rates between the highest and lowest quintiles of the predicted risks (Fig. 2). For proteins, these ratios were 7.0- and 5.2-fold in the two initial validation datasets and 6.7-fold in the meta-cohort of 11,609 participants in all six validation datasets with 4-year outcomes combined.

Because the standard pooled cohort equation (PCE) was not qualified for our higher-risk study populations, we trained and validated a clinical model, refitting the same cardiovascular risk factors (age, sex, race, total cholesterol, high-density lipoprotein (HDL) cholesterol, blood pressure, diabetes, and smoking) in the same datasets as proteins, for the same time horizon, study populations, and composite end point. Compared to the protein model, the clinical model's dynamic range of stratification was more constrained. The ratios of observed events in clinical model–predicted quintile 5 versus clinical model–predicted quintile 1 were 4.3- and 1.5-fold in initial validation datasets; in the subsequent validation meta-cohort, there were 5593 participants with all available clinical model components, and in this subset, the ratio was 2.9-fold for the clinical model versus 8.4-fold for the protein model. Net continuous risk reclassification indexes (event-NRI [total NRI]) for proteins versus the clinical model were also highly positive at +61% [0.60] and +49% [0.53] in the two initial validation datasets and +42% [0.43] in the subset of the meta-cohort with available clinical model components.

Predicted and observed event rates and timing of events

Figure 2 shows the relation between predicted and observed event rates by decile in the training and initial validation datasets in secondary populations. For potential clinical applications, four protein-defined risk categories were created, with cutoffs that could be applied consistently across studies with very different risk spectra. Four-year observed event rates were 5.6% ("low" risk, predicted 0 to 7.5%, n = 1677), 11.2% ("low-medium" risk, predicted 7.6 to 25%, n = 4720), 20.0% ("medium-high" risk, predicted 26 to 50%, n = 3064), and 43.4% ("high" risk, predicted 51 to 100%, n = 2148) in the six studies in the validation meta-cohort. Kaplan-Meier survival estimates are also shown in Fig. 3. A simple recalibration process mitigated the tendency of the model to overpredict risk at the higher end of the prediction range, particularly in populations with low observed event rates; this was applied post hoc and is described in fig. S1. The median time to event for people who had events within the prediction period was 1.71 years overall, 1.48 years for the top quintile, and 1.40 years for the top decile.

Consistency of performance measured using discrimination metrics

Although discrimination measures are not thought to be the most reflective of a model's application in clinical practice (*22*, *23*), we used them to test the consistency of the protein model across ethnicities/races, demographics, multimorbidities, age ranges, and geographic regions (Fig. 4). Area under the curve (AUC) values for the protein model in the initial validations in HUNT3 and ARIC were 0.77 and 0.74; *c*-statistics were 0.73 and 0.70, respectively. For the six-study independent validation meta-cohort, the 4-year AUC for the protein model was 0.73 (with 95% confidence intervals of 0.72 to 0.74), and the *c*-statistic was 0.71 (0.69 to 0.72), generally consistent across all the different populations [men, women, Caucasian, Black, Japanese, those with known coronary heart disease (CHD) present or absent, elderly without known CHD, those with prevalent heart failure, those with diabetes, and those with suspected chronic coronary syndromes], whereas the AUC and *c*-statistics for the refit clinical model in the 5593 participants with all components available were 0.64 (0.62 to 0.65) and 0.62 (0.60 to 0.63). The discrimination performance of the standard PCE in these participants was similar to the optimized clinical model, with AUC and *c*-statistics of 0.67 (0.65 to 0.68) and 0.63 (0.62 to 0.64), respectively. Thus, refitting of the optimized clinical model enabled the calculation of the NRI but did not improve discrimination performance versus the standard PCE.

In addition, in a post hoc analysis, combining the PCE and the protein model resulted in AUC and *c*-statistics of 0.75 (0.74 to 0.77) and 0.73 (0.71 to 0.74)—0.02 and 0.01 incremental improvements, respectively, versus the protein model alone. Furthermore, systolic blood pressure (in millimeters of mercury) did not vary significantly across protein-predicted risk categories: low, 133.7; low-medium, 132.6; medium-high, 133.2; and high, 132.8 [analysis of variance (ANOVA), P = 0.18].

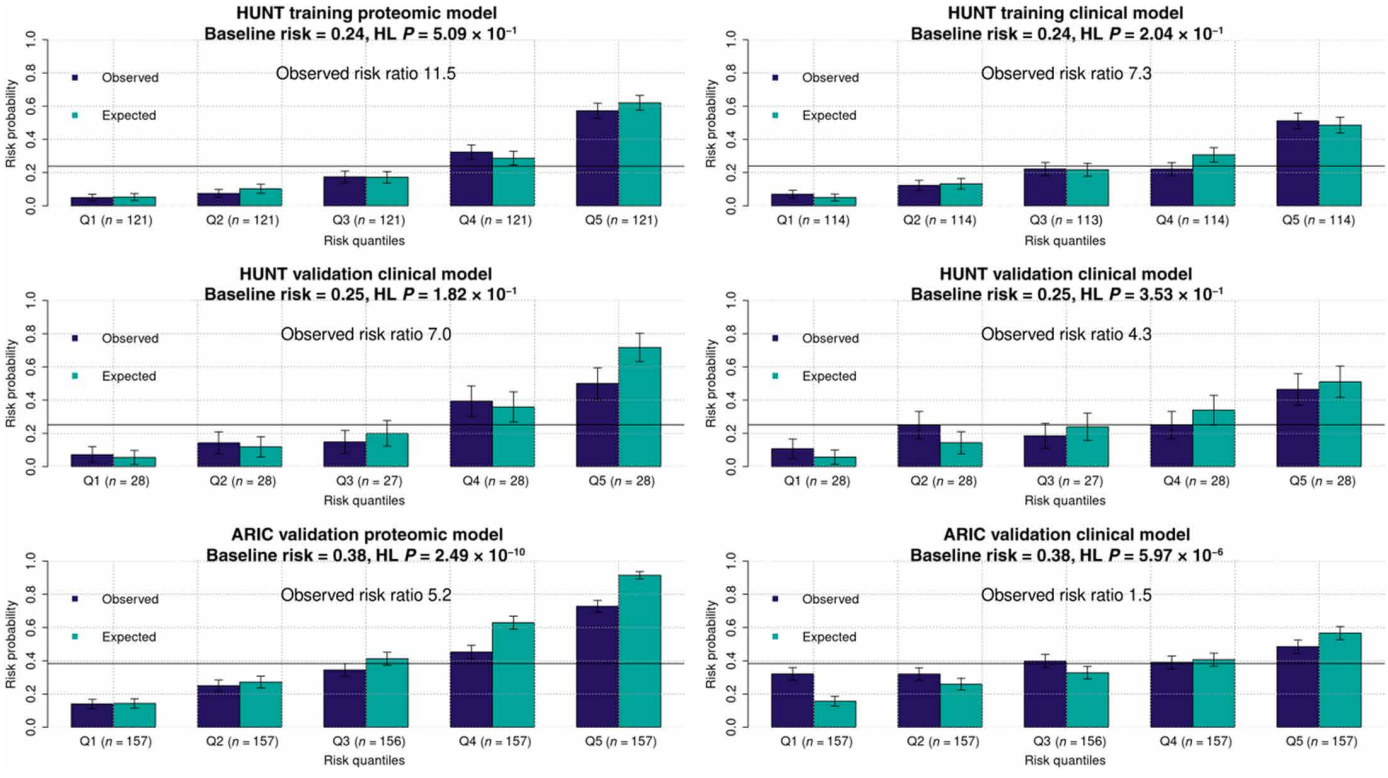

Fig. 2. Calibration and dynamic range of stratification in secondary populations for the protein model versus the clinical model. The left panels represent the protein model, whereas the right panels represent the clinical model. The top row shows model training in 80% of HUNT3; the middle row shows model validation in 20% holdout fraction of HUNT; the bottom row shows model validation in 80% of ARIC secondary (visit 5). The slightly lower participant numbers for training the clinical model arose because of some missing clinical components. HL, Hosmer-Lemeshow goodness-of-fit statistic. Significant *P* values denote imperfect calibration.

Sensitivity of the model to adverse, neutral, and beneficial changes in risk in paired samples

Table 3 shows the eight evaluations of changes in protein-predicted absolute risk within paired longitudinal samples, as well as by several commonly favored individual protein biomarkers that were available on the SomaScan platform. Approaching an adverse event that occurred after the second paired sample was correctly heralded by a greater increase in the calculated 27-protein model risk than in those without subsequent recorded events: a 2.9% greater increment over 1 year in the EXSCEL trial ($P < 0.01$) and a 6% greater increment over 2 years in the ACCORD trial (*24*) ($P < 0.05$). Adverse changes were also correctly detected from anthracycline chemotherapy in the PRADA trial (*25*): an increase of +6.2% versus baseline ($P < 0.01$) within about 3 months. In addition, the expected adverse effect of aging on cardiovascular risk was also manifest [+2.25% ($P < 0.01$) in 1 year in the placebo group of the EXSCEL trial, +4.89% ($P < 0.01$) over 1 year in the control group in the DiRECT trial, and +7.5% ($P < 0.01$) over 2 years in the ACCORD trial]. Neutral effects of treatment were correctly predicted by the 27-protein model for the effect of more intensive diabetic control in the fraction of the ACCORD population studied here, where there was no change in observed event rates, and in PRADA, where angiotensin receptor blockers or beta blockers were neither observed nor predicted by the 27-protein score to individually reduce the impact of cardiotoxic anthracycline chemotherapy (*26*). Beneficial predicted changes were evident in the EXSCEL trial (*27*) where the effect of GLP-1 RA exenatide was predicted by the 27-protein model to result in an absolute 4-year risk reduction compared to placebo of −1.49% at 1 year into the randomized study ($P < 0.01$), compared with observed risk reduction of −0.8% ($P = 0.06$) at a median follow-up of 3.2 years. Beneficial changes were also predicted in the DiRECT trial (*28*) where remission of type 2 diabetes was achieved in nearly half the participants at 12 months, with effective caloric restriction and mean weight loss of 10 kg. The protein model predicted an absolute risk difference of −6.7% versus the standard diet group ($P < 0.01$). In contrast, several individual protein biomarkers that were measurable on the SomaScan platform and known to be prognostic for cardiovascular events showed inconsistent or insensitive relations to the same interventions (Table 3).

Detecting elevated risks from different conditions known to be associated with higher observed event rates

The ability of the protein model to further detect mechanistically distinct drivers of risk was evaluated for conditions that have well-accepted epidemiologically observed elevated cardiovascular risks (*29*, *30*). In PRADA, in breast cancer survivors 1 year after follow-up, the predicted cardiovascular risks were 13.5 versus 4.9% from our earlier analysis of a subset of matched women from the Fenland study (*31*). In ARIC visit 3, not used for training or validation ($n = 11{,}301$), shown in Fig. 5, the protein model correctly detected elevated risks in people with prior events, with history of cancer, currently smoking, with diabetes, and with rheumatoid arthritis, although the latter group was too small to detect the increased observed event rates known epidemiologically (*32*), because there were only two events in the 39 participants with RA. In the subgroup of individuals without

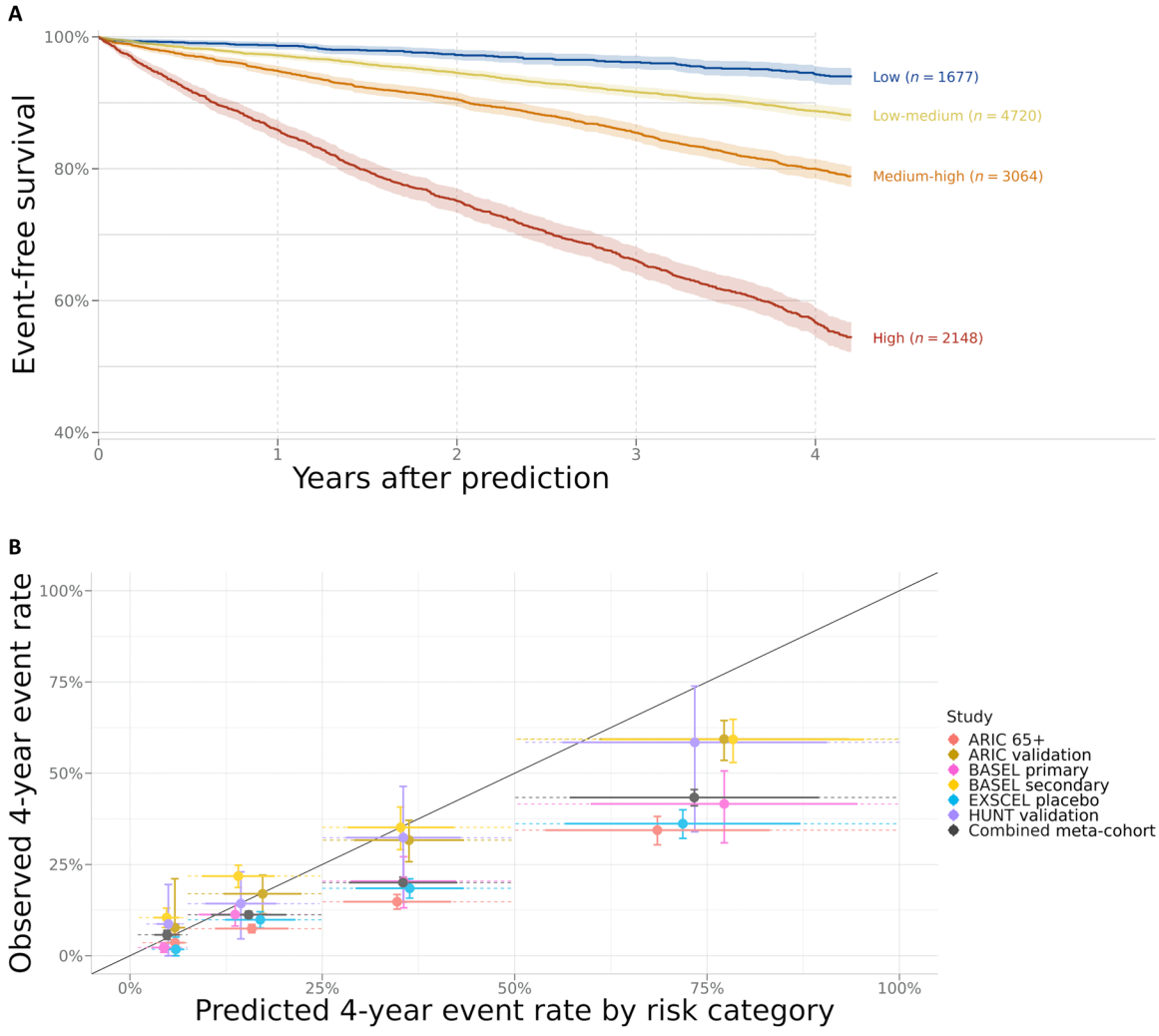

Fig. 3. Stratification and calibration of protein predictions in defined risk categories in the validation meta-cohort for all studies with 4-year follow-up. (**A**) Kaplan-Meier survival curves in the merged meta-cohort ($n = 11,609$) illustrating stratification between the four risk categories, median plus 95% confidence intervals in shading. (**B**) Predicted risks [mean, SD (solid line) and range (dashed line)] and observed risks in individual studies and in the meta-cohort for each risk category. Observed rates are in Kaplan-Meier estimates to account for censoring.

diagnosed or treated hypertension, people with above-median systolic blood pressure at the time of ARIC visit 3 had both higher observed (+2.2%, $P < 0.01$) and protein-predicted (+1.2%, $P < 0.01$) event rates than participants in the same subgroup with equal-to- or below-median systolic blood pressure; similarly, participants without diagnosed or treated hyperlipidemia but with above-median total cholesterol/HDL ratio had higher observed (+3.0%, $P < 0.01$) and protein-predicted (+2.3%, $P < 0.01$) event rates within 4 years. Across all these conditions, there was a significant relationship ($r = 0.83$, $P < 0.04$) between the observed event rate differences in cases versus controls and protein-predicted event rate differences. This proportionality relation is shown in fig. S2.

DISCUSSION

Despite the development of lipid lowering (*2*, *5*, *8*), anti-inflammatory (*4*, *6*), antithrombotic (*33*), dual antiplatelet (*34*), and antidiabetic (*7*, *35*) treatments, cardiovascular disease remains the leading cause of death and disability worldwide (*12*). This suggests that there is considerable scope for both precision use of existing medicines and the continuing development of further mechanisms of risk reduction.

Here, we reported the results from a large-scale prospectively designed surrogate endpoint discovery and validation study based on proteomics. In 32,130 samples from 22,849 participants in nine clinical studies with >170,000 participant-years of follow-up and more than 150 million individual protein measurements, we trained and validated a 27-protein prognostic cardiovascular model and evaluated it for the technical features of surrogacy as defined by the FDA biomarker qualification framework (*16*) in subsequent paragraphs.

Whereas the potential benefits of a surrogate end point are high, including accelerating patient access to effective drugs and terminating ineffective or unsafe drugs before large-scale exposure, the consequences of false results are also serious and include the approval of an ineffective or unsafe drug. The weight of evidence required for qualification of a surrogate end point is therefore high.

Biologic plausibility for the 27-protein model is high. Proteins regulate biological processes and integrate the effects of genes with

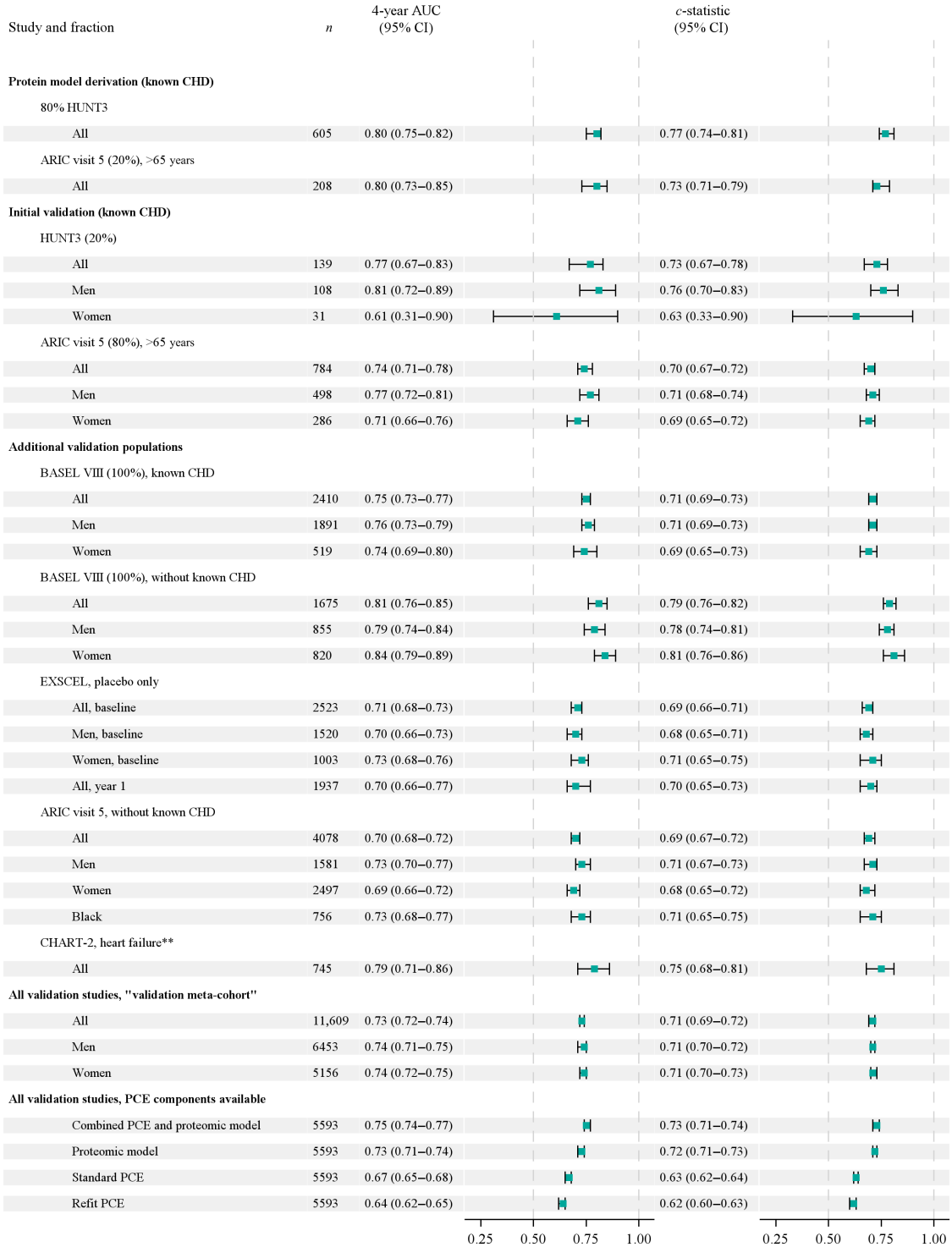

Fig. 4. Consistency of discrimination performance across groups. The forest plot shows the value of the 4-year AUC and the time-independent *c*-statistic across subgroups; means and 95% confidence intervals are shown. **For all studies, 4-year outcome data were available except for CHART-2, which is therefore not included in the meta-cohort. For that study alone, the 2-year AUC is shown instead of the 4-year AUC. All *c*-statistics used all available follow-up data described in Table 1. In addition to the combined-group comparisons with the clinical risk factors shown above, in the initial ARIC and HUNT validation studies (in participants where the component data were available), the clinical risk factor model had a *c*-statistic of 0.63 and 0.56, and the PCE had a *c*-statistic of 0.64 in the over-65 participants without known disease (where application of the equation is valid) in ARIC.

Table 3. Longitudinal sensitivity to change in risk for the 27-protein model and common individual prognostic biomarkers. Percentages for the protein model are of the change in absolute risk predictions within individuals (or the difference in risk increment across groups for randomized trials) from paired samples in the same participants over intervals of 1 or 2 years. "Event" and "no event" participants are from combined randomized groups. Upward arrows represent statistically significant effects in the direction of adverse risk and downward arrows in the direction of reduced risk (for biomarkers, $P < 0.05$ corrected for six multiple comparisons). NS, not statistically significant; ACCORD, the Action to Control Cardiovascular Risk in Diabetes; DiRECT, Diabetes Remission Clinical Trial; PRADA, Prevention of Cardiac Dysfunction During Adjuvant Breast Cancer Therapy. Biomarkers were measured from the SomaScan assay; correlations with standard laboratory measures (except myeloperoxidase) are shown in table S9.

	Condition	27 Proteins, absolute change in risk		CRP	Cystatin -C	GDF-15	Myelo-peroxidase	NT-proBNP	Troponin T
Expected Adverse Change	Approaching an event, 1-year change vs. no event (EXSCEL)	+2.9%	↑	↑	NS	NS	NS	↑	↑
	Approaching an event, 2-year change vs. no event (ACCORD)	+6.0%	↑	NS	NS	NS	NS	NS	NS
	Anthracycline chemotherapy, 3-month change (PRADA)	+6.2%	↑	NS	↑	↑	↑	NS	NS
Expected Neutral Change	Intensive diabetic control, vs. standard control (ACCORD)	NS	NS	↑	↑	NS	↑	NS	NS
	Angiotensin receptor blocker vs. placebo (PRADA)	NS	NS	NS	NS	NS	NS	NS	NS
	Beta blocker in chemotherapy vs. placebo (PRADA)	NS	NS	NS	↓	↓	NS	NS	NS
Expected Beneficial Change	Exenatide, within-subject change vs. placebo (EXSCEL)	-1.5%	↓	↓	NS	↓	NS	↓	NS
	Diet in diabetes in one year vs. standard diet (DiRECT)	-6.7%	↓	↓	NS	↓	NS	↑	NS

the environment, age, comorbidities, lifestyle, and drugs (*11*, *31*); proteins are highly mutable as conditions change (*18*, *19*, *36*, *37*), encode demographic information (*31*), and are the targets of 95% of all known drugs (*38*). All 27 proteins in the model related to biologic processes that are generally reported to be risk prognostic for cardiovascular outcomes. In addition, 16 of the 27 model proteins were included in a proteo-genomic database (*21*), and 12 of these proteins (75%) had at least one genetic causal association with cardiovascular disease or its risk factors.

The prognostic performance of the 27-protein model in independent validation cohorts was superior to clinical models in every standard metric: dynamic range (about twofold greater), discrimination (about 0.1 greater AUC), and risk reclassification (an event NRI of more than 40%). It was also robust to sex, geographic, racial/ethnic, demographic, and morbidity differences across studies and does not use race as a variable. In addition, the reliable identification of individuals with an observed event rate of >50% and a median time to event of 18 months is of clinical and economic relevance.

Combining the protein model with the standard PCE or the optimized clinical model added very little discrimination performance (0.01 gain in *c*-statistic with overlapping confidence intervals versus the protein model alone). This implied that the protein model had already encoded the biological information conferred by the clinical model.

The protein model demonstrated universally concordant responses to known changes in risk when evaluated using paired samples from the same individuals. In people approaching events after the second sample in the EXSCEL and ACCORD trials, greater predicted incremental risk elevations heralded subsequent events. For aging, protein-predicted risk elevations of +2.25 and +4.89% within 1 year in the placebo group of the EXSCEL study and the control group in the DiRECT study and +7.5% in 2 years in the ACCORD study were observed. For drug toxicity, protein-predicted risks increased by +6.2% over about 3 months in the PRADA trial because left ventricular ejection fractions fell (*25*) during cardiotoxic anthracycline chemotherapy (*39*). For neutral effects, in the ACCORD trial of intensive glycemic control, in the fraction of the participants with paired samples that we studied (and without hypoglycemic episodes that affected the larger trial), observed event rates were unchanged across groups, as were the protein predictions. In PRADA, the use of neither angiotensin receptor blockers nor beta blockers had an observed long-term cardioprotective effect (*26*), again matching the protein-predicted lack of benefit for those groups. For protein-predicted improvements, in EXSCEL, at first glance, the −1.49% ($P < 0.01$) protein-predicted benefit of exenatide, even when discounted for the shorter study follow-up than the prediction period, appears somewhat more beneficial and statistically significant than the observed 0.6% ($P = 0.06$) absolute reduction in cardiovascular event rates in 3.2 years (*27*). However, declining adherence to exenatide over the median follow-up of 3.2 years may have reduced the observed benefits (*27*), and our study only evaluated changes within the first year. This, coupled with the finding that four of seven outcome studies with other GLP-1 RAs showed

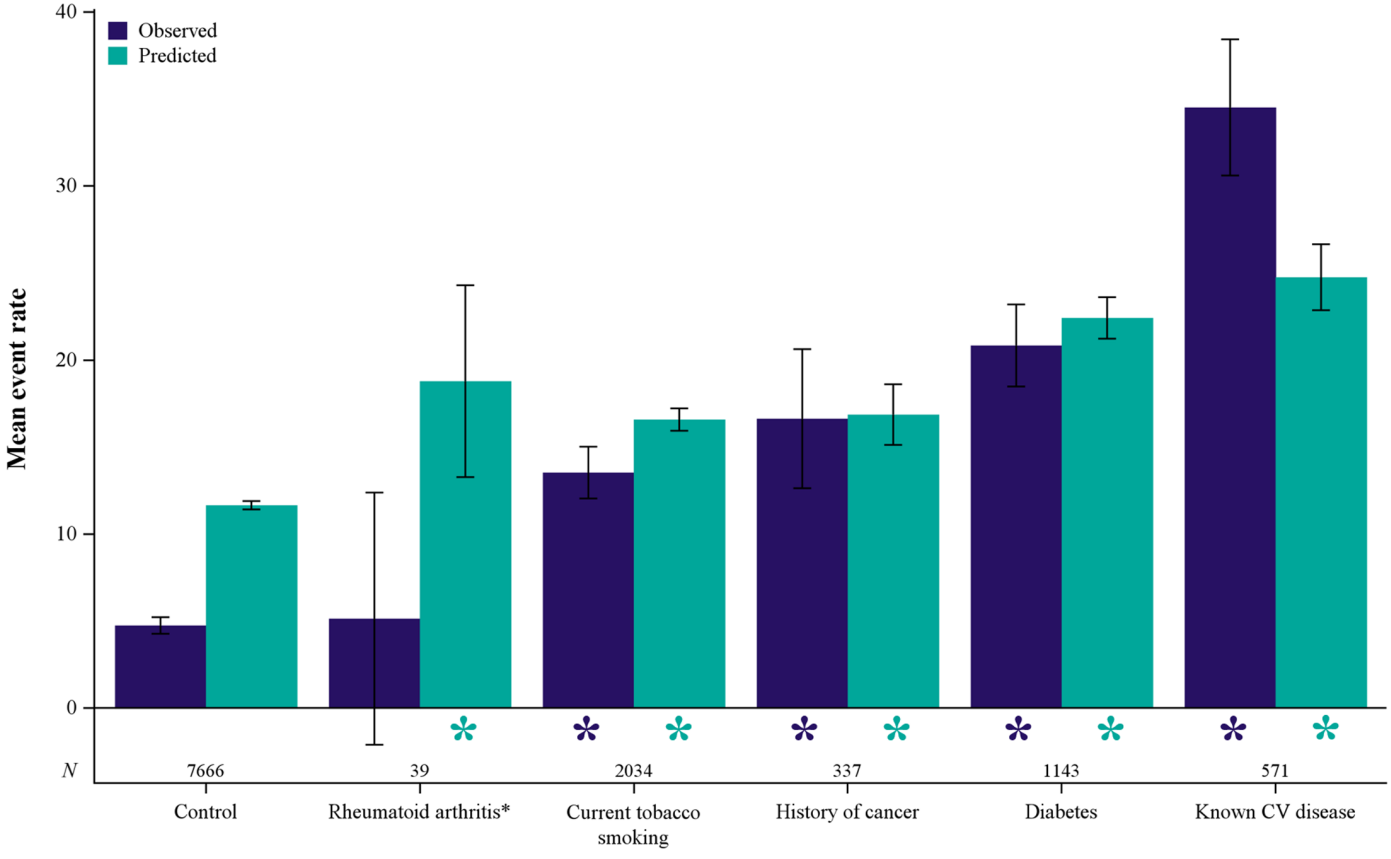

Fig. 5. Protein cardiovascular risk prediction and observed event rates under conditions with known higher cardiovascular risk. Analysis of 11,301participants in ARIC visit 3, which was not used for model training or model validation. Except for the control group, participants may be in more than one group. Groups with a dark blue asterisk have observed event rates significantly higher than the control group ($P < 0.05$, Dunnett's correction for multiple comparisons). Groups with a teal asterisk have predicted event rates significantly higher than the control group ($P < 0.05$, Dunnett's correction for multiple comparisons). *The rheumatoid arthritis group has 39 participants and only two (5.1%) cardiovascular events. Note that the overprediction in the controls is to be expected because they do not have any of the higher-risk conditions that the model was trained on.

cardioprotective superiority over placebo (*35*), suggests that a real exenatide benefit in EXSCEL was detected by the protein model and that the significance ($P = 0.06$) of observed changes in outcomes was likely obscured by an insufficiently powered trial with a hierarchical statistical design, testing first for noninferiority. The predicted improved cardiovascular risk in DiRECT is also likely to reflect a true benefit. Although 4-year event rates are not yet available, events were numerically fewer over the first 2 years (*28*), and reanalysis of the Look-AHEAD diabetes trial showed that individuals who lost ≥10% of their bodyweight had proportionately 21% lower incidence of adverse cardiovascular outcomes (*40*). The same 27-protein model has also been evaluated in more than 800 patients presenting with coronavirus disease 2019; predicted risks increased rapidly during active infection and declined during recovery, and it predicted all-cause mortality within 28 days with an AUC of 0.83 (*41*).

We also evaluated the model cross-sectionally under conditions known to have elevations in cardiovascular event rates. Figure 5 introduces such tests in more than 11,000 participants from ARIC visit 3 for four further risk-changing mechanisms not explicitly included in the longitudinal studies: currently smoking, presence of diabetes, cancer history, and rheumatoid arthritis, plus known cardiovascular disease, all consistent with epidemiologic observations (*29*, *30*, *42*). In addition, the model correctly predicted elevated risks of undiagnosed, untreated above-median systolic blood pressure and undiagnosed, untreated above-median hyperlipidemia in this population, and last, consistent with epidemiology and the cancer survival evidence from ARIC, the predicted proteomic risks for cancer survivors in PRADA at long-term follow-up were also about threefold higher than in matched women without cancer (*31*).

The rationale for testing the putative surrogate's responses to multiple mechanistic drivers or remediations of cardiovascular risk is to test empirically the concept of "universality" (that the surrogate will respond to the net change in outcome that results from any mechanism). Universality was added to the surrogate endpoint evidentiary framework as an alternative to historically requiring causality (*43*) because under complex conditions, it is not known. Although some of the proteins in the risk model may be causal mediators, it is unlikely that the protein model captured all such causal factors for cardiovascular disease.

A multivariate model also appears to be a more universal risk integrator than individual cardiovascular biomarkers, none of which were as reliably sensitive to change in risk across mechanisms as the 27-protein model (as shown in Table 3). For example, NTproBNP is known to increase during dietary weight loss (*44*), and this was observed in the DiRECT cohort. If this protein was viewed as a univariate surrogate, then this would wrongly have suggested increased risk

from diet and resolution of diabetes. Although this protein is 1 of the 27 proteins in the multivariate model, the net multivariate prediction was correct.

Our study and findings have some important limitations. First, the calibration of the model trained in secondary populations overpredicts risk at the high end of the range in primary or mixed populations with lower event rates (similar to reported calibration errors for the PCE and Framingham risk scores) (*45*), although predictions in the lower and mid-range, where accuracy is most important, were reasonably aligned with observed event rates (Figs. 2 and 3) despite a fivefold difference in these rates across studies. Improvement by post hoc recalibration (*46*) could be obtained (fig. S1). Second, although an AUC of 0.73 in validation cannot be viewed as anywhere approaching perfection, it is unknowable what the achievable "ceiling" for a perfect prognostic test is, because changes in environmental exposure, medication, and behavior after the sample is taken, but within the 4-year prediction period, can easily change the risks. Third, although the sensitivity of protein models to universally detect changes (or lack of changes) in cardiovascular risk from any mechanism has been tested across 15 different mechanisms, there are notable drug class omissions, including PCSK9 inhibitors and SGLT2i. In addition, the cross-sectional epidemiologic evaluations of risks from different diseases in Fig. 5 do not represent a fully adjusted multiparametric comparison with existing risk factors. Last, the determination as to whether the weight of evidence is yet sufficient for qualification as a drug development tool for any category of surrogate validation has not yet been made by the FDA.

In conclusion, these findings suggest that agnostic, multidimensional machine learning, applied to a highly multiplexed information source and without favoring any historically used biomarkers, is a useful approach toward enhancing risk model performance. This study provides initial evidence that multiprotein models, exemplified by our 27-protein model, may provide a much sought-after universal cardiovascular surrogate end point.

MATERIALS AND METHODS

Study design

The proteomic program (Fig. 1) was initiated in secondary populations, because they have higher event rates and are likely more enriched for relevant biologic mechanisms of risk than healthier groups. Having met predefined performance metrics, the program continued to validate the model and to extend the intended use population to other higher-risk populations, all of which are known to be difficult to stratify (bottom left side of Fig. 1). Specifically excluded was primary prevention without known drivers of elevated risk. Then, the sensitivity of the model to within-participant change in risks from multiple mechanisms was tested using longitudinal blood samples (bottom center of Fig. 1). Last, the model was tested across groups with different epidemiologically observed elevated risks, again across multiple mechanisms (bottom right side of Fig. 1). Each study evaluation was predefined in a statistical analysis plan; technical reports of the results were documented, and these were all filed to an auditable regulatory filing system (Arena Inc.).

Definition of the composite outcome

The outcome in this study was defined as the first event, subsequent to the blood sample, of myocardial infarction, stroke, heart failure hospitalization, or all-cause death. This composite end point used in our prior studies (*18*–*20*, *31*) includes a broad range of events important to patients, their providers, and in clinical trials (*18*, *47*, *48*). We chose all-cause death because of mounting evidence that clinical adjudication of cardiovascular from noncardiovascular causes tends to be inaccurate (*49*) and because clinical drug safety evaluations, for which a surrogate will be used, should be more encompassing rather than more exclusive. Revascularization procedures were excluded from the composite because of the contemporary changing frequency of nonacute procedures (*50*, *51*) and because of likely geographic variation. Our inclusion of heart failure alongside atherosclerotic events resembles the "global" (*48*) or "general" (*47*) cardiovascular outcome used by ARIC and Framingham investigators, respectively. Combining such atherosclerotic and heart failure events is facilitated by their clinical predictors being virtually identical (*52*).

Proteomic platform

The SomaScan assay (*53*) and its performance characteristics (*54*) have been described previously. It used DNA-based binding reagents (modified aptamers) (*55*) to quantify the availability of binding epitopes on plasma proteins for about 5000 proteins, with high specificity and limits of detection largely comparable to antibody-based assays (tables S8 and S9). Briefly, the SomaScan assay started as a mix of thousands of fluorophore-labeled SOMAmer reagents immobilized on streptavidin-coated beads and incubated with 55 μl of EDTA plasma. Samples were run at three different dilutions to expand dynamic range to about 10 logs. After a series of washing steps and the use of a polyanionic competitor to negate nonspecific binding and a second capture step, SOMAmer reagents were hybridized to complementary sequences on a DNA microarray chip and quantified by fluorescence, which was related to the relative availability of the three-dimensional shape-charge epitope on each protein in the original sample. This integrates each protein's abundance, shape, charge, and availability of the binding epitope.

Statistical methods for training and validation of the 27-protein model

Univariate Cox proportional hazards models were designed for proteins associated with the composite outcome and assessed for all individual proteins in the training set [80% of HUNT3 (*56*) (Norway) and 20% of ARIC visit 5 (*48*) (United States)], in participants with known cardiovascular disease. One hundred and forty-four of the top 400 most significant ($P < 0.1$, false discovery rate adjusted) proteins in each univariate list were common to both training datasets. These were input into a Cox regularized regression model using 10-fold cross-validation in the training set and penalized as a least absolute shrinkage and selection operator (LASSO) model (*18*) for the purpose of selecting a stable set of informative proteins. The resulting 27 proteins with nonzero coefficients from the cross-validated LASSO model using the minimum lambda (0.3) were then used to train a parametric AFT final model to provide additional time-varying flexibility for a model predictive of absolute risk. Four-year risk of a cardiovascular disease event and four categorical ranges (derived to have nonoverlapping confidence intervals and simple-to-communicate multiples of risk) of predicted event rates were the outputs.

Performance metrics

For prognosis in survival models, the most commonly evaluated metrics are for discrimination: the *c*-statistic (the ability of a model to discriminate between random pairs of individuals with different time

to event) or the AUC (the ability of the model to dichotomize the population into two classes at a defined time point). Accordingly, we used these measures as metrics, especially for testing consistency across populations and as comparators to the optimized clinical model. However, because neither of these assessments fully reflects the use of a surrogate with a continuous likelihood output in medical practice (*22*) or in clinical trials, we also placed emphasis on the dynamic range of these predictions.

Comparator clinical model

In the absence of clinical comparators previously validated for the composite cardiovascular end point, time horizon, and study populations, we selected the same measures used by the Framingham investigators to create a model for high-risk populations (*57*) and the 2013 ACC/AHA PCE (*58*): age, sex, race, total cholesterol, HDL cholesterol, blood pressure, diabetes, and smoking. We retrained their coefficients in the same populations used for training the protein model, as we have done previously (*18*) and as is necessary to calculate an NRI. Comparative performance was assessed by discrimination (*c*-statistic and 4-year receiver operating characteristic AUC) and by the NRI in the first two validation datasets (HUNT3 and ARIC visit 5). The refit clinical model and the standard PCE were also applied to 5593 participants in the meta-cohort for which all the components were available.

Cardiovascular risk under conditions with epidemiologically elevated event rates

For the diseases and conditions known to be associated with elevated cardiovascular event rates shown in Fig. 5, 11,301 participants' samples and data from ARIC visit 3 were used in a series of case-control designs; each case [current diabetes (n = 1143), history of cancer (n = 337), rheumatoid arthritis (n = 39), known cardiovascular disease (n = 571), and current smokers (n = 2034)] was compared against the remainder of the cohort without those conditions (n = 7666). All were around the same age when enrolled, and some participants with more than one condition were used in more than one group. In a second approach, individuals with diagnosed hypertension or diagnosed hyperlipidemia were eliminated from either group. Then, each group was split at the median measurement (elevated systolic blood pressure and high total cholesterol/HDL ratio, respectively), and the predicted and observed event rates for individuals with directionally adverse measures (n = 2907 for systolic blood pressure and n = 4182 for lipids) were counted as cases and directionally beneficial measures (n = 3041 for systolic blood pressure and n = 4176 for lipids) as controls. The predicted and observed case-control differences were statistically evaluated. For comparison with breast cancer in PRADA, which did not have noncancer controls, we used prior results from women in Fenland matched for age and cardiovascular history (n = 500).

Statistical analysis

The median reproducibility of the 27-protein model in nine replicates of the same 10 samples was 4.98% (table S10). Specificity testing for each of the 27 modified aptamer reagents used in the model is shown in table S11, and interference testing showing no issues for common substances relevant to cardiovascular disease is in tables S12 and S13.

The final model was initially validated in study subsets, still with known cardiovascular disease and high event rates: 20% (139 participants) of HUNT3 and 80% (784 participants) of ARIC visit 5 not used in model training. Subsequent independent validation was performed in datasets with populations of varying morbidities: BASEL VIII (*59*) (Switzerland, suspicion of chronic coronary syndromes), the placebo group of EXSCEL (*27*) (35 countries, type 2 diabetes), ARIC visit 5 (United States, elderly without known disease), and CHART-2 (*60*) (Japan, heart failure). Participants in EXSCEL assigned to placebo but who received a GLP-1 RA (n = 88) or SGLT2i (n = 169) treatments after baseline were excluded from this validation.

Participants from the six validation datasets with 4-year follow-up shown in Fig. 1 (n = 11,609) were pooled into a single meta-cohort. Baseline characteristics of each of the seven cohorts for which discrimination metrics were calculated are shown in Table 1, and additional description of each cohort is provided in section S1. Where the *c*-statistic and AUC are calculated for these datasets, the variability of these metrics is also derived using 95% confidence intervals from 200 bootstraps of each respective dataset.

Longitudinal sensitivity to change was assessed by comparing the mean change of protein-predicted cardiovascular disease risk for each respective group using one-tailed paired t testing. Parametric testing was deemed sufficient because of distribution of changes among, typically, hundreds of paired samples. Additional paired t testing was applied to measured abundances of six chosen biomarkers, with those results Bonferroni-adjusted for the six comparisons, and results considered significant if $P < 0.05$. For the cross-sectional case-control analyses in Fig. 5, t tests with Dunnett's correction for multiple comparisons were used. Analysis was performed using R version 4.1.0, with model training and analysis using the following packages and respective versions: pROC version 1.17.0.1, survival version 3.2-11, glmnet version 4.1-1, and rms version 6.2-0.

SUPPLEMENTARY MATERIALS

www.science.org/doi/10.1126/scitranslmed.abj9625
Materials and Methods
Figs. S1 and S2
Tables S1 to S13
MDAR Reproducibility Checklist
References (*61*–*110*)

View/request a protocol for this paper from Bio-protocol.

REFERENCES AND NOTES

1. U.S. Food and Drug Administration, "Innovation or stagnation; challenge and opportunity on the critical path to new products" (U.S. Food and Drug Administration, 2004).
2. D. L. Bhatt, P. G. Steg, M. Miller, E. A. Brinton, T. A. Jacobson, S. B. Ketchum, R. T. Doyle Jr., R. A. Juliano, L. Jiao, C. Granowitz, J. C. Tardif, C. M. Ballantyne; REDUCE-IT Investigators, Cardiovascular risk reduction with icosapent ethyl for hypertriglyceridemia. *N. Engl. J. Med.* **380**, 11–22 (2019).
3. D. Fitchett, S. E. Inzucchi, C. P. Cannon, D. K. McGuire, B. M. Scirica, O. E. Johansen, S. Sambevski, S. Kaspers, E. Pfarr, J. T. George, B. Zinman, Empagliflozin reduced mortality and hospitalization for heart failure across the spectrum of cardiovascular risk in the EMPA-REG OUTCOME trial. *Circulation* **139**, 1384–1395 (2019).
4. P. M. Ridker, B. M. Everett, T. Thuren, J. G. MacFadyen, W. H. Chang, C. Ballantyne, F. Fonseca, J. Nicolau, W. Koenig, S. D. Anker, J. J. P. Kastelein, J. H. Cornel, P. Pais, D. Pella, J. Genest, R. Cifkova, A. Lorenzatti, T. Forster, Z. Kobalava, L. Vida-Simiti, M. Flather, H. Shimokawa, H. Ogawa, M. Dellborg, P. R. F. Rossi, R. P. T. Troquay, P. Libby, R. J. Glynn; CANTOS Trial Group, Antiinflammatory therapy with canakinumab for atherosclerotic disease. *N. Engl. J. Med.* **377**, 1119–1131 (2017).
5. M. S. Sabatine, R. P. Giugliano, A. C. Keech, N. Honarpour, S. D. Wiviott, S. A. Murphy, J. F. Kuder, H. Wang, T. Liu, S. M. Wasserman, P. S. Sever, T. R. Pedersen; FOURIER Steering Committee and Investigators, Evolocumab and clinical outcomes in patients with cardiovascular disease. *N. Engl. J. Med.* **376**, 1713–1722 (2017).
6. J. C. Tardif, S. Kouz, D. D. Waters, O. F. Bertrand, R. Diaz, A. P. Maggioni, F. J. Pinto, R. Ibrahim, H. Gamra, G. S. Kiwan, C. Berry, J. Lopez-Sendon, P. Ostadal, W. Koenig, D. Angoulvant, J. C. Gregoire, M. A. Lavoie, M. P. Dube, D. Rhainds, M. Provencher, L. Blondeau, A. Orfanos,

P. L. L'Allier, M. C. Guertin, F. Roubille, Efficacy and safety of low-dose colchicine after myocardial infarction. *N. Engl. J. Med.* **381**, 2497–2505 (2019).
7. S. L. Zheng, A. J. Roddick, R. Aghar-Jaffar, M. J. Shun-Shin, D. Francis, N. Oliver, K. Meeran, Association between use of sodium-glucose cotransporter 2 inhibitors, glucagon-like peptide 1 agonists, and dipeptidyl peptidase 4 inhibitors with all-cause mortality in patients with type 2 diabetes: A systematic review and meta-analysis. *JAMA* **319**, 1580–1591 (2018).
8. R. S. Rosenson, L. J. Burgess, C. F. Ebenbichler, S. J. Baum, E. S. G. Stroes, S. Ali, N. Khilla, R. Hamlin, R. Pordy, Y. Dong, V. Son, D. Gaudet, Evinacumab in patients with refractory hypercholesterolemia. *N. Engl. J. Med.* **383**, 2307–2319 (2020).
9. P. J. Barter, M. Caulfield, M. Eriksson, S. M. Grundy, J. J. Kastelein, M. Komajda, J. Lopez-Sendon, L. Mosca, J. C. Tardif, D. D. Waters, C. L. Shear, J. H. Revkin, K. A. Buhr, M. R. Fisher, A. R. Tall, B. Brewer; ILLUMINATE Investigators, Effects of torcetrapib in patients at high risk for coronary events. *N. Engl. J. Med.* **357**, 2109–2122 (2007).
10. T. Regan, FDA mea culpa part of cautionary tale. *Am. J. Manag. Care* **SP7**, 242–243 (2013).
11. P. Ganz, R. Deo, R. F. Dubin, Proteomics for personalized cardiovascular risk assessment: In pursuit of the holy grail. *Eur. Heart J.* **41**, 4008–4010 (2020).
12. D. S. Dhindsa, P. B. Sandesara, M. D. Shapiro, N. D. Wong, The evolving understanding and approach to residual cardiovascular risk management. *Front. Cardiovasc. Med.* **7**, 88 (2020).
13. R. L. Coleman, R. J. Stevens, R. Retnakaran, R. R. Holman, Framingham, SCORE, and DECODE risk equations do not provide reliable cardiovascular risk estimates in type 2 diabetes. *Diabetes Care* **30**, 1292–1293 (2007).
14. J. E. Dalton, M. B. Rothberg, N. V. Dawson, N. I. Krieger, D. A. Zidar, A. T. Perzynski, Failure of traditional risk factors to adequately predict cardiovascular events in older populations. *J. Am. Geriatr. Soc.* **68**, 754–761 (2020).
15. M. Nanna, M. W. Rich, "Failure of traditional risk factors to adequately predict cardiovascular events in older populations" (Journal of the American College of Cardiology, 2020).
16. C. D. Lathia, D. Amakye, W. Dai, C. Girman, S. Madani, J. Mayne, P. MacCarthy, P. Pertel, L. Seman, A. Stoch, P. Tarantino, C. Webster, S. Williams, J. A. Wagner, The value, qualification, and regulatory use of surrogate end points in drug development. *Clin. Pharmacol. Ther.* **86**, 32–43 (2009).
17. D. M. Lloyd-Jones, L. T. Braun, C. E. Ndumele, S. C. Smith Jr., L. S. Sperling, S. S. Virani, R. S. Blumenthal, Use of risk assessment tools to guide decision-making in the primary prevention of atherosclerotic cardiovascular disease: A special report from the American Heart Association and American College of Cardiology. *Circulation* **139**, e1162–e1177 (2019).
18. P. Ganz, B. Heidecker, K. Hveem, C. Jonasson, S. Kato, M. R. Segal, D. G. Sterling, S. A. Williams, Development and validation of a protein-based risk score for cardiovascular outcomes among patients with stable coronary heart disease. *JAMA* **315**, 2532–2541 (2016).
19. S. A. Williams, A. C. Murthy, R. K. DeLisle, C. Hyde, A. Malarstig, R. Ostroff, S. J. Weiss, M. R. Segal, P. Ganz, Improving assessment of drug safety through proteomics: Early detection and mechanistic characterization of the unforeseen harmful effects of torcetrapib. *Circulation* **137**, 999–1010 (2018).
20. J. Yang, E. N. Brody, A. C. Murthy, R. E. Mehler, S. J. Weiss, R. K. DeLisle, R. Ostroff, S. A. Williams, P. Ganz, Impact of kidney function on the blood proteome and on protein cardiovascular risk biomarkers in patients with stable coronary heart disease. *J. Am. Heart Assoc.* **9**, e016463 (2020).
21. J. Zheng, V. Haberland, D. Baird, V. Walker, P. C. Haycock, M. R. Hurle, A. Gutteridge, P. Erola, Y. Liu, S. Luo, J. Robinson, T. G. Richardson, J. R. Staley, B. Elsworth, S. Burgess, B. B. Sun, J. Danesh, H. Runz, J. C. Maranville, H. M. Martin, J. Yarmolinsky, C. Laurin, M. V. Holmes, J. Z. Liu, K. Estrada, R. Santos, L. McCarthy, D. Waterworth, M. R. Nelson, G. D. Smith, A. S. Butterworth, G. Hemani, R. A. Scott, T. R. Gaunt, Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat. Genet.* **52**, 1122–1131 (2020).
22. M. S. Sabatine, Using aptamer-based technology to probe the plasma proteome for cardiovascular disease prediction. *JAMA* **315**, 2525–2526 (2016).
23. N. R. Cook, Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation* **115**, 928–935 (2007).
24. Action to Control Cardiovascular Risk in Diabetes Study Group, H. C. Gerstein, M. E. Miller, R. P. Byington, D. C. Goff Jr., J. T. Bigger, J. B. Buse, W. C. Cushman, S. Genuth, F. Ismail-Beigi, R. H. Grimm Jr., J. L. Probstfield, D. G. Simons-Morton, W. T. Friedewald, Effects of intensive glucose lowering in type 2 diabetes. *N. Engl. J. Med.* **358**, 2545–2559 (2008).
25. G. Gulati, S. L. Heck, A. H. Ree, P. Hoffmann, J. Schulz-Menger, M. W. Fagerland, B. Gravdehaug, F. von Knobelsdorff-Brenkenhoff, A. Bratland, T. H. Storas, T. A. Hagve, H. Rosjo, K. Steine, J. Geisler, T. Omland, Prevention of cardiac dysfunction during adjuvant breast cancer therapy (PRADA): A 2 × 2 factorial, randomized, placebo-controlled, double-blind clinical trial of candesartan and metoprolol. *Eur. Heart J.* **37**, 1671–1680 (2016).
26. S. L. Heck, A. Mecinaj, A. H. Ree, P. Hoffmann, J. E. Schulz-Menger, M. W. Fagerland, B. Gravdehaug, H. Rosjo, K. Steine, J. Geisler, G. Gulati, T. Omland, Prevention of cardiac dysfunction during adjuvant breast cancer therapy (PRADA): Extended follow-up of a 2x2 factorial, randomized, placebo-controlled, double-blind clinical trial of candesartan and metoprolol. *Circulation* **143**, 2431–2440 (2021).
27. R. R. Holman, M. A. Bethel, R. J. Mentz, V. P. Thompson, Y. Lokhnygina, J. B. Buse, J. C. Chan, J. Choi, S. M. Gustavson, N. Iqbal, A. P. Maggioni, S. P. Marso, P. Ohman, N. J. Pagidipati, N. Poulter, A. Ramachandran, B. Zinman, A. F. Hernandez; EXSCEL Study Group, Effects of once-weekly exenatide on cardiovascular outcomes in type 2 diabetes. *N. Engl. J. Med.* **377**, 1228–1239 (2017).
28. M. E. J. Lean, W. S. Leslie, A. C. Barnes, N. Brosnahan, G. Thom, L. McCombie, C. Peters, S. Zhyzhneuskaya, A. Al-Mrabeh, K. G. Hollingsworth, A. M. Rodrigues, L. Rehackova, A. J. Adamson, F. F. Sniehotta, J. C. Mathers, H. M. Ross, Y. McIlvenna, P. Welsh, S. Kean, I. Ford, A. McConnachie, C. M. Messow, N. Sattar, R. Taylor, Durability of a primary care-led weight-management intervention for remission of type 2 diabetes: 2-year results of the DiRECT open-label, cluster-randomised trial. *Lancet Diabetes Endocrinol.* **7**, 344–355 (2019).
29. M. T. Nurmohamed, Cardiovascular risk in rheumatoid arthritis. *Autoimmun. Rev.* **8**, 663–667 (2009).
30. H. Strongman, S. Gadd, A. Matthews, K. E. Mansfield, S. Stanway, A. R. Lyon, I. Dos-Santos-Silva, L. Smeeth, K. Bhaskaran, Medium and long-term risks of specific cardiovascular diseases in survivors of 20 adult cancers: A population-based cohort study using multiple linked UK electronic health records databases. *Lancet* **394**, 1041–1054 (2019).
31. S. A. Williams, M. Kivimaki, C. Langenberg, A. D. Hingorani, J. P. Casas, C. Bouchard, C. Jonasson, M. A. Sarzynski, M. J. Shipley, L. Alexander, J. Ash, T. Bauer, J. Chadwick, G. Datta, R. K. DeLisle, Y. Hagar, M. Hinterberg, R. Ostroff, S. Weiss, P. Ganz, N. J. Wareham, Plasma protein patterns as comprehensive indicators of health. *Nat. Med.* **25**, 1851–1857 (2019).
32. S. S. Dhawan, A. A. Quyyumi, Rheumatoid arthritis and cardiovascular disease. *Curr. Atheroscler. Rep.* **10**, 128–133 (2008).
33. J. W. Eikelboom, S. J. Connolly, J. Bosch, G. R. Dagenais, R. G. Hart, O. Shestakovska, R. Diaz, M. Alings, E. M. Lonn, S. S. Anand, P. Widimsky, M. Hori, A. Avezum, L. S. Piegas, K. R. H. Branch, J. Probstfield, D. L. Bhatt, J. Zhu, Y. Liang, A. P. Maggioni, P. Lopez-Jaramillo, M. O'Donnell, A. K. Kakkar, K. A. A. Fox, A. N. Parkhomenko, G. Ertl, S. Stork, M. Keltai, L. Ryden, N. Pogosova, A. L. Dans, F. Lanas, P. J. Commerford, C. Torp-Pedersen, T. J. Guzik, P. B. Verhamme, D. Vinereanu, J. H. Kim, A. M. Tonkin, B. S. Lewis, C. Felix, K. Yusoff, P. G. Steg, K. P. Metsarinne, N. C. Bruns, F. Misselwitz, E. Chen, D. Leong, S. Yusuf; COMPASS Investigators, Rivaroxaban with or without aspirin in stable cardiovascular disease. *N. Engl. J. Med.* **377**, 1319–1330 (2017).
34. M. P. Bonaca, D. L. Bhatt, M. Cohen, P. G. Steg, R. F. Storey, E. C. Jensen, G. Magnani, S. Bansilal, M. P. Fish, K. Im, O. Bengtsoon, T. O. Ophuis, A. Budaj, P. Theroux, M. Ruda, C. Hamm, S. Goto, J. Spinar, J. C. Nicolau, R. G. Kiss, S. A. Murphy, S. D. Wiviott, P. Held, E. Braunwald, M. S. Sabatine, Long-term use of ticagrelor in patients with prior myocardial infarction. *N. Engl. J. Med.* **372**, 1791–1800 (2015).
35. K. H. Sheahan, E. A. Wahlberg, M. P. Gilbert, An overview of GLP-1 agonists and recent cardiovascular outcomes trials. *Postgrad. Med. J.* **96**, 156–161 (2020).
36. B. Lehallier, D. Gate, N. Schaum, T. Nanasi, S. E. Lee, H. Yousef, P. Moran Losada, D. Berdnik, A. Keller, J. Verghese, S. Sathyan, C. Franceschi, S. Milman, N. Barzilai, T. Wyss-Coray, Undulating changes in human plasma proteome profiles across the lifespan. *Nat. Med.* **25**, 1843–1850 (2019).
37. E. Ferrannini, A. C. Murthy, Y. H. Lee, E. Muscelli, S. Weiss, R. M. Ostroff, N. Sattar, S. A. Williams, P. Ganz, Mechanisms of sodium-glucose cotransporter 2 inhibition: Insights from large-scale proteomics. *Diabetes Care* **43**, 2183–2189 (2020).
38. R. Santos, O. Ursu, A. Gaulton, A. P. Bento, R. S. Donadi, C. G. Bologa, A. Karlsson, B. Al-Lazikani, A. Hersey, T. I. Oprea, J. P. Overington, A comprehensive map of molecular drug targets. *Nat. Rev. Drug Discov.* **16**, 19–34 (2017).
39. D. Cardinale, F. Iacopo, C. M. Cipolla, Cardiotoxicity of anthracyclines. *Front. Cardiovasc. Med.* **7**, 26 (2020).
40. A. R. G. Look, E. W. Gregg, J. M. Jakicic, G. Blackburn, P. Bloomquist, G. A. Bray, J. M. Clark, M. Coday, J. M. Curtis, C. Egan, M. Evans, J. Foreyt, G. Foster, H. P. Hazuda, J. O. Hill, E. S. Horton, V. S. Hubbard, R. W. Jeffery, K. C. Johnson, A. E. Kitabchi, W. C. Knowler, A. Kriska, W. Lang, C. E. Lewis, M. G. Montez, D. M. Nathan, R. H. Neiberg, J. Patricio, A. Peters, X. Pi-Sunyer, H. Pownall, B. Redmon, J. Regensteiner, J. Rejeski, P. M. Ribisl, M. Safford, K. Stewart, D. Trence, T. A. Wadden, R. R. Wing, S. Z. Yanovski, Association of the magnitude of weight loss and changes in physical fitness with long-term cardiovascular disease outcomes in overweight or obese people with type 2 diabetes: A post-hoc analysis of the Look AHEAD randomised clinical trial. *Lancet Diabetes Endocrinol.* **4**, 913–921 (2016).
41. H. Y. e. a. Paterson C, Application of a 27-protein candidate cardiovascular surrogate endpoint to predict COVID-19 outcomes, and to track risk ascendancy and resolution. *BioRxViv* (2021).
42. L. S. Mehta, K. E. Watson, A. Barac, T. M. Beckie, V. Bittner, S. Cruz-Flores, S. Dent, L. Kondapalli, B. Ky, T. Okwuosa, I. L. Pina, A. S. Volgman; American Heart Association Cardiovascular Disease in Women and Special Populations Committee of the Council on

Clinical Cardiology; Council on Cardiovascular and Stroke Nursing; Council on Quality of Care and Outcomes Research, Cardiovascular disease and breast cancer: Where these entities intersect: A scientific statement from the american heart association. *Circulation* **137**, e30–e66 (2018).
43. T. R. Fleming, D. L. DeMets, Surrogate end points in clinical trials: Are we being misled? *Ann. Intern. Med.* **125**, 605–613 (1996).
44. D. Fedele, V. Bicchiega, A. Collo, F. Barutta, E. Pistone, G. Gruden, G. Bruno, Short term variation in NTproBNP after lifestyle intervention in severe obesity. *PLOS ONE* **12**, e0181212 (2017).
45. J. A. Damen, R. Pajouheshnia, P. Heus, K. G. M. Moons, J. B. Reitsma, R. Scholten, L. Hooft, T. P. A. Debray, Performance of the Framingham risk models and pooled cohort equations for predicting 10-year risk of cardiovascular disease: A systematic review and meta-analysis. *BMC Med.* **17**, 109 (2019).
46. L. Pennells, S. Kaptoge, A. Wood, M. Sweeting, X. Zhao, I. White, S. Burgess, P. Willeit, T. Bolton, K. G. M. Moons, Y. T. van der Schouw, R. Selmer, K. T. Khaw, V. Gudnason, G. Assmann, P. Amouyel, V. Salomaa, M. Kivimaki, B. G. Nordestgaard, M. J. Blaha, L. H. Kuller, H. Brenner, R. F. Gillum, C. Meisinger, I. Ford, M. W. Knuiman, A. Rosengren, D. A. Lawlor, H. Volzke, C. Cooper, A. M. Ibanez, E. Casiglia, J. Kauhanen, J. A. Cooper, B. Rodriguez, J. Sundstrom, E. Barrett-Connor, R. Dankner, P. J. Nietert, K. W. Davidson, R. B. Wallace, D. G. Blazer, C. Bjorkelund, C. Donfrancesco, H. M. Krumholz, A. Nissinen, B. R. Davis, S. Coady, P. H. Whincup, T. Jorgensen, P. Ducimetiere, M. Trevisan, G. Engstrom, C. J. Crespo, T. W. Meade, M. Visser, D. Kromhout, S. Kiechl, M. Daimon, J. F. Price, A. G. de la Camara, J. W. Jukema, B. Lamarche, A. Onat, L. A. Simons, M. Kavousi, Y. Ben-Shlomo, J. Gallacher, J. M. Dekker, H. Arima, N. Shara, R. W. Tipping, R. Roussel, E. J. Brunner, W. Koenig, M. Sakurai, J. Pavlovic, R. T. Gansevoort, D. Nagel, U. Goldbourt, E. L. M. Barr, L. Palmieri, I. Njolstad, S. Sato, W. M. M. Verschuren, C. V. Varghese, I. Graham, O. Onuma, P. Greenland, M. Woodward, M. Ezzati, B. M. Psaty, N. Sattar, R. Jackson, P. M. Ridker, N. R. Cook, R. B. D'Agostino, S. G. Thompson, J. Danesh, E. Di Angelantonio; Emerging Risk Factors Collaboration, Equalization of four cardiovascular risk algorithms after systematic recalibration: Individual-participant meta-analysis of 86 prospective studies. *Eur. Heart J.* **40**, 621–631 (2019).
47. R. B. D'Agostino Sr., R. S. Vasan, M. J. Pencina, P. A. Wolf, M. Cobain, J. M. Massaro, W. B. Kannel, General cardiovascular risk profile for use in primary care: The Framingham Heart study. *Circulation* **117**, 743–753 (2008).
48. A. Saeed, V. Nambi, W. Sun, S. S. Virani, G. E. Taffet, A. Deswal, E. Selvin, K. Matsushita, L. E. Wagenknecht, R. Hoogeveen, J. Coresh, J. A. de Lemos, C. M. Ballantyne, Short-term global cardiovascular disease risk prediction in older adults. *J. Am. Coll. Cardiol.* **71**, 2527–2536 (2018).
49. Z. H. Tseng, J. E. Olgin, E. Vittinghoff, P. C. Ursell, A. S. Kim, K. Sporer, C. Yeh, B. Colburn, N. M. Clark, R. Khan, A. P. Hart, E. Moffatt, Prospective countywide surveillance and autopsy characterization of sudden cardiac death: POST SCD study. *Circulation* **137**, 2689–2700 (2018).
50. D. J. Maron, J. S. Hochman, H. R. Reynolds, S. Bangalore, S. M. O'Brien, W. E. Boden, B. R. Chaitman, R. Senior, J. Lopez-Sendon, K. P. Alexander, R. D. Lopes, L. J. Shaw, J. S. Berger, J. D. Newman, M. S. Sidhu, S. G. Goodman, W. Ruzyllo, G. Gosselin, A. P. Maggioni, H. D. White, B. Bhargava, J. K. Min, G. B. J. Mancini, D. S. Berman, M. H. Picard, R. Y. Kwong, Z. A. Ali, D. B. Mark, J. A. Spertus, M. N. Krishnan, A. Elghamaz, N. Moorthy, W. A. Hueb, M. Demkow, K. Mavromatis, O. Bockeria, J. Peteiro, T. D. Miller, H. Szwed, R. Doerr, M. Keltai, J. B. Selvanayagam, P. G. Steg, C. Held, S. Kohsaka, S. Mavromichalis, R. Kirby, N. O. Jeffries, F. E. Harrell Jr., F. W. Rockhold, S. Broderick, T. B. Ferguson Jr., D. O. Williams, R. A. Harrington, G. W. Stone, Y. Rosenberg; ISCHEMIA Research Group, Initial invasive or conservative strategy for stable coronary disease. *N. Engl. J. Med.* **382**, 1395–1407 (2020).
51. W. E. Boden, R. A. O'Rourke, K. K. Teo, P. M. Hartigan, D. J. Maron, W. J. Kostuk, M. Knudtson, M. Dada, P. Casperson, C. L. Harris, B. R. Chaitman, L. Shaw, G. Gosselin, S. Nawaz, L. M. Title, G. Gau, A. S. Blaustein, D. C. Booth, E. R. Bates, J. A. Spertus, D. S. Berman, G. B. Mancini, W. S. Weintraub; C. T. R. Group, Optimal medical therapy with or without PCI for stable coronary disease. *N. Engl. J. Med.* **356**, 1503–1516 (2007).
52. S. S. Khan, H. Ning, S. J. Shah, C. W. Yancy, M. Carnethon, J. D. Berry, R. J. Mentz, E. O'Brien, A. Correa, N. Suthahar, R. A. de Boer, J. T. Wilkins, D. M. Lloyd-Jones, 10-Year risk equations for incident heart failure in the general population. *J. Am. Coll. Cardiol.* **73**, 2388–2397 (2019).
53. E. Brody, L. Gold, M. Mehan, R. Ostroff, J. Rohloff, J. Walker, D. Zichi, Life's simple measures: Unlocking the proteome. *J. Mol. Biol.* **422**, 595–606 (2012).
54. C. H. Kim, S. S. Tworoger, M. J. Stampfer, S. T. Dillon, X. Gu, S. J. Sawyer, A. T. Chan, T. A. Libermann, A. H. Eliassen, Stability and reproducibility of proteomic profiles measured with an aptamer-based platform. *Sci. Rep.* **8**, 8382 (2018).
55. J. C. Rohloff, A. D. Gelinas, T. C. Jarvis, U. A. Ochsner, D. J. Schneider, L. Gold, N. Janjic, Nucleic acid ligands with protein-like side chains: Modified aptamers and their use as diagnostic and therapeutic agents. *Mol. Ther. Nucleic Acids* **3**, e201 (2014).
56. S. Krokstad, A. Langhammer, K. Hveem, T. L. Holmen, K. Midthjell, T. R. Stene, G. Bratberg, J. Heggland, J. Holmen, Cohort profile: the HUNT Study, Norway. *Int. J. Epidemiol.* **42**, 968–977 (2013).
57. R. B. D'Agostino, M. W. Russell, D. M. Huse, R. C. Ellison, H. Silbershatz, P. W. Wilson, S. C. Hartz, Primary and subsequent coronary risk appraisal: New results from the Framingham study. *Am. Heart J.* **139**, 272–281 (2000).
58. D. C. Goff Jr., D. M. Lloyd-Jones, G. Bennett, S. Coady, R. B. D'Agostino, R. Gibbons, P. Greenland, D. T. Lackland, D. Levy, C. J. O'Donnell, J. G. Robinson, J. S. Schwartz, S. T. Shero, S. C. Smith Jr., P. Sorlie, N. J. Stone, P. W. Wilson, H. S. Jordan, L. Nevo, J. Wnek, J. L. Anderson, J. L. Halperin, N. M. Albert, B. Bozkurt, R. G. Brindis, L. H. Curtis, D. DeMets, J. S. Hochman, R. J. Kovacs, E. M. Ohman, S. J. Pressler, F. W. Sellke, W. K. Shen, S. C. Smith Jr., G. F. Tomaselli; G. American College of Cardiology/American Heart Association Task Force on Practice, 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* **129**, S49–S73 (2014).
59. J. Walter, J. du Fay de Lavallaz, L. Koechlin, T. Zimmermann, J. Boeddinghaus, U. Honegger, I. Strebel, R. Twerenbold, M. Amrein, T. Nestelberger, D. Wussler, C. Puelacher, P. Badertscher, M. Zellweger, G. Fahrni, R. Jeger, C. Kaiser, T. Reichlin, C. Mueller, Using high-sensitivity cardiac troponin for the exclusion of inducible myocardial ischemia in symptomatic patients: A cohort study. *Ann. Intern. Med.* **172**, 175–185 (2020).
60. N. Shiba, K. Nochioka, M. Miura, H. Kohno, H. Shimokawa; CHART-2 Investigators, Trend of westernization of etiology and clinical characteristics of heart failure patients in Japan–first report from the CHART-2 study. *Circ. J.* **75**, 823–833 (2011).
61. The Atherosclerosis Risk in Communities (ARIC) study: Design and objectives. The ARIC investigators. *Am. J. Epidemiol.* **129**, 687–702 (1989).
62. D. Mueller, C. Puelacher, U. Honegger, J. E. Walter, P. Badertscher, N. Schaerli, I. Strebel, R. Twerenbold, J. Boeddinghaus, T. Nestelberger, C. Hollenstein, J. du Fay de Lavallaz, R. Jeger, C. Kaiser, D. Wild, K. Rentsch, A. Buser, M. Zellweger, T. Reichlin, C. Mueller, Direct comparison of cardiac troponin T and I using a uniform and a sex-specific approach in the detection of functionally relevant coronary artery disease. *Clin. Chem.* **64**, 1596–1606 (2018).
63. R. R. Holman, M. A. Bethel, J. George, H. Sourij, Z. Doran, J. Keenan, N. S. Khurmi, R. J. Mentz, A. Oulhaj, J. B. Buse, J. C. Chan, N. Iqbal, S. Kundu, A. P. Maggioni, S. P. Marso, P. Ohman, M. J. Pencina, N. Poulter, L. E. Porter, A. Ramachandran, B. Zinman, A. F. Hernandez, Rationale and design of the EXenatide Study of Cardiovascular Event Lowering (EXSCEL) trial. *Am. Heart J.* **174**, 103–110 (2016).
64. M. E. Lean, W. S. Leslie, A. C. Barnes, N. Brosnahan, G. Thom, L. McCombie, C. Peters, S. Zhyzhneuskaya, A. Al-Mrabeh, K. G. Hollingsworth, A. M. Rodrigues, L. Rehackova, A. J. Adamson, F. F. Sniehotta, J. C. Mathers, H. M. Ross, Y. McIlvenna, R. Stefanetti, M. Trenell, P. Welsh, S. Kean, I. Ford, A. McConnachie, N. Sattar, R. Taylor, Primary care-led weight management for remission of type 2 diabetes (DiRECT): An open-label, cluster-randomised trial. *Lancet* **391**, 541–551 (2018).
65. W. Hosmer, T. Hosmer, S. Le Cessie, S. Lemeshow, A comparison of goodness-of-fit tests for the logistic regression model. *Stat. Med.* **16**, 965–980 (1997).
66. A. Tin, B. Yu, J. Ma, K. Masushita, N. Daya, R. C. Hoogeveen, C. M. Ballantyne, D. Couper, C. M. Rebholz, M. E. Grams, A. Alonso, T. Mosley, G. Heiss, P. Ganz, E. Selvin, E. Boerwinkle, J. Coresh, Reproducibility and variability of protein analytes measured using a multiplexed modified aptamer assay. *J. Appl. Lab. Med.* **4**, 30–39 (2019).
67. V. Emilsson, M. Ilkov, J. R. Lamb, N. Finkel, E. F. Gudmundsson, R. Pitts, H. Hoover, V. Gudmundsdottir, S. R. Horman, T. Aspelund, L. Shu, V. Trifonov, S. Sigurdsson, A. Manolescu, J. Zhu, O. Olafsson, J. Jakobsdottir, S. A. Lesley; J. To, J. Zhang, T. B. Harris, L. J. Launer, B. Zhang, G. Eiriksdottir, X. Yang, A. P. Orth, L. L. Jennings, V. Gudnason, Co-regulatory networks of human serum proteins link genetics to disease. *Science* **361**, 769–773 (2018).
68. B. B. Sun, J. C. Maranville, J. E. Peters, D. Stacey, J. R. Staley, J. Blackshaw, S. Burgess, T. Jiang, E. Paige, P. Surendran, C. Oliver-Williams, M. A. Kamat, B. P. Prins, S. K. Wilcox, E. S. Zimmerman, A. Chi, N. Bansal, S. L. Spain, A. M. Wood, N. W. Morrell, J. R. Bradley, N. Janjic, D. J. Roberts, W. H. Ouwehand, J. A. Todd, N. Soranzo, K. Suhre, D. S. Paul, C. S. Fox, R. M. Plenge, J. Danesh, H. Runz, A. S. Butterworth, Genomic atlas of the human plasma proteome. *Nature* **558**, 73–79 (2018).
69. L. Gold, D. Ayers, J. Bertino, C. Bock, A. Bock, E. N. Brody, J. Carter, A. B. Dalby, B. E. Eaton, T. Fitzwater, D. Flather, A. Forbes, T. Foreman, C. Fowler, B. Gawande, M. Goss, M. Gunn, S. Gupta, D. Halladay, J. Heil, J. Heilig, B. Hicke, G. Husar, N. Janjic, T. Jarvis, S. Jennings, E. Katilius, T. R. Keeney, N. Kim, T. H. Koch, S. Kraemer, L. Kroiss, N. Le, D. Levine, W. Lindsey, B. Lollo, W. Mayfield, M. Mehan, R. Mehler, S. K. Nelson, M. Nelson, D. Nieuwlandt, M. Nikrad, U. Ochsner, R. M. Ostroff, M. Otis, T. Parker, S. Pietrasiewicz, D. I. Resnicow, J. Rohloff, G. Sanders, S. Sattin, D. Schneider, B. Singer, M. Stanton, A. Sterkel, A. Stewart, S. Stratford, J. D. Vaught, M. Vrkljan, J. J. Walker, M. Watrobka, S. Waugh, A. Weiss, S. K. Wilcox, A. Wolfson, S. K. Wolk, C. Zhang, D. Zichi, Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLOS ONE* **5**, e15004 (2010).

70. T. J. Wang, M. G. Larson, D. Levy, E. J. Benjamin, E. P. Leip, T. Omland, P. A. Wolf, R. S. Vasan, Plasma natriuretic peptide levels and the risk of cardiovascular events and death. *N. Engl. J. Med.* **350**, 655–663 (2004).
71. A. S. Maisel, J. McCord, R. M. Nowak, J. E. Hollander, A. H. Wu, P. Duc, T. Omland, A. B. Storrow, P. Krishnaswamy, W. T. Abraham, P. Clopton, G. Steg, M. C. Aumont, A. Westheim, C. W. Knudsen, A. Perez, R. Kamin, R. Kazanegra, H. C. Herrmann, P. A. McCullough; Breathing Not Properly Multinational Study Investigators, Bedside B-type natriuretic peptide in the emergency diagnosis of heart failure with reduced or preserved ejection fraction. Results from the Breathing Not Properly Multinational Study. *J. Am. Coll. Cardiol.* **41**, 2010–2017 (2003).
72. P. Ponikowski, A. A. Voors, S. D. Anker, H. Bueno, J. G. Cleland, A. J. Coats, V. Falk, J. R. Gonzalez-Juanatey, V. P. Harjola, E. A. Jankowska, M. Jessup, C. Linde, P. Nihoyannopoulos, J. T. Parissis, B. Pieske, J. P. Riley, G. M. Rosano, L. M. Ruilope, F. Ruschitzka, F. H. Rutten, P. van der Meer; 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. J. Heart Fail.* **18**, 891–975 (2016).
73. C. W. Yancy, M. Jessup, B. Bozkurt, J. Butler, D. E. Casey Jr., M. H. Drazner, G. C. Fonarow, S. A. Geraci, T. Horwich, J. L. Januzzi, M. R. Johnson, E. K. Kasper, W. C. Levy, F. A. Masoudi, P. E. McBride, J. J. McMurray, J. E. Mitchell, P. N. Peterson, B. Riegel, F. Sam, L. W. Stevenson, W. H. Tang, E. J. Tsai, B. L. Wilkoff; American College of Cardiology Foundation; American Heart Association Task Force on Practice Guidelines, 2013 ACCF/AHA guideline for the management of heart failure: A report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *J. Am. Coll. Cardiol.* **62**, e147–e239 (2013).
74. S. Suzuki, M. Yoshimura, M. Nakayama, Y. Mizuno, E. Harada, T. Ito, S. Nakamura, K. Abe, M. Yamamuro, T. Sakamoto, Y. Saito, K. Nakao, H. Yasue, H. Ogawa, Plasma level of B-type natriuretic peptide as a prognostic marker after acute myocardial infarction: A long-term follow-up analysis. *Circulation* **110**, 1387–1391 (2004).
75. I. Goncalves, E. Bengtsson, H. M. Colhoun, A. C. Shore, C. Palombo, A. Natali, A. Edsfeldt, P. Duner, G. N. Fredrikson, H. Bjorkbacka, G. Ostling, K. Aizawa, F. Casanova, M. Persson, K. Gooding, D. Strain, F. Khan, H. C. Looker, F. Adams, J. Belch, S. Pinnoli, E. Venturi, M. Kozakova, L. M. Gan, V. Schnecke, J. Nilsson; Summit Consortium, Elevated plasma levels of MMP-12 are associated with atherosclerotic burden and symptomatic cardiovascular disease in subjects with type 2 diabetes. *Arterioscler. Thromb. Vasc. Biol.* **35**, 1723–1731 (2015).
76. R. P. Iyer, N. L. Patterson, F. A. Zouein, Y. Ma, V. Dive, L. E. de Castro Bras, M. L. Lindsey, Early matrix metalloproteinase-12 inhibition worsens post-myocardial infarction cardiac dysfunction by delaying inflammation resolution. *Int. J. Cardiol.* **185**, 198–208 (2015).
77. F. A. van Nieuwenhoven, C. Munts, R. C. Op't Veld, A. Gonzalez, J. Diez, S. Heymans, B. Schroen, M. van Bilsen, Cartilage intermediate layer protein 1 (CILP1): A novel mediator of cardiac extracellular matrix remodelling. *Sci. Rep.* **7**, 16042 (2017).
78. C. L. Hung, T. C. Hung, Y. H. Lai, C. S. Lu, Y. J. Wu, H. I. Yeh, Beyond malignancy: The role of carbohydrate antigen 125 in heart failure. *Biomark Res.* **1**, 25 (2013).
79. K. Hakki, K. Recep, B. Osman, Cancer antigen 125 is associated with length of stay in patients with acute heart failure. *Tex. Heart Inst. J.* **44**, 22–28 (2017).
80. F. Falcao, F. R. A. de Oliveira, M. da Silva, D. C. Sobral Filho, Carbohydrate antigen 125: A promising tool for risk stratification in heart diseases. *Biomark. Med* **12**, 367–381 (2018).
81. P. Yu, J. Zhao, H. Jiang, M. Liu, X. Yang, B. Zhang, Y. Yu, L. Zhang, R. Tong, G. Liu, R. Chen, Y. Zou, J. Ge, Neural cell adhesion molecule-1 may be a new biomarker of coronary artery disease. *Int. J. Cardiol.* **257**, 238–242 (2018).
82. M. A. Ackermann, J. M. Petrosino, H. R. Manring, P. Wright, V. Shettigar, A. Kilic, P. M. L. Janssen, M. T. Ziolo, F. Accornero, TGF-β1 affects cell-cell adhesion in the heart in an NCAM1-dependent mechanism. *J. Mol. Cell. Cardiol.* **112**, 49–57 (2017).
83. M. Stenemo, C. Nowak, L. Byberg, J. Sundstrom, V. Giedraitis, L. Lind, E. Ingelsson, T. Fall, J. Arnlov, Circulating proteins as predictors of incident heart failure in the elderly. *Eur. J. Heart Fail.* **20**, 55–62 (2018).
84. B. T. Santema, M. Kloosterman, I. C. Van Gelder, I. Mordi, C. C. Lang, C. S. P. Lam, S. D. Anker, J. G. Cleland, K. Dickstein, G. Filippatos, P. Van der Harst, H. L. Hillege, J. M. Ter Maaten, M. Metra, L. L. Ng, P. Ponikowski, N. J. Samani, D. J. Van Veldhuisen, A. H. Zwinderman, F. Zannad, K. Damman, P. Van der Meer, M. Rienstra, A. A. Voors, Comparing biomarker profiles of patients with heart failure: A trial fibrillation vs. sinus rhythm and reduced vs. preserved ejection fraction. *Eur. Heart J.* **39**, 3867–3875 (2018).
85. F. Obendorf, C. Herz, C. Hobaus, Abstract 16877: High serum levels of trefoil factor 3 are associated with an increased risk for cardiovascular events. *Circulation* **132**, A16877 (2015).
86. M. Brankovic, K. Martijn Akkerhuis, H. Mouthaan, A. Constantinescu, K. Caliskan, J. van Ramshorst, T. Germans, V. Umans, I. Kardys, Utility of temporal profiles of new cardio-renal and pulmonary candidate biomarkers in chronic heart failure. *Int. J. Cardiol.* **276**, 157–165 (2019).
87. X. Tang, X. F. Chen, N. Y. Wang, X. M. Wang, S. T. Liang, W. Zheng, Y. B. Lu, X. Zhao, D. L. Hao, Z. Q. Zhang, M. H. Zou, D. P. Liu, H. Z. Chen, SIRT2 acts as a cardioprotective deacetylase in pathological cardiac hypertrophy. *Circulation* **136**, 2051–2067 (2017).
88. M. Sarikhani, S. Maity, S. Mishra, A. Jain, A. K. Tamta, V. Ravi, M. S. Kondapalli, P. A. Desingu, D. Khan, S. Kumar, S. Rao, M. Inbaraj, A. S. Pandit, N. R. Sundaresan, SIRT2 deacetylase represses NFAT transcription factor to maintain cardiac homeostasis. *J. Biol. Chem.* **293**, 5281–5294 (2018).
89. G. Chen, X. Zhou, P. Nicolaou, P. Rodriguez, G. Song, B. Mitton, A. Pathak, A. Zachariah, G. C. Fan, G. W. Dorn II, E. G. Kranias, A human polymorphism of protein phosphatase-1 inhibitor-1 is associated with attenuated contractile response of cardiomyocytes to beta-adrenergic stimulation. *FASEB J.* **22**, 1790–1796 (2008).
90. P. Nicolaou, R. J. Hajjar, E. G. Kranias, Role of protein phosphatase-1 inhibitor-1 in cardiac physiology and pathophysiology. *J. Mol. Cell. Cardiol.* **47**, 365–371 (2009).
91. S. Rubattu, G. Bigatti, A. Evangelista, C. Lanzani, R. Stanzione, L. Zagato, P. Manunta, S. Marchitti, V. Venturelli, G. Bianchi, M. Volpe, P. Stella, Association of atrial natriuretic peptide and type a natriuretic peptide receptor gene polymorphisms with left ventricular mass in human essential hypertension. *J. Am. Coll. Cardiol.* **48**, 499–505 (2006).
92. E. Barbato, J. Bartunek, S. Marchitti, F. Mangiacapra, R. Stanzione, L. Delrue, M. Cotugno, S. Di Castro, B. De Bruyne, W. Wijns, M. Volpe, S. Rubattu, NT-proANP circulating level is a prognostic marker in stable ischemic heart disease. *Int. J. Cardiol.* **155**, 311–312 (2012).
93. M. S. Sabatine, D. A. Morrow, J. A. de Lemos, T. Omland, S. Sloan, P. Jarolim, S. D. Solomon, M. A. Pfeffer, E. Braunwald, Evaluation of multiple biomarkers of cardiovascular stress for risk prediction and guiding medical therapy in patients with stable coronary disease. *Circulation* **125**, 233–240 (2012).
94. Y. Saito, Roles of atrial natriuretic peptide and its therapeutic use. *J. Cardiol.* **56**, 262–270 (2010).
95. N. O. Stitziel, K. E. Stirrups, N. G. Masca, J. Erdmann, P. G. Ferrario, I. R. Konig, P. E. Weeke, T. R. Webb, P. L. Auer, U. M. Schick, Y. Lu, H. Zhang, M. P. Dube, A. Goel, M. Farrall, G. M. Peloso, H. H. Won, R. Do, E. van Iperen, S. Kanoni, J. Kruppa, A. Mahajan, R. A. Scott, C. Willenberg, P. S. Braund, J. C. van Capelleveen, A. S. Doney, L. A. Donnelly, R. Asselta, P. A. Merlini, S. Duga, N. Marziliano, J. C. Denny, C. M. Shaffer, N. E. El-Mokhtari, A. Franke, O. Gottesman, S. Heilmann, C. Hengstenberg, P. Hoffman, O. L. Holmen, K. Hveem, J. H. Jansson, K. H. Jockel, T. Kessler, J. Kriebel, K. L. Laugwitz, E. Marouli, N. Martinelli, M. I. McCarthy, N. R. Van Zuydam, C. Meisinger, T. Esko, E. Mihailov, S. A. Escher, M. Alver, S. Moebus, A. D. Morris, M. Muller-Nurasyid, M. Nikpay, O. Olivieri, L. P. L. Perreault, A. AlQarawi, N. R. Robertson, K. O. Akinsanya, D. F. Reilly, T. F. Vogt, W. Yin, F. W. Asselbergs, C. Kooperberg, R. D. Jackson, E. Stahl, K. Strauch, T. V. Varga, M. Waldenberger, L. Zeng, A. T. Kraja, C. Liu, G. B. Ehret, C. Newton-Cheh, D. I. Chasman, R. Chowdhury, M. Ferrario, I. Ford, J. W. Jukema, F. Kee, K. Kuulasmaa, B. G. Nordestgaard, M. Perola, D. Saleheen, N. Sattar, P. Surendran, D. Tregouet, R. Young, J. M. Howson, A. S. Butterworth, J. Danesh, D. Ardissino, E. P. Bottinger, R. Erbel, P. W. Franks, D. Girelli, A. S. Hall, G. K. Hovingh, A. Kastrati, W. Lieb, T. Meitinger, W. E. Kraus, S. H. Shah, R. McPherson, M. Orho-Melander, O. Melander, A. Metspalu, C. N. Palmer, A. Peters, D. Rader, M. P. Reilly, R. J. Loos, A. P. Reiner, D. M. Roden, J. C. Tardif, J. R. Thompson, N. J. Wareham, H. Watkins, C. J. Willer, S. Kathiresan, P. Deloukas, N. J. Samani, H. Schunkert, Coding variation in ANGPTL4, LPL, and SVEP1 and the risk of coronary disease. *N. Engl. J. Med.* **374**, 1134–1144 (2016).
96. T. A. Nakada, W. Takahashi, E. Nakada, T. Shimada, J. A. Russell, K. R. Walley, Genetic polymorphisms in sepsis and cardiovascular disease: Do similar risk genes suggest similar drug targets? *Chest* **155**, 1260–1271 (2019).
97. Y. A. Senis, Protein-tyrosine phosphatases: A new frontier in platelet signal transduction. *J. Thromb. Haemost.* **11**, 1800–1813 (2013).
98. T. Takahashi, K. Takahashi, P. L. S. John, P. A. Fleming, T. Tomemori, T. Watanabe, D. R. Abrahamson, C. J. Drake, T. Shirasawa, T. O. Daniel, A mutant receptor tyrosine phosphatase, CD148, causes defects in vascular development. *Mol. Cell. Biol.* **23**, 1817–1831 (2003).
99. A. Edsfeldt, M. Nitulescu, H. Grufman, C. Gronberg, A. Persson, M. Nilsson, M. Persson, H. Bjorkbacka, I. Goncalves, Soluble urokinase plasminogen activator receptor is associated with inflammation in the vulnerable human atherosclerotic plaque. *Stroke* **43**, 3305–3312 (2012).
100. S. Lyngbaek, J. L. Marott, T. Sehestedt, T. W. Hansen, M. H. Olsen, O. Andersen, A. Linneberg, S. B. Haugaard, J. Eugen-Olsen, P. R. Hansen, J. Jeppesen, Cardiovascular risk prediction in the general population with use of suPAR, CRP, and Framingham risk score. *Int. J. Cardiol.* **167**, 2904–2911 (2013).
101. R. L. Prentice, S. Paczesny, A. Aragaki, L. M. Amon, L. Chen, S. J. Pitteri, M. McIntosh, P. Wang, T. Buson Busald, J. Hsia, R. D. Jackson, J. E. Rossouw, J. E. Manson, K. Johnson, C. Eaton, S. M. Hanash, Novel proteins associated with risk for coronary heart disease or stroke among postmenopausal women identified by in-depth plasma proteome profiling. *Genome Med.* **2**, 48 (2010).
102. K. T. Kouassi, P. Gunasekar, D. K. Agrawal, G. P. Jadhav, TREM-1; is it a pivotal target for cardiovascular diseases? *J. Cardiovasc. Dev. Dis.* **5**, 45 (2018).

103. A. Boufenzer, J. Lemarie, T. Simon, M. Derive, Y. Bouazza, N. Tran, F. Maskali, F. Groubatch, P. Bonnin, C. Bastien, P. Bruneval, P. Y. Marie, R. Cohen, N. Danchin, J. S. Silvestre, H. Ait-Oufella, S. Gibot, TREM-1 mediates inflammatory injury and cardiac remodeling following myocardial infarction. *Circ. Res.* **116**, 1772–1782 (2015).
104. N. Adhikari, D. L. Basi, D. Townsend, M. Rusch, A. Mariash, S. Mullegama, A. Watson, J. Larson, S. Tan, B. Lerman, J. D. Esko, S. B. Selleck, J. L. Hall, Heparan sulfate Ndst1 regulates vascular smooth muscle cell proliferation, vessel size and vascular remodeling. *J. Mol. Cell. Cardiol.* **49**, 287–293 (2010).
105. M. A. Sonneveld, O. H. Franco, M. A. Ikram, A. Hofman, M. Kavousi, M. P. de Maat, F. W. Leebeek, Von Willebrand factor, ADAMTS13, and the risk of mortality: The Rotterdam study. *Arterioscler. Thromb. Vasc. Biol.* **36**, 2446–2451 (2016).
106. M. A. Sonneveld, M. P. de Maat, M. L. Portegies, M. Kavousi, A. Hofman, P. L. Turecek, H. Rottensteiner, F. Scheiflinger, P. J. Koudstaal, M. A. Ikram, F. W. Leebeek, Low ADAMTS13 activity is associated with an increased risk of ischemic stroke. *Blood* **126**, 2739–2746 (2015).
107. K. A. Olson, A. L. Beatty, B. Heidecker, M. C. Regan, E. N. Brody, T. Foreman, S. Kato, R. E. Mehler, B. S. Singer, K. Hveem, H. Dalen, D. G. Sterling, R. M. Lawn, N. B. Schiller, S. A. Williams, M. A. Whooley, P. Ganz, Association of growth differentiation factor 11/8, putative anti-ageing factor, with cardiovascular outcomes and overall mortality in humans: Analysis of the Heart and Soul and HUNT3 cohorts. *Eur. Heart J.* **36**, 3426–3434 (2015).
108. F. S. Loffredo, M. L. Steinhauser, S. M. Jay, J. Gannon, J. R. Pancoast, P. Yalamanchi, M. Sinha, C. Dall'Osso, D. Khong, J. L. Shadrach, C. M. Miller, B. S. Singer, A. Stewart, N. Psychogios, R. E. Gerszten, A. J. Hartigan, M. J. Kim, T. Serwold, A. J. Wagers, R. T. Lee, Growth differentiation factor 11 is a circulating factor that reverses age-related cardiac hypertrophy. *Cell* **153**, 828–839 (2013).
109. A. Meloux, L. Rochette, M. Maza, F. Bichat, L. Tribouillard, Y. Cottin, M. Zeller, C. Vergely, Growth differentiation factor-8 (GDF8)/myostatin is a predictor of troponin I peak and a marker of clinical severity after acute myocardial infarction. *J. Clin. Med.* **9**, 116 (2020).
110. N. Biesemann, L. Mendler, A. Wietelmann, S. Hermann, M. Schafers, M. Kruger, T. Boettger, T. Borchardt, T. Braun, Myostatin regulates energy homeostasis in the heart and prevents heart failure. *Circ. Res.* **115**, 296–310 (2014).

Funding: HUNT is a collaboration between HUNT Research Centre (Faculty of Medicine and Health Sciences, Norwegian University of Science, and Technology NTNU), Trøndelag County Council, Central Norway Regional Health Authority, and the Norwegian Institute of Public Health. The ARIC study has been funded in whole or in part with federal funds from the National Heart, Lung, and Blood Institute, NIH, Department of Health and Human Services (contract numbers HHSN268201700001I, HHSN268201700002I, HHSN268201700003I, HHSN268201700004I, and HHSN268201700005I), R01HL087641, R01HL086694; National Human Genome Research Institute contract U01HG004402; and NIH contract HHSN268200625226C. J.W. is supported by Swiss Heart Foundation grants FF19097 and F18111 and grants by the Swiss Academy of Medical Sciences/Gottrfied and Julia Bangerter-Rhyner Foundation. The DiRECT trial (co-PIs Lean and Taylor) is funded by Diabetes UK. The CHART-2 study was funded by Japan Agency for Medical Research and Development (AMED). The PRADA study received funding or research support from the University of Oslo, The Extra Foundation for Health and Rehabilitation, The Norwegian Cancer Society, Akershus University Hospital, Abbott Diagnostics, and AstraZeneca. C.P. is supported by NIH grants R01HL-085757 and UO1DK106962. P.G.'s proteomic research is supported by NIH grants RO1HL129856, UO1DK108809, and R01AG052964. **Author contributions:** Conceptualization: S.A.W., R.O., M.A.H., P.G., J. Chadwick, N.S., M.E.L., and N.P.T. Machine learning, model building, and bioinformatics: M.A.H., Y.H., J.P., N.K., S.K., and D.A. Clinical parent study conceptualization, design, execution, and databasing: J. Coresh, C.M.B., K.M., C.E.M., J.W., C.J., R.R.H., S.H.S., N.S., R.T., M.E.L., H.S., Y.S., K.N., C.P., S.G.C., C.R.P., and T.O. Clinical data/analysis management and outcomes research: J. Chadwick, K.L., M.S., and M.A.H. Writing—original draft: S.A.W., P.G., R.O., and M.A.H. Writing—review and editing: S.A.W., P.G., R.O., M.A.H., J. Coresh, C.M.B., K.M., C.E.M., J.W., C.J., R.R.H., S.H.S., N.S., R.T., M.E.L., S.K., H.S., Y.S., K.N., C.P., S.G.C., T.O., and N.P.T. **Competing interests:** SomaLogic Inc. funded the costs of the proteomic assays, and S.A.W., R.O., M.A.H., J. Chadwick, D.A., Y.H., C.P., M.S., K.L., J.P., and N.K. are employees of the company. S.K. is an employee of NEC Solution Innovators Ltd., which has a formal business relationship with SomaLogic Inc. C.P. is on the advisory board of RenalytixAI and DSMB committee of Genfit Pharma. The EXSCEL trial (co-PI RH) was funded by Amylin Pharmaceuticals, a wholly owned subsidiary of AstraZeneca. AstraZeneca also made a contribution to EXSCEL proteomic assay costs. R.R.H. also reports research support from AstraZeneca, Bayer and Merck, and Sharp and Dohme and personal fees from Bayer, Novartis, and Novo Nordisk. S.G.C. serves on the advisory board and owns equity in Renalytix. S.G.C. has reported receiving consulting fees from Bayer, Boehringer-Ingelheim, Nuwellis, Axon Therapies, Reprieve Cardiovascular, 3ive, Renalytix, and Vifor in the past 3 years. N.S. has consulted for Afimmune, Amgen, AstraZeneca, Boehringer Ingelheim, Eli Lilly, Hanmi Pharmaceuticals, Merck Sharp & Dohme, Novartis, Novo Nordisk, Pfizer, and Sanofi and received grant support paid to his university from AstraZeneca, Boehringer Ingelheim, Novartis, and Roche Diagnostics outside the submitted work. P.G. serves on a medical advisory board to SomaLogic Inc., for which he accepts no salary, honoraria, or any other financial incentives. The following authors declare that they have no competing interests: J.C., C.M.B., K.M., C.E.M., J.W., C.J., S.H.S., R.T., M.E.L., H.S., Y.S., K.N., T.O., and N.P.T. **Data and materials availability:** All data associated with this study are present in the paper or the Supplementary Materials. Preexisting data access policies for each of the parent cohort studies specify that research data requests can be submitted to each steering committee; these will be reviewed promptly for confidentiality or intellectual property restrictions and will not unreasonably be refused. Individual patient or protein data may further be restricted by consent, confidentiality, or privacy laws/considerations. These policies apply to both clinical and proteomic data.

Submitted 16 June 2021
Resubmitted 27 October 2021
Accepted 18 March 2022
Published 6 April 2022
10.1126/scitranslmed.abj9625

Science Translational Medicine

A proteomic surrogate for cardiovascular outcomes that is sensitive to multiple mechanisms of change in risk

Stephen A. WilliamsRachel OstroffMichael A. HinterbergJosef CoreshChristie M. BallantyneKunihiro MatsushitaChristian E. MuellerJoan WalterChristian JonassonRury R. HolmanSvati H. ShahNaveed SattarRoy TaylorMichael E. LeanShintaro KatoHiroaki ShimokawaYasuhiko SakataKotaro NochiokaChirag R. ParikhSteven G. CocaTorbjørn OmlandJessica ChadwickDavid AstlingYolanda HagarNatasha KureshiKelsey LoupyClare PatersonJeremy PrimusMissy SimpsonNelson P. TrujilloPeter Ganz

Sci. Transl. Med., 14 (639), eabj9625. • DOI: 10.1126/scitranslmed.abj9625

Creating a surrogate for cardiovascular risk
Clinical trials can be limited by the lack of surrogates for cardiovascular risk, leading to increased costs and potentially delaying important results. Here, Williams *et al.* used proteomics and machine learning to derive a 27-protein model that could predict the 4-year likelihood of myocardial infarction, heart failure, stroke, or death better than a clinical model. The proteins included in the model represented 10 mechanistic pathways, and 12 were associated with causal genetic traits. This model was validated across more than 11,000 participants from multiple large studies and was sensitive to both adverse and beneficial changes in outcome, suggesting that it has potential as a surrogate end point for use in phase 2 trials.