

A Preliminary Predictive Panel for Pre-Frailty Based on Serum Proteomic Biomarkers: A Two-Phase Cross-Sectional Study

Yu Ye^{1,2,*}, Jinwei Liu^{1,2,*}, Zhen Zhang^{1,2,*}, Shuaixuan Xu², Chenghao Chang^{1,2}, Mengyu Cao², Yongyi Zhang^{1,3}, Fang Wang², Nihui Zhang², Guibin Wang⁴, Nan Peng^{1,2}

¹Medical School, Chinese PLA General Hospital, Beijing, People's Republic of China; ²Department of Rehabilitation Medicine, second Medical Center of Chinese PLA General Hospital, Beijing, People's Republic of China; ³Department of Anaesthesiology, No. 962 Hospital of the PLA Joint Logistic Support Force, Harbin, Heilongjiang, People's Republic of China; ⁴State Key Laboratory of Medical Proteomics, Beijing Institute of Lifeomics, Beijing, People's Republic of China

*These authors have contributed equally to this work

Correspondence: Nan Peng; Guibin Wang, Email Pengnan301@163.com; wangguibin85@163.com

Background: Prefrailty is associated with anomalies in protein metabolism; however, the serum proteomic signatures remain unclear. This study investigated protein profiles across different health statuses and evaluated their potential for the early identification of prefrailty.

Methods: Older adults were categorized as robust, prefrail, and frail. Untargeted proteomic screening in a discovery cohort (n = 30) was followed by parallel reaction monitoring (PRM) validation (n = 99). Multidimensional clinical parameters and differentially expressed proteins were integrated within machine learning pipelines to refine the search for characteristic features of prefrailty.

Results: 166 proteins were found to be differentially expressed across frailty statuses, with 15 significantly prefrailty-associated proteins subsequently confirmed by PRM validation. These proteins were functionally enriched in pathways related to cell signaling, protein metabolism, immune regulation, and skeletal muscle function maintenance. A Random Forest model, further assembled from gait speed, skeletal muscle mass, E3-independent E2 ubiquitin-conjugating enzyme (UBE2O), Timed Up and Go test time, alpha-actinin-3 (ACTN3), and Mini-Mental State Examination score, exhibited the most robust performance for early frailty identification among multiple algorithms compared.

Conclusion: This exploratory study identified candidate serum protein biomarkers associated with prefrailty. Preliminary machine learning models incorporating UBE2O and ACTN3 suggested the feasibility of discriminating prefrailty from robust status, reflecting underlying proteomic heterogeneity among community-dwelling older adults.

Keywords: frailty, pre-frailty, older adults, proteomics, machine learning

Introduction

Frailty is a geriatric syndrome characterized by decreased physiological reserve, increased vulnerability, and reduced resistance to stressors due to multi-system dysregulation.¹ Currently, approximately 18% of older adults meet the criteria for frailty, while the proportion of prefrailty, the transitional phase from robustness to frailty, reaches 49%.² This transition poses a significant threat to healthy aging and represents one of the most serious public health challenges of the 21st century.^{3,4} Notably, prefrailty is a dynamic and reversible state, making its effective identification and timely intervention crucial for improving health outcomes in older populations.⁵

Prefrailty typically manifests as exhaustion, reduced gait speed, low grip strength, decreased physical activity, and unintentional weight loss. These clinical features largely overlap with sarcopenia, a condition primarily characterized by impaired muscle strength.⁶ Leveraging the concept of progressive loss in skeletal muscle mass and function is essential for elucidating the early progression of frailty. This musculoskeletal deterioration is fundamentally driven by disrupted

proteostasis and widespread proteomic remodeling.^{7,8} Studies indicate that the aberrant expression of pro-inflammatory and lipid-related proteins can accelerate functional decline by promoting catabolism, perturbing energy homeostasis, and engaging in multi-organ crosstalk.^{9,10} Such proteomic shifts likely serve as the pivotal mechanisms driving the transition toward the prefrailty phenotype.

Characterized as a multidimensional decline in physiological reserve, prefrailty is not attributable to fluctuations in individual proteins but rather reflects the coordinated dysregulation of multiple biological processes, including chronic inflammation, metabolic imbalance, cellular senescence, and anabolic resistance.^{11–13} A recent population-based cohort study identified that various circulating proteins remain persistently associated with the frailty index. The involvement of diverse biological pathways ranging from renal and skeletal systems to immune and apoptotic signaling underscores the multi-system deterioration characteristic of frailty.¹⁴ Consequently, compared to single-biomarker analyses, exploring interactions within protein networks offers a more comprehensive explanation for the onset and evolution of prefrailty. Furthermore, integrating rapid clinical assessments with multi-omic features via advanced machine-learning algorithms holds significant promise for enhancing the diagnostic precision and clinical utility of early frailty identification.^{15,16}

In this study, mass spectrometry-based proteomics was employed to screen for key proteins associated with prefrailty. By integrating differentially expressed proteins with clinical features into algorithmic models, we established a preliminary interpretable classification framework for prefrailty, providing a reference for its precise identification.

Methods

Study Design

This two-stage cross-sectional observational study was conducted between December 2022 and June 2023. Participants were recruited from community-based settings, including health service centers, seniors' activity centers, and via public platform advertisements in Haidian District, Beijing, China, using convenience sampling. The overall workflow is illustrated in Figure 1. Clinical evaluations and proteomic analyses were performed by independent personnel blinded to each other's data. The study framework comprised a discovery phase for untargeted proteomics screening and a validation phase using an independent cohort for targeted proteomics. Feature selection and comparative analyses of multiple machine learning algorithms were performed to identify the best combination for early frailty identification. All eligible participants provided written informed consent. The protocol adhered to the Declaration of Helsinki and was approved by the Ethics Committee of Chinese PLA General Hospital (Approval No. S2019-140-03).

Participants

A total of 129 community-dwelling older adults aged 60 years or older were eligible for inclusion in this study. Participants were categorized into robust, prefrail, and frail groups based on the Fried Physical Frailty Phenotype.¹⁷

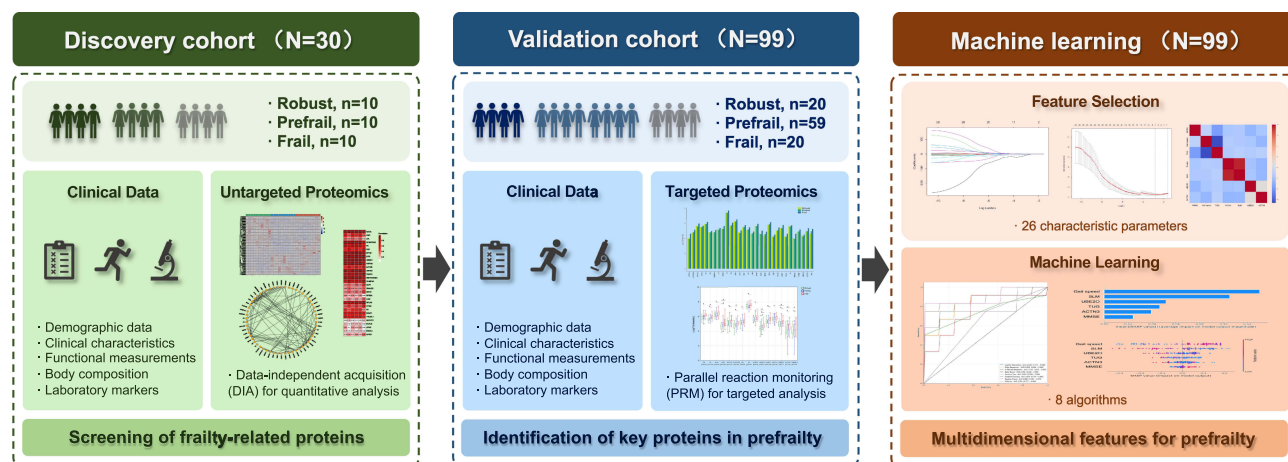


Figure 1 Flowchart of the study design.

Exclusion criteria included uncontrolled hypertension, acute cardiovascular or cerebrovascular events, active infection, or postoperative status within the preceding 3 months. Potential participants were also excluded for unstable functional status, severe pain, impairments in cognition, speech, or hearing, a history of substance abuse, and end-stage organ failure. Additional exclusion criteria involved contraindications to exercise, severe malnutrition defined by a Mini-Nutritional Assessment-Short Form (MNA-SF) score below 7, and major diseases affecting systemic protein metabolism such as malignancy, autoimmune diseases, or decompensated cirrhosis, along with the use of related medications. Individuals with pacemakers or metal implants that precluded body composition assessment were also excluded.

In the discovery phase, 30 participants were included based on common design strategies for exploratory proteomics studies, with 10 individuals in each group.¹⁸ In the validation phase, an independent cohort was recruited. Sample size was estimated using G*Power 3.1, assuming a medium-to-large effect size (Cohen's $f = 0.50$),^{19,20} a significance level of $\alpha = 0.05$, and a statistical power of $1 - \beta = 0.80$, resulting in a minimum required sample size of 42 for comparisons among three groups. Considering the prevalence of prefrailty in community populations and potential data loss,²¹ 99 participants were ultimately included in the validation cohort, consisting of 20 robust individuals, 59 prefrail individuals, and 20 frail individuals.

Assessment of Pre-Frailty

Prefrailty was defined using the Fried Physical Frailty Phenotype, which evaluates five clinical criteria.¹⁷

1. Weight loss: Unintentional loss of >4.5 kg or $>5\%$ of body weight within the preceding year.
2. Gait speed: walking time ≥ 6 s or ≥ 7 s over a 4.57-meter course, adjusted for sex and height.
3. Grip strength: Measured using a JAMAR Plus handheld dynamometer (Sammons Preston, USA). Participants were seated with their dominant shoulder in a neutral position and elbow flexed at 90° . After two 5-second maximal voluntary contractions, the average value was recorded. Thresholds were stratified by sex and body mass index (BMI), ranging from 29–32 kg for men and 17–21 kg for women.
4. Physical activity: Identified by a weekly energy expenditure below 383 kcal for men or 270 kcal for women.
5. Exhaustion: Scored 1 if either “everything I did was an effort” or “I could not get going” persisted for ≥ 3 days during the past week.

The total score ranged from 0 to 5. Participants were categorized as robust with a score of 0, prefrail with scores of 1 to 2, and frail with scores of 3 or higher.

Clinical Data Collection

Demographic data included gender (male/female), age (years), education (years), BMI (kg/m^2), smoking history (yes/no), alcohol consumption (yes/no), insomnia status (yes/no), sedentary behavior (yes/no) and the Mini-Mental State Examination (MMSE). MMSE scores range from 0 to 30, with cognitive impairment cut-offs set at ≤ 17 , ≤ 20 , and ≤ 24 for uneducated, primary, and middle school or above education levels, respectively.^{22,23}

Clinical characteristics were evaluated using the Fried Physical Frailty Phenotype (Fried score). The Number of chronic diseases was quantified by identifying clinical items defined in the Charlson Comorbidity Index framework.²⁴ Pain intensity was determined by Visual Analogue Scale (VAS) on a 10-cm scale where 0 represents no pain and 10 represents unbearable pain.²⁵ Nutritional status was assessed using the MNA-SF, where scores of 12–14 indicate normal status, 8–11 indicate risk of malnutrition, and 0–7 indicate malnutrition.²⁶ Depressive symptoms were screened via the Geriatric Depression Scale-15 (GDS-15), where a score of 5 or higher suggests depression.²⁷ The Modified Barthel Index (MBI) evaluated activities of daily living on a scale of 0 to 100, where lower scores indicate greater dependence.²⁸

Functional measurements consisted of gait speed, grip strength, Timed Up and Go Test (TUG) time, and the Short Physical Performance Battery (SPPB) score. For the TUG, the average time was recorded for participants to rise from a 46-cm chair, walk 3 meters, and return. The SPPB integrated balance, a 4-meter walk, and five-times sit-to-stand tests, with a maximum score of 12 where lower scores indicate poorer physical function.²⁹

Body composition was measured via direct segmental multi-frequency bioimpedance analysis (DSM-BIA) using the InBody 770 device (Korea). Parameters obtained included total body protein (Protein), body fat mass (BFM), skeletal muscle mass (SLM). The standard ranges for these parameters were defined in kg as 90%–110% of the predicted ideal values based on individual age, gender, and height.

Laboratory markers and their normal reference ranges were as follows: 25-hydroxyvitamin D (Vit D) was 30 to 60 ng/mL, cystatin C (CysC) was 0.053 to 0.095 mg/dL, insulin-like growth factor-1 (IGF-1) was 60 to 200 ng/mL, c-reactive protein (CRP) was 0 to 0.9 mg/L, creatine kinase (CK) was 2 to 200 U/L, creatinine (Cr) was 30 to 110 μ mol/L, and triglycerides (TG) was 0.4 to 1.7 mmol/L.

Serum Sample Preparation

Blood samples were collected from participants under fasting conditions using vacuum tubes, with a volume of 5 mL per sample. Within one hour after collection, the blood was centrifuged at 3,000 rpm for 10 minutes to separate the serum. The supernatant was carefully aliquoted and stored at -80°C until analysis. For subsequent testing, frozen serum samples were thawed to room temperature prior to processing.

Untargeted Proteomic

Untargeted proteomic profiling was performed using data-independent acquisition (DIA) for quantitative analysis.³⁰ After protein extraction with lysis buffer (8 M urea containing protease inhibitors), high-abundance proteins were removed using an Agilent MARS Hu-14 immunoaffinity column to enrich low-abundance proteins from the flow-through fraction. Protein concentration was determined by the bicinchoninic acid assay. Subsequent sample processing included reduction (10 mM dithiothreitol, 56°C), alkylation (55 mM iodoacetamide, in the dark), and digestion via the filter-aided sample preparation method. The resulting peptides were desalted using a C18 column prior to analysis.

Peptide separation was carried out on an EASY-nLC 1200 nano-LC system equipped with a C18 analytical column, using mobile phase A (0.1% formic acid in water) and mobile phase B (0.1% formic acid in 80% acetonitrile). The gradient elution program was as follows: 7–12% B, 12–30% B, 30–45% B, 45–95% B, and maintained until 60 min.

Mass spectrometry analysis was performed on a Q Exactive HF mass spectrometer equipped with a Nanospray Flex ion source. The spray voltage was set to 2.2 kV and the ion transfer temperature to 320°C . DIA acquisition parameters included a full-scan range of 350–1150 m/z at a resolution of 120,000, followed by 30 MS/MS windows acquired at a resolution of 30,000 with a normalized collision energy of 33%.

DIA data were processed using Spectronaut software against the UniProt human database. Search settings included trypsin digestion with up to two missed cleavages, fixed modification of carbamidomethylation on cysteine, and variable modifications of methionine oxidation and N-terminal acetylation.

Targeted Proteomic

Targeted proteomic quantification was performed using mass spectrometry-based parallel reaction monitoring (PRM). This method enables both quantitative and qualitative analysis of specific proteins by monitoring the ion signals of predetermined target peptides, with at least two unique peptides selected per protein.³¹

Following tryptic digestion, peptides were separated by liquid chromatography and subsequently analyzed by mass spectrometry. PRM acquisition comprises two main steps: Precursor Ion Selection—specific precursor ions corresponding to the target peptides were isolated based on their predefined mass-to-charge (m/z) ratios; and Fragment Ion Monitoring—after fragmentation, the resulting product ions were recorded. Quantification was performed based on the signal intensity of these fragment ions, allowing relative or absolute protein abundance to be determined.

Analyses were carried out on a Q-Exactive high-resolution mass spectrometer, which offers precise mass accuracy and high sensitivity to ensure reliable detection of low-abundance proteins. Peak identification was performed using chromatographic and spectral data generated from the instrument. The signal intensity (peak area) of each peptide was extracted for quantification. Absolute quantification was performed for target peptides using calibration curves constructed from synthetic peptide standards of known concentrations.

Statistical Analysis

Data were processed and analyzed using SPSS Statistics, version 27.0 (IBM Corp., Armonk, NY, USA). All statistical tests were two-sided, with $P < 0.05$ considered statistically significant. Missing data for all variables were $< 5\%$ and addressed via Markov Chain Monte Carlo (MCMC) multiple imputation with five iterations. Outlier treatment was performed in conjunction with the actual pathophysiological status of the participants. When outliers were consistent with pathological upregulation, they were considered valid data and retained for analysis.

Normality of continuous variables was assessed using the Shapiro–Wilk test. Continuous variables were expressed as mean $\pm$ standard deviation (Mean $\pm$ SD) for normally distributed data, or as median with interquartile range [M (P25–P75)] for non-normally distributed data. Comparisons among groups were performed using one-way analysis of variance (ANOVA) for normally distributed data, and the Kruskal–Wallis test for non-normally distributed data. Categorical variables were expressed as frequencies and percentages and analyzed using the chi-squared test.

For proteomics data, protein abundance was $\log_2$ -transformed and normalized. In the discovery phase, differentially expressed proteins among the three groups were identified using one-way ANOVA or the Kruskal–Wallis test according to data distribution, with the false discovery rate (FDR) controlled using the Benjamini–Hochberg method. Differential expression was defined as $|\log_2 \text{ fold change (FC)}| > 0.58$ and $\text{FDR} < 0.05$. Mfuzz clustering and Spearman correlation analysis were applied to identify core proteins, followed by gender-specific filtering.

In the validation phase, PRM data were analyzed using the same statistical criteria. When a global test was significant ($P < 0.05$), post-hoc pairwise comparisons were conducted with Bonferroni correction to control for Type I error.

Bioinformatics Analysis

Functional annotation and enrichment analyses were performed to explore the biological significance of differentially expressed proteins. Gene Ontology (GO) terms, including biological process, cellular component, and molecular function, were obtained from the GO database (<http://www.geneontology.org/>). Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was conducted using the KEGG database (<https://www.kegg.jp/kegg/pathway.html>). Enrichment analysis was performed using a hypergeometric test, and the Benjamini–Hochberg method was applied to control FDR. Enrichment was considered statistically significant at $P < 0.05$ and $\text{FDR} < 0.05$.

Protein–protein interaction (PPI) networks were constructed by mapping differentially expressed proteins to the STRING database (<https://string-db.org/>) with a medium confidence score threshold of 0.400. The network was visualized using Cytoscape, and hub proteins were identified based on topological analysis.

Feature Selection

The dependent variable was defined as prefrailty, with a Fried score of 1–2 serving as the outcome of interest. Robust and frail groups were included as controls to help exclude false-positive proteins potentially arising from non-specific aging in the older adult population. A total of 26 candidate parameters were considered as independent variables, among which those showing statistically significant differences across the three groups in the validation cohort were selected ($P < 0.05$). These included 11 clinical parameters, namely sedentary behavior, MMSE, MBI, gait speed, grip strength, TUG, SPPB, protein, SLM, CysC, and IGF-1, and 15 differentially expressed proteins, including CP, F8, HPX, VWF, RPLP0, ATP5F1B, LPL, F5, MYH9, HSPG2, SVEP1, BPIFB1, UBE2O, ACTN3, and TREML1.

First, the least absolute shrinkage and selection operator (LASSO) regression algorithm was applied for feature standardization by adding an L1 norm to the objective function to penalize the regression coefficients. This resulted in the shrinkage of regression coefficients of redundant variables to zero, thereby achieving feature selection, reducing model complexity, and addressing multicollinearity.³²

Subsequently, Kendall's rank correlation coefficient (τ) was used to detect monotonic correlations among features to ensure model parsimony and stability.³³ Pairwise Kendall correlation coefficients were calculated for all features to generate a symmetric matrix. A threshold of $|\tau| \geq 0.8$ was set, where feature pairs with absolute correlation values above this threshold were considered redundant. In such cases, the feature with a stronger correlation to the target variable was retained, and the other was removed.

Machine Learning

Model development and comparison were performed in the Python 3.7 programming environment using eight machine learning algorithms based on the selected important features. These included ensemble learning algorithms suitable for high-dimensional data: random forest (RF), gradient boosting (GB), and eXtreme gradient boosting (XGBoost);^{34–36} linear algorithms with good interpretability: logistic regression (LR) and ridge regression (RR);^{37,38} a probabilistic algorithm: naive Bayes (NB);³⁹ an instance-based learning method: K-nearest neighbors (KNN);⁴⁰ and a single-tree algorithm: decision tree (DT).⁴¹ During the modeling process, five rounds of cross-validation and grid search for parameter optimization were repeated to prevent overfitting. Hyperparameter tuning for all algorithms was performed on the training set using nested cross-validation. The final results were summarized across five rounds of validation, and the average performance of each algorithm was calculated. Model performance was evaluated primarily using the receiver operating characteristic (ROC) curve and the area under the curve (AUC) along with its 95% confidence interval (CI), supplemented by accuracy, precision, recall, and F1 score for comprehensive assessment of the classification performance of the eight machine learning algorithms.

Model Interpretability

Shapley Additive Explanations (SHAP) were applied to interpret the optimal model. Rooted in cooperative game theory, SHAP assigns each feature an importance value by averaging its marginal contribution across all possible feature subsets. This approach provides a mathematically grounded, transparent explanation of how each feature influences model predictions, thereby enhancing the interpretability and credibility of the model.⁴²

Results

Clinical Characteristics in Discovery Cohort

In the discovery phase, the mean age of the participants was 77.00 ± 3.17 years, and 15 participants were male (50%). No statistically significant differences were observed in general demographic characteristics among the three groups (all $P > 0.05$). In terms of clinical functional performance, Fried score ($P < 0.001$), VAS score ($P = 0.010$), GDS-15 score ($P = 0.013$), and TUG time increased with the severity of frailty ($P = 0.009$). MNA-SF score ($P = 0.024$), MBI ($P = 0.009$), gait speed ($P = 0.003$), grip strength ($P = 0.013$), and SPPB score ($P = 0.010$) were all decreased in both the prefrail and frail groups. Circulating levels of CK ($P = 0.038$) also decreased progressively from the robust to the frail state. No significant inter-group differences were observed for number of chronic diseases, protein, BFM, SLM, Vit D, CysC, IGF-1, CRP, Cr, or TG (all $P > 0.05$, Table 1).

Proteins in the Discovery Cohort

In the discovery cohort, a total of 1,145 raw proteins were identified at quantifiable levels. Comparative analysis revealed 166 proteins that were differentially expressed among the robust, prefrail, and frail groups (Figure 2A). To elucidate the biological functions associated with frailty progression, enrichment analyses were performed using the GO and KEGG databases. Significantly enriched GO terms included vesicle-mediated transport, regulation of response to stimulus, regulation of immune system processes, platelet activation, regulation of cell communication, proteolysis, muscle structure development, and muscle cell differentiation (Figure 2B).

KEGG pathway analysis revealed that the transition from robust to frail states involved complement and coagulation cascades, ferroptosis, focal adhesion, riboflavin metabolism, platelet activation, cholesterol metabolism, glutathione metabolism, glycerolipid metabolism, and ubiquitin-mediated proteolysis (Figure 2C).

PPI was constructed to visualize functional connectivity among the differentially expressed proteins (Figure 2D). Clustering analysis highlighted 57 key proteins involved in biological processes such as regulation of adaptive immune response, regulation of cytokine production, protein polyubiquitination, regulation of skeletal muscle tissue development, lipid metabolism, and proteolysis.

Based on the bioinformatic analysis of differentially expressed proteins, candidate proteins were prioritized according to their correlation with clinical phenotypes (Figure 2E); the six clinical indicators were selected based on their

Table 1 Comparison of Clinical Characteristics in the Discovery Cohort (N=30)

Characteristic	Robust (n=10)	Prefrail (n=10)	Frail (n=10)	F/Z/ χ^2	P value
Demographic characteristics					
Gender (male/female, n)	5/5	5/5	5/5	<0.001	1.000
Age (years)	75.60±1.17	78.10±4.58	77.30±2.54	1.697	0.202
Education (years)*	12.00(9.00–16.00)	9.00(8.25–16.00)	11.32(9.00–12.00)	0.476	0.788
BMI (kg/m ²)	23.86±1.01	24.07±3.46	24.16±4.80	0.020	0.980
Smoking history (yes/no)	4/6	2/8	4/6	1.200	0.549
Alcohol consumption (yes/no)	6/4	2/8	3/7	3.732	0.155
Insomnia (yes/no)	3/7	7/3	8/2	5.833	0.054
Sedentary behavior (yes/no)	2/8	5/5	7/3	5.089	0.079
MMSE (score)*	27.50(25.00–29.00)	26.50(25.00–29.00)	27.50(25.00–28.25)	0.327	0.849
Clinical characteristics					
Fried (score)*	0.00(0.00–0.00)	1.00(1.00–1.25)	3.00(3.00–3.25)	27.878	<0.001
Chronic diseases(count)*	2.00(1.00–2.50)	2.22(1.00–4.02)	3.50(1.75–4.00)	2.421	0.298
VAS (score)*	0.00(0.00–0.00)	0.00(0.00–5.00)	4.00(0.28–5.25)	9.280	0.010
MNA-SF (score)*	14.00(14.00–14.00)	12.50(10.00–14.00)	13.00(10.75–14.00)	7.441	0.024
GDS-15 (score)*	1.00(0.00–2.00)	1.00(1.00–2.00)	3.20(1.75–6.25)	8.730	0.013
MBI (score)*	100.00(99.50–100.00)	99.00(94.25–100.00)	93.50(86.75–98.50)	9.368	0.009
Gait speed (m/s)	1.45±0.56	1.21±0.37	0.76±0.25	7.066	0.003
Grip strength (kg)	31.93±9.46	29.75±6.42	22.03±5.35	5.093	0.013
TUG (time)*	8.49(7.07–10.40)	8.96(8.10–12.51)	11.82(10.71–13.24)	9.363	0.009
SPPB (score)*	12.00(10.75–12.00)	11.00(9.75–11.00)	9.00(7.75–10.25)	9.206	0.010
Protein (kg)	9.16±1.58	8.75±1.64	8.21±1.58	0.887	0.424
BFM (kg)	17.75±2.06	21.00±7.30	22.07±9.93	0.973	0.391
SLM (kg)	44.18±7.47	42.47±7.94	40.25±7.78	0.648	0.531
Vit D (ng/mL)	22.55±6.52	23.19±13.19	17.68±4.21	1.163	0.328
CysC (mg/dL)	0.09±0.01	0.11±0.02	0.11±0.03	2.940	0.070
IGF-I (ng/mL)*	13.99(10.90–22.02)	29.41(18.96–33.26)	21.02(15.37–27.80)	5.772	0.056
CRP (mg/L)*	0.92(0.13–1.66)	0.49(0.35–0.63)	0.63(0.16–2.80)	0.321	0.852
CK (U/L)*	89.91(64.75–110.88)	87.80(67.10–138.48)	57.95(53.30–84.37)	6.521	0.038
Cr (μmol/L)	71.99±12.11	70.15±15.64	70.01±13.41	0.064	0.938
TG (mmol/L)*	1.10(0.97–1.86)	1.27(0.86–1.85)	1.18(1.07–1.71)	0.156	0.925

Notes: Qualitative data are presented as frequency and analyzed by chi-squared test; Normally distributed quantitative data are expressed as mean ± SD and analyzed by ANOVA; *, Non-normally distributed data are expressed as M (P25–P75) and analyzed by Kruskal–Wallis test.

Abbreviations: BMI, body mass index; MMSE, Mini-Mental State Examination; Fried, Fried Physical Frailty Phenotype; VAS, Visual Analogue Scale; MNA-SF, Mini-Nutritional Assessment-Short Form; GDS-15, Geriatric Depression Scale; MBI, Modified Barthel Index; TUG, Timed Up and Go Test; SPPB, Short Physical Performance Battery; Protein, total body protein; BFM, body fat mass; SLM, skeletal muscle mass; Vit D, 25-hydroxyvitamin D; CysC, cystatin C; IGF I, insulin-like growth factor I; CRP, c-reactive protein; CK, creatine kinase; Cr, creatinine; TG, triglycerides.

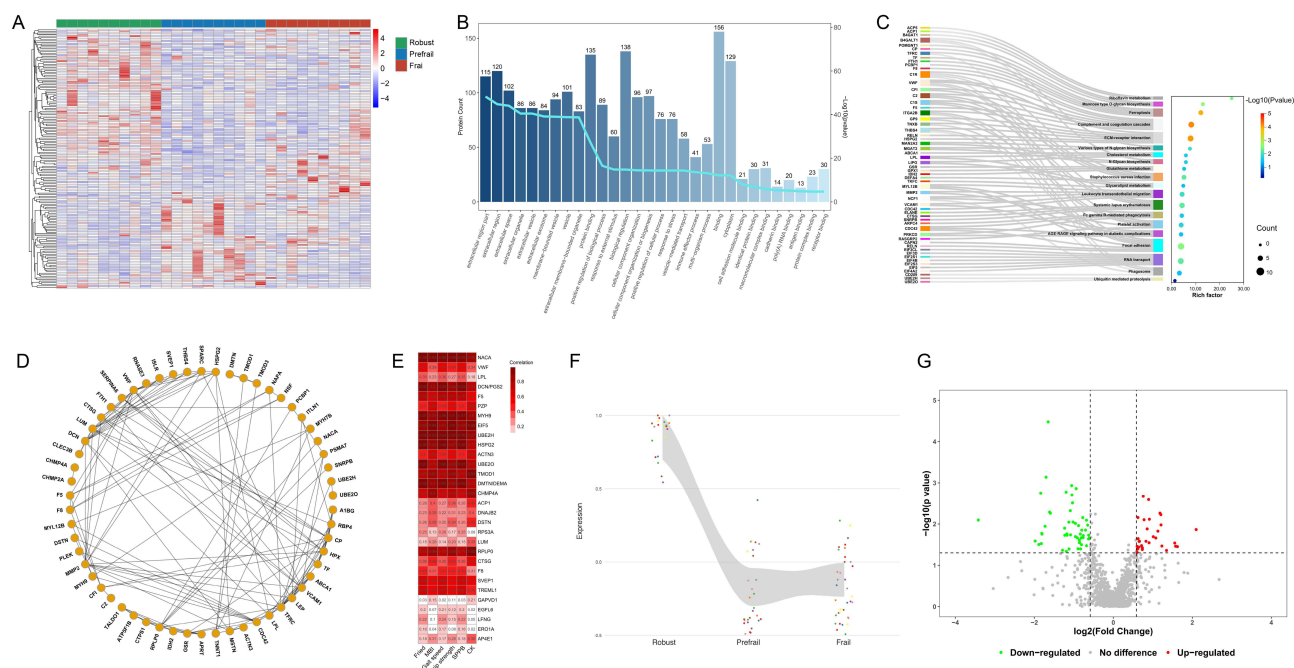


Figure 2 Discovery phase untargeted proteomic analysis (N = 30). **(A)** Heatmap of 166 differentially expressed proteins across three groups; color scale indicates relative protein abundance (blue, low; red, high). **(B)** Gene Ontology (GO) enrichment analysis; bars represent protein counts; line indicates enrichment score. **(C)** Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment. **(D)** Protein-protein interaction (PPI) network; edges represent interactions. **(E)** Correlation matrix of key proteins and clinical phenotypes; color intensity reflects correlation strength. **(F)** Expression pattern map. **(G)** Volcano plot of sex-specific differences; red, upregulated; green, downregulated. Differentially expressed proteins were identified using one-way ANOVA or Kruskal-Wallis test, Benjamini-Hochberg-adjusted $P < 0.05$, and $|\log_2 \text{fold change}| > 0.58$. Protein abbreviations in Table 2. Fried, Fried Physical Frailty Phenotype; MBI, Modified Barthel Index; SPPB, Short Physical Performance Battery; CK, creatine kinase.

significant baseline differences and clinical representativeness in assessing physical performance. These protein targets were further filtered by evaluating their fold-change expression trends (Figure 2F) and sex-specific distributions (Figure 2G). These analyses led to the identification of 30 candidate proteins as potential prefrailty biomarkers for subsequent PRM-based validation in the verification cohort (all FDR < 0.05; Table 2).

Clinical Characteristics in Validation Cohort

In the validation phase, the mean age of participants was 75.32 ± 5.09 years, with 27 males (27.3%). The prevalence of sedentary behavior was significantly higher in the prefrail and frail groups (47.5% and 95.0%, respectively; $P = 0.001$) compared to the robust group (45.0%). MMSE scores decreased progressively with worsening frailty status ($P = 0.004$). No significant differences were observed among the three groups in gender, age, years of education, BMI, smoking history, alcohol consumption use, or insomnia (all $P > 0.05$).

Fried scores differed significantly among groups ($P < 0.001$). Compared to the robust group, the frail group showed significantly lower scores in the MBI ($P < 0.001$), gait speed ($P < 0.001$), grip strength ($P = 0.010$), and SPPB ($P < 0.001$). TUG time was significantly longer in both the prefrail and frail groups than in the robust group ($P < 0.001$). Protein and SLM were notably lower in the prefrail group compared to the robust group ($P = 0.017$ and $P = 0.009$, respectively). CysC levels increased significantly only at the frail stage ($P = 0.040$), whereas IGF-1 followed a non-linear trend, decreasing slightly in prefrailty before peaking in the frail group ($P = 0.004$). In contrast, no significant inter-group differences were detected for number of chronic diseases, VAS, MNA-SF, GDS-15, BFM, Vit D, CRP, CK, Cr, or TG levels (all $P > 0.05$; Table 3).

Targeted Proteins in the Validation Cohort

Targeted proteomic analysis via PRM was performed in the validation cohort to quantify the expression levels of 30 candidate proteins previously identified in the discovery phase (Figure 3A). Among these, 15 proteins exhibited

Table 2 Candidate Proteins Identified by ANOVA for Parallel Reaction Monitoring (PRM) Validation (N=30)

UniProt ID	Protein	Name	F	FDR
E9PAV3	NACA	Nascent-Polypeptide-Associated Complex Alpha Polypeptide	3.592	0.043
P00450	CP	Ceruloplasmin	3.919	0.030
P00451	F8	Coagulation Factor VIII	4.579	0.031
P02787	TF	Transcription Factor	3.923	0.023
P02790	HPX	Hemopexin	3.861	0.024
P04275	VWF	Von Willebrand Factor	4.771	0.017
P05156	CFI	Complement Factor I	3.749	0.039
P05388	RPLP0	60S Acidic Ribosomal Protein P0	3.448	0.048
P06576	ATP5F1B	ATP Synthase Subunit Beta, Mitochondrial	6.353	0.003
P06681	C2	Complement component 2	4.247	0.043
P06858	LPL	Lipoprotein Lipase	4.594	0.019
P07585	DCN	Decorin	17.270	<0.001
P08311	CTSG	Cathepsin G	3.752	0.027
P12259	F5	Coagulation Factor V	4.579	0.042
P20742	PZP	Pregnancy Zone Protein	3.927	0.024
P28289	TMOD1	Tropomodulin-1	3.358	0.048
P35443	THBS4	Thrombospondin 4	4.743	0.025
P35579	MYH9	Myosin-9	6.566	0.006
P51884	LUM	Lumican	3.758	0.029
P55010	EIF5	Eukaryotic Translation Initiation Factor 5	6.952	0.003
P62256	UBE2H	Ubiquitin-conjugating Enzyme E2 H	5.162	0.010
P98160	HSPG2	Basement Membrane-Specific Heparan Sulfate Proteoglycan Core Protein	4.872	0.022
Q08043	ACTN3	Alpha-Actinin-3	4.855	0.013
Q08495	DEMA	Dynamin-Related Protein 1 / Dermatological Evaluation of Mucocutaneous Adverse Events	3.919	0.027
Q4LDE5	SVEP1	Sushi, Von Willebrand Factor Type A, EGF and Pentraxin Domain-containing Protein 1	3.878	0.023
Q86YW5	TREML1	Trem-like Transcript 1 Protein	3.542	0.035
Q8TDL5	BPIFB1	BPI Fold-containing Family B Member 1	6.241	0.006
Q8WWA0	ITLN1	Intelectin 1	3.424	0.041
Q9BY43	CHMP4A	Charged Multivesicular Body Protein 4A	4.253	0.015
Q9C0C9	UBE2O	E3-independent E2 Ubiquitin-conjugating Enzyme	4.273	0.021

Abbreviation: FDR, false discovery rate.

statistically significant differential expression across the robust, prefrail, and frail groups ($P < 0.05$; [Figure 3B](#)), including ceruloplasmin (CP), coagulation factor VIII (F8), hemopexin (HPX), von Willebrand factor (VWF), 60S acidic ribosomal protein P0 (RPLP0), ATP synthase subunit beta mitochondrial (ATP5F1B), lipoprotein lipase (LPL), coagulation factor V (F5), myosin-9 (MYH9), basement membrane-specific heparan sulfate proteoglycan core protein (HSPG2), sushi, von Willebrand factor type A, EGF and pentraxin domain-containing protein 1 (SVEP1), BPI fold-containing family B member 1 (BPIFB1), E3-independent E2 ubiquitin-conjugating enzyme (UBE2O), alpha-actinin-3 (ACTN3), and triggering receptor expressed on myeloid cells like transcript-1 (TREML1).

Feature Selection

A total of 26 independent variables showing statistical significance in the validation phase were subjected to LASSO regression analysis ([Figure 4A](#)). Feature selection was performed using LASSO with cross-validation, and seven features with non-zero coefficients corresponding to the optimal lambda on the left side of the curve were selected ([Figure 4B](#)). Based on the LASSO results, Kendall correlation analysis ([Figure 4C](#)) revealed high collinearity between SLM and total protein content ($\tau = 0.95$). The redundant feature with relatively lower nonlinear association, “Protein” was removed. The remaining six features, namely MMSE, gait speed, TUG, SLM, UBE2O, and ACTN3, were subsequently used to construct the machine learning model.

Table 3 Comparison of Clinical Characteristics in the Validation Cohort (N=99)

Characteristic	Robust (n=20)	Prefrail (n=59)	Frail (n=20)	FIZI χ^2	P value
Demographic characteristics					
Gender (male/female, n)	7/13	12/47	8/12	3.666	0.160
Age (years)	75.55±4.20	74.71±4.97	76.90±6.06	1.419	0.247
Education (years)*	11.28(9.00–16.00)	12.00(9.00–14.00)	12.00(9.00–16.00)	0.452	0.798
BMI (kg/m ²)	25.70±3.65	23.91±3.02	23.69±3.48	2.616	0.078
Smoking history (yes/no)	1/19	8/51	2/18	1.139	0.566
Alcohol consumption (yes/no)	6/14	13/46	2/18	2.452	0.293
Insomnia (yes/no)	10/10	32/27	11/9	0.129	0.937
Sedentary behavior (yes/no)	9/11	28/31	19/1	15.106	0.001
MMSE (score)*	28.00(27.00–29.00)	28.00(26.30–29.00)	26.00(25.00–27.00)	10.941	0.004
Clinical characteristics					
Fried (score)*	0.00(0.00–0.00)	1.00(1.00–2.00)	3.00(3.00–4.00)	61.308	<0.001
Chronic diseases (count)*	2.50(2.00–4.00)	2.00(1.00–4.00)	3.00(2.00–5.50)	4.454	0.110
VAS (score)*	0.00(0.00–1.00)	0.00(0.00–5.00)	3.00(0.00–5.50)	5.557	0.060
MNA-SF (score)*	14.00(13.00–14.00)	13.00(12.00–14.00)	13.00(12.00–14.00)	4.895	0.090
GDS-15 (score)*	1.00(0.00–2.00)	2.00(1.00–3.50)	2.00(1.00–4.50)	5.494	0.060
MBI (score)*	100.00(98.00–100.00)	100.00(96.00–100.00)	88.00(83.5–94.00)	30.948	<0.001
Gait speed (m/s)*	1.13(0.90–1.40)	1.09(0.97–1.30)	0.69(0.47–0.81)	23.608	<0.001
Grip strength (kg)*	25.45(23.60–35.35)	24.10(21.95–29.55)	21.70(18.80–25.10)	9.785	0.010
TUG (time)*	8.74(8.15–9.64)	9.00(7.80–11.51)	15.42(13.44–25.72)	24.813	<0.001
SPPB (score)*	11.00(9.00–11.50)	11.00(10.00–12.00)	7.00(3.50–8.50)	28.535	<0.001
Protein (kg)*	8.60(7.80–10.55)	7.60(7.25–8.50)	8.40(7.30–9.25)	8.205	0.017
BFM (kg)	22.08±6.52	20.59±5.86	20.23±5.67	0.589	0.557
SLM (kg)*	41.70(38.05–51.15)	37.40(35.00–41.35)	41.15(35.75–45.85)	9.403	0.009
Vit D (ng/mL)*	17.93(11.26–27.93)	17.28(13.45–24.64)	16.77(11.04–23.03)	0.983	0.612
CysC (mg/dL)*	0.10(0.09–0.12)	0.10(0.09–0.11)	0.12(0.10–0.14)	6.441	0.040
IGF-I (ng/mL)*	20.83(12.92–34.12)	18.97(13.62–27.64)	38.75(21.71–49.42)	11.007	0.004
CRP (mg/L)*	0.52(0.30–1.33)	0.54(0.30–1.22)	0.73(0.20–1.22)	0.072	0.965
CK (U/L)*	85.15(72.25–114.30)	75.80(58.10–100.00)	62.65(44.05–110.90)	4.887	0.087
Cr (μmol/L)*	65.53(48.20–81.25)	65.00(53.65–77.55)	70.65(59.85–91.00)	4.767	0.092
TG (mmol/L)*	1.55(1.22–2.17)	1.23(0.90–1.63)	1.34(1.12–1.96)	4.859	0.088

Notes: Qualitative data are presented as frequency and analyzed by chi-squared test; Normally distributed quantitative data are expressed as mean ± SD and analyzed by ANOVA; *, Non-normally distributed data are expressed as M (P25–P75) and analyzed by Kruskal–Wallis test.

Abbreviations: BMI, body mass index; MMSE, Mini-Mental State Examination; Fried, Fried Physical Frailty Phenotype; VAS, Visual Analogue Scale; MNA-SF, Mini-Nutritional Assessment-Short Form; GDS-15, Geriatric Depression Scale; MBI, Modified Barthel Index; TUG, Timed Up and Go Test; SPPB, Short Physical Performance Battery; Protein, total body protein; BFM, body fat mass; SLM, skeletal muscle mass; Vit D, 25-hydroxyvitamin D; CysC, cystatin C; IGF I, insulin-like growth factor I; CRP, c-reactive protein; CK, creatine kinase; Cr, creatinine; TG, triglycerides.

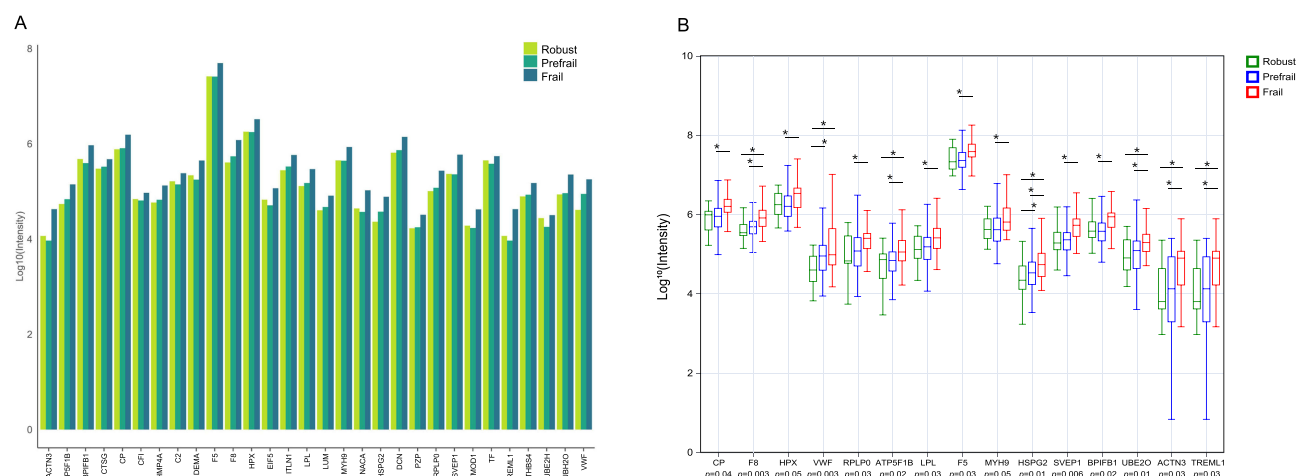


Figure 3 Validation phase targeted proteomic analysis (N = 99). **(A)** Quantitative expression of candidate proteins; relative abundance of 30 proteins. **(B)** Differential protein expression among the three groups; *, Bonferroni-corrected post-hoc pairwise comparisons, $P < 0.05$. Protein abbreviations in Table 2.

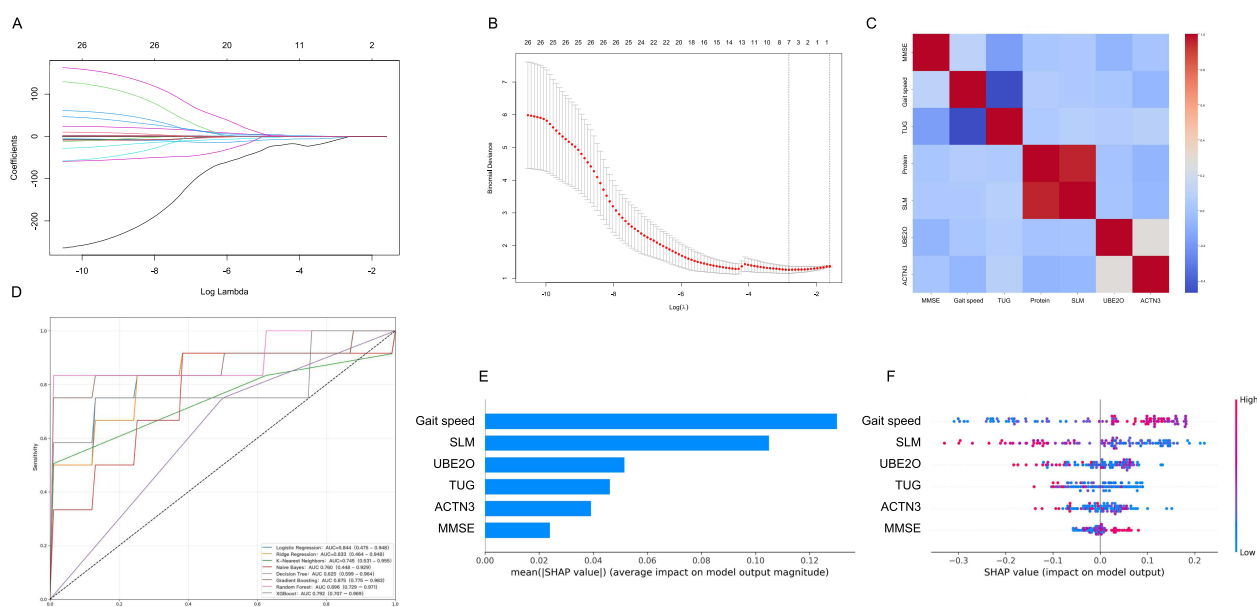


Figure 4 Machine learning analysis (N = 99). **(A)** LASSO coefficient profiles; variable selection across $\log(\lambda)$. **(B)** Cross validation curve; selection of optimal λ based on binomial deviance. **(C)** Feature correlation heatmap; relationships among selected variables. **(D)** Receiver operating characteristic (ROC) curves of the eight machine learning models. **(E)** SHAP feature importance; ranking based on mean SHAP values in the random forest model. **(F)** SHAP summary plot.

Abbreviations: MMSE, Mini-Mental State Examination; TUG, Timed Up and Go Test; Protein, total body protein; SLM, skeletal muscle mass; UBE2O, E3-independent E2 ubiquitin-conjugating enzyme; ACTN3, alpha-actinin-3.

Performance Comparison of Machine Learning

Based on the six selected important features, the study systematically evaluated the classification performance of eight widely used machine learning algorithms, including RF, GB, XGBoost, LR, RR, NB, KNN, and DT (Figure 4D). The results showed that ensemble learning methods exhibited the most prominent overall performance. The RF algorithm achieved the highest AUC (0.896, 95% CI: 0.792–0.971), indicating excellent discriminative ability and stability. GB ranked second (AUC = 0.875, 95% CI: 0.775–0.983), followed by XGBoost (AUC = 0.792, 95% CI: 0.707–0.969), which demonstrated relatively narrow CIs and good generalizability. Among traditional models, LR (AUC = 0.844, 95% CI: 0.475–0.948) and RR (AUC = 0.833, 95% CI: 0.464–0.948) showed moderate performance, yet their wide CIs suggested

Table 4 Comparison of Predictive Performance Among the Eight Model (N=99)

Algorithm	AUC (95% CI)	Accuracy	Precision	Recall	F1 Score
Random Forest	0.896(0.792–0.971)	0.750	0.733	0.917	0.815
Gradient Boosting	0.875(0.775–0.983)	0.700	0.714	0.833	0.769
eXtreme Gradient Boosting	0.792(0.707–0.969)	0.600	0.625	0.833	0.714
Logistic Regression	0.844(0.475–0.948)	0.707	0.694	0.834	0.732
Ridge Regression	0.833(0.464–0.948)	0.658	0.641	0.825	0.723
Naive Bayes	0.760(0.448–0.929)	0.713	0.696	0.880	0.778
K-Nearest Neighbors	0.745(0.531–0.955)	0.650	0.667	0.833	0.741
Decision Tree	0.625(0.599–0.964)	0.650	0.692	0.750	0.720

Abbreviations: AUC, area under the curve; CI, confidence interval.

sensitivity to data distribution. The probabilistic model NB (AUC = 0.760, 95% CI: 0.448–0.929) and distance-sensitive KNN (AUC = 0.745, 95% CI: 0.531–0.955) performed similarly, with NB displaying greater variability. In contrast, DT yielded the poorest results (AUC = 0.625, 95% CI: 0.599–0.964), reflecting its instability and limited applicability in this classification task (Table 4).

SHAP Analysis

The best-performing RF algorithm was further interpreted using SHAP analysis. Figure 4E illustrates the mean absolute SHAP values for each feature, ranked in descending order of importance: gait speed, SLM, UBE2O, TUG, ACTN3, and MMSE. Figure 4F further reveals the contribution magnitude and direction of each feature to the classification results in the RF model. Gait speed exhibited the strongest predictive influence. SLM and the protein markers UBE2O and ACTN3 showed more complex, non-linear relationships, where both high and low levels could contribute to prefrailty risk. TUG time, while less influential, was positively associated with prefrailty risk, whereas MMSE score only contributed substantially under conditions of pronounced cognitive decline.

Discussion

To our knowledge, this is the first study to integrate high-throughput proteomic data with clinical phenotypes to systematically characterize the proteomic profiles across different stages of frailty, with prefrailty investigated as a distinct clinical status. The results revealed that among 166 significantly differentially expressed proteins detected in older adults across different health statuses, 15 core proteins associated with prefrailty were validated by targeted proteomics, including CP, F8, HPX, VWF, RPLP0, ATP5F1B, LPL, F5, MYH9, HSPG2, SVEP1, BPIFB1, UBE2O, ACTN3, and TREML1. Through machine learning combined with SHAP-based interpretability analysis, UBE2O and ACTN3, together with clinical features including gait speed, SLM, TUG time, and MMSE score, were further identified as the most promising indicators for early frailty detection. Our findings provide preliminary insights into the molecular alterations associated with prefrailty and suggest potential protein candidates that may inform future clinical identification strategies.

Clinical characteristics, including gait speed, SLM, TUG time, and MMSE scores, have long served as key indicators for distinguishing prefrailty in older adults.⁴³ Our findings align with established research on frailty risk prediction, where functional indicators, socio-psychological factors, and physiological parameters remain the most frequently utilized variables.^{43,44} Indeed, a large-scale Chinese cohort study validated that predictive models incorporating cognitive function, lifestyle behaviors, and body composition indices, specifically muscle mass and visceral fat achieved significant efficacy in early screening and risk stratification.⁴⁵ Such clinical indicators are considered indispensable for prefrailty discrimination due to their high operability in diverse settings and well-defined biological mechanisms.

Integrating two independent cohorts, our results revealed widespread proteomic alterations across different frailty statuses. A total of 15 proteins, primarily involved in intercellular signaling, protein homeostasis, immune regulation, and skeletal muscle maintenance, showed significant differences. To further explore the clinical translational potential of

these markers, an interpretable machine learning framework highlighted ACTN3 and UBE2O as key protein molecules for identifying prefrailty.

In this study, the altered expression of ACTN3 in the peripheral circulation of prefrail individuals may reflect early-stage skeletal muscle distress. As a fundamental structural component of type II muscle fibers, ACTN3 is typically sequestered within the myofibrillar apparatus.⁴⁶ Its upregulation in serum likely signifies stress-induced leakage from damaged myocytes, a process increasingly attributed to extracellular vesicle-mediated transport in degenerative muscle pathologies.⁴⁷ While ACTN3 elevation serves as a marker of structural vulnerability, UBE2O, a multifaceted E3 ubiquitin ligase, represents a more direct enzymatic driver of muscle attrition. We hypothesize that these two proteins converge on the AMPK/mTOR signaling axis, which is a master regulator of muscle proteostasis.^{48,49} Specifically, excessive UBE2O accelerates the ubiquitination and degradation of the AMPK α 2 subunit, thereby blunting mTOR-mediated protein synthesis.^{50,51} This pathological synergy, supported by our PPI analysis showing indirect interactions between ACTN3 and UBE2O, suggests that the transition to frailty involves a fundamental shift in muscle protein turnover. The compensatory increase in ACTN3 and the proteolytic activity of UBE2O collectively disrupt the balance between muscle synthesis and breakdown. These proteomic trends indicate that the onset of prefrailty typically emerges gradually under the influence of multi-system dysregulation, while providing measurable biological corroboration for conventional clinical assessments.

Physical functional performance, muscle mass, and cognitive status primarily reflect the macroscopic manifestations of physical frailty. In the context of the growing emphasis on precision medicine, integrating these core clinical indicators with key protein biomarkers may provide a more refined characterization of prefrailty by capturing molecular remodeling associated with proteostasis imbalance. Such a multidimensional approach may help reveal heterogeneity in prefrailty that is not detectable using conventional assessment tools alone, thereby offering additional insights for personalized risk stratification and targeted interventions. Nevertheless, certain non-linear trends in our data warrant further consideration. The relatively lower levels of total protein, IGF-1, and SLM in the prefrail group compared to the frail group may mirror an early catabolic or compensatory dysregulation phase preceding later-stage stabilization.⁵² Concurrently, the universal hypovitaminosis D among participants serves as a systemic background factor, reflecting a widespread vulnerability in protein synthesis and calcium signaling pathways inherent to the aging process.⁵³

This study has several limitations. First, its cross-sectional design prevents observing the dynamic progression of prefrailty and precludes establishing causal relationships. Second, the modest sample size may result in limited statistical power, potentially masking subtle proteomic alterations or associations in secondary analyses. Third, the single-center cohort may not fully capture the heterogeneity of older populations; potential confounding factors such as comorbidities and complex drug histories were not fully controlled, which may influence protein expression. Finally, while external validation is required to confirm the generalizability of our machine-learning models, computational constraints limited the breadth of our algorithmic comparisons. Future multi-center, longitudinal studies with larger cohorts and more rigorous control for clinical confounders are necessary to validate these exploratory findings.

Conclusions

This study revealed distinct proteomic shifts and identified 15 proteins potentially associated with prefrailty, primarily involved in cell signaling, protein metabolism, immune regulation, and skeletal muscle function maintenance. Based on the random forest algorithm, a preliminary model combining clinical indicators and protein markers was established for prefrailty identification. Notably, ACTN3 and UBE2O emerged as promising candidate markers associated with prefrailty. These findings provide exploratory evidence for the potential of integrating proteomic data into multidimensional prefrailty assessment. However, further large-scale and longitudinal validation is essential to confirm the clinical utility and robustness of these candidate markers.

Ethics Approval and Informed Consent

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethics Committee of the Chinese PLA General Hospital (Approval No. S2019-140-03). All participants provided written informed consent for participation and data collection.

Acknowledgments

An unauthorized version of the Chinese MMSE was used by the study team without permission, however this has now been rectified with PAR. The MMSE is a copyrighted instrument and may not be used or reproduced in whole or in part, in any form or language, or by any means without written permission of PAR (www.parinc.com).

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This study was supported by the National Key Research and Development Program of China (Grant No. 2018YFC2002004).

Disclosure

The authors report no conflicts of interest in this work.

References

- Orkaby AR, Schwartz AW, Callahan KE. Frailty. *Ann Intern Med.* 2026;179;(2):ITC17–ITC32. doi:10.7326/ANNALS-25-04412
- Cai S, Li J, Fang Y, et al. Frailty and pre-frailty prevalence in community-dwelling elderly with multimorbidity: a systematic review and meta-analysis. *Arch Gerontol Geriatr.* 2025;132:105782. doi:10.1016/j.archger.2025.105782
- Lapi F, Marconi E, Aprile PL, et al. Frailty, comorbidity, and multimorbidity and their relation with medications adherence in primary care older adults. *Eur Geriatr Med.* 2024;16(2):583–603. doi:10.1007/s41999-024-01098-4
- Sato R, Vatic M, Peixoto da Fonseca GW, Anker SD, von Haehling S. Biological basis and treatment of frailty and sarcopenia. *Cardiovasc Res.* 2024;120(9):982–998. doi:10.1093/cvr/cvae073
- Serra-Prat M, Fortuny Borsot A, Burdoy E, et al. Frailty reversal and its main determinants: a population-based observational and longitudinal study. *Fam Med Comm Health.* 2025;13(2):e003250. doi:10.1136/fmch-2024-003250
- Aliberti SM, Menini A, Sacco AM, et al. A pragmatic two-step screening algorithm for sarcopenia and frailty in community-dwelling older adults: a cross-sectional population-based study. *Life.* 2026;16(1):106. doi:10.3390/life16010106
- Jude M, Rajasingh S, Alway S, Rajasingh J. Molecular basis of muscle wasting and emerging therapeutic strategies targeting degenerative muscle disorders. *Exp Cell Res.* 2026;457(2):114969. doi:10.1016/j.yexcr.2026.114969
- Kemper EJ, Blaak EE, Adriaens ME, Meex RCR. Muscle and liver insulin resistance are associated with distinct plasma protein profiles: data from the DiOGenes trial. *J Clin Endocrinol Metab.* 2026;111(4):1031–1039. doi:10.1210/clinem/dgaf562
- Liu G, Jiang S, Xie W, et al. Biomarkers for sarcopenia, muscle mass, muscle strength, and physical performance: an umbrella review. *J Transl Med.* 2025;23(1):650. doi:10.1186/s12967-025-06575-3
- Sathyan S, Ayers E, Gao T, Milman S, Barzilai N, Verghese J. Plasma proteomic profile of frailty. *Aging Cell.* 2020;19(9):e13193. doi:10.1111/ace1.13193
- Wilkinson DJ, Piasecki M, Atherton PJ. The age-related loss of skeletal muscle mass and function: measurement and physiology of muscle fibre atrophy and muscle fibre loss in humans. *Ageing Res Rev.* 2018;47:123–132. doi:10.1016/j.arr.2018.07.005
- Hodgins ML, Maxwell SP, Howlett SE, Rockwood K. Blood biomarkers of frailty and cognition: a scoping review. *Neurobiol Aging.* 2026;161:14–30. doi:10.1016/j.neurobiolaging.2026.01.003
- Sciancalepore F, Salemm S, Valletta M, et al. Frailty and the brain: a narrative review of functional and pathological correlates. *Ageing Res Rev.* 2025;113:102945. doi:10.1016/j.arr.2025.102945
- Mitchell A, Malmgren L, Bartosch P, McGuigan FE, Akesson KE. Pro-inflammatory proteins associated with frailty and its progression: a longitudinal study in community-dwelling women. *J Bone Miner Res.* 2023;38(8):1076–1091. doi:10.1002/jbmr.4861
- Pedro KM, Alvi MA, Hejrati N, et al. Machine learning-based cluster analysis identifies four unique phenotypes of patients with degenerative cervical myelopathy with distinct clinical profiles and long-term functional and neurological outcomes. *EBioMedicine.* 2024;106:105226. doi:10.1016/j.ebiom.2024.105226
- Niu L, Thiele M, Geyer PE, et al. Noninvasive proteomic biomarkers for alcohol-related liver disease. *Nat Med.* 2022;28(6):1277–1287. doi:10.1038/s41591-022-01850-y
- Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. *J Gerontol a Biol Sci Med Sci.* 2001;56(3):M146–M156. doi:10.1093/gerona/56.3.M146
- Hochrainer K, Yang W. Stroke proteomics: from discovery to diagnostic and therapeutic applications. *Circ Res.* 2022;130(8):1145–1166. doi:10.1161/CIRCRESAHA.122.320110
- Masood A, Benabdelkamel H, Alfadda AA, et al. Proteomic profiling reveals novel molecular insights into dysregulated proteins in established cases of rheumatoid arthritis. *Proteomes.* 2025;13(3):32. doi:10.3390/proteomes13030032
- Lim SM, Ng YL, Majeed ABA, et al. Differential proteomic profiles between cognitive frail and robust older adults from the MELoR cohort. *GeroScience.* 2024;47(3):5069–5088. doi:10.1007/s11357-024-01462-z

21. Yuan Y, Lin S, Huang X, et al. The identification and prediction of frailty based on Bayesian network analysis in a community-dwelling older population. *BMC Geriatr.* **2022**;22(1):847. doi:10.1186/s12877-022-03520-7
22. Katzman R, Zhang MY, Ou-Yang YQ, et al. A Chinese version of the mini-mental state examination: impact of illiteracy in a shanghai dementia survey. *J Clin Epidemiol.* **1988**;41(10):971–978. doi:10.1016/0895-4356(88)90034-0
23. Folstein MF, Folstein SE, McHugh PR. Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res.* **1975**;12(3):189–198. doi:10.1016/0022-3956(75)90026-6
24. Charlson M, Szatrowski TP, Peterson J, Gold J. Validation of a combined comorbidity index. *J Clin Epidemiol.* **1994**;47(11):1245–1251. doi:10.1016/0895-4356(94)90129-5
25. Anbarasan D, Merchant RA. Association between chronic pain severity, falls, frailty and perceived health in older adults at risk of falls. *Eur J Med Res.* **2025**;30(1):1047. doi:10.1186/s40001-025-03304-w
26. Ishiwari M, Kono Y, Togashi Y, et al. Mini nutritional assessment short-form as a risk factor for mortality in patients with respiratory disease undergoing urgent hospitalization. *BMC Geriatr.* **2025**;25(1):113. doi:10.1186/s12877-025-05767-2
27. Park SH, Kwak MJ. Performance of the geriatric depression scale-15 with older adults aged over 65 years: an updated review 2000–2019. *Clin Gerontol.* **2020**;44(2):83–96. doi:10.1080/07317115.2020.1839992
28. Sainsbury A, Seebass G, Bansal A, Young JB. Reliability of the Barthel index when used with older people. *Age Ageing.* **2005**;34(3):228–232. doi:10.1093/ageing/afi063
29. Houston DK, Marsh AP, Neiberg RH, et al. Vitamin D supplementation and muscle power, strength and physical performance in older adults: a randomized controlled trial. *Am J Clin Nutr.* **2023**;117(6):1086–1095. doi:10.1016/j.ajcnut.2023.04.021
30. Li K, Teo GC, Yang KL, Yu F, Nesvizhskii AI. diaTracer enables spectrum-centric analysis of diaPASEF proteomics data. *Nat Commun.* **2025**;16(1):95. doi:10.1038/s41467-024-55448-8
31. Chambers AG, Sweet SMM, Chain D, Kim YJ. Global extraction from parallel reaction monitoring to quantify background peptides for improved normalization and quality control in targeted proteomics. *Anal Chem.* **2021**;93(40):13434–13440. doi:10.1021/acs.analchem.1c01375
32. Deng M, Sun S, Zhao R, et al. The pyroptosis-related gene signature predicts prognosis and indicates immune activity in hepatocellular carcinoma. *Mol Med.* **2022**;28(1):16. doi:10.1186/s10020-022-00445-0
33. Takefuji Y. Reevaluating feature importance in machine learning: concerns regarding SHAP interpretations in the context of the EU artificial intelligence act. *Water Res.* **2025**;280:123514. doi:10.1016/j.watres.2025.123514
34. Owens CD, Pinto CB, Mukli P, et al. Neurovascular coupling, functional connectivity, and cerebrovascular endothelial extracellular vesicles as biomarkers of mild cognitive impairment. *Alzheimers Dement.* **2024**;20(8):5590–5606. doi:10.1002/alz.14072
35. Yi Y, Tae M, Shin S, Choi SI. Predicting acute kidney injury in trauma using an extreme gradient boosting model. *Clin Kidney J.* **2025**;18(4):sfaf002. doi:10.1093/ckj/sfaf002
36. Liang D, Wang L, Zhong P, et al. Perspective: global burden of iodine deficiency: insights and projections to 2050 using XGBoost and SHAP. *Adv Nutr.* **2025**;16(3):100384. doi:10.1016/j.advnut.2025.100384
37. Dey D, Haque MS, Islam MM, et al. The proper application of logistic regression model in complex survey data: a systematic review. *BMC Med Res Methodol.* **2025**;25(1):15. doi:10.1186/s12874-024-02454-5
38. Akhtar N, Alharthi MF. Enhancing accuracy in modelling highly multicollinear data using alternative shrinkage parameters for ridge regression methods. *Sci Rep.* **2025**;15(1):10774. doi:10.1038/s41598-025-94857-7
39. Sugahara S, Ueno M. Exact learning augmented naive Bayes classifier. *Entropy.* **2021**;23(12):1703. doi:10.3390/e23121703
40. Wang J, Geng X. Large margin weighted k-nearest neighbors label distribution learning for classification. *IEEE Trans Neural Netw Learn Syst.* **2024**;35(11):16720–16732. doi:10.1109/TNNLS.2023.3297261
41. Hou R, Wang X, Zhang W, et al. Decision tree extraction for clinical decision support system with if-else pseudocode and PlanSelect strategy. *IEEE J Biomed Health Inform.* **2025**;29(5):3642–3653. doi:10.1109/JBHI.2025.3529682
42. Cao S, Hu Y. Creating machine learning models that interpretably link systemic inflammatory index, sex steroid hormones, and dietary antioxidants to identify gout using the SHAP (SHapley additive exPlanations) method. *Front Immunol.* **2024**;15:1367340. doi:10.3389/fimmu.2024.1367340
43. Perdiguerro GG, Spina JC, Martínez J, et al. Enhancing ACLF prediction by integrating sarcopenia assessment and frailty in liver transplant candidates on the waiting list. *JHEP Rep.* **2024**;6(3):100985. doi:10.1016/j.jhepr.2023.100985
44. Shah SJ, Oreper S, Jeon SY, Boscardin WJ, Fang MC, Covinsky KE. Social frailty index: development and validation of an index of social attributes predictive of mortality in older adults. *Proc Natl Acad Sci U S A.* **2023**;120(7):e2209414120. doi:10.1073/pnas.2209414120
45. Dong Y, Wang Q, Zhang K, et al. Development and validation of a prediction model of frailty risk in community-dwelling older adults: from a national longitudinal survey. *Public Health.* **2025**;240:63–70. doi:10.1016/j.puhe.2024.12.055
46. El Ouali EM, Barthelemy B, Del Coso J, et al. A systematic review and meta-analysis of the association between ACTN3 R577X genotypes and performance in endurance versus power athletes and non-athletes. *Sports Med Open.* **2024**;10(1):37. doi:10.1186/s40798-024-00711-x
47. Giza S, Mojica-Santiago JA, Parafati M, et al. Microphysiological system for studying contractile differences in young, active, and old, sedentary adult derived skeletal muscle cells. *Aging Cell.* **2022**;21(7):e13650. doi:10.1111/ace1.13650
48. Seto JT, Roeszler KN, Meehan LR, et al. ACTN3 genotype influences skeletal muscle mass regulation and response to dexamethasone. *Sci Adv.* **2021**;7(27):eabg0088. doi:10.1126/sciadv.abg0088
49. Huang H, Zhu W, Huang B, et al. Structural insights into the biochemical mechanism of the E2/E3 hybrid enzyme UBE2O. *Structure.* **2025**;33(2):274–288.e4. doi:10.1016/j.str.2024.12.002
50. Vila IK, Park MK, Setijono SR, et al. A muscle-specific UBE2O/AMPKα2 axis promotes insulin resistance and metabolic syndrome in obesity. *JCI Insight.* **2019**;4(13):e128269. doi:10.1172/jci.insight.128269
51. de Carvalho LGA, Komoto TT, Moreno DA, et al. USP15-USP7 axis and UBE2T differential expression may predict pathogenesis and poor prognosis in de novo myelodysplastic neoplasm. *Int J Mol Sci.* **2023**;24(12):10058. doi:10.3390/ijms241210058
52. Jia X, Gao W, Hagelin H, et al. Plasma proteomic signature of frailty in 50,506 adults. *Cell Metab.* **2026**;S1550-4131(26). doi:10.1016/j.cmet.2026.02.013
53. Voulgaridou G, Papadopoulou SD, Spanoudaki M, et al. Increasing muscle mass in elders through diet and exercise: a literature review of recent RCTs. *Foods.* **2023**;12(6):1218. doi:10.3390/foods12061218

Clinical Interventions in Aging**Dovepress**

Taylor & Francis Group

Publish your work in this journal

Clinical Interventions in Aging is an international, peer-reviewed journal focusing on evidence-based reports on the value or lack thereof of treatments intended to prevent or delay the onset of maladaptive correlates of aging in human beings. This journal is indexed on PubMed Central, MedLine, CAS, Scopus and the Elsevier Bibliographic databases. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/clinical-interventions-in-aging-journal>