medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

A novel plasma proteomic signature predicts long-term incident heart failure risk among persons living with HIV

Saate S Shakil MD MS,[1]*
Tristan Grogan MS,[1]
Kaku So-Armah PhD,[2]
Peter Ganz MD,[3]
Matthew S Freiberg MD MSc5,[4]
Priscilla Y Hsue MD[1]

[1] Department of Medicine, University of California Los Angeles, CA, USA
[2] Department of Medicine, Boston University, Boston, MA, USA
[3] Department of Medicine, University of California San Francisco, CA, USA
[4] Department of Medicine, Vanderbilt University, Nashville, TN, USA
* Corresponding author: SSS, Division of Cardiology, 1100 Glendon Ave Ste 1820, Los Angeles, CA 90401 | sshakil@mednet.ucla.edu

Word count: 4249 | Tables: 3 | Figures: 5

Disclaimer: The views and opinions expressed in this manuscript are those of the authors and do not necessarily represent those of the Department of Veterans Affairs nor the United States government.

Funding: This study is funded by the following grants: United States National Institutes of Health (NIH)/National Heart, Lung, and Blood Institute (NHLBI) K24AI112393 (PYH), R01HL129856 (PG, MF, PYH), NHLBI R01HL170600-03S1 (PYH, SSS), and American Heart Association 24SCEFIA1258761 (SSS). Parent study funded by NIH P01AA029545 (The HIV and Alcohol Research center focused on Polypharmacy [HARP]), U24AA020794 (COMpAAAS coordinating center), and U10AA013566.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Background: Persons with HIV (PWH) have an increased risk of heart failure compared with the general population despite effective viral suppression. No methods currently exist to predict heart failure risk in PWH. We sought to develop a proteomics-based risk model of incident heart failure in a longitudinal cohort of PWH in the Veterans Aging Cohort Study Biomarker Cohort.
Methods: We characterized the baseline plasma proteome of 1,398 PWH using the SomaScan 4.0 aptamer-based platform. We incorporated traditional risk factors, HIV-specific variables, and inflammatory/coagulation markers into Cox proportional hazards models of incident heart failure. For protein risk models, we used least absolute shrinkage and selection operator (LASSO) regression with 10-fold cross-validation to identify the optimal set of predictors. Model performance was evaluated using Harrell's Concordance Index (C-index).
Results: Among 1,398 PWH, 240 (17.1%) developed incident heart failure over 13.1 years of median follow-up. The sample was randomly divided 2:1 into training (n = 932) and test (n = 466) sets. In the training set, 4,248 proteins were screened, identifying 227 significant markers (Bonferroni-corrected p-value <1E-5); 108 of these were validated in the test set (Bonferroni-corrected p <2E-4) and included as candidate predictors in downstream LASSO models. Using LASSO-Cox regression, we identified a subset of predictor proteins alone and in combination with clinical and laboratory risk factors. Among the clinical risk factor models, individual traditional risk factors in combination with HIV disease markers ($CD4^+$ T-cell count, viral load) exhibited the best predictive performance (C-index 0.656 [95% confidence interval 0.591 – 0.720]). Addition of circulating inflammatory and coagulation markers (soluble CD-14, interleukin-6, D-dimer) did not substantially improve performance (C-index 0.657 [0.594 – 0.720]). By contrast, a protein-only model incorporating 8 markers achieved an improved C-index of 0.725 [0.597 – 0.825]; proteins vs clinical model, p = 0.033), while the addition of $CD4^+$ count and viral load performed similarly (C-index 0.715 [0.587 – 0.816]).
Conclusion: An 8-protein signature demonstrated superior predictive performance for long-term incident heart failure risk among PWH compared with baseline traditional risk factors and inflammatory markers. These findings highlight the potential of a proteomics-based approach for risk stratification in this population, which provides additive value beyond traditional or HIV-specific clinical risk factors.

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

Despite effective viral suppression with antiretroviral therapy, persons with HIV (PWH) suffer from an increased risk of heart failure compared with the general population, independently of the presence of coronary artery disease.[1,2] While traditional cardiometabolic comorbidities contribute to this risk,[3–5] HIV-specific immune dysregulation also predisposes to cardiovascular disease and mortality, with increased monocyte and T-cell activation and other markers of immune dysfunction, inflammation, and coagulation (e.g., sCD14, interleukin-6, D-dimer).[6–9] These pathways are believed to lead to myocardial inflammation, steatosis, and fibrosis predisposing to adverse remodeling and both systolic and diastolic clinical heart failure.[2,10–14] However, the exact mechanisms underlying heart failure in PWH remain largely uncharacterized.

PWH with heart failure experience higher all-cause mortality compared with those without HIV with heart failure and may therefore benefit from tailored screening and risk stratification.[15] However, current heart failure risk prediction tools have largely been developed in the general population, including the Pooled Cohort Equations to Prevent Heart Failure (PCP-HF) and the American Heart Association Predicting Risk of CVD EVENTs (PREVENT) equations.[16,17] Further, multiple studies have demonstrated that cardiovascular risk prediction algorithms derived in the general population perform poorly in PWH.[18–20] Currently, there are no available risk calculators to stratify PWH at highest risk of developing heart failure. Precision approaches that account for unique HIV-specific mechanisms are needed to optimize screening and risk stratification in this population.

A recent large-scale proteomic study in the general population identified 37 novel heart failure biomarkers that are independent of clinical risk factors, 10 of which are thought to be causal.[21] Two recent cross-sectional discovery studies of echocardiographic indices identified a signature of subclinical left atrial remodeling that was heightened in adults with HIV compared with persons without HIV, and a 4-protein signature of abnormal myocardial performance index in pediatric PWH.[22,23] Although protein-based risk prediction may capture HIV disease-specific information beyond traditional and non-traditional risk factors, no studies to date have evaluated proteomic signatures of incident heart failure in a longitudinal cohort of PWH. In the present

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

study, we sought to develop a novel protein-based risk model to predict incident heart failure events among PWH in the Veterans Aging Cohort Study HIV Biomarker Cohort. We hypothesized that protein-based risk models would demonstrate superior performance compared with models based on clinical, inflammatory, or HIV-specific risk factors.

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Methods

Study population

The Veterans Aging Cohort Study (VACS) HIV is a longitudinal cohort of veterans living with HIV and uninfected veterans matched on age, sex, race-ethnicity, and geographic region that has been previously described.[24] The VACS Biomarker Cohort (VACS BC) is a subset of VACS HIV, which prospectively enrolled veterans with (n = 1525) and without HIV (n = 843). From 2005 to 2007, VACS BC participants provided informed consent for blood samples. These specimens were collected using a serum separator and EDTA blood collection and shipped to a central repository at the Massachusetts Veterans Epidemiology Research and Information Center (Boston, MA). The date of the blood draw was used as the baseline date for each participant in VACS BC. Participants were followed from their baseline date until death or until censored. The original VACS BC study was approved by the Institutional Review Boards (IRB) at all VACS locations including Yale University (#1506016006) and VA Connecticut Healthcare System (#AJ0013), which included written consent prior to enrollment. As the present study is based on secondary analysis of de-identified data, no additional informed consent or IRB approval was required.

Data collection and laboratory procedures

Clinical and demographic data closest to the date of blood sample collection were extracted from VACS BC participants' electronic health records. Variables of interest included age, sex, self-reported race and ethnicity, medical conditions, medication use, smoking status, blood pressure, height and weight, HIV-1 RNA viral load, and $CD4^+$ T-cell count, among others. Baseline circulating inflammatory and coagulation markers were previously measured in VACS BC.[25] Levels of the soluble cluster-of-differentiation 14 (sCD14) protein were measured using ELISA (Quantikine sCD14 immunoassay, R&D Systems, Minneapolis, MN). Interleukin-6 (IL-6) levels were measured using a chemiluminescent immunoassay (QuantiGlo IL-6 immunoassay; R&D Systems, Minneapolis, MN). D-dimer was measured using an immuno-turbidimetric assay (Liatest D-DI; Diagnostica Stago, Parsippany, NJ).

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Proteomic profiling by modified aptamers

We used the SomaScan v4.0 assay (5,284 SOMAmer reagents), a single-stranded DNA aptamer-based platform, to quantify proteins in VACS BC specimens collected at study baseline. Methods for modified aptamer-based protein quantification have been described previously.[26] The SomaScan platform is an assay based on modified aptamers, which are chemically modified single-strand deoxyribonucleic acid approximately 40 nucleotides in length that act as binding reagents for target proteins. Prior studies assessing fidelity of aptamer-protein binding have demonstrated that >95% of aptamers correctly targeted the intended proteins, provided the proteins are present in sufficient concentration to be detected by mass spectrometry.[26] Samples are assayed at three different dilutions on the SomaScan platform to ensure each analyte is assayed within its linear range of concentrations. Results are then quantified on a hybridization microarray and reported in relative fluorescence units (RFU). We followed typical procedures for calibration, standardization, and internal controls.[27] After quality control (QC) filtering performed by SomaLogic (4,268 SOMAmers passing the QC ratio), removal of non-human and control aptamers, and collapsing duplicate reagents, 4,248 protein targets remained for analysis.

Outcomes

The primary outcome was time to incident heart failure event. As we and others have done previously, heart failure was defined by the presence of at least one inpatient (discharge diagnosis) or at least two outpatient VA, Medicare, or VA fee-for-service International Classification of Diseases (ICD) -9 or ICD-10 codes.[28] We excluded prevalent cases of heart failure at baseline. We evaluated 5-year heart failure-free survival as a secondary endpoint in our calibration analyses.

Data processing and statistical analysis

We used descriptive statistics to compare demographic and clinical variables between PLWH who did and did not have an incident heart failure event. We compared continuous variables using two-sample t-tests or Wilcoxon rank-sum tests for mean and median values, respectively. We used Fisher's exact test to compare binary variables and Chi-squared tests for categorical

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

variables. A p-value of <0.05 was considered statistically significant for our descriptive analyses.

Proteomics data underwent several preprocessing steps, including outlier capping using a 5x median absolute deviation (MAD) threshold (i.e., MAD-based winsorization), followed by $\log_2$ transformation and median-centered standardization (MAD-scaling). This approach ensured robust normalization and comparability across proteins, reducing the influence of extreme values. The final dataset used for modeling was log2-transformed and MAD-scaled. Clinical heart failure risk factors and HIV-specific variables were incorporated into Cox proportional hazards models for incident heart failure. For risk models involving proteins, we used least absolute shrinkage and selection operator (LASSO) regression with 10-fold cross-validation to identify the optimal set of predictors. The regularization parameter (λ) was selected using the minimum cross-validation partial likelihood deviance. LASSO served as a variable selection tool, shrinking the coefficients of less informative markers to zero while retaining the most predictive markers for inclusion in the final model. For models combining proteins with other clinical or laboratory parameters, we constructed joint predictor pools consisting of the relevant non-protein covariates (e.g., clinical risk factors, $CD4^+$ cell count, viral load, inflammatory/coagulation markers) plus the validated proteins, to which we applied LASSO; thus, both protein and non-protein predictors were subjected to penalization in the combined models. Model performance was evaluated using Harrell's Concordance Index (C-index), which quantifies the discriminative ability of the Cox proportional hazards model. The C-index was computed on the held-out test dataset to assess model fit and predictive accuracy.

For the primary comparisons between the protein-only model and the clinical models, we formally compared C-indices using nonparametric bootstrapping. We generated 1,000 bootstrap resamples of the test dataset, refit both models in each resample, and calculated the distribution of the C-index difference to obtain two-sided p-values and 95% confidence intervals. Other model contrasts were not subjected to formal hypothesis testing. Model calibration was assessed in the test dataset by grouping individuals into quintiles of predicted risk and estimating observed 5-year heart failure-free survival within each quintile using Kaplan-Meier curves. Observed survival estimates (with 95% confidence intervals based on Greenwood's formula) were plotted against mean predicted survival to generate a quintile-based calibration plot.

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Pathway analysis

We performed exploratory pathway analyses on the subset of proteins that were included as candidate predictors in LASSO-Cox regression models (m = 108). Using their UniProt ID numbers, we used the biomaRt R package to map these to Entrez Gene ID numbers to facilitate the use of pathway enrichment tools.[29–31] After deduplication, enrichment was performed with clusterProfiler and ReactomePA across Gene Ontology (GO) Biological Process, Kyoto Encyclopedia of Genes and Genomes (KEGG), and Reactome pathways.[32–34] All analyses used a Benjamini-Hochberg false discovery rate (FDR) correction ($q < 0.05$). Results were exported, sorted by FDR-adjusted p-value, and visualized as dot plots (top 20 categories per database) to illustrate the top pathways and number of key proteins were represented in each.

Data availability

Due to US Department of Veterans Affairs (VA) regulations and ethics agreements, the analytic data sets used for this study are not permitted to leave the VA firewall without a Data Use Agreement. This limitation is consistent with other studies based on VA data. However, VA data are made freely available to researchers with an approved VA study protocol. For more information, please visit https://www.virec.research.va.gov or contact the VA Information Resource Center at VIReC@va.gov.

Code availability

Only publicly available statistical packages were used in the analyses for this study as detailed in the Methods section. Statistical code and details of packages employed are available from the corresponding author upon reasonable request.

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Results

VACS cohort and heart failure outcomes

After excluding baseline prevalent cases of heart failure, a total of 240 incident heart failure events occurred among 1,398 PWH over 13.1 years of median follow-up (2005 to 2019). **Table 1** shows baseline characteristics of the study population, stratified by whether participants did (n = 240) or did not (n = 1158) have an incident heart failure event. The mean age and standard deviation (SD) of PLWH in VACS was 52.1 (SD 8.1) years, and participants with heart failure were older at baseline (51.5 [SD 8.1] vs 54.9 [7.6] years, p <0.001). The baseline prevalence of traditional cardiovascular risk factors was higher among persons with incident heart failure, including hypertension (65% vs 43.4%, p <0.001), diabetes (33.3% vs 14.6%, p <0.001), hyperlipidemia (40.4% vs 30.7%, p = 0.003), and cocaine use (42.9% vs 33.6%, p = 0.006). Circulating levels of soluble CD14 (sCD14) and D-dimer were also higher among those who had an incident heart failure event (p <0.001 and p = 0.006, respectively). Mean systolic blood pressure was also higher among those with an incident heart failure event (130.8 [SD 15.5] vs 127.8 [13.9] mmHg, p = 0.003). There was no significant difference in $CD4^+$ T-cell count, HIV viral load, antiretroviral therapy regimen, or cholesterol levels between those who did or did not have an incident heart failure event. There were no significant differences in clinical characteristics between testing and training sets, except for the distribution of female participants (3.6% vs 1.5% in training and testing sets, respectively; p = 0.025; **Supplementary Table S1**).

Development and performance of clinical, inflammatory, and protein risk models

The sample was divided randomly 2:1 into training (n = 932) and test (n = 466) sets, each of which included 164 and 76 participants with incident heart failure events, respectively. We developed clinical, inflammatory marker, and protein-based risk models in the training set and evaluated model fit in the test set (**Table 2**). Among the clinical models, the ASCVD 10-year risk score by the Pooled Cohort Equations (Model 1) demonstrated a test C-index of 0.607 (95% confidence interval [CI] 0.544 – 0.670), while individual clinical heart failure risk factors (Model 2; age, sex, race, systolic blood pressure, antihypertensive use, diabetes, smoking status, body mass index, total cholesterol, high-density lipoprotein cholesterol) yielded a C-index of 0.639

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

(0.574 – 0.704). Among the clinical risk factor models, individual traditional risk factors in combination with HIV disease markers ($CD4^+$ T-cell count, viral load) exhibited the best predictive performance (Model 5b; C-index 0.656 [95% CI 0.591 – 0.720]). Addition of circulating inflammatory and coagulation markers (soluble CD-14 [sCD14], interleukin-6 [IL-6], D-dimer) did not substantially improve performance (Model 8; C-index 0.657 [0.594 – 0.720]).

To develop protein risk models, we initially screened all 4,248 proteins in the training set in univariable single-protein Cox proportional hazards models of incident heart failure. After stringent Bonferroni correction (p <1E-5), we identified 227 of 4,248 proteins that predicted incident heart failure; these 227 were then validated in the test set in univariable single-protein Cox models. This two-step selection process yielded 108 of 227 candidate proteins that passed the second Bonferroni threshold (p <2E-4). These 108 proteins were included as candidate predictors for downstream multivariable LASSO-Cox regression models (**Figure 1**).

LASSO-Cox regression identified a subset of 8 protein markers in a protein-only risk model of incident heart failure (Model 15): pancreatic ribonuclease (RNASE1), WAP four-disulfide core domain protein 2 (HE4), transgelin (TAGLN), fibulin 5 (FBLN5), hepatitis A viral cellular receptor 1, also known as T-cell immunoglobulin and mucin domain 1 (TIM-1), adipocyte fatty acid-binding protein (FABPA), macrophage metalloelastase (MMP-12), and transmembrane glycoprotein nonmetastatic melanoma protein B (GPNMB; **Tables 2 & 3, Figure 2**). Plasma concentrations of all proteins differed between PWH who did and did not have an incident heart failure event (p <0.001 for all; **Supplementary Table S2**). The protein-only model (Model 15) outperformed the clinical risk model (Model 2) in prediction of incident heart failure events in the test set (C-indices 0.725 [95% CI 0.597 – 0.825] vs 0.639 [0.574 – 0.704]; p = 0.033). In combined models, LASSO-Cox regression selected between 6 and 8 proteins in addition to clinical and/or inflammatory markers, all of which included RNASE1, TAGLN, and FBLN5. C-indices for all models incorporating protein markers were $\geq$0.710 (**Table 2**); the protein-only model (Model 15) exhibited the highest C-index among all the combined protein models. There was no meaningful improvement in model performance when inflammatory markers and HIV-specific variables (sCD14, IL-6, D-dimer, CD4+ count, viral load) were added to the protein model (Model 17; C-index 0.720 [0.591 – 0.820]; p = 0.0592 vs Model 15). In the LASSO

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

selection process, none of these variables were retained in the final penalized model, indicating limited incremental predictive value beyond the protein signature.

We compared heart failure event-free survival in the test set (n = 466) stratified into risk quartiles using Model 2 (clinical risk factors) and Model 15 (proteins only; **Figure 3**). The clinical risk factor model did not discriminate between risk quartiles in terms of incident heart failure risk (p = 0.97), while the protein-only model demonstrated divergence between risk quartiles within the first 5 years of follow-up (p <0.0001). Participants in the highest-risk quartile based on the protein model had a Kaplan-Meier-estimated 10-year cumulative incidence of heart failure of approximately 33%, whereas the incidence in the lowest-risk quartile was approximately 2%, demonstrating the broad dynamic range of the protein model.

As a validation exercise, we fit models derived in the PWH training set to VACS participants without HIV (PWoH; n = 715). Their baseline clinical and demographic characteristics are provided in **Supplementary Table S3**. Over 12.9 years of median follow-up, there were 143 incident heart failure events among 715 PWoH in VACS. Model performance in PWoH is summarized in **Supplementary Table S4**. PWH-derived models that included clinical risk scores or individual clinical risk factors performed better in PWoH. The PWH-derived inflammatory marker model (Model 6) performed similarly in PWoH as in PWH (C-indices: 0.589 [95% CI 0.542 – 0.636] vs 0.611 [0.542 – 0.680], respectively). Among the PWH-derived protein-based models, the combined clinical risk factor and protein model (Model 13) had the highest C-index (0.730 [0.639 – 0.806]). This combined PWH-derived model selected 6 proteins, of which 5 were shared with the PWH-derived protein-only model (RNASE1, HE4, TAGLN, FBLN5, TIM-1), with the addition of growth arrest-specific protein 1 (GAS1). The PWH-derived protein-only model (Model 15) had the lowest C-index in PWoH (0.699 [0.608 – 0.777]) among the protein models.

Model calibration

Figure 4 shows a quintile-based calibration curve for the 8-protein incident heart failure risk model in the test set (n = 466). A total of 76 incident heart failure events occurred in the test cohort over the full follow-up period; calibration was evaluated specifically at the 5-year time

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

horizon using Kaplan-Meier estimates. Predicted and observed 5-year survival were closely aligned across quintiles, with overlapping 95% confidence intervals and no evidence of systematic over- or under-prediction, indicating reliable model calibration.

Protein correlations with clinical risk factors and exploratory pathway analysis

Figure 5 shows the correlation matrix between the 8 proteins selected in the final model and established clinical heart failure risk factors, with the corresponding hazard ratio (HR) for each protein shown at the bottom. Positive correlations (purple) indicate that higher protein levels were associated with higher values of the corresponding clinical variable, whereas negative correlations (green) indicate inverse associations. The HR color scale illustrates each protein's effect size on incident heart failure in the test set. Correlations between protein levels and clinical risk factors were generally modest ($|r| < 0.3$), suggesting that these biomarkers capture complementary information beyond that afforded solely by clinical risk factors.

To gain potential mechanistic insights into protein networks involved in incident heart failure, we performed an exploratory pathway enrichment analysis. We based this analysis on the 108 candidate markers for LASSO-Cox regression models from the 4,248 background proteins. This analysis demonstrated significant enrichment for proteins involved in a variety of pathways, including cytokine/immunoregulatory signaling and cellular response to viruses; cellular programming and intracellular signaling; chemotaxis; and extracellular matrix organization (**Supplementary Figures S1-S3**).

The full list of proteins significantly associated with incident heart failure validated in the test set ($m = 108$) and included as candidates in the predictive protein models is presented in **Supplementary Table S5**. Among these, 44 were druggable targets whose protein structures have been delineated and published based on the Therapeutic Target Database.[35]

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Discussion

In this study of proteomic signatures of heart failure in PWH, we quantified 4,248 plasma proteins in a prospective longitudinal cohort of United States Veterans with and without HIV. We identified an 8-protein plasma signature at baseline that predicted incident heart failure in PWH and outperformed traditional, inflammatory, and HIV-specific clinical risk models. Our protein model was well-calibrated and correlations between individual proteins and risk factors were modest, suggesting that the selected protein biomarkers capture biology that is largely independent of established clinical heart failure risk factors and HIV-specific parameters. Among PWH with heart failure, we noted upregulated pathways involved in immune and intracellular signaling, cellular programming, chemotaxis, and extracellular matrix organization, among others. Our findings point to potential biological pathways of heart failure in HIV that can be leveraged for risk stratification and as potential treatment targets in this population.

Protein markers predict clinical heart failure in the general population

To our knowledge, this is the first large-scale proteomic profiling study of incident heart failure in a longitudinal cohort of persons living with HIV. Notably, despite deriving models in PWH, 6 of the 8 proteins in our final model have been described in association with clinical heart failure in the general population without HIV. These prior studies evaluated development of incident heart failure and prognosis in patients with existing heart failure. Three markers—FBLN5, MMP-12, and TAGLN—predicted incident heart failure in a recent proteomic profiling analysis in the general population.[21] Similarly, in patients with chronic kidney disease, another large-scale analysis externally validated three of our markers (FBLN5, TAGLN, HE4) in incident heart failure; in a separate targeted study, HE4 predicted death and readmission in patients with chronic kidney disease hospitalized with acute heart failure.[36,37] Five of the markers we identified have been associated with increased mortality in heart failure with reduced ejection fraction (FABPA, HE4, RNASE1, TAGLN) and heart failure with preserved ejection fraction (FABPA, HE4, MMP-12) in the general population.[38–42] While we derived our final protein model exclusively in PWH, taken together, our results are reinforced by other studies reporting several of the same proteomic markers of clinical heart failure in the general population. Further,

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

the consistency of the results despite the various proteomic platforms used throughout the literature points toward a true association of these proteins with clinical heart failure.

Protein markers of subclinical myocardial remodeling in HIV

None of our final markers of incident heart failure in PWH overlapped with two recent discovery investigations of echocardiographic myocardial remodeling in adults and children with HIV, though proteins identified by both studies have been linked with cardiovascular disease in persons without HIV.[22,23] HIV is an independent risk factor for myocardial disease; while unique HIV-specific mechanisms are thought to mediate this risk, the overlap in our PWH-derived heart failure markers with those published in the general population suggests shared biology of heart failure between PWH and the general population.[2] The consistency of predictive markers across various studies, proteomic platforms, and subpopulations indicates that clinical heart failure pathology favors a final common pathway—regardless of etiology, HIV, or systemic disease substrate—in which these proteins play a biological role and/or reflect heart failure physiology. Indeed, HIV may accelerate or amplify these core heart failure pathways, or act further upstream via HIV-specific mechanisms; this is further supported by the lack of strong correlation between our final proteins and clinical risk factors, including HIV-specific parameters. Similarly, our PWH-derived protein model did not perform as well when fit to VACS BC participants without HIV. While the selected protein markers may be common to both persons with and without HIV, their weighting coefficients likely differ, thus driving the relatively poorer fit in those without HIV. This highlights a limitation in our sample size, as the number of VACS BC participants without HIV was insufficient to support independent model derivation and validation. Further, we used a stringent two-step Bonferroni correction to identify the candidate pool of protein markers for our final models, which may have screened out important markers through an overly conservative selection process.

Potential biological functions of identified proteins and implications in HIV-specific pathways

While our exploratory analyses identified several important pathways upregulated in heart failure, the exact biological role of our final 8 predictors in PWH remains unclear. In addition to

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

clinical prognostication, several of these markers have been associated with structural and functional myocardial abnormalities, including FBLN5 with right ventricular dysfunction, TAGLN with pulmonary arteriolar smooth muscle cell proliferation in congenital heart disease-related pulmonary hypertension, MMP-12 and GPNMB with post-infarct myocardial remodeling, and FABPA with subclinical myocardial abnormalities in obesity and its associated metabolic complications.[39,43–54]

Whether these markers have additional specificity or a role in HIV-associated heart failure pathogenesis remains to be determined. Circulating FABPA, for instance, did not differ by HIV status in one study, but was associated with age, fat composition, BMI, lipids, and cardiovascular risk.[55] Similarly, hepatitis A viral cellular receptor 1, also known as T-cell immunoglobulin and mucin domain 1 (TIM-1), belongs to the family of cell surface glycoproteins that mediate entry of various viruses into T-helper cells, including hepatitis A. Notably, TIM-1 has been shown to block HIV release from infected $CD4^+$ T-cells.[56,57] TIM-1 has been implicated in incident ischemic stroke in a Swedish cohort without HIV, with genetic analyses pointing to a potential causal role.[58] Murine studies suggest TIM-1 has a role in cerebral ischemia-reperfusion injury and atherosclerosis, potentially driven by an imbalance in the Th1/Th2 cell ratio.[59,60] Prospective studies in PWH also signal that higher $CD4^+$ Th1 proportion may be associated with incident heart failure.[61] Whether TIM-1's potential role in heart failure is mediated via inflammation, and whether HIV-specific mechanisms interact with this pathway, remain to be characterized.

Further mechanistic studies on the heart failure markers we identified can also clarify the directionality of risk. Of our 8 final markers, GPNMB was the only one that has not been validated in proteome-wide outcomes studies of clinical heart failure. However, our finding that GPNMB confers increased risk conflicts with clinical studies demonstrating lower circulating levels in patients with heart failure compared to heart failure-free controls, and with other translational studies pointing to its possible role in myocardial regeneration and repair post-infarction.[46] A recent comprehensive translational study found that GPNMB expression is increased in failing human hearts and post-infarct murine hearts; its deficiency led to rapid post-infarct left ventricular dysfunction, cardiac rupture, and increased mortality. Conversely, the same study noted that by increasing circulating GPNMB levels, there was enhanced myocyte contraction, reduced fibroblast activation, and improved post-infarct myocardial function.[47] We

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

also noted an increased risk of heart failure associated with MMP-12, consistent with another large-scale proteomics analysis in the general population and in line with other clinical and translational studies demonstrating association with atherosclerosis development and incident coronary events.[21,62,63] This is in contrast with murine models showing that MMP-12 may have a role in post-infarct myocardial recovery by inhibiting neutrophil infiltration.[54,64] In aggregate, these observations suggest that rather than having causal or harmful functions predisposing to heart failure, GPNMB and MMP-12 are upregulated due to their potential protective role in myocardial tissue repair. Further mechanistic and translational studies are needed to clarify these pathways, particularly if their pharmacologic amplification could improve heart failure outcomes in PWH.

Risk prediction and implications for PWH

While the exact biological roles of our final proteins are not yet defined, our protein model outperformed clinical models in heart failure risk prediction. The broader implications for PWH are significant, especially as atherosclerotic cardiovascular risk prediction algorithms developed in the general population translate poorly in PWH.[18–20] Further, all-cause mortality in heart failure is higher among PWH than their peers without HIV, highlighting the need for tailored risk stratification tools in this population.[15] In our study, a proteomic signature measured at a single baseline time point effectively predicted incident heart failure events over long-term follow-up, outperforming clinical and HIV-specific risk models. This is in line with other proteomic profiling studies of heart failure in special populations, including a recent study in patients with chronic kidney disease, which showed that protein risk prediction models consistently outperformed the 2019 Pooled Cohort Equations to Prevent Heart Failure (PCP-HF).[16,36] Importantly, the input range of clinical parameters for PCP-HF and the 2024 American Heart Association Predicting Risk of CVD EVENTs (PREVENT) Heart Failure equations precluded estimation of risk for up to a quarter of our VACS BC participants; the most commonly out-of-range input variables were body mass index and total cholesterol. While this highlights a limitation of our approach—inability to evaluate performance of contemporary heart failure risk prediction algorithms in comparison to our protein models—the narrow range of inputs likely

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

limits the applicability of these tools to many populations even beyond PWH and lends further support for a protein-based approach.

Limitations

Despite the strength of our sample and the rigorous study design, our findings must be interpreted in the context of several limitations. Firstly, this was a cohort composed of US Veterans, and consequently primarily male. This raises a second limitation—the lack of an external validation cohort—to assess the generalizability of our results to broader populations of PWH, including females. We also lacked statistical power to examine subtypes of heart failure, such as heart failure with reduced and preserved ejection fraction, individually.

Conclusions

In summary, we derived a baseline 8-protein plasma signature in PWH that predicted incident heart failure events over 13 years of median follow-up, outperforming clinical, HIV-specific, and inflammatory/coagulation markers. These proteins have been described in relation to heart failure or myocardial remodeling in persons without HIV, suggesting that their biological role may be core to heart failure physiology, regardless of HIV status. Pathways of extracellular matrix organization and cellular and immune signaling were upregulated in our exploratory analysis of 108 candidate markers, but the exact biology of the final 8 protein markers in our predictive model and their mechanistic role in heart failure remain to be determined. To our knowledge, this is the first study to report proteomic markers of incident heart failure in a longitudinal cohort of PWH. While further validation in external cohorts is needed, protein prediction can be used to identify PWH at highest risk for heart failure and druggable targets can be leveraged to develop novel therapeutic targets with the ultimate goal of preventing clinical disease. Future translational research should aim to bridge clinical risk markers with mechanistic pathways underlying heart failure in this population.

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

References

1. Butt AA, Chang CC, Kuller L, et al. Risk of Heart Failure With Human Immunodeficiency Virus in the Absence of Prior Diagnosis of Coronary Heart Disease. *Arch Intern Med*. 2011;171(8):737-743. doi:10.1001/archinternmed.2011.151

2. Freiberg MS, Chang CCH, Skanderson M, et al. Association Between HIV Infection and the Risk of Heart Failure With Reduced Ejection Fraction and Preserved Ejection Fraction in the Antiretroviral Therapy Era: Results From the Veterans Aging Cohort Study. *JAMA Cardiology*. 2017;2(5):536-546. doi:10.1001/jamacardio.2017.0264

3. Triant VA, Lee H, Hadigan C, et al. Increased Acute Myocardial Infarction Rates and Cardiovascular Risk Factors among Patients with Human Immunodeficiency Virus Disease. *The Journal of Clinical Endocrinology & Metabolism*. 2007;92(7):2506-2512. doi:10.1210/jc.2006-2190

4. Helleberg M, Afzal S, Kronborg G, et al. Mortality attributable to smoking among HIV-1-infected individuals: a nationwide, population-based cohort study. *Clin Infect Dis*. 2013;56(5):727-734. doi:10.1093/cid/cis933

5. Wong C, Gange SJ, Moore RD, et al. Multimorbidity Among Persons Living with Human Immunodeficiency Virus in the United States. *Clin Infect Dis*. 2018;66(8):1230-1238. doi:10.1093/cid/cix998

6. Duprez DA, Neuhaus J, Kuller LH, et al. Inflammation, coagulation and cardiovascular disease in HIV-infected individuals. *PLoS One*. 2012;7(9):e44454. doi:10.1371/journal.pone.0044454

7. Armah KA, McGinnis K, Baker J, et al. HIV Status, Burden of Comorbid Disease, and Biomarkers of Inflammation, Altered Coagulation, and Monocyte Activation. *Clin Infect Dis*. 2012;55(1):126-136. doi:10.1093/cid/cis406

8. Tenorio AR, Zheng Y, Bosch RJ, et al. Soluble markers of inflammation and coagulation but not T-cell activation predict non-AIDS-defining morbid events during suppressive antiretroviral treatment. *J Infect Dis*. 2014;210(8):1248-1259. doi:10.1093/infdis/jiu254

9. Baker JV, Sharma S, Grund B, et al. Systemic Inflammation, Coagulation, and Clinical Risk in the START Trial. *Open Forum Infect Dis*. 2017;4(4):ofx262. doi:10.1093/ofid/ofx262

10. Thiara DK, Liu CY, Raman F, et al. Abnormal Myocardial Function Is Related to Myocardial Steatosis and Diffuse Myocardial Fibrosis in HIV-Infected Adults. *J Infect Dis*. 2015;212(10):1544-1551. doi:10.1093/infdis/jiv274

11. Luetkens JA, Doerner J, Schwarze-Zander C, et al. Cardiac Magnetic Resonance Reveals Signs of Subclinical Myocardial Inflammation in Asymptomatic HIV-Infected Patients. *Circulation: Cardiovascular Imaging*. 2016;9(3):e004091. doi:10.1161/CIRCIMAGING.115.004091

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

12. Kingery JR, Goyal P, Hosalli R, et al. Human Immunodeficiency Virus–Associated Myocardial Diastolic Dysfunction and Soluble ST2 Concentration in Tanzanian Adults: A Cross-Sectional Study. *J Infect Dis*. 2020;223(1):83-93. doi:10.1093/infdis/jiaa328

13. Shuldiner SR, Wong LY, Peterson TE, et al. Myocardial Fibrosis Among Antiretroviral Therapy-Treated Persons With Human Immunodeficiency Virus in South Africa. *Open Forum Infectious Diseases*. 2021;8(1):ofaa600. doi:10.1093/ofid/ofaa600

14. Al-Kindi SG, ElAmm C, Ginwalla M, et al. Heart failure in patients with human immunodeficiency virus infection: Epidemiology and management disparities. *Int J Cardiol*. 2016;218:43-46. doi:10.1016/j.ijcard.2016.05.027

15. Erqou S, Jiang L, Choudhary G, et al. Heart Failure Outcomes and Associated Factors Among Veterans With Human Immunodeficiency Virus Infection. *JACC: Heart Failure*. 2020;8(6):501-511. doi:10.1016/j.jchf.2019.12.007

16. Khan SS, Ning H, Shah SJ, et al. 10-Year Risk Equations for Incident Heart Failure in the General Population. *Journal of the American College of Cardiology*. 2019;73(19):2388-2397. doi:10.1016/j.jacc.2019.02.057

17. Khan SS, Matsushita K, Sang Y, et al. Development and Validation of the American Heart Association's PREVENT Equations. *Circulation*. 2024;149(6):430-449. doi:10.1161/CIRCULATIONAHA.123.067626

18. Triant VA, Perez J, Regan S, et al. Cardiovascular Risk Prediction Functions Underestimate Risk in HIV Infection. *Circulation*. 2018;137(21):2203-2214. doi:10.1161/CIRCULATIONAHA.117.028975

19. Triant VA, Lyass A, Hurley LB, et al. Cardiovascular Risk Estimation Is Suboptimal in People With HIV. *Journal of the American Heart Association*. 2024;13(10):e029228. doi:10.1161/JAHA.123.029228

20. Grinspoon SK, Zanni MV, Triant VA, et al. Performance of the pooled cohort equations and D:A:D risk scores among individuals with HIV in a global cardiovascular disease prevention trial: a cohort study leveraging data from REPRIEVE. *The Lancet HIV*. 2025;12(2):e118-e129. doi:10.1016/S2352-3018(24)00276-5

21. Shah AM, Myhre PL, Arthur V, et al. Large scale plasma proteomics identifies novel proteins and protein networks associated with heart failure development. *Nat Commun*. 2024;15(1):528. doi:10.1038/s41467-023-44680-3

22. Peterson TE, Hahn VS, Moaddel R, et al. Proteomic signature of HIV-associated subclinical left atrial remodeling and incident heart failure. *Nat Commun*. 2025;16(1):610. doi:10.1038/s41467-025-55911-0

23. Harrington J, McCrary AW, Nguyen M, et al. Proteomics discovery in children and young adults with HIV identifies fibrosis, inflammatory, and immune biomarkers associated with myocardial impairment. *AIDS*. 2024;38(7):1090. doi:10.1097/QAD.0000000000003879

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

24. Justice AC, Dombrowski E, Conigliaro J, et al. Veterans Aging Cohort Study (VACS): Overview and description. *Med Care*. 2006;44(8 Suppl 2):S13-24. doi:10.1097/01.mlr.0000223741.02074.66

25. Justice AC, Freiberg MS, Tracy R, et al. Does an Index Composed of Clinical Data Reflect Effects of Inflammation, Coagulation, and Monocyte Activation on Mortality Among Those Aging With HIV? *Clin Infect Dis*. 2012;54(7):984-994. doi:10.1093/cid/cir989

26. Sun BB, Maranville JC, Peters JE, et al. Genomic atlas of the human plasma proteome. *Nature*. 2018;558(7708):73-79. doi:10.1038/s41586-018-0175-2

27. Williams SA, Ostroff R, Hinterberg MA, et al. A proteomic surrogate for cardiovascular outcomes that is sensitive to multiple mechanisms of change in risk. *Science Translational Medicine*. 2022;14(639):eabj9625. doi:10.1126/scitranslmed.abj9625

28. Saczynski JS, Andrade SE, Harrold LR, et al. A systematic review of validated methods for identifying heart failure using administrative data. *Pharmacoepidemiol Drug Saf*. 2012;21 Suppl 1(0 1):129-140. doi:10.1002/pds.2313

29. The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2025. *Nucleic Acids Res*. 2025;53(D1):D609-D617. doi:10.1093/nar/gkae1010

30. Durinck S, Moreau Y, Kasprzyk A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*. 2005;21(16):3439-3440. doi:10.1093/bioinformatics/bti525

31. Durinck S, Spellman PT, Birney E, et al. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc*. 2009;4(8):1184-1191. doi:10.1038/nprot.2009.97

32. Thomas PD, Ebert D, Muruganujan A, et al. PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Science*. 2022;31(1):8-22. doi:10.1002/pro.4218

33. Kanehisa M, Furumichi M, Sato Y, et al. KEGG: biological systems database as a model of the real world. *Nucleic Acids Res*. 2025;53(D1):D672-D677. doi:10.1093/nar/gkae909

34. Milacic M, Beavers D, Conley P, et al. The Reactome Pathway Knowledgebase 2024. *Nucleic Acids Res*. 2024;52(D1):D672-D678. doi:10.1093/nar/gkad1025

35. Zhou Y, Zhang Y, Zhao D, et al. TTD: Therapeutic Target Database describing target druggability information. *Nucleic Acids Res*. 2024;52(D1):D1465-D1477. doi:10.1093/nar/gkad751

36. Dubin RF, Deo R, Ren Y, et al. Incident heart failure in chronic kidney disease: proteomics informs biology and risk stratification. *Eur Heart J*. 2024;45(30):2752-2767. doi:10.1093/eurheartj/ehae288

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

37. Tang Y, Hu Z, Liu Z, et al. HE4 Serum Levels are Associated with Poor Prognosis in Patients with Acute Heart Failure Combined with Chronic Kidney Disease. *Int J Gen Med*. 2024;17:1273-1280. doi:10.2147/IJGM.S444680

38. de Boer RA, Cao Q, Postmus D, et al. The WAP Four-Disulfide Core Domain Protein HE4: A Novel Biomarker for Heart Failure. *JACC: Heart Failure*. 2013;1(2):164-169. doi:10.1016/j.jchf.2012.11.005

39. Piek A, Meijers WC, Schroten NF, et al. HE4 Serum Levels Are Associated with Heart Failure Severity in Patients With Chronic Heart Failure. *J Card Fail*. 2017;23(1):12-19. doi:10.1016/j.cardfail.2016.05.002

40. Chirinos JA, Orlenko A, Zhao L, et al. Multiple Plasma Biomarkers for Risk Stratification in Patients With Heart Failure and Preserved Ejection Fraction. *Journal of the American College of Cardiology*. 2020;75(11):1281-1295. doi:10.1016/j.jacc.2019.12.069

41. Dib M, Levin MG, Zhao L, et al. Proteomic Associations of Adverse Outcomes in Human Heart Failure. *Journal of the American Heart Association*. 2024;13(5):e031154. doi:10.1161/JAHA.123.031154

42. Chadwick J, Hinterberg MA, Asselbergs FW, et al. Harnessing the Plasma Proteome to Predict Mortality in Heart Failure Subpopulations. *Circulation: Heart Failure*. 2025;18(4):e011208. doi:10.1161/CIRCHEARTFAILURE.123.011208

43. Běhounek M, Lipcseyová D, Vít O, et al. Biomarkers of RV Dysfunction in HFrEF Identified by Direct Tissue Proteomics: Extracellular Proteins Fibromodulin and Fibulin-5. *Circulation: Heart Failure*. 2025;18(3):e011984. doi:10.1161/CIRCHEARTFAILURE.124.011984

44. Huang L, Li L, Yang T, et al. Transgelin as a potential target in the reversibility of pulmonary arterial hypertension secondary to congenital heart disease. *J Cell Mol Med*. 2018;22(12):6249-6261. doi:10.1111/jcmm.13912

45. Zhou J jing, Yang J, Li L, et al. Transgelin exacerbates pulmonary artery smooth muscle cell dysfunction in shunt-related pulmonary arterial hypertension. *ESC Heart Failure*. 2022;9(5):3407-3417. doi:10.1002/ehf2.14080

46. Lin LY, Chun Chang S, O'Hearn J, et al. Systems Genetics Approach to Biomarker Discovery: GPNMB and Heart Failure in Mice and Humans. *G3 Genes|Genomes|Genetics*. 2018;8(11):3499-3506. doi:10.1534/g3.118.200655

47. Ramadoss S, Qin J, Tao B, et al. Bone-marrow macrophage-derived GPNMB protein binds to orphan receptor GPR39 and plays a critical role in cardiac repair. *Nature Cardiovascular Research*. 2024;3(11):1356-1373. doi:10.1038/s44161-024-00555-4

48. Fuseya T, Furuhashi M, Yuda S, et al. Elevation of circulating fatty acid-binding protein 4 is independently associated with left ventricular diastolic dysfunction in a general population. *Cardiovasc Diabetol*. 2014;13:126. doi:10.1186/s12933-014-0126-7

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

49. Liu M, Zhou M, Bao Y, et al. Circulating adipocyte fatty acid-binding protein levels are independently associated with heart failure. *Clin Sci (Lond)*. 2013;124(2):115-122. doi:10.1042/CS20120004

50. Balci MM, Arslan U, Firat H, et al. Serum levels of adipocyte fatty acid-binding protein are independently associated with left ventricular mass and myocardial performance index in obstructive sleep apnea syndrome. *J Investig Med*. 2012;60(7):1020-1026. doi:10.2310/JIM.0b013e31826868f2

51. Engeli S, Utz W, Haufe S, et al. Fatty acid binding protein 4 predicts left ventricular mass and longitudinal function in overweight and obese women. *Heart*. 2013;99(13):944-948. doi:10.1136/heartjnl-2013-303735

52. Huang CL, Wu YW, Wu CC, et al. Association between serum adipocyte fatty-acid binding protein concentrations, left ventricular function and myocardial perfusion abnormalities in patients with coronary artery disease. *Cardiovasc Diabetol*. 2013;12:105. doi:10.1186/1475-2840-12-105

53. Baessler A, Lamounier-Zepter V, Fenk S, et al. Adipocyte fatty acid-binding protein levels are associated with left ventricular diastolic dysfunction in morbidly obese subjects. *Nutr Diabetes*. 2014;4(2):e106. doi:10.1038/nutd.2014.3

54. Iyer RP, Patterson NL, Zouein FA, et al. Early matrix metalloproteinase-12 inhibition worsens post-myocardial infarction cardiac dysfunction by delaying inflammation resolution. *International Journal of Cardiology*. 2015;185:198-208. doi:10.1016/j.ijcard.2015.03.054

55. Morieri ML, Guardigni V, Sanz JM, et al. Adipokines levels in HIV infected patients: lipocalin-2 and fatty acid binding protein-4 as possible markers of HIV and antiretroviral therapy-related adipose tissue inflammation. *BMC Infectious Diseases*. 2018;18(1):10. doi:10.1186/s12879-017-2925-4

56. Li M, Ablan SD, Miao C, et al. TIM-family proteins inhibit HIV-1 release. *Proc Natl Acad Sci U S A*. 2014;111(35):E3699-3707. doi:10.1073/pnas.1404851111

57. Vlach J, Saad JS. HIV: a vicTIM. *Trends in Microbiology*. 2014;22(11):603-604. doi:10.1016/j.tim.2014.09.008

58. Song L, Sun J, Söderholm M, et al. Association of TIM-1 (T-Cell Immunoglobulin and Mucin Domain 1) With Incidence of Stroke. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 2020;40(7):1777-1786. doi:10.1161/ATVBAHA.120.314269

59. Zheng Y, Wang L, Chen M, et al. Inhibition of T cell immunoglobulin and mucin-1 (TIM-1) protects against cerebral ischemia-reperfusion injury. *Cell Communication and Signaling*. 2019;17(1):103. doi:10.1186/s12964-019-0417-4

60. Douna H, Smit V, van Puijvelde GHM, et al. Tim-1 mucin domain-mutant mice display exacerbated atherosclerosis. *Atherosclerosis*. 2022;352:1-9. doi:10.1016/j.atherosclerosis.2022.05.017

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

61. Kundu S, Freiberg MS, Tracy RP, et al. Circulating T Cells and Cardiovascular Risk in People With and Without HIV Infection. *Journal of the American College of Cardiology*. 2022;80(17):1633-1644. doi:10.1016/j.jacc.2022.08.756

62. Amor M, Bianco V, Buerger M, et al. Genetic deletion of MMP12 ameliorates cardiometabolic disease by improving insulin sensitivity, systemic inflammation, and atherosclerotic features in mice. *Cardiovascular Diabetology*. 2023;22(1):327. doi:10.1186/s12933-023-02064-3

63. Goncalves I, Bengtsson E, Colhoun HM, et al. Elevated Plasma Levels of MMP-12 Are Associated With Atherosclerotic Burden and Symptomatic Cardiovascular Disease in Subjects With Type 2 Diabetes. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 2015;35(7):1723-1731. doi:10.1161/ATVBAHA.115.305631

64. Kubota A, Suto A, Suzuki K, et al. Matrix metalloproteinase-12 produced by Ly6Clow macrophages prolongs the survival after myocardial infarction by preventing neutrophil influx. *Journal of Molecular and Cellular Cardiology*. 2019;131:41-52. doi:10.1016/j.yjmcc.2019.04.007

65. Cabrera-Fuentes HA, Ruiz-Meana M, Simsekyilmaz S, et al. RNase1 prevents the damaging interplay between extracellular RNA and tumour necrosis factor-α in cardiac ischaemia/reperfusion injury. *Thromb Haemost.* 2014;112(6):1110-1119. doi:10.1160/TH14-08-0703

66. Assinder SJ, Stanton JAL, Prasad PD. Transgelin: an actin-binding protein and tumour suppressor. *Int J Biochem Cell Biol*. 2009;41(3):482-486. doi:10.1016/j.biocel.2008.02.011

67. Yanagisawa H, Schluterman MK, Brekken RA. Fibulin-5, an integrin-binding matricellular protein: its function in development and disease. *J Cell Commun Signal*. 2009;3(3-4):337-347. doi:10.1007/s12079-009-0065-3

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Acknowledgments

The authors acknowledge the veterans who participated in the Veterans Aging Cohort Study and the study coordinators and staff at each respective enrollment and coordination sites, without whom this research would not be possible.

Author contributions

SSS, PYH, MF, and PG designed the study. PG supervised the laboratory work and quality control of the proteomics assays. MF and KS provided access to the clinical data. TG performed the data analysis. SSS drafted the manuscript. All authors reviewed and critically revised the final manuscript and are accountable for all aspects of the work presented.

Competing interests

PG serves on the Medical Advisory Board to SomaLogic without renumeration. PYH has received honoraria from Genentech, Viiv Healthcare, Gilead, Pfizer, and Merck. PYH has received grant support from Abbott and Excision Biotherapeutics and serves on their Scientific Advisory Boards. The other authors report no competing interests.

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Tables and Figures

Table 1: Baseline demographic and clinical characteristics among PLWH in the Veterans Aging Cohort Study Biomarker Cohort, stratified by incident heart failure event

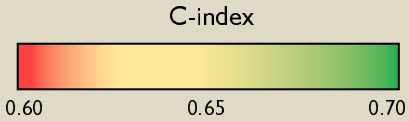
		Incident Heart Failure		
	Overall	No	Yes	P-value
	(N=1398)	(N=1158)	(N=240)	
Mean age, years (SD)	52.1 (8.1)	51.5 (8.1)	54.9 (7.6)	**<0.001**
Female, n (%)	41 (2.9%)	38 (3.3%)	3 (1.2%)	0.090
Race/Ethnicity, n (%)				0.246
White	266 (19.0%)	216 (18.7%)	50 (20.8%)	
Black	960 (68.7%)	805 (69.5%)	155 (64.6%)	
Hispanic	118 (8.4%)	97 (8.4%)	21 (8.8%)	
Other	54 (3.9%)	40 (3.5%)	14 (5.8%)	
Medical history, n (%)				
Hypertension	659 (47.1%)	503 (43.4%)	156 (65.0%)	**<0.001**
On antihypertensive treatment	952 (68.1%)	759 (65.5%)	193 (80.4%)	<0.001
Diabetes	249 (17.8%)	169 (14.6%)	80 (33.3%)	**<0.001**
Hyperlipidemia	452 (32.3%)	355 (30.7%)	97 (40.4%)	**0.003**
On lipid-lowering therapy	402 (28.8%)	317 (27.4%)	85 (35.4%)	0.012
Overweight/obese (body mass index >25 kg/m^2)	736 (52.8%)	604 (52.3%)	132 (55.0%)	0.453
Current smoker	689 (49.3%)	557 (48.1%)	132 (55.0%)	0.052
Cocaine or crack use	294 (22.1%)	233 (21.1%)	61 (27.2%)	0.120
Methamphetamine and other stimulant use	59 (4.5%)	50 (4.5%)	9 (4.1%)	0.820
Prior ASCVD event	69 (4.9%)	45 (3.9%)	24 (10.0%)	**<0.001**
Mean (SD) systolic blood pressure, mmHg	128.3 (14.2)	127.8 (13.9)	130.8 (15.5)	**0.003**
Mean (SD) body mass index, kg/m²	25.9 (4.8)	25.8 (4.7)	26.2 (5.5)	0.261
On antiretroviral regimen at baseline, n (%)	1144 (81.8%)	943 (81.4%)	201 (83.8%)	0.397
NRTI	1156 (82.7%)	950 (82.0%)	206 (85.8%)	0.157
NNRTI	499 (35.7%)	418 (36.1%)	81 (33.8%)	0.490
Mean (SD) laboratory values				
Total Cholesterol, mg/dL	176.6 (43.1)	177.2 (44.2)	173.7 (37.3)	0.259
LDL Cholesterol, mg/dL	99.9 (34.6)	100.4 (34.8)	97.4 (34.0)	0.244
HDL Cholesterol, mg/dL	44.5 (16.8)	44.6 (16.6)	43.9 (17.3)	0.548
Estimated glomerular filtration rate, mL/min/1.73 m²	99.0 (30.4)	99.7 (28.9)	95.6 (36.8)	0.054
CD4$^+$ count, cells/mm³	446 (286)	451 (289)	417 (270)	0.090
Nadir CD4$^+$ count, cells/mm³	229 (178)	232 (179)	218 (175)	0.269
HIV viral load, copies/mL	20623 (75828)	20441 (77896)	21499 (65085)	0.844
Soluble CD14 (sCD14), ng/mL	1810 (543)	1777 (529)	1971 (579)	**<0.001**
D-dimer, ng/mL	0.5 (1.0)	0.4 (0.7)	0.6 (1.8)	**0.006**
Interleukin-6 (IL-6), pg/mL	3.2 (7.5)	3.2 (7.7)	3.6 (6.4)	0.425

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2: Predictive performance of multivariable models for incident heart failure among PLWH

Model	Covariates included in model	Proteins selected in model	Event-free survival C-index (95% confidence interval) Training set (n = 932)	Test set (n = 466)
	CLINICAL RISK MODELS			
1	ASCVD 10-year risk score[1]	--	0.681 (0.640 - 0.722)	0.607 (0.544 - 0.670)
2	Clinical heart failure risk factors[2]	--	0.724 (0.687 - 0.761)	0.639 (0.574 - 0.704)
3a	Model 1 + eGFR[3]	--	0.661 (0.616 - 0.706)	0.607 (0.544 - 0.670)
3b	Model 2 + eGFR	--	0.724 (0.685 - 0.763)	0.640 (0.575 - 0.705)
4a	Model 1 + eGFR + stimulant use	--	0.685 (0.640 - 0.730)	0.630 (0.567 - 0.693)
4b	Model 2 + eGFR + stimulant use	--	0.738 (0.699 - 0.777)	0.647 (0.580 - 0.714)
5a	Model 3a + $CD4^+$ count + viral load	--	0.707 (0.666 - 0.748)	0.637 (0.576 - 0.698)
5b	Model 3b + $CD4^+$ count + viral load	--	0.754 (0.717 - 0.791)	0.656 (0.591 - 0.721)
	INFLAMMATORY MARKER RISK MODELS			
6	sCD14 + IL-6 + D-dimer	--	0.638 (0.593 - 0.683)	0.611 (0.542 - 0.680)
7	ASCVD score + sCD14 + IL-6 + D-dimer	--	0.711 (0.672 - 0.750)	0.632 (0.567 - 0.697)
8	Clinical heart failure risk factors[2] + sCD14 + IL-6 + D-dimer	--	0.749 (0.712 - 0.786)	0.645 (0.580 - 0.710)
9	Model 6 + $CD4^+$ count + viral load	--	0.641 (0.598 - 0.684)	0.600 (0.529 - 0.671)
10	Model 7 + $CD4^+$ count + viral load		0.720 (0.683 - 0.757)	0.632 (0.567 - 0.697)
11	Model 8 + $CD4^+$ count + viral load		0.757 (0.722 - 0.792)	0.657 (0.594 - 0.720)
	PROTEIN RISK MODELS			
12	ASCVD score + **proteins**	RNASE1, NBL1, HE4, GAS1, FBLN3, TAGLN, FBLN5, TIM-1	0.740 (0.651 - 0.812)	0.713 (0.582 - 0.815)
13	Clinical heart failure risk factors[2] + **proteins**	RNASE1, HE4, GAS1, TAGLN, FBLN5, TIM-1	0.753 (0.666 - 0.824)	0.717 (0.587 - 0.819)
14	ASCVD score + sCD14 + D-dimer + IL-6 + **proteins**	RNASE1, NBL1, HE4, GAS1, TAGLN, FBLN5, TIM-1	0.740 (0.651 - 0.812)	0.710 (0.579 - 0.814)
15	**Proteins only**	RNASE1, HE4, TAGLN, FBLN5, TIM-1, FABPA, MMP-12, GPNMB	0.719 (0.632 - 0.792)	0.725 (0.597 - 0.825)
16	**Proteins** + $CD4^+$ count + viral load	RNASE1, HE4, PXDN, TAGLN, FBLN5, TIM-1, GPNMB	0.716 (0.629 - 0.789)	0.715 (0.587 - 0.816)
17	**Proteins** + $CD4^+$ count + viral load + sCD14 + IL-6 + D-dimer	RNASE1, HE4, PXDN, TAGLN, FBLN5, TIM-1, GPNMB	0.714 (0.627 - 0.789)	0.720 (0.591 - 0.820)

[1] ACC/AHA Pooled Cohort Equations
[2] Age; sex; race; systolic blood pressure; antihypertensive medication use; diabetes; current smoker; body mass index; total cholesterol; high-density lipoprotein cholesterol
[3] eGFR, estimated glomerular filtration rate
[4] sCD14, soluble CD14; IL-6, interleukin-6



medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 3: Major functions of individual proteins in the final 8-protein incident heart failure model.

Protein	Penalized Hazard ratio (LASSO)	Biological function(s)	Relevant drug
Pancreatic ribonuclease (**RNASE1**)	1.19	Digests extracellular RNA; potential protective effect in myocardial ischemia-reperfusion injury in mouse models;[65] associated with increased mortality in HFrEF in humans[41]	2-(4-Fluoro-indole-1-sulfonyl)-benzoic acid (investigative small molecule)
WAP four-disulfide core domain protein 2 (**HE4**)	1.47	Also known as human epididymis protein 4; correlated with heart failure severity, renal function, and adverse heart failure outcomes[39]	--
Transgelin (**TAGLN**)	1.08	Actin-binding protein that regulates smooth muscle cell contraction;[66] upregulated in pulmonary arterioles in pulmonary arterial hypertension[44]	--
Fibulin-5 (**FBLN5**)	2.26	Extracellular matrix protein secreted by vascular smooth muscle cells, fibroblasts, and endothelial cells;[67] predicts heart failure in the general population[21]	--
Hepatitis A virus cellular receptor 2, or T-cell immunoglobulin and mucin domain 1 (**TIM-1**)	1.19	Mediates entry of viruses into T-helper cells (including hepatitis A) but blocks HIV release from infected $CD4^+$ T-cells;[56,57] predicts ischemic stroke and may be involved in cerebral ischemia-reperfusion injury and atherosclerosis[58–60]	--
Adipocyte fatty acid binding protein (**FABPA**)	1.01	Fat-derived circulating protein associated with adverse myocardial remodeling in obesity and heart failure prognosis [48–53]	--
Macrophage metalloelastase 12 (**MMP-12**)	1.01	Macrophage-derived protease involved in extracellular matrix remodeling and elastin degradation; involved in post-infarct myocardial remodeling in murine models (Iyer, Kubota),[54,64] incident heart failure,[21] and HFpEF outcomes[42]	FP-025 (phase I, asthma)
Transmembrane glycoprotein nonmetastatic melanoma protein B (**GPNMB**)	1.14	Type I transmembrane protein implicated in cancer progression and immune regulation and organ tissue repair; decreased circulating levels in humans with heart failure,[46] with mouse and in vivo models pointing to a role in myocardial regeneration and repair[47]	--

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

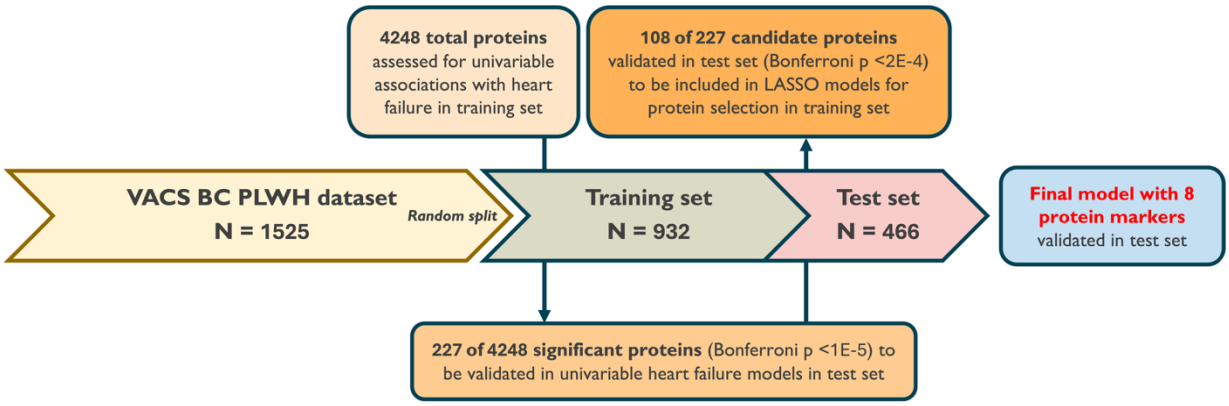


Figure 1: Candidate protein selection process for inclusion in incident heart failure models.

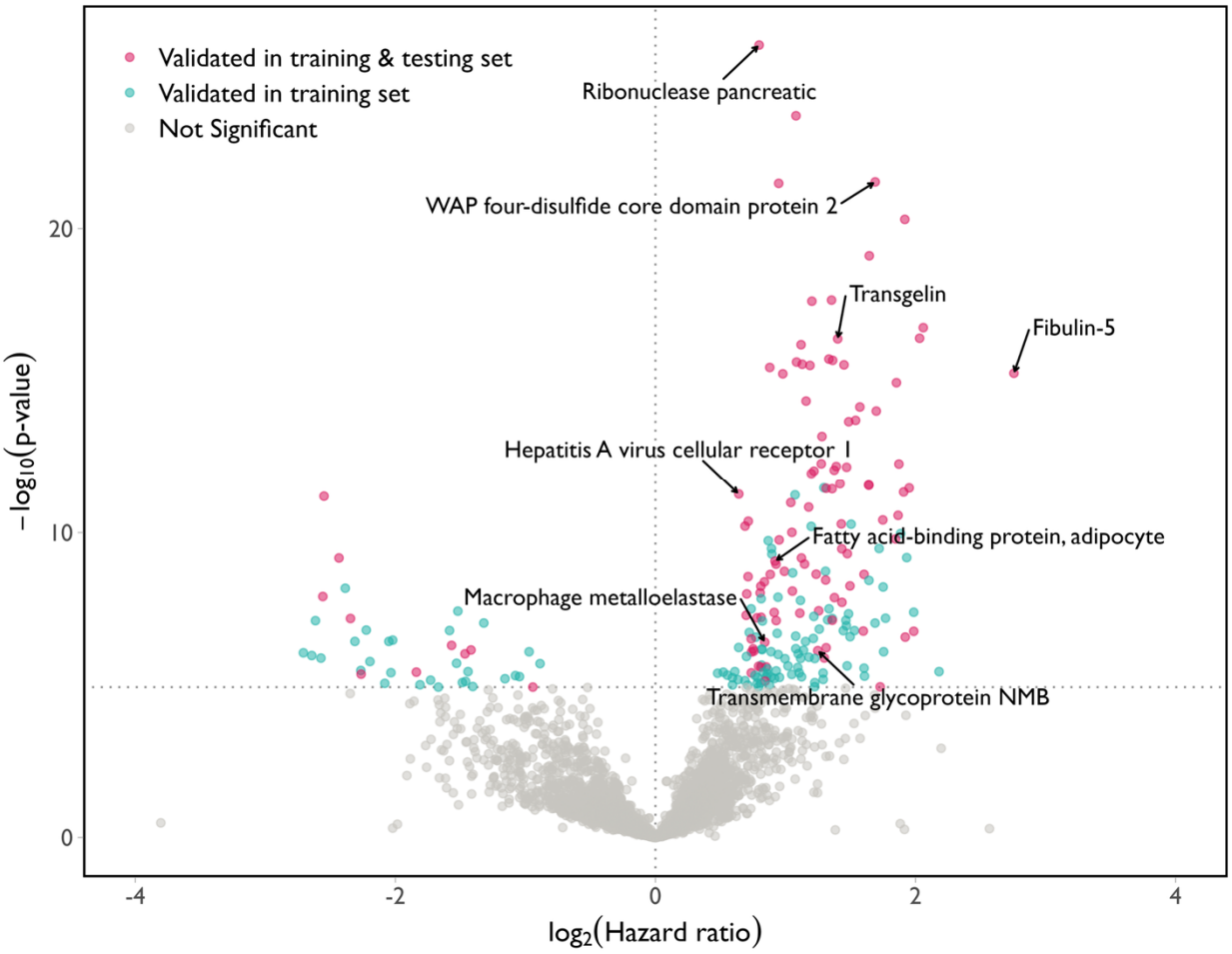


Figure 2: Proteins in sequential phases of selection and model development. Horizontal and vertical dotted lines represent first-round Bonferroni-corrected significance threshold (p <1E-5) and hazard ratio of 1.0, respectively. Markers selected in final protein-only model of incident heart failure are labeled.

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

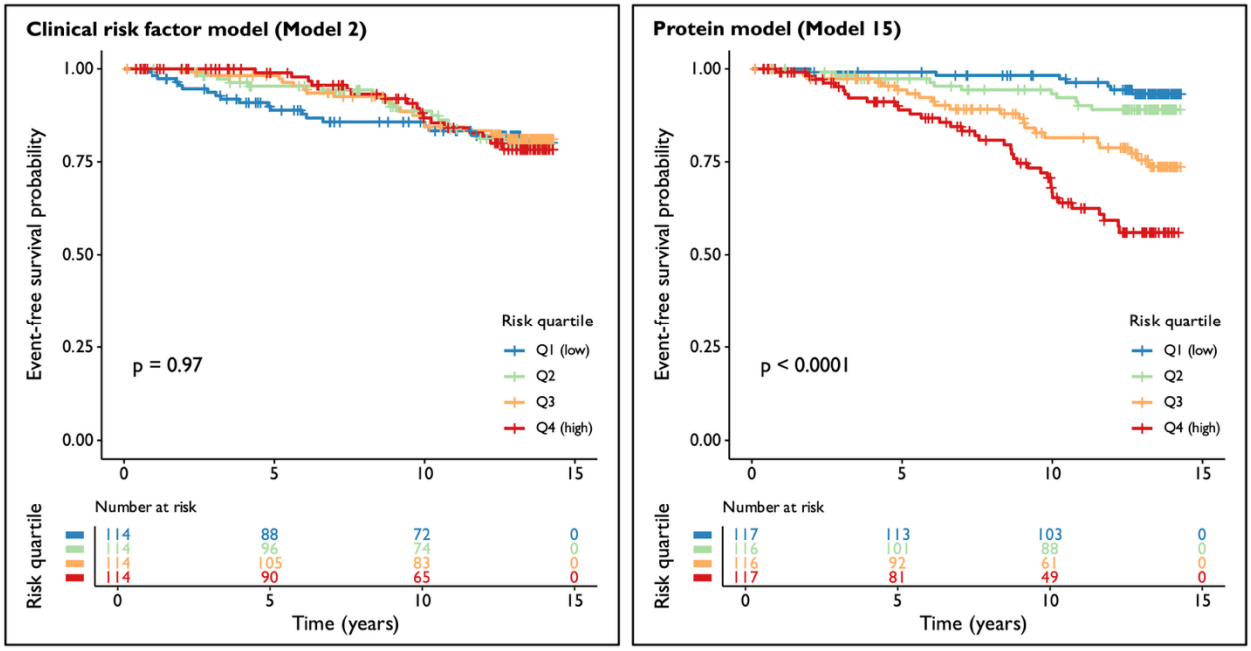


Figure 3: Quartile-based Kaplan-Meier curves for incident heart failure-free survival in test set of PWH (n = 466). Individuals were stratified into predicted risk quartiles based on models incorporating clinical risk factors (Model 2) and proteins only (Model 15). The clinical model included 455 total participants, as 10 participants were missing one or more clinical risk factor parameters, and thus could not be risk-stratified. By contrast, protein data was available for all 466 participants. The protein-only model demonstrated improved risk discrimination compared with the clinical model, particularly in the highest-risk quintile.

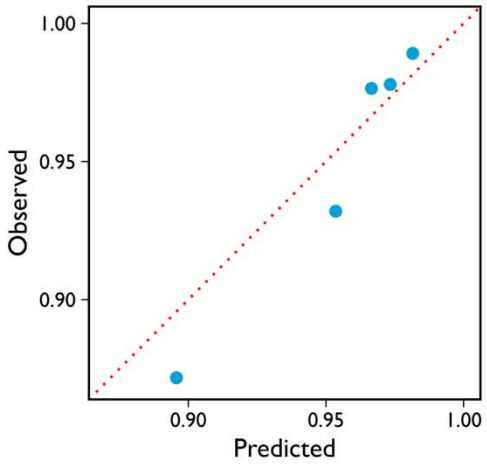


Figure 4: Quintile-based calibration curve in test set (n = 466) for observed vs predicted 5-year heart failure-free survival based on protein-only model. All 95% confidence intervals for each point (not shown) overlap with the dashed red line, which represents perfect calibration between observed and predicted events.

medRxiv preprint doi: https://doi.org/10.64898/2026.01.08.25343149; this version posted January 11, 2026. The copyright holder for this preprint
(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

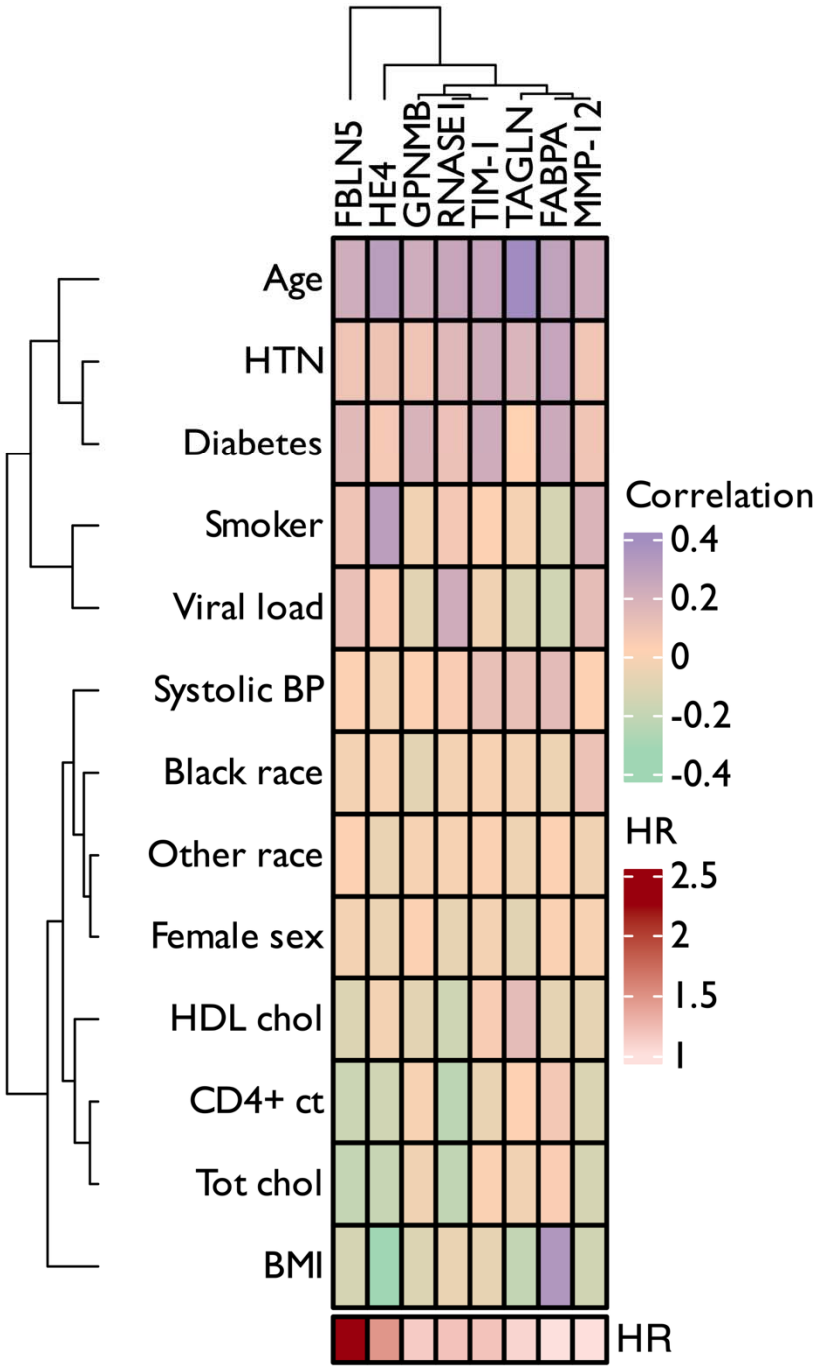


Figure 4: Correlation matrix between clinical risk factors and final 8 selected protein markers in incident heart failure models. Proteins are ordered by hierarchical clustering of their penalized hazard ratios, and clinical variables are hierarchically clustered based on similarity of their correlation profiles across the 8 proteins, thereby grouping variables with similar correlation patterns.