

Role of Plasma Proteomics in Predicting the Prognosis of Older Adult Patients with Chronic Coronary Syndrome

Yulun Cai

Chinese PLA General Hospital

Benchuan Hao

Chinese PLA General Hospital

Jianqiao Chen

Chinese PLA General Hospital

Yuerui Li

Chinese PLA General Hospital

Hongbin Liu (✉ liuhbcad301@163.com)

Chinese PLA General Hospital

Research

Keywords: Proteomics, Chronic coronary syndrome, Prognosis, Endothelial protein C receptor, Cholesteryl ester transfer protein

Posted Date: September 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-868543/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Chronic coronary syndrome (CCS) is a newly proposed concept and is hallmarked by more long-term major adverse cardiovascular events (MACEs), calling for accurate prognostic biomarkers for initial risk stratification.

Methods: Data-independent acquisition liquid chromatography tandem mass spectrometry (DIA LC-MS/MS) quantitative proteomics was performed on 38 patients with CCS; 19 in the CCS events group and 19 in the non-events group as the controls. We also developed a machine-learning-based pipeline to identify proteins as potential biomarkers and validated the target proteins by enzyme-linked immunosorbent assay (ELISA) in an independent prospective cohort (n = 352).

Results: Fifty-seven differentially expressed proteins were identified by quantitative proteomics and three final biomarkers were preliminarily selected from the machine-learning-based pipeline. Further validation with the prospective cohort showed that endothelial protein C receptor (EPCR) and cholesteryl ester transfer protein (CETP) levels at admission were significantly higher in the CCS events group than they were in the non-events group, whereas the carboxypeptidase B2 (CPB2) level was similar in the two groups. A correlation analysis showed that CETP was positively related to high-density lipoprotein cholesterol and triglyceride, and EPCR was positively related to fibrinogen. In the Cox survival analysis, EPCR and CETP were independent risk factors for MACEs. The cumulative risk duration of patients with high EPCR and CETP levels was significantly shorter than that of patients with low EPCR and CETP levels. We constructed a new prognostic model by combining the Framingham coronary heart disease (CHD) risk model with EPCR and CETP levels. This new model significantly improved the C-statistics for MACE prediction compared with that of the Framingham CHD risk model alone (AUC 0.732 vs. 0.684, $p < 0.05$).

Conclusions: Plasma proteomics was used to find biomarkers of predicting MACEs in patients with CCS. EPCR and CETP were identified as promising prognostic biomarkers for CCS. The Framingham CHD risk model combined with EPCR and CETP levels was shown to be a high-performance prognostic model for CCS.

Introduction

The Guidelines for Chronic Coronary Syndrome (CCS) were announced at the annual meeting of the European Society of Cardiology (ESC) in August 2019[1]. The ESC updated the “Guidelines for Treatment of Stable Coronary Artery Diseases (SCAD)” released in 2013[2] and defined the concept of CCS. This changed the previous inherent concept of SCAD and reflected the current deeper understanding of the pathophysiological mechanism of coronary artery disease. Traditionally, the term SCAD was used to describe CCS which often shaped the disease as “stable”. Although CCS is often “relatively stable” compared with acute coronary artery disease, the underlying pathophysiological state can become “unstable” at any time, causing plaque rupture or erosion, leading to acute thrombosis. Therefore, the risk

of future cardiovascular events in patients with CCS can be different; hence, predicting major adverse cardiovascular events (MACEs) will have great clinical significance for patients with CCS.

Genomics and transcriptomics have been widely used, but, in recent years, many researchers have found that the results for proteins are often not highly consistent with those for genomes or transcriptomes[3]. This is partly because the products of transcription or translation are usually metabolized or modified, thereby changing the downstream protein abundance[4]. Proteins are key regulators of many biological processes, and are directly related to the occurrence of many diseases and their clinical prognosis[5]. Proteomics approaches have become increasingly mature, from the discovery of a single biomarker for early disease to the comprehensive characterization of protein abundance profiles of specific diseases. Indeed, some diseases are affected by more than one biological pathway. The advent of high-throughput proteomics has made research on such processes possible, and the study of differentially expressed proteins (DEPs) has provided insights into the molecular mechanisms of many human diseases[6–8]. Currently, liquid chromatography tandem mass spectrometry (LC-MS/MS) is the main tool used to analyze whole proteomes, and it has been applied in studies of cardiovascular diseases, including the recently redefined CCS.

In this study, we performed a proteomics analysis to discover potential biomarkers of CCS, used machine learning methods to screen the identified biomarkers[9, 10], and validated the selected biomarkers in an independent prospective cohort. Two proteins were identified as new biomarkers for predicting the risk of adverse cardiovascular events in patients with CCS. First, we used discovery mass spectrometry to quantify thousands of different proteins without the need for previous knowledge, and thus identify proteins not previously associated with CCS. After protein relative quantification between the CSS events and non-events groups, the classification power was evaluated by machine learning-based selection. On the basis of the results and clinical relevance, we identified three proteins, endothelial protein C receptor (EPCR), carboxypeptidase B2 (CPB2), and cholesteryl ester transfer protein (CETP) as candidate biomarkers. Finally, in the independent validation cohort, we found that EPCR and CETP were better than CPB2 for predicting MACEs in patients with CCS by enzyme-linked immunosorbent assay (ELISA) and, when combined with the Framingham coronary heart disease (CHD) risk model, they improved the risk prediction beyond the Framingham CHD risk model alone.

Methods

Study design and patient enrollment

The overall design of this study is shown in Figure 1. The discovery cohort was a retrospective cohort. We selected 38 patients who had undergone a physical examination conducted at the People's Liberation Army (PLA) General Hospital from April to July 2015. The inclusion criteria were: 1) Patients who were asymptomatic or had stable symptoms within one year after onset of acute coronary syndromes; 2) Patients who were stable more than one year after initial diagnosis or revascularization regardless of symptoms; 3) Patients with angina pectoris, suspected vasospasm, or microvascular disease; and 4)

Asymptomatic patients screened for CHD. The exclusion criteria were: 1) Patients with severe heart failure; 2) Patients with CHD in the acute phase; and 3) Patients with other diseases that made them unsuitable for this study. According to whether MACEs (including cardiovascular related death, non-fatal myocardial infarction, unstable angina, and heart failure) occurred until June 2019, the participants were divided into a CCS events group (Group A, n=19) as the cases and a non-events group (Group B, n=19) as the controls for the proteomics analysis. The validation cohort was an independent prospective cohort that included 352 patients who were recruited from those who had undergone a routine physical examination at the PLA General Hospital from April to July 2017. The inclusion and exclusion criteria were the same as those for the discovery cohort. This cohort was followed up until April 2021. This study was approved by the Ethics Board of the Chinese PLA General Hospital and written informed consent was obtained from each patient. This study conforms to the principles outlined in the Declaration of Helsinki.

Blood Sampling

Blood samples were collected after fasting 12 hours. The samples were stored at -80°C with ethylenediaminetetraacetate until analysis.

Proteomics analysis

Quantitative proteomics analysis was performed using liquid chromatography tandem mass spectrometry (LC-MS/MS) to identify potential protein biomarker candidates among those proteins differing in abundance between event group and non-event group.

We used an integrated approach involving DIA strategy, HPLC fractionation to quantify the dynamic changes of the whole proteome.

Protein Extraction

Firstly, the cellular debris of plasma sample was removed by centrifugation at 12,000 g at 4°C for 10 min. Then, the supernatant was transferred to a new centrifuge tube. The top 12 high abundance proteins were removed by Pierce™ Top 12 Abundant Protein Depletion Spin Columns Kit (Thermo Fisher). Finally, the protein concentration was determined with BCA kit according to the manufacturer's instructions.

Trypsin Digestion

For digestion, the protein solution was reduced with 5 mM dithiothreitol for 30 min at 56°C and alkylated with 11 mM iodoacetamide for 15 min at room temperature in darkness. The protein sample was then diluted by adding 100 mM TEAB to urea concentration less than 2M. Finally, trypsin was added at 1:50

trypsin-to-protein mass ratio for the first digestion overnight and 1:100 trypsin-to-protein mass ratio for a second 4 h-digestion.

HPLC Fractionation

The tryptic peptides were fractionated into fractions by high pH reverse-phase HPLC using Agilent 300Extend C18 column (5 µm particles, 4.6 mm ID, 250 mm length). Briefly, peptides were first separated with a gradient of 8% to 32% acetonitrile (pH 9.0) over 60 min into 60 fractions. Then, the peptides were combined into 18 fractions and dried by vacuum centrifuging.

Data-independent Acquisition (DIA)–LC-MS/MS Analysis

The iRT kit was added to all the samples according to manufacturer's instructions. The tryptic peptides were dissolved in solvent A (0.1% formic acid, 2% acetonitrile), directly loaded onto a home-made reversed-phase analytical column (25-cm length, 100 µm i.d.). Peptides were separated with a gradient from 4% to 32% solvent B (0.1% formic acid in 90% acetonitrile) over 114 min, and climbing to 80% in 3 min then holding at 80% for the last 3 min, all at a constant flowrate of 450 nL/min on an EASY-nLC 1200 UPLC system (Thermo Fisher Scientific). The separated peptides were analyzed in DDA mode by Q Exactive™ HF-X (Thermo Fisher Scientific) with a nano-electrospray ion source.

The separated peptides were analyzed in Q Exactive™ HF-X (Thermo Fisher Scientific) with a nano-electrospray ion source. The full MS scan resolution was set to 120,000 for a scan range of 385–1200 m/z. The data acquisition was performed in DIA mode. Each cycle contains one full scan followed by 70 DIA MS/MS scans with predefined precursor m/z range. The HCD fragmentation was performed at a normalized collision energy (NCE) of 27%. The fragments were detected in the Orbitrap at a resolution of 15,000. Fixed first mass was set as 200 m/z. Automatic gain control (AGC) target was set at 5E5.

Data Analysis

Spectral library generation: The resulting DDA data were processed using MaxQuant search engine (v.1.6.6.0). Tandem mass spectra were searched against the human SwissProt database (20387 entries) concatenated with reverse decoy database. Trypsin/P was specified as cleavage enzyme allowing up to 2 missing cleavages. The mass tolerance for precursor ions was set as 20 ppm in First search and 4.5 ppm in Main search, and the mass tolerance for fragment ions was set as 0.02 Da. Carbamidomethyl on Cys was specified as fixed modification. Acetylation on protein N-terminal and oxidation on Met were specified as variable modifications. FDR was adjusted to < 1%. The false discovery rates of the PSMs and proteins were set to less than 1%.

DIA data analysis: All DIA data were analyzed in Skyline (v 4.1.0). The DDA search results were imported to Skyline to generate the spectral library, and the retention times were aligned to iRT reference values.

Transition settings: precursor charges were set as 2, 3, 4, 5, ion charges were set as 1, 2. The ion match tolerance was set as 0.02 Da. Six most intense fragment ions from the spectral library were selected for each precursor. Decoy generation was based on shuffled sequences, and the FDR was estimated with the mProphet approach and set to 1%. Relative quantification of proteins was performed using MSstats package.

Bioinformatics analysis

For the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses, we used the two-tailed Fisher's exact test to determine the significance of the functional enrichment of the DEPs against all the identified proteins. A corrected p-value <0.05 was considered significant. Gene Set Enrichment Analysis (GSEA) is an aggregate score and running-sum statistic approach for molecular signature-based statistical significance testing[11]. The entire gene set containing a ranked list of all expression values in a dataset is considered without requiring a cutoff of differentially expressed values for functional analysis. For the GSEA, we used the entire proteomics abundance profiles dataset. Gene sets from the Molecular Signatures Database (MSigDB) v7.2 were used (H: hallmark gene sets; C2: KEGG pathway database; C5: GO terms database).

Machine learning-based selection of biomarkers

Construction of voting classifier

We used three machine learning classification algorithms, Logistic Regression, Support Vector Machine, and Random Forest, as the base classifiers (Supplementary figure S1). On the basis of these classifiers, we built a voting classifier. When a new sample had to be assigned to a category, each base classifier was used to predict the probability that the new sample belonged to a particular category. The final classification result was determined by the weighted value of the predicted probability of each category by all three base classifiers. This is an integrated method that used the Voting Classifier model in Python.

Feature ranking

Each sample was represented by a feature vector composed of numerous expression data. To quantify the ability of these expression features to distinguish different samples, we performed univariate feature analysis using a variance test to calculate the correlation between each feature and the sample category one by one. In this way, the ability of each feature to distinguish the sample category was obtained and the score and the corresponding p-value was calculated. The expression features were sorted according to the calculated p-value, and used in the subsequent analysis.

Accuracy evaluation index

To compare the difference between the category predicted by the model and the actual sample category, we calculated an accuracy index using the Matthews coefficient value as an indicator of the accuracy of the predictive power of the model as

$$\text{Matthews coefficient} = (TP \times TN - FP \times FN) / \sqrt{((TP + FP)(TP + FN)(TN + FN)(TN + FP))}$$

where, TP, TN, FP, and FN are true positive, true negative, false positive, and false negative, respectively.

Feature selection

We plotted the calculated accuracy index results against the number of features in the expression feature subset as the incremental feature selection curve. When the Matthews coefficient reached the maximum value, we considered the expression feature subset corresponding to the model as the optimal expression feature subset.

Cross-validation

The expression feature subset and sample category were used as input of the Voting Classifier model, and 10-fold cross-validation was used to calculate the prediction accuracy of the local optimal expression feature subset for a sample. The results were expressed by the receiver operating characteristic curve (ROC). This is a dynamic validation that reduced the impact of data partitioning[12].

Statistics analysis

Data are presented as numbers and frequencies for categorical variables and as means \pm standard deviation for continuous variables. Baseline characteristics were compared using the chi-square test for categorical variables and analysis of variance test for continuous variables. The effect of the candidate biomarkers was evaluated using a Cox proportional hazards model, and p-values were calculated using the log-rank test. Kaplan-Meier plots were used to compare the cumulative risk. The ROC was compared using a z test (DeLong's method[13]) between the classic Framingham CHD risk model alone and the Framingham CHD model combined with the candidate biomarkers EPCR and CETP. The Framingham CHD model included age, sex, total cholesterol, high-density lipoprotein cholesterol (HDL-C), systolic blood pressure, current smoking, and diabetes status as confounding factors[14].

A two-tailed p-value of <0.05 was considered to indicate statistically significant difference for all the analyses. The statistical analyses were performed using SPSS (version 23.0), STATA (version 12.0), and MedCalc Software.

Results

Baseline characteristics

The baseline clinical data of the discovery and validation cohorts are shown in Table 1. The discovery cohort comprised 19 cases and 19 controls, with an average follow-up time of 47.6 months. The use of clopidogrel, statins, and nitrate esters were significantly higher in the non-events group than they were in the CCS events group. In the validation cohort, after an average of 41.6 months of follow-up, 86 patients (24.3%) had MACEs. Age, fibrinogen, D-dimer, glycosylated hemoglobin, and NT-proBNP (N-terminal-pro-brain natriuretic peptide) were significantly higher in the events group than they were in the non-events group. Current smoking, clopidogrel use, diastolic blood pressure, and HDL-C were significantly higher in the non-events group than they were in the events group.

Proteomics analysis

Quality control results for the proteomics data are shown in Supplementary figure S2. The length distribution showed that 70% of the identified peptides were 7–20 amino acids long, which is consistent with the general rules of trypsin enzymatic hydrolysis and higher-energy collisional dissociation (HCD). Peptides shorter than 5 amino acids cannot be effectively identified, and peptides longer than 20 amino acids are not suitable for HCD because of their high mass and charge. The distribution of peptides per protein showed that there were more than two peptides for most proteins. In general, proteins that have multiple corresponding specific peptides increase the precision and accuracy of the quantification results. The mass distribution of the proteins showed that the mass was relatively well-distributed, indicating that there was no significant molecular weight bias for proteins during sample preparation. These results confirm that the overall proteomics data meet the quality control requirements.

A total of 5480 peptides were identified in the mass spectrum. By comparing the peptides back to proteins, we identified a total of 1120 proteins, of which 783 were quantified.

We used the proteomic data to identify signatures of MACEs by analyzing the plasma proteins that underwent significant fold changes (FCs) between Group A and Group B ($FC > 1.2$ or $FC < 0.8$; unpaired two-sided Welch's t test; $p < 0.05$). As shown in the volcano map (Figure 2), a total of 57 DEPs were identified under this condition. As expected, the patients in the two groups were distinguished in an unsupervised clustering analysis based on overall abundance trends of the DEPs (Supplementary figure S3). The principal component analysis (Supplementary figure S4) showed that the events and non-events groups were unambiguously distinguished by the DEPs.

Functional analysis of the DEPs

The DEPs were annotated with GO terms and assigned KEGG pathways for the functional enrichment analyses (Figure 3). The highly enriched processes included cellular lipid metabolic process, acylglycerol transport, inflammatory response, response to bacterium, glycerolipid metabolism, hypertrophic cardiomyopathy, complement and coagulation cascades, and PI3K-Akt signaling pathway. These results are consistent with previous reports that lipid metabolic, coagulation, inflammation, and ventricular remodeling processes were associated with MACEs. To gain a better understanding of the functional differences between the CCS events group and the non-events group and elucidate potentially unique protein signatures, we performed a GSEA of the hallmark, KEGG, and GO gene sets from the MSigDB (Supplementary figure S5). Coagulation and lipid metabolic process were the top differential pathways between the two groups.

Together, these analyses pinpointed specific pathways (lipid metabolic, coagulation, inflammation) that may operate in the events group compared with the non-events group, and highlighted CCS-related pathways for further functional investigation.

Machine learning-based selection of biomarkers for prognosis of CCS

We used the plasma proteomic data of the discovery cohort and developed a series of algorithms to identify potential biomarker combinations to classify CCS cases. The feature analysis was used to rank the expression features based on scores and p-values, as shown in Supplementary figure S6. To visualize the ranking results of the protein expression characteristics, we plotted a ranking histogram of the top 30 features with the highest scores (Supplementary figure S7).

The incremental feature selection curve (Supplementary figure S8) shows that when the top eight ranked features were selected, the Matthews coefficient value of the model reached the maximum for the first time. These eight features were analyzed further. Next, we selected the protein combinations with the highest area under the curve (AUC) value from the 10-fold cross-validation. The ROC curve (Figure 4) showed that the AUC of the first six proteins in the final trained model were >0.8. These six proteins were identified as candidate proteins and the detailed information is shown in Table 2. The distribution of these six proteins between the events and non-events groups was significantly different as shown in the box plots in Supplementary figure S9.

Validation of the biomarkers of different CCS outcomes

By combining the results of the DEP selection, GO and KEGG pathway analyses, machine learning, previous knowledge and clinical relevance, we selected three proteins EPCR (Q9UNN8), CPB2 (Q961Y4), and CETP (P11597) as the target proteins for validation. The expression levels of the three proteins are shown in Figure 5. EPCR and CETP levels at admission were significantly higher in the CCS events group than in the control non-events group, whereas the CPB2 levels were similar. CETP was positively related to

HDL-C and triglyceride, and EPCR was positively related to fibrinogen as shown in the correlation scatter plot (Supplementary figure S10), which is consistent with our previous expectations. The Cox survival analysis showed that EPCR and CETP were risk factors for MACEs (Table 3). After correcting for the confounding factors in the Framingham CHD risk model, EPCR and CETP were still found to be independent risk factors for MACEs. Additionally, when the Youden index reached the maximum, we used the expression levels of the three proteins as the cutoff values. Accordingly, the patients were divided into a high protein level group and a low protein level group, and Kaplan-Meier cumulative risk curves were compared between the two groups. As shown in Figure 6, the cumulative risk duration of patients with high EPCR and CETP levels was significantly shorter than that of patients with low EPCR and CETP levels. A new prognostic model was constructed by combining the Framingham CHD risk model with the candidate biomarkers EPCR and CETP. This new model significantly improved the C-statistics for MACE prediction compared with that of the Framingham CHD risk model alone (AUC 0.732 vs. 0.684, $p < 0.05$) (Figure 6). These results showed that EPCR and CETP were independent risk factors for MACEs in patients with CCS, and combined with the classic Framingham model, EPCR and CETP provided better prediction metrics than the Framingham model alone.

Discussion

We performed a series of studies on the plasma proteins of patients with CCS, followed by global proteomic mass spectrometry identification, machine learning-based selection of biomarkers, and ELISA for prospective validation of expanded samples.

Cardiovascular disease is one of the leading causes of death worldwide[15]. One of the main risks of developing cardiovascular disease is vascular endothelial dysfunction[16]. In a healthy state, endothelial cells maintain a balanced hemostatic state by producing procoagulants and anticoagulants, as well as proinflammatory and anti-inflammatory cytokines. In the disease state, endothelial cells are activated and exert procoagulant and proinflammatory effects. Endothelial cell dysfunction leads to thrombosis and coagulation imbalance[17]. The new international guidelines consider that CHD is a dynamic process of atherosclerotic plaque accumulation and changes in coronary circulatory function. Plaques can show the following trends: gradually increased instability or even rupture, stability maintained for a long time, and gradual shrinking. The composition of plaques also continues to change. However, about one-third of patients with a cardiovascular disease have angina pectoris, but have no obstructive coronary artery disease[18]. Therefore, patients with CCS can have a relatively stable period, but the relatively stable vascular environment and circulatory function may become unstable because of inflammatory reactions, vulnerable plaques, and abnormal lipid metabolism[19]. Therefore, “stability” is only temporary and relative, not absolute. The clinical evaluation and management of such seemingly “stable” patients with CCS is of great significance to improve the prognosis. In this study, we found EPCR and CETP were closely related to vascular endothelial homeostasis.

Currently, precise and personalized medicine is limited mainly to genetic methods and still needs to be effectively integrated with individual characteristics at the transcriptome and proteome levels. Recent

advances in MS-based proteomics methodology may help to ensure the high accuracy and sensitivity required for single plasma analysis[20]. Therefore, in the near future, plasma proteomics is expected to play a strategic role in identifying new biomarkers. Plasma proteomics conforms to the three pillars of personalized medicine: accurate molecular maps, non-invasive samples, and endotype characterization[21].

There are two variants of EPCR: mEPCR (membrane EPCR), which is present on endothelial cell membranes, and soluble EPCR (soluble EPCR), which circulates in the blood[22]. Protein C is a vitamin K-dependent serine protease that is synthesized mainly in the liver and circulates in the plasma. Protein C binds to mEPCR with high affinity, and is converted to activated protein C (APC) by the thrombin–thrombomodulin complex on the surface of endothelial cells through a limited proteolytic process[23]. The mEPCR variant can bind to APC and plays important roles in anticoagulation, anti-inflammatory, cell protection (anti-apoptosis), protecting endothelial barrier function, and promoting neovascularization[24–28]. When activated by the thrombin–thrombomodulin complex, APC dissociates from the membrane-bound receptor mEPCR, and functions as an anticoagulant by inactivating coagulation factors Va and VIIIa[29]. When APC is combined with mEPCR, it shows strong anti-inflammatory and cytoprotective activities. The cytoprotective signal activity of APC is mediated by protease activated receptor 1 (PAR1) on endothelial cells bound by mEPCR[30]. The APC–mEPCR complex relies on the anti-inflammatory activity of PAR1 to mediate the inhibition of inflammatory gene expression, including *c-Fos* and *FosB*, which belong to the activator protein 1 (AP-1) family. Protective signals can also inhibit the release of inflammatory cytokines (such as IL-1 β , IL-6, and tumor necrosis factor- α) and the nuclear translocation of NF- κ B, and downregulate the expression of genes that encode endothelial cell adhesion proteins (such as ICAM1, VCAM1, and E-selectin), thereby restricting the penetration of white blood cells through the vascular system, protecting the endothelial barrier function, and inhibiting inflammation[31]. Conversely, the sEPCR variant detaches from the cell membrane surface through shedding and enters the circulation. Possible reasons for shedding include a systemic inflammatory response and vascular endothelial damage[32]. The sEPCR plays a negative competitive role in blood circulation. It binds protein C and APC with similar affinity and inhibits protein C activation on the endothelium. It also inhibits the anticoagulant activity of APC by blocking the binding of APC to phospholipids. It is the sEPCR variant in blood circulation that was tested in this study. Elevated levels of sEPCR can disturb vascular homeostasis, promote coagulation, aggravate inflammation, and accelerate endothelial cell apoptosis, which is associated with increased risk of thrombosis[33].

CETP is a plasma protein secreted by the liver. It is one of the most effective endogenous regulators of plasma HDL-C, which protects the cardiovascular system in many ways[34]. It can promote the transfer of cholesterol ester from HDL-C to apolipoprotein B (ApoB). In addition to removing excess cholesterol from the arterial wall, HDL-C can also inhibit lipid oxidation, restore endothelial function, and exert anti-inflammatory and anti-apoptotic effects[35, 36]. CETP leads to a net reduction of HDL-C in plasma, which increases the risk of atherosclerosis development. Some large randomized controlled trials have studied the effects of CETP inhibitors on lipid metabolism and MACEs, but the results are inconsistent[37–39]. Possible reasons include incomplete function of the CETP inhibitors and side effects of drugs. A study

explored whether CETP was related to atherosclerosis through its role in HDL-C and low-density lipoprotein metabolism. In a case–control study of 50 patients with coronary atherosclerosis and 50 controls, no significant difference was detected in the lipid profiles between the two groups, even though the serum CETP level of the case group was significantly higher than that of control group[40]. This finding indicated that CETP may have atherogenic effects. However, no further studies have reported the long-term adverse prognosis risk of CHD. In our study, we included both case–control and prospective cohorts, which more fully illustrated the relationship between CETP and the poor prognosis of patients with CCS.

To our knowledge, this is the first study to evaluate the relationship between plasma proteomics and the prognosis of patients with CCS. This study was layered involving three different methods, and the results are accurate and reliable. In addition, we analyzed more cases with longer follow-up times than most of the other studies. Our study has some limitations. First, the discovery and validation cohorts were from a single center and had relatively fewer older adult female patients. Therefore, to reduce bias, only male patients were included in this study. Second, we identified 57 DEPs that could predict the occurrence of MACEs, which may provide a rich biomarker pool for CCS, but further refinement of the diagnostic biomarkers is needed. Third, whether EPCR and CETP will act as prognostic predictors of CCS caused by other etiologies needs further study.

Conclusions

We performed plasma proteomics analysis to find biomarkers for predicting MACEs in patients with CCS. We identified EPCR and CETP as independent risk factors for MACEs. The Framingham CHD risk model combined with EPCR and CETP was found to be a high-performance prognostic model for CCS. EPCR and CETP may be associated with vascular homeostasis, involving lipid metabolism, and inflammatory, coagulation, and cell protection processes. Further investigations are needed to understand the specific mechanisms involved.

Declarations

Acknowledgements

Not applicable.

Authors' contributions

YC and HL designed the protocol; YC drafted the manuscript; HL supervised patient recruitment and study procedures; YC, BH, JC and YL conducted study procedures; all authors carried out study procedures, critically revised the manuscript for important intellectual content and approved the final manuscript.

Funding

This work was supported by the National Key Research Program of China (2020YFC2008304).

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

This study was approved by the Ethics Board of the Chinese PLA General Hospital and written informed consent was obtained from each patient. This study conforms to the principles outlined in the Declaration of Helsinki.

Competing interests

The authors declare that they have no competing interests.

References

1. Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C, Prescott E, Storey RF, Deaton C, Cuisset T, et al. 2019 ESC Guidelines for the diagnosis and management of chronic coronary syndromes. *Eur Heart J*. 2020;41(3):407–77.
2. Task Force M, Montalescot G, Sechtem U, Achenbach S, Andreotti F, Arden C, Budaj A, Bugiardini R, Crea F, Cuisset T, et al. 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. *Eur Heart J*. 2013;34(38):2949–3003.
3. Zhang B, Wang J, Wang X, Zhu J, Liu Q, Shi Z, Chambers MC, Zimmerman LJ, Shaddox KF, Kim S, et al. Proteogenomic characterization of human colon and rectal cancer. *Nature*. 2014;513(7518):382–7.
4. Zhang H, Liu T, Zhang Z, Payne SH, Zhang B, McDermott JE, Zhou JY, Petyuk VA, Chen L, Ray D, et al. Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. *Cell*. 2016;166(3):755–65.

5. Papaioannou MD, Djuric U, Kao J, Karimi S, Zadeh G, Aldape K, Diamandis P. Proteomic analysis of meningiomas reveals clinically distinct molecular patterns. *Neuro Oncol.* 2019;21(8):1028–38.
6. Djomehri SI, Gonzalez ME, da Veiga Leprevost F, Tekula SR, Chang HY, White MJ, Cimino-Mathews A, Burman B, Basrur V, Argani P, et al. Quantitative proteomic landscape of metaplastic breast carcinoma pathological subtypes and their relationship to triple-negative tumors. *Nat Commun.* 2020;11(1):1723.
7. Shu T, Ning W, Wu D, Xu J, Han Q, Huang M, Zou X, Yang Q, Yuan Y, Bie Y, et al. Plasma Proteomics Identify Biomarkers and Pathogenesis of COVID-19. *Immunity.* 2020;53(5):1108–22 e1105.
8. Wu D, Zhang S, Xie Z, Chen E, Rao Q, Liu X, Huang K, Yang J, Xiao L, Ji F, et al. Plasminogen as a prognostic biomarker for HBV-related acute-on-chronic liver failure. *J Clin Invest.* 2020;130(4):2069–80.
9. Chua W, Purmah Y, Cardoso VR, Gkoutos GV, Tull SP, Neculau G, Thomas MR, Kotecha D, Lip GYH, Kirchhof P, et al. Data-driven discovery and validation of circulating blood-based biomarkers associated with prevalent atrial fibrillation. *Eur Heart J.* 2019;40(16):1268–76.
10. Hoshino A, Kim HS, Bojmar L, Gyan KE, Cioffi M, Hernandez J, Zambirinis CP, Rodrigues G, Molina H, Heissel S, et al. Extracellular Vesicle and Particle Biomarkers Define Multiple Human Cancers. *Cell.* 2020;182(4):1044–61 e1018.
11. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102(43):15545–50.
12. Lanfear DE, Gibbs JJ, Li J, She R, Petucci C, Culver JA, Tang WHW, Pinto YM, Williams LK, Sabbah HN, et al. Targeted Metabolomic Profiling of Plasma and Survival in Heart Failure Patients. *JACC Heart Fail.* 2017;5(11):823–32.
13. Nishi I, Seo Y, Hamada-Harimura Y, Yamamoto M, Ishizu T, Sugano A, Sato K, Sai S, Obara K, Suzuki S, et al. Geriatric nutritional risk index predicts all-cause deaths in heart failure with preserved ejection fraction. *ESC Heart Fail.* 2019;6(2):396–405.
14. D'Agostino RB, Sr., Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, Kannel WB. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation.* 2008;117(6):743–53.
15. Balakumar P, Maung UK, Jagadeesh G. Prevalence and prevention of cardiovascular disease and diabetes mellitus. *Pharmacol Res.* 2016;113(Pt A):600–9.
16. Zhang P, Liu Y, Su J, Bai J, Zhao S, Zhao S. Resistin impairs activation of protein C by suppressing EPCR and increasing SP1 expression. *Biomed Pharmacother.* 2019;109:930–7.
17. Nagareddy P, Smyth SS. Inflammation and thrombosis in cardiovascular disease. *Curr Opin Hematol.* 2013;20(5):457–63.
18. Ford TJ, Corcoran D, Berry C. Stable coronary syndromes: pathophysiology, diagnostic advances and therapeutic need. *Heart.* 2018;104(4):284–92.

19. van Diepen JA, Berbee JF, Havekes LM, Rensen PC. Interactions between inflammation and lipid metabolism: relevance for efficacy of anti-inflammatory drugs in the treatment of atherosclerosis. *Atherosclerosis*. 2013;228(2):306–15.
20. Shen B, Yi X, Sun Y, Bi X, Du J, Zhang C, Quan S, Zhang F, Sun R, Qian L, et al. Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell*. 2020;182(1):59–72 e15.
21. Ponzini E, Santambrogio C, De Palma A, Mauri P, Tavazzi S, Grandori R. **Mass spectrometry-based tear proteomics for noninvasive biomarker discovery**. *Mass Spectrom Rev* 2021.
22. Ku SK, Han MS, Lee MY, Lee YM, Bae JS. Inhibitory effects of oroxylin A on endothelial protein C receptor shedding in vitro and in vivo. *BMB Rep*. 2014;47(6):336–41.
23. Sriwastva MK, Kunjunni R, Andrabi M, Prasad K, Saxena R, Subbiah V. **Neuroprotective Effects of Activated Protein C Involve the PARP/AIF Pathway against Oxygen-Glucose Deprivation in SH-SY5Y Cells**. *Brain Sci* 2020, 10(12).
24. Minhas N, Xue M, Jackson CJ. Activated protein C binds directly to Tie2: possible beneficial effects on endothelial barrier function. *Cell Mol Life Sci*. 2017;74(10):1895–906.
25. Xue M, Dervish S, Chan B, Jackson CJ. The Endothelial Protein C Receptor Is a Potential Stem Cell Marker for Epidermal Keratinocytes. *Stem Cells*. 2017;35(7):1786–98.
26. Catieau B, Devos V, Chtourou S, Borgel D, Plantier JL. Endothelial cell surface limits coagulation without modulating the antithrombin potency. *Thromb Res*. 2018;167:88–95.
27. Shavit Stein E, Ben Shimon M, Artan Furman A, Golderman V, Chapman J, Maggio N. Thrombin Inhibition Reduces the Expression of Brain Inflammation Markers upon Systemic LPS Treatment. *Neural Plast*. 2018;2018:7692182.
28. Wang H, Wang P, Liang X, Li W, Yang M, Ma J, Yue W, Fan S. Down-regulation of endothelial protein C receptor promotes preeclampsia by affecting actin polymerization. *J Cell Mol Med*. 2020;24(6):3370–83.
29. Lopez-Ramirez MA, Pham A, Girard R, Wyseure T, Hale P, Yamashita A, Koskimaki J, Polster S, Saadat L, Romero IA, et al. Cerebral cavernous malformations form an anticoagulant vascular domain in humans and mice. *Blood*. 2019;133(3):193–204.
30. Kondreddy V, Wang J, Keshava S, Esmon CT, Rao LVM, Pendurthi UR. Factor VIIa induces anti-inflammatory signaling via EPCR and PAR1. *Blood*. 2018;131(21):2379–92.
31. Healy LD, Fernandez JA, Mosnier LO, Griffin JH. Activated protein C and PAR1-derived and PAR3-derived peptides are anti-inflammatory by suppressing macrophage NLRP3 inflammasomes. *J Thromb Haemost*. 2021;19(1):269–80.
32. Lecuyer H, Virion Z, Barnier JP, Matczak S, Bourdoulous S, Bianchini E, Saller F, Borgel D, Nassif X, Coureuil M. An ADAM-10 dependent EPCR shedding links meningococcal interaction with endothelial cells to purpura fulminans. *PLoS Pathog*. 2018;14(4):e1006981.
33. Tanalp AC, Oduncu V, Erkol A, Gozubuyuk G, Ozveren O, Dundar C, Canbay A, Kirma C. Soluble endothelial protein C receptor levels and protein C activity in patients with acute ST-segment elevation myocardial infarction. *Coron Artery Dis*. 2013;24(3):209–16.

34. Kontush A. HDL-mediated mechanisms of protection in cardiovascular disease. *Cardiovasc Res.* 2014;103(3):341–9.
35. Chapman MJ, Le Goff W, Guerin M, Kontush A. Cholesteryl ester transfer protein: at the heart of the action of lipid-modulating therapy with statins, fibrates, niacin, and cholesteryl ester transfer protein inhibitors. *Eur Heart J.* 2010;31(2):149–64.
36. Kettunen J, Holmes MV, Allara E, Anufrieva O, Ohukainen P, Oliver-Williams C, Wang Q, Tillin T, Hughes AD, Kahonen M, et al. Lipoprotein signatures of cholesteryl ester transfer protein and HMG-CoA reductase inhibition. *PLoS Biol.* 2019;17(12):e3000572.
37. Lincoff AM, Nicholls SJ, Riesmeyer JS, Barter PJ, Brewer HB, Fox KAA, Gibson CM, Granger C, Menon V, Montalescot G, et al. Evacetrapib and Cardiovascular Outcomes in High-Risk Vascular Disease. *N Engl J Med.* 2017;376(20):1933–42.
38. Tall AR, Rader DJ. Trials and Tribulations of CETP Inhibitors. *Circ Res.* 2018;122(1):106–12.
39. Armitage J, Holmes MV, Preiss D. Cholesteryl Ester Transfer Protein Inhibition for Preventing Cardiovascular Events: JACC Review Topic of the Week. *J Am Coll Cardiol.* 2019;73(4):477–87.
40. Devi A, Singh R, Dawar R, Tyagi S. Association of Cholesteryl Ester Transfer Protein (CETP) Gene – 629C/A Polymorphism with Angiographically Proven Atherosclerosis. *Indian J Clin Biochem.* 2017;32(2):235–8.

Tables

Table 1

Baseline clinical and laboratory characteristics of the study patients

	Discovery Cohort		Validation Cohort	
	Event (n=19)	No event (n=19)	Event (n=86)	No event (n=266)
Age, years	79.13±12.12	78.39±6.55	83.78±9.66	80.41±9.15*
Waistline (cm)	91.00±10.18	90.20±7.45	91.75±12.2	92.85±11.64
BMI (kg/m ²)	23.89±1.94	23.70±1.66	24.26±2.90	24.72±3.41
Current smokers, n (%)	1 (5.3)	1 (5.3)	5 (5.8)	38 (14.3)*
Hypertension, n (%)	14 (73.7)	16 (84.2)	86 (75.6)	194 (72.9)
Diabetes mellitus, n (%)	12 (63.2)	9 (47.4)	35 (40.7)	97 (36.5)
Stroke, n (%)	3 (15.8)	3 (15.8)	10 (11.6)	28 (10.5)
Systolic pressure (mmHg)	129.84±17.98	137.73±15.31	134.30±16.5	132.92±17.54
Diastolic pressure (mmHg)	64.42±11.49	67.15±11.85	65.72±9.82	68.43±9.59*
Fibrinogen (g/L)	3.46±0.66	3.46±0.52	3.52±0.63	3.35±0.61*
D-dimer (mmol/L)	0.66±0.62	0.26±0.23	0.96±1.79	0.64±0.65*
Total cholesterol (mmol/L)	3.97±0.76	3.78±0.67	3.97±0.77	4.03±0.79
Triglyceride (mmol/L)	1.44±0.79	1.19±0.35	1.42±0.75	1.26±0.70
HDL-C (mmol/L)	1.24±0.28	1.28±0.38	1.25±0.40	1.41±0.50*
LDL-C (mmol/L)	2.55±0.67	2.25±0.55	2.49±0.67	2.44±0.70
HBA1c (%)	6.27±1.14	6.24±0.86	6.36±1.17	6.10±0.79*
NT-proBNP (pg/ml)	225.04±230.82	201.21±181.24	344.21±499.52	213.06±400*
Medications				
Aspirin, n (%)	7 (36.8)	8 (42.1)	49 (57.0)	122 (45.9)
Clopidogrel, n (%)	4 (21.1)	11 (57.9)*	36 (41.9)	72 (27.1)*
Statins, n (%)	8 (42.1)	14 (73.7)*	70 (81.4)	196 (73.7)
Nitrate esters, n (%)	6 (31.6)	14 (73.7)*	52 (60.4)	138 (51.8)
<p>Note: BMI: Body mass index; HDL-C: High-density lipoprotein cholesterol; LDL-C: Low-density lipoprotein cholesterol; HBA1c: Hemoglobin A1c; NT-proBNP: N-terminal-pro-brain natriuretic peptide;</p> <p>*p value <0.05.</p>				

Table 2**Information of candidate proteins selected by machine learning**

Protein Gene	Protein Accession	Protein Description
C1R	P00736	Complement C1r subcomponent OS=Homo sapiens OX=9606 GN=C1R PE=1 SV=2
BCHE	P06276	Cholinesterase OS=Homo sapiens OX=9606 GN=BCHE PE=1 SV=1
CETP	P11597	Cholesteryl ester transfer protein OS=Homo sapiens OX=9606 GN=CETP PE=1 SV=2
CPB2	Q96IY4	Carboxypeptidase B2 OS=Homo sapiens OX=9606 GN=CPB2 PE=1 SV=2
EPCR	Q9UNN8	Endothelial protein C receptor OS=Homo sapiens OX=9606 GN=PROCR PE=1 SV=1
PLXDC2	Q6UX71	Plexin domain-containing protein 2 OS=Homo sapiens OX=9606 GN=PLXDC2 PE=1 SV=1

Table 3**Relation of target proteins and MACEs in univariate and multivariate survival analysis**

	Univariate models		Multivariate models	
	HR (95%CI)	P-value	HR (95%CI)	P-value
CPB2	0.843 (0.690-1.030)	0.096	0.824 (0.676-1.004)	0.055
CETP	1.063 (1.005-1.125)	0.034	1.058 (1.000-1.120)	0.048
EPCR	1.086 (1.021-1.155)	0.008	1.075 (1.013-1.142)	0.017

Note: CPB2: Carboxypeptidase B2; CETP: Cholesteryl ester transfer protein; EPCR: Endothelial protein C receptor.

Multivariate model included age, sex, total cholesterol, high-density lipoprotein cholesterol, systolic blood pressure, current smoking, and diabetes status. CETP is calculated as per 100mmol/L and EPCR is calculated as per 10mmol/L.

Figures

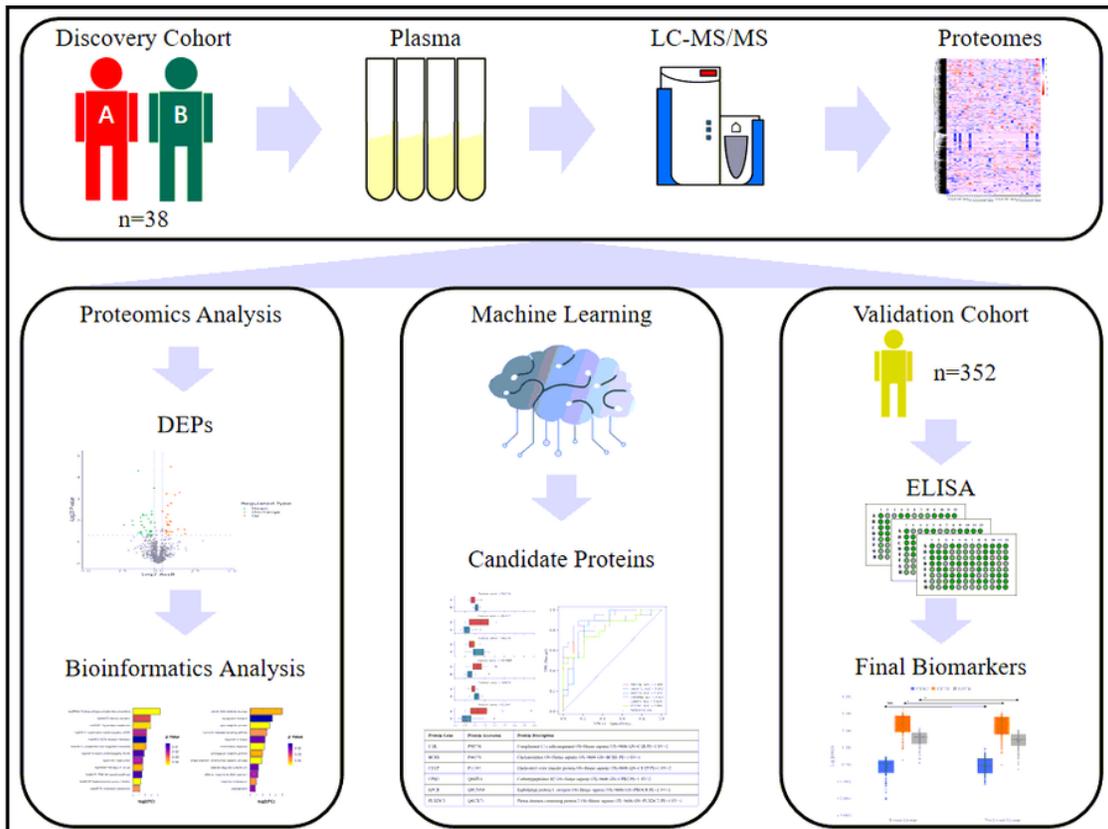


Figure 1

Study design

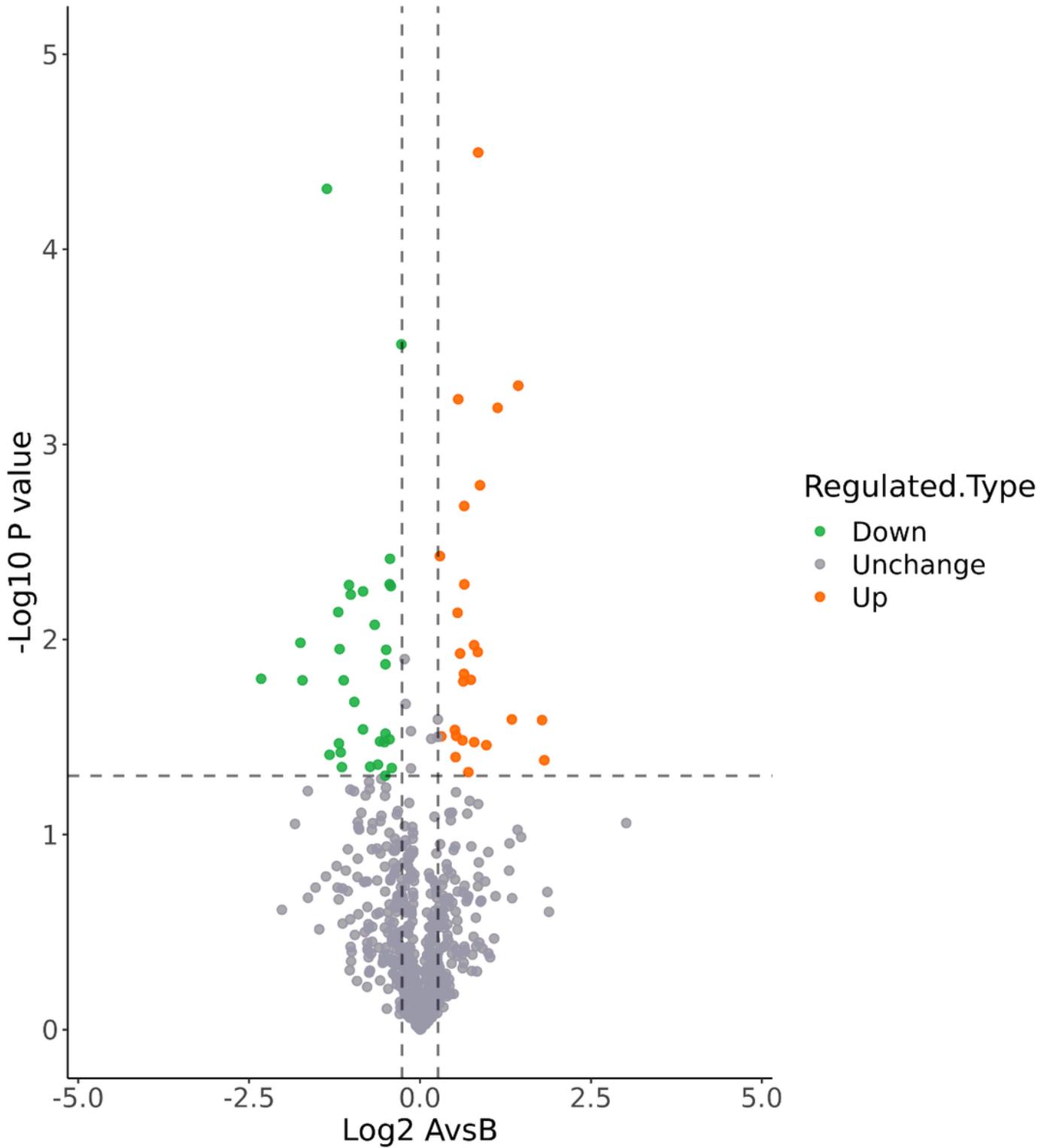


Figure 2

Volcano plot of differentially expressed proteins Note: Plasma proteins that underwent significant fold changes (FCs) between Group A and Group B (FC >1.2 or FC <0.8; unpaired two-sided Welch's t test; p <0.05).

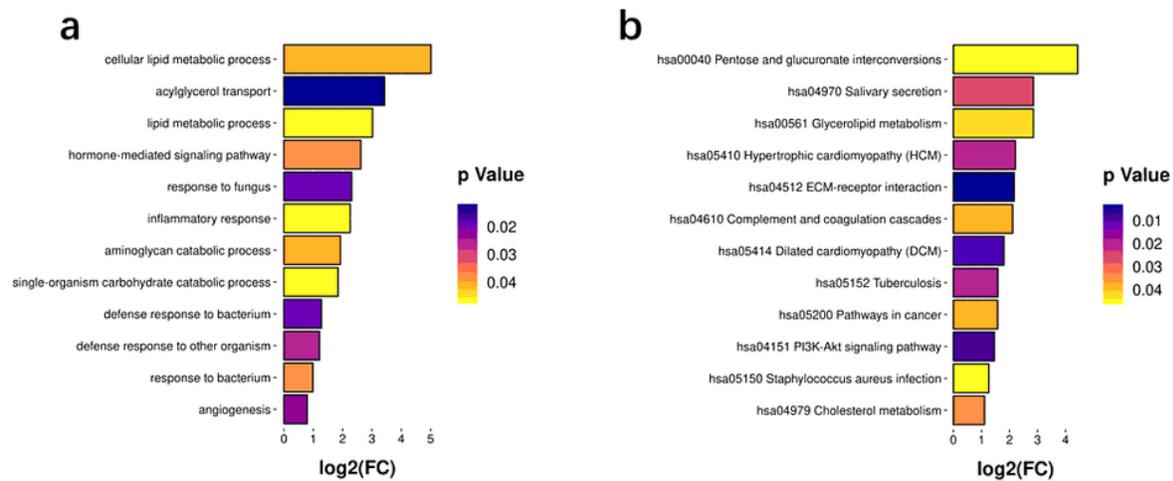


Figure 3

GO (a) and KEGG (b) pathway enrichment analyses Note: GO:Gene Ontology; KEGG:Kyoto Encyclopedia of Genes and Genomes; FC:Fold Change.

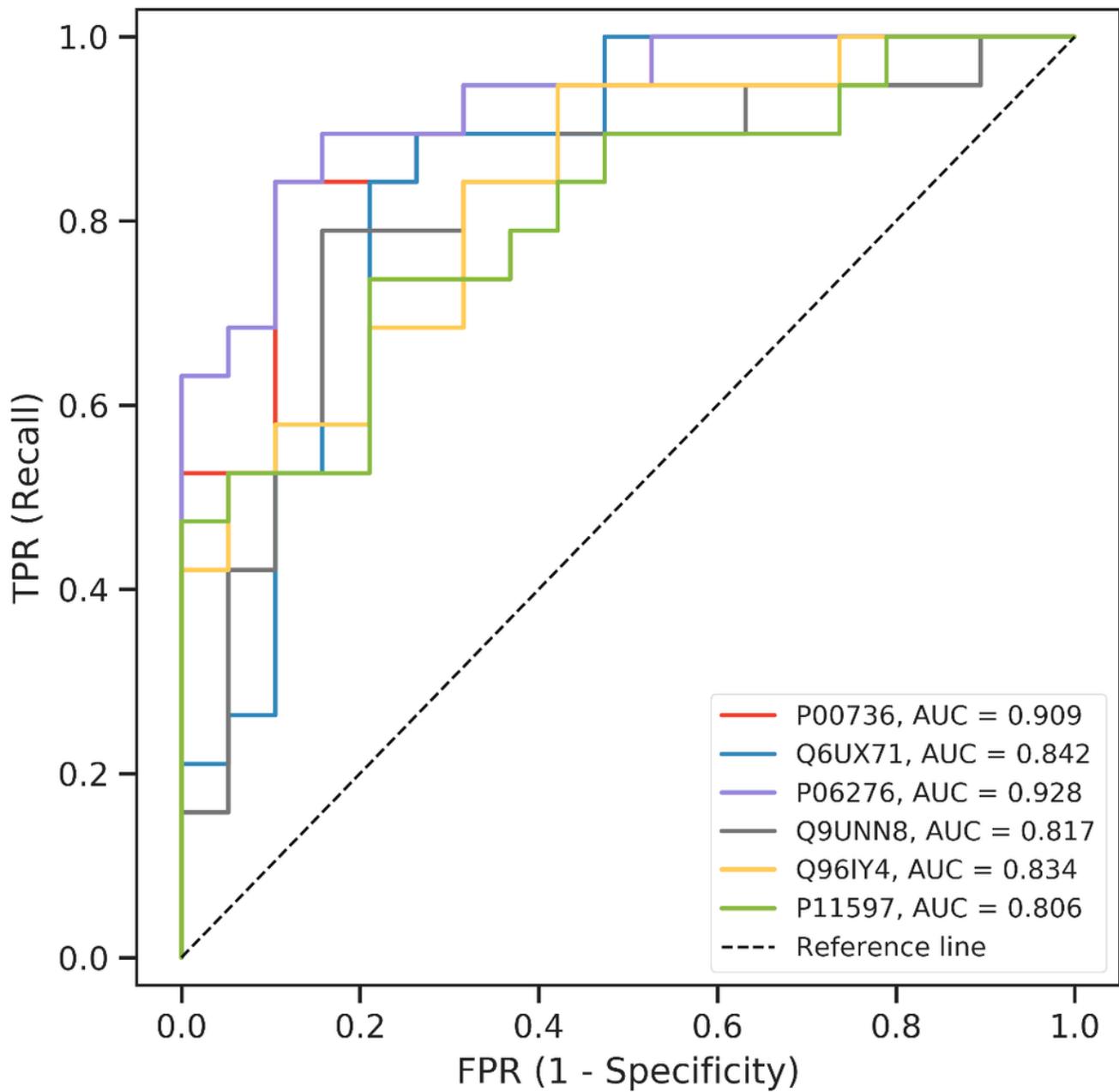


Figure 4

ROC curve of the model using the candidate proteins in machine learning Note: TPR: True Positive Rate; FPR: False Positive Rate; AUC: Area Under Curve.

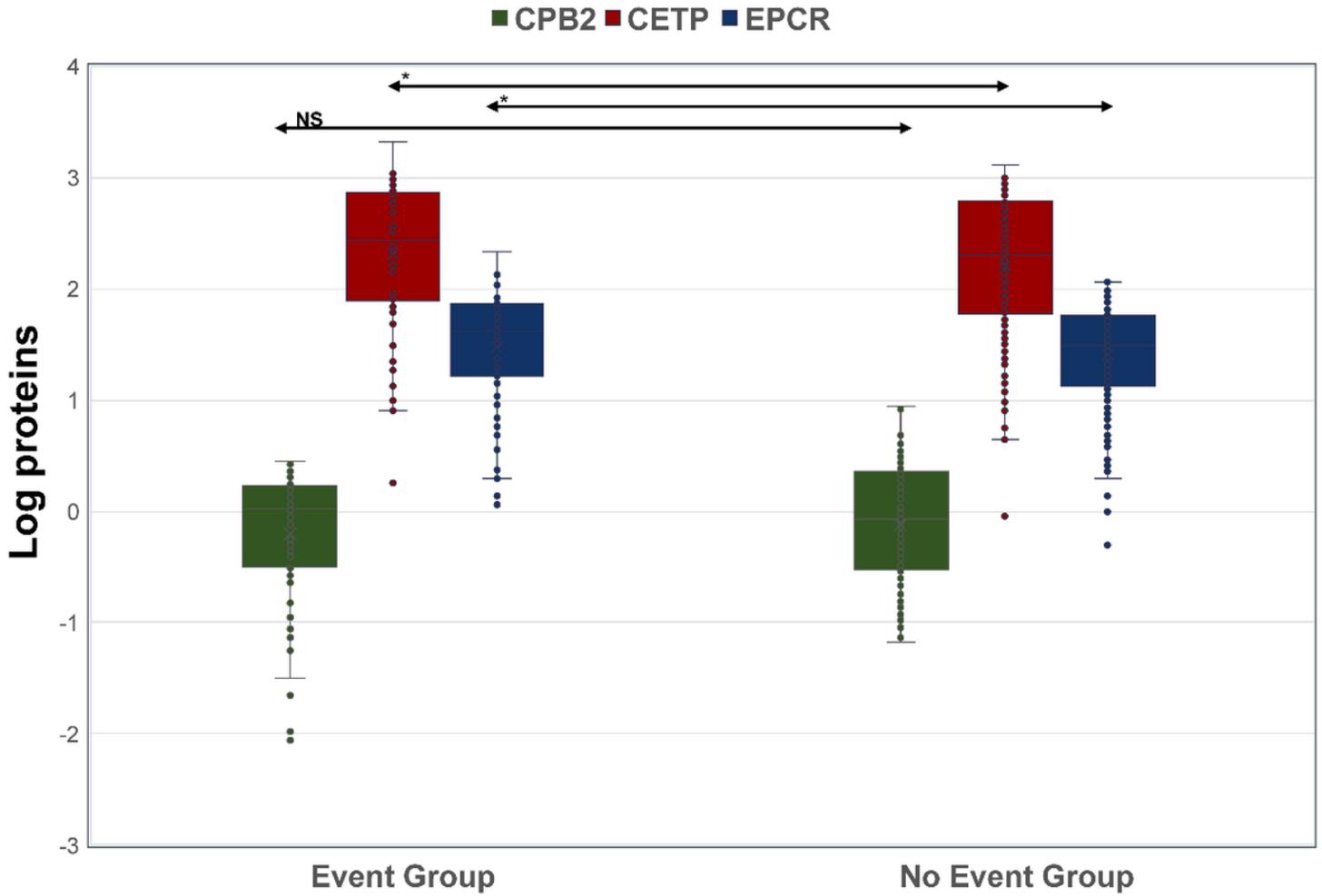


Figure 5

Expression levels of target proteins (CPB2, EPCR and CETP) in Validation Cohort Note: Compared between event group and no event group: *P value < 0.05; NS: No significance.

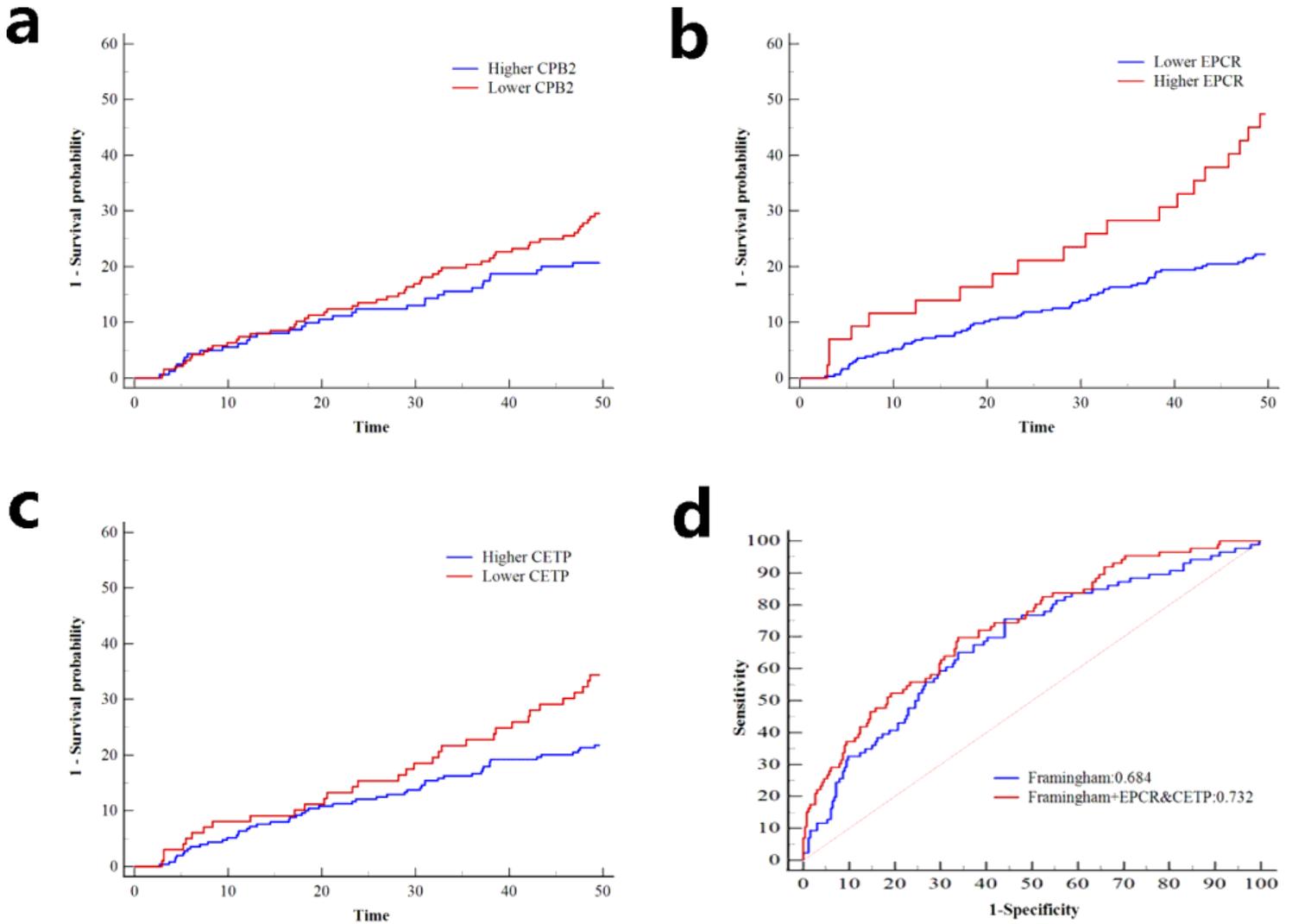


Figure 6

Kaplan-Meier cumulative risk curves of CPB2 (a), EPCR (b) and CETP (c) and comparative ROC curves of Framingham CHD risk model combined with CETP and EPCR (d) Note: Compared between two groups: p value (a)=0.086, p value (b)<0.001, p value (c)=0.023; Compared between two models: p value (d)=0.037.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [11.Supplementarymaterial.docx](#)