Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Large-scale cross-species proteomics and machine learning identify biomarkers and peripheral therapeutic targets for delirium

Xinwei Huang

huanggenetics@tongji.edu.cn

Tongji University

Jingxin Yan

South China University of Technology

Yinggang Zheng

Tongji University

Peilin Cong

Tongji University

Qingyuan Miao

Tongji University

Liuqing Yang

Tongji University

Jing Wang

Tongji University

Ban Feng

Tongji University

Hui Zhang

Tongji University

Xinhua Wei

South China University of Technology

Jing Wang

Tongji University

Lize Xiong

Tongji University

Article

Keywords: Delirium, UKB, C2, SPON2, APLP1

Posted Date: November 10th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10672346/v1

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

Additional Declarations: No competing interests reported.

Large-scale cross-species proteomics and machine learning identify biomarkers and peripheral therapeutic targets for delirium

Xinwei Huang[1,†,*] Ph.D, Jingxin Yan[2,†] Ph.D, MD, Yinggang Zheng[1,†] Ph.D, Peilin Cong[1] Ph.D, Qingyuan Miao[1] Ph.D, Liuqing Yang[1] Ph.D, Jing Wang[3] Ph.D, Ban Feng[1] Ph.D, Hui Zhang[1] Ph.D, Xinhua Wei[2,*] MD, Jing Wang[1,*] Ph.D, Lize Xiong[1,*] Ph.D

[1]Shanghai Key Laboratory of Anesthesiology and Brain Functional Modulation, Clinical Research Center for Anesthesiology and Perioperative Medicine, Translational Research Institute of Brain and Brain-Like Intelligence, Department of Anesthesiology and perioperative Medicine, Shanghai Fourth People's Hospital, School of Medicine, Tongji University, Shanghai 200434, China

[2]Department of Radiology, School of Medicine, the Second Affiliated Hospital of South China University of Technology, Guangzhou 510180, China

[3]Department of Radiology, Shanghai Fourth People's Hospital, School of Medicine, Tongji University, Shanghai 200434, China

[†]Xinwei Huang, Jingxin Yan and Yinggang Zheng share co-first authorship.

[*]Corresponding author: Xinwei Huang, huanggenetics@tongji.edu.cn; Xinhua Wei, eyxinhuawei@163.com; Jing Wang, wangjing2021@tongji.edu.cn; Lize Xiong, lizexiong@tongji.edu.cn or mzkxlz@126.com.

Abstract

Delirium is a frequent neurocognitive disorder in critically ill patients. We aimed to identify circulating biomarkers and therapeutic targets for delirium. We conducted a multi-tiered study integrating: (1) UK Biobank plasma proteomics (52,936 participants, 2,911 proteins, 17.5-year follow-up) with machine learning for risk prediction; (2) two-sample Mendelian randomization using protein QTL (N=35,559) and delirium GWAS (4,381 cases vs. 469,981 controls) for causal inference; (3) cross-species proteomic profiling in human and mouse postoperative delirium (POD) cohorts; and (4) functional validation in POD mice via liver-specific knockdown of candidate proteins or systemic supplementation, with cognitive tests, histopathology, and cytokine assays. We identified 1,470 plasma proteins associated with incident delirium. A 21-feature predictive panel (20 proteins + age) achieved consistent performance (AUC 0.77-0.82) for delirium risk identification. Cross-species proteomic profiling converged on three peripheral proteins C2 and SPON2 (risk protein) and APLP1 (protective protein) which were validated across human POD, mouse POD, and UK Biobank cohorts. In POD mice, liver-specific knockdown of C2 or SPON2, as well as systemic supplementation of recombinant human APLP1, significantly ameliorated cognitive deficits, attenuated hippocampal neuroinflammation, reduced glial activation, and preserved synaptic integrity, without overt organ toxicity. Our findings identify a series of peripheral blood biomarkers for delirium risk prediction, and demonstrate that peripheral modulation of circulating C2, SPON2, and APLP1 confers therapeutic benefit. This study highlights the peripheral-to-central axis as a critical node in delirium risk and provides insight into the etiology and therapeutic strategies in delirium patients.

Keywords: Delirium; UKB; C2; SPON2; APLP1

Introduction

Delirium is an acute and frequently reversible neurocognitive disorder marked by inattention, reduced consciousness, and disorganized thought[1]. It is a common clinical condition, affecting over 20% of hospitalized patients and up to 74% of individuals in intensive care units (ICU)[2,3]. Delirium is associated with a range of adverse clinical outcomes, containing prolonged mechanical ventilation, extended stays in both ICU and hospital, increased reliance on benzodiazepines, long-term cognitive decline, diminished daily activity, and elevated mortality rates[4]. Key risk factors for delirium include both baseline patient characteristics, such as advanced age, pre-existing dementia, frailty, and multimorbidity, and acute health events such as surgical trauma[5]. Surgical intervention is a major trigger, making postoperative delirium (POD) a significant clinical issue. Consequently, animal models of POD serve as classic paradigms in preclinical research for investigating this condition[2]. Despite its high prevalence and clinical significance, delirium pathophysiology remains poorly understood, owing to its fluctuating course and multifactorial etiology. With the ongoing global demographic shift toward an aging population, delirium has emerged as a significant and growing public health concern, particularly among older adults[6].

To date, most preclinical investigations, including those from our own team, have concentrated on the central nervous system (CNS) pathology of delirium models. Although those work has identified multiple potential CNS-based targets for intervention, subsequent clinical trials of pharmacological agents directed at these targets have largely failed to demonstrate efficacy in preventing delirium[2,5,7,8]. A key factor underlying this translational gap may lie in the predominant reliance on rodent models for target discovery[4,9]. Given the intrinsic differences in the neurobiological mechanisms underlying delirium between animal and human, therapeutic targets derived solely from rodent studies may have limited clinical applicability. Therefore, a more promising strategy may involve first identifying key pathological targets directly from human associated with delirium, followed by validation in appropriate animal

models. This approach could yield more translatable targets and inform more effective clinical strategies for delirium prevention and treatment.

Our prior work has demonstrated that perioperative alterations of proteins (eg. complement component C3) in the blood or liver are closely related to cognitive impairment in surgical patients and POD mice, while maintaining peripheral protein homeostasis can effectively prevent cognitive impairment in POD mice, indicating their potential for predicting and preventing postoperative cognitive impairment and POD[10]. However, the peripheral proteome comprises thousands of proteins with complex inter-regulatory networks, and the candidate-driven approach testing one or a few preselected proteins is inherently limited in its ability to uncover novel mediators. Traditional small-scale case-control studies are also susceptible to false discoveries due to limited statistical power and lack of replication. Therefore, a systematic, unbiased, and adequately powered screening strategy in large-scale human populations is critically needed to comprehensively map peripheral protein-delirium associations and to prioritize the most promising candidates for subsequent causal validation.

In the present study, we used an integrated, cross-species, multi-omics approach to systematically identify peripheral protein drivers of delirium and evaluate their therapeutic potential. Through large-scale proteomic analyses in the UK Biobank cohort, we identified plasma proteins associated with incident delirium and developed machine learning-based prediction models. By integrating Mendelian randomization for causal inference and cross-species proteomic profiling in an independent POD patient cohort and mouse models, we pinpointed three circulating proteins including C2, SPON2, and APLP1 as key mediators of delirium risk. Functional validation in POD mice demonstrated that liver-specific knockdown of C2 or SPON2, as well as systemic supplementation of human APLP1, significantly ameliorated cognitive deficits, attenuated hippocampal neuroinflammation, and preserved synaptic integrity, without overt toxicity. Collectively, our findings demonstrate peripheral-to-central axes

as critical nodes in delirium pathogenesis and circulating C2, SPON2, and APLP1 as promising biomarkers and therapeutic targets for delirium.

Results

The overall study design is summarized in **Figure 1**. We followed a six-stage strategy that progressed from population-scale plasma proteomic discovery and risk prediction to genetic causal prioritization, cross-species validation, tissue-source inference, and functional testing in POD mice.

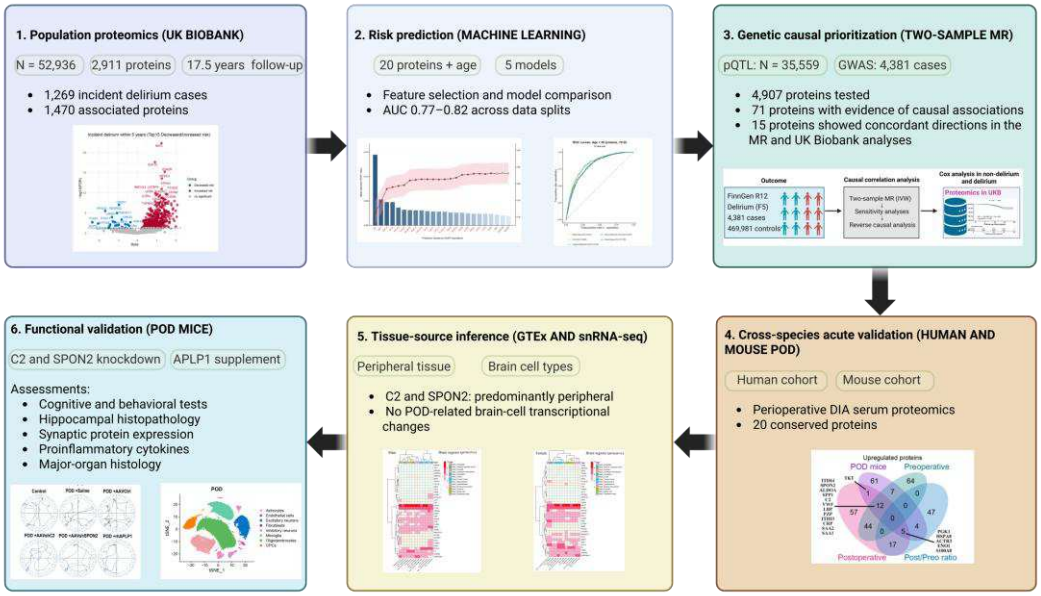


Figure 1. Overview of the study design and analytical workflow. (Created in BioRender. https://BioRender.com/dtx4shq)

The study followed a six-stage strategy integrating population-scale proteomics, risk prediction, genetic causal inference, cross-species validation, tissue-source inference, and experimental validation. First, 2,911 plasma proteins were examined in 52,936 UK Biobank participants over a median follow-up of 17.5 years, during which 1,269 incident delirium cases were identified. A total of 1,470 proteins were associated with incident delirium. Second, feature selection and comparison of five machine-learning models produced a prediction panel comprising 20 proteins and age, achieving AUCs of 0.77 – 0.82 across data splits. Third, two-sample MR integrating plasma pQTL data from 35,559 individuals with delirium GWAS data identified 71 proteins with evidence

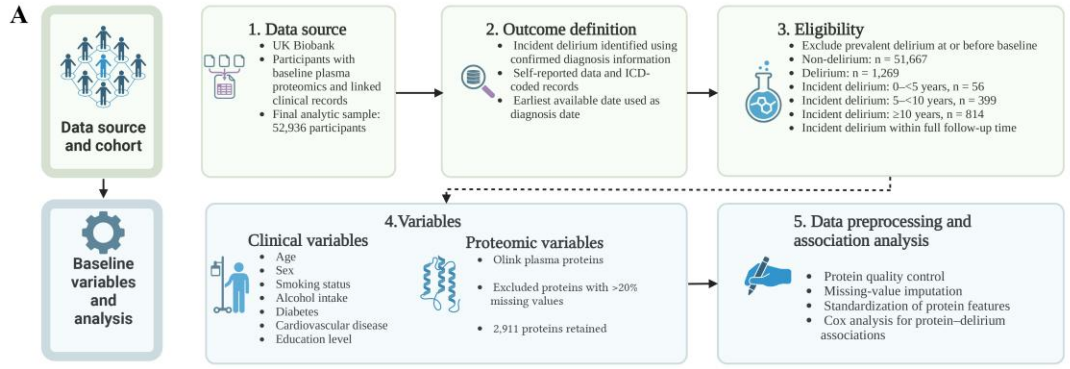
supporting causal associations; 15 showed concordant directions in the MR and UK Biobank analyses. Fourth, perioperative DIA serum proteomics in a human POD cohort and parallel profiling in a mouse POD model identified 20 conserved proteins across species. Fifth, GTEx tissue-expression data and brain snRNA-seq data were used to infer the potential tissue origins of the prioritized proteins, supporting predominantly peripheral origins for C2 and SPON2. Finally, the functional effects of peripheral protein modulation were evaluated in POD mice using liver-specific knockdown of C2 or SPON2 and systemic supplementation with recombinant human APLP1, followed by behavioral, hippocampal histopathological, synaptic, inflammatory, and major-organ safety assessments.

Protein risk factors for delirium

As shown in **Figure 2A**, this study contained 52,936 European participants from the UK Biobank with baseline plasma proteomics data and clinical records[11]. The median follow-up duration was 17.51 years (IQR, 16.68-18.3). During follow-up, 1,269 incident delirium cases were identified (incidence rate 2.4%), of which 56 occurred within 5 years, 399 between 5 and 10 years, and 814 after 10 years or more. A total of 2,911 proteins were finally included in this study after quality control. Baseline characteristics are summarized in **Figure 2B**. Compared with the non-delirium group, participants who developed delirium were significantly older (63.91±5.22 vs. 56.63±8.19 years, $P < 2\times10^{-16}$), more likely to be male (54% vs. 45.9%, $P = 1.16\times10^{-8}$), and had a lower educational level (low education: 76% vs. 66.1%, $P = 2.91\times10^{-16}$). In addition, compared with the non-delirium group, the delirium group had a slightly lower proportion of alcohol intake, but significantly higher proportions of smoking, diabetes, and cardiovascular disease (all $P < 0.001$).

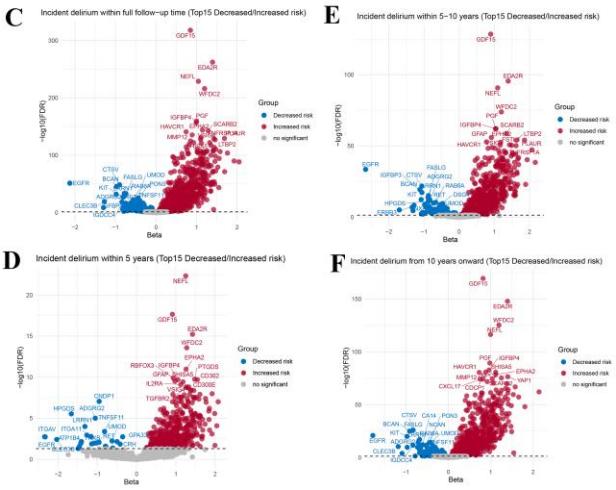
To investigate the prospective associations between circulating plasma proteins and the risk of incident delirium, we performed Cox proportional-hazards models adjusted for age, sex, education level, smoking status, alcohol intake, diabetes, and cardiovascular disease. Over the entire follow-up period, a total of 1,470 proteins were significantly associated with incident delirium. Among these, 306 proteins were associated with a

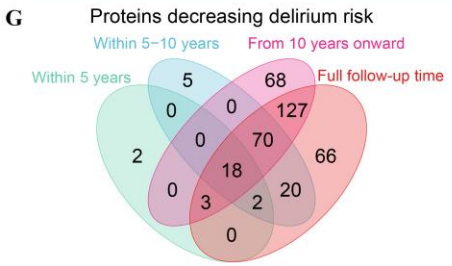
reduced future risk of delirium, while 1,164 proteins were associated with an increased future risk (**Figure 2C** and **Table S1**). To explore the underlying biological pathways, we conducted Gene Set Enrichment Analysis (GSEA) on the protein association results. Those results revealed significant enrichment of pathways related to immune response (adaptive and humoral), inflammation, chemotaxis (cell, leukocyte, neutrophil), and T cell-mediated processes (**Figure S1**). We further stratified our analysis by time to diagnosis to examine the temporal dynamics of these associations (**Figure 2D-F**). For delirium occurring within 5 years of protein measurement (short-term risk), 454 proteins were significant, of which 28 were unique to this time window (26 associated with increased risk and 2 with decreased risk). For delirium diagnosed between 5 and 10 years (medium-term risk), 1,046 proteins were significant, with 15 unique to this period (10 associated with increased risk and 5 with decreased risk). For delirium occurring beyond 10 years (long-term risk), 1,308 proteins were significant, with 73 unique to this period (5 associated with increased risk and 68 with decreased risk) and (**Table S2-4**). Finally, to identify proteins consistently associated with delirium risk across all timeframes, we examined the intersection of significant proteins from the full follow-up period and each of the three time-stratified analyses. This intersection yielded 401 proteins that were significantly associated with delirium across all time points. Of these, 18 proteins exhibited a consistent negative effect on incident delirium risk, indicating that higher plasma levels of these proteins were associated with a reduced future risk. Conversely, 383 proteins showed a consistent positive effect, meaning that higher plasma levels were associated with an increased future risk of delirium (**Figure 2G, H**).



B

Characteristics	Non-delirium (n=5,1667)	Delirium (n=1,269)	*P* value
Age, years	56.63 ± 8.19	63.91 ± 5.22	< 2E-16
Sex			1.16E-08
Female	27953 (54.1%)	587 (46%)	
Male	23714 (45.9%)	689 (54%)	
Education level			2.91E-16
High (College degree)	16637 (32.2%)	272 (21.3%)	
Low	34158 (66.1%)	970 (76%)	
Unknown	872 (1.7%)	34 (2.7%)	
Smoking			9.58E-04
No	46000 (89%)	1094 (85.7%)	
Yes	5420 (10.5%)	175 (13.7%)	
Unknown	247 (0.5%)	7 (0.5%)	
Alcohol intake			3.96E-13
No	4387 (8.5%)	184 (14.4%)	
Yes	47160 (91.3%)	1087 (85.2%)	
Unknown	223 (0.4%)	6 (0.5%)	
Diabetes			< 2E-16
No	48631 (94.1%)	1073 (84.1%)	
Yes	2813 (5.4%)	197 (15.4%)	
Unknown	223 (0.4%)	6 (0.5%)	
Cardiovascular disease			< 2E-16
No	35747 (69.2%)	595 (46.6%)	
Yes	15725 (30.4%)	672 (52.7%)	
Unknown	195 (0.4%)	9 (0.7%)	





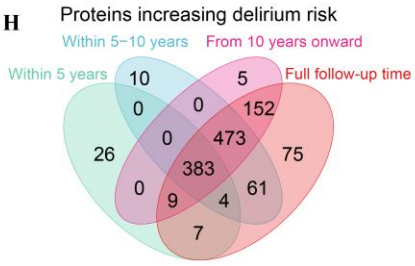


Figure 2 Plasma proteomic profiling reveals time-dependent protein associations with incident delirium in UKB cohort. **A** Schematic overview of the study design. A total of 52,936 European participants from the UK Biobank with baseline plasma proteomics data and clinical records were included. After quality control, 2,911 proteins were retained for Cox regression analyses. **B** Baseline characteristics of participants stratified by delirium status (non-delirium: n = 51,667; delirium: n = 1,269). **C** Number of proteins significantly associated with incident delirium over the entire follow-up period. **D-F** Temporal dynamics of protein-delirium associations stratified by time to diagnosis: within 5 years (D), between 5 and 10 years (E), and beyond 10 years (F). **G, H** Intersection analysis

of significant proteins across all follow-up time. HR hazard ratio, ICD International Classification of Diseases, IQR interquartile range.

Predictive model development and evaluation for future delirium risk

To construct a predictive model for incident delirium, we integrated the 401 delirium-associated proteins with all clinical covariates using LASSO regression, random forest, and SHAP analyses for feature selection (**Figure 3A**). A total of 21 features, including 20 proteins and age, were identified as the optimal feature set, as the cumulative AUC plateaued at approximately 0.81 beyond this threshold (**Figure 3B** and **Table S5**). SHAP analysis revealed that age exerted the greatest influence on delirium risk, with all 20 proteins emerging as risk factors (**Figure 3C**). Across three random training-testing splits (80:20, 70:30, and 60:40), five machine learning models achieved consistent predictive performance, with AUCs ranging from 0.77 to 0.82 (**Figure 3D-F** and **Table S6**). These findings suggest that the set of plasma proteins, combined with age, can effectively stratify individuals at risk for incident delirium, offering a potential tool for early risk assessment for this disorder.

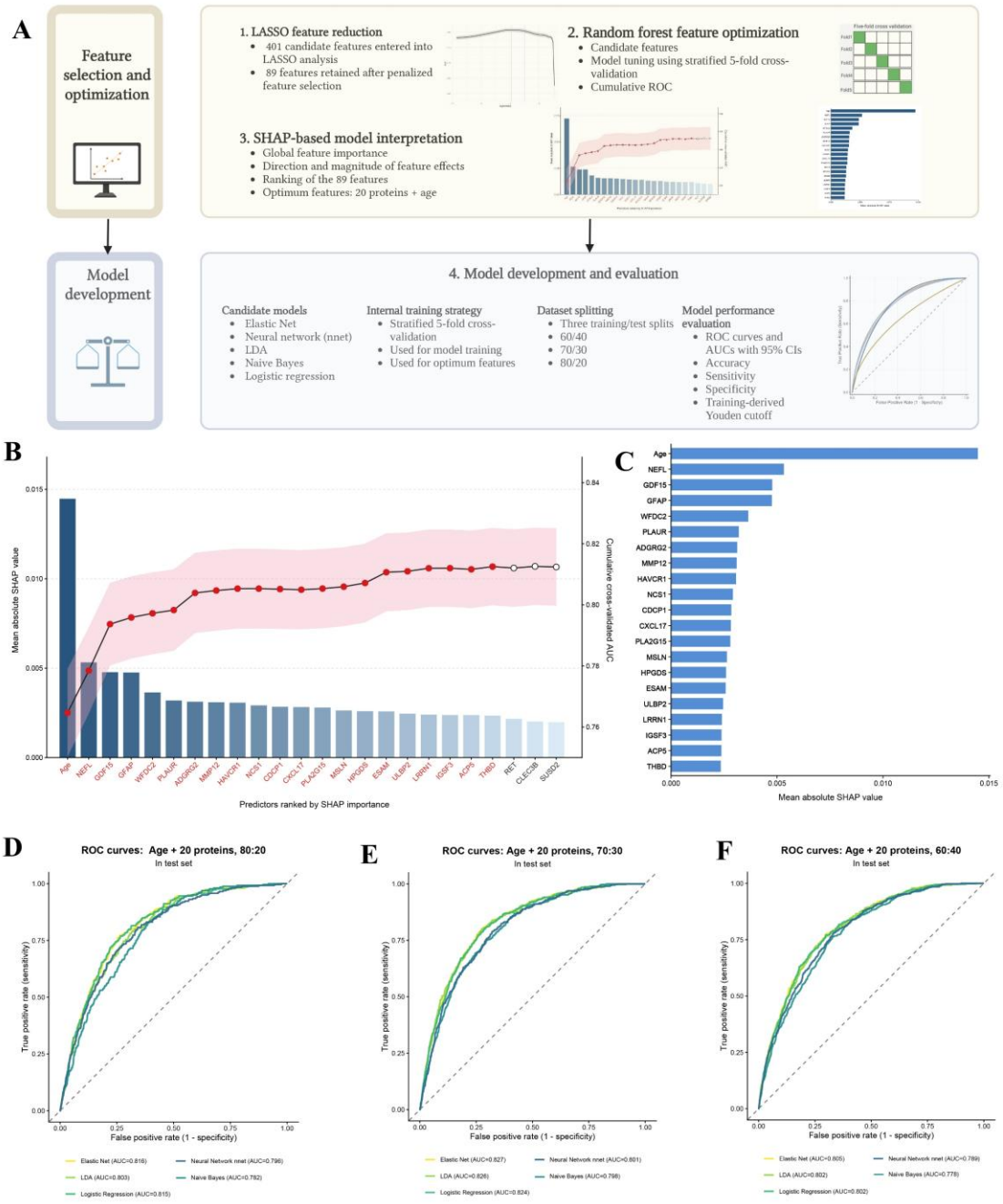


Figure 3 Machine learning-based feature selection and predictive performance of models for incident delirium. **A** Schematic workflow of feature selection and model optimization, including LASSO regression for initial feature reduction, random forest for further feature optimization with stratified 5-fold cross-validation, and SHAP analysis for model interpretation. **B** Cumulative AUC curve showing that the optimal feature set (20 proteins and age) achieved a plateaued cumulative AUC of approximately 0.81, with no meaningful improvement upon addition of more features. **C** SHAP summary plot illustrating global feature importance, with age as the most influential predictor

and all 20 proteins identified as risk factors for delirium. **D-F** Receiver operating characteristic (ROC) curves for five machine learning models evaluated under three random training-testing splits: 80:20 (d), 70:30 (e), and 60:40 (f), demonstrating consistent AUCs ranging from 0.77 to 0.82 across all splits. AUC area under the curve, LASSO least absolute shrinkage and selection operator, ROC receiver operating characteristic, SHAP SHapley Additive exPlanations.

Causal protein-delirium associations identified by Mendelian randomization

To further determine whether the observed associations between plasma proteins and delirium risk reflect causal relationships, we performed a two-sample Mendelian randomization (MR) analysis integrating GWAS summary data for 4,907 plasma proteins (35,559 individuals) with delirium GWAS data (4,381 cases vs. 469,981 controls) (**Figure 4A** and **Table S7-12**). MR analysis identified 71 circulating proteins causally associated with delirium risk. Among these, increased concentrations of 36 proteins were associated with reduced delirium risk (**Figure 4B**), whereas elevated levels of the remaining 35 proteins correlated with increased susceptibility to delirium (**Figure 4C**). Notably, the MR results confirmed that higher plasma levels of APLP1, VAV3, and CDON were causally associated with reduced delirium risk, consistent with the negative associations observed in our UK Biobank Cox regression analyses (**Figure 4D**). Conversely, MR confirmed that elevated levels of C2, SPON2, APCS, CA1, SNAP25, TPP1, CHGB, KIR2DL3, LAG3, GFRA2, HAVCR1, and CST3 were causally associated with increased delirium risk, corroborating the positive associations from the UK Biobank Cox analyses (**Figure 4E**). Functional enrichment analysis of above 15 proteins were functionally enriched in immune-related pathways (C2, SPON2, and APCS), phagocytosis (C2 and SPON2), and complement activation (C2 and APCS) (**Figure 4F**).

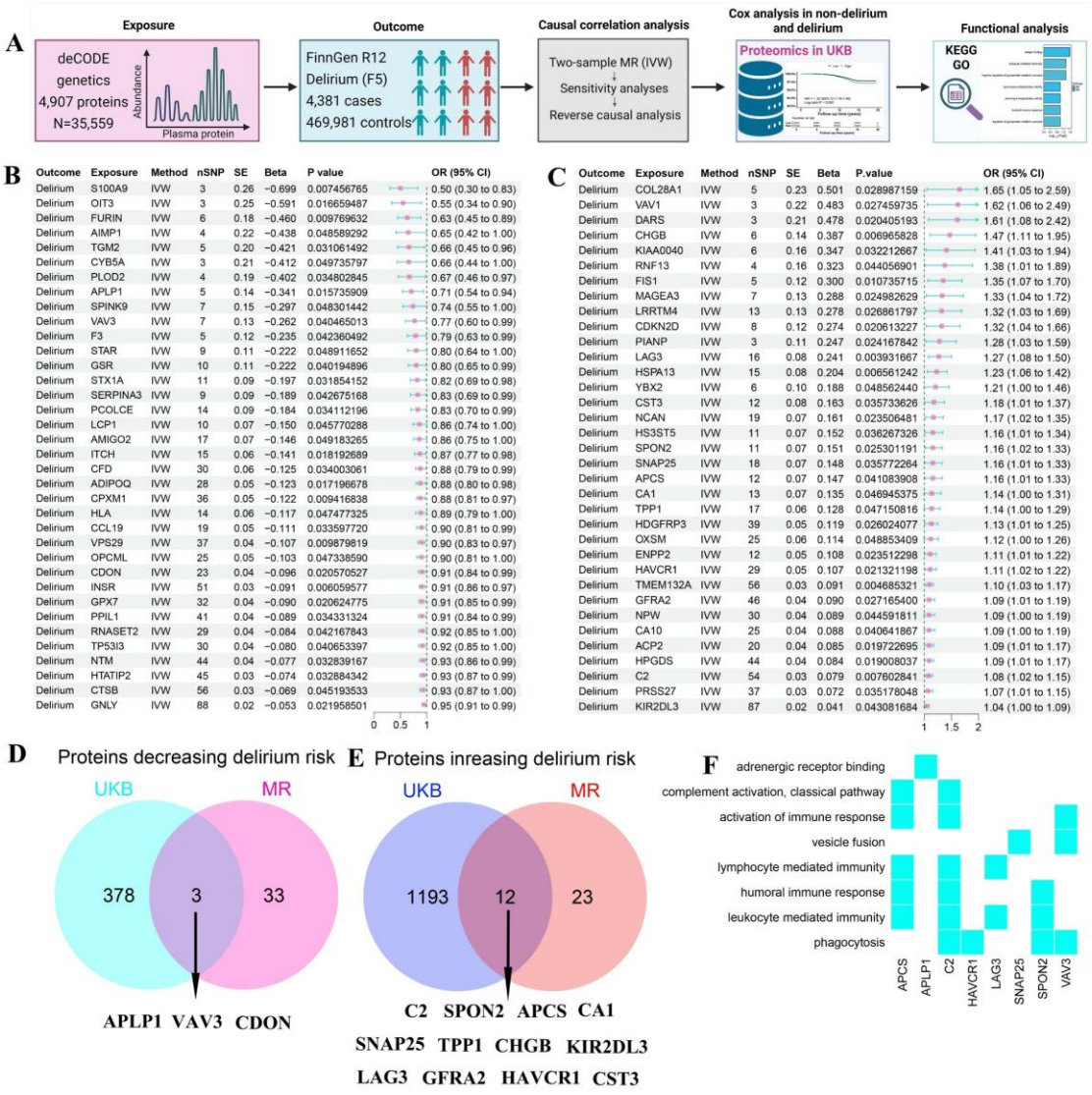


Figure 4 Mendelian randomization analysis of genetically predicted plasma proteins and incident delirium risk. **A** Schematic overview of the two-sample MR design, integrating GWAS summary data for 4,907 plasma proteins (N = 35,559) with delirium GWAS data (4,381 cases vs. 469,981 controls). **B** MR estimates for 36 proteins whose genetically predicted higher concentrations were associated with reduced delirium risk. **C** MR estimates for 35 proteins whose genetically predicted higher concentrations were associated with increased delirium risk. **D** Comparison of MR and UK Biobank Cox regression results for three protective proteins (APLP1, VAV3, CDON), showing consistent negative associations with delirium risk. **E** Comparison of MR and UK Biobank Cox regression results for 12 risk proteins (C2, SPON2, APCS, CA1, SNAP25, TPP1, CHGB, KIR2DL3, LAG3, GFRA2, HAVCR1, CST3), showing consistent positive associations with delirium risk. **F**

Functional enrichment of 15 proteins. MR Mendelian randomization, GWAS genome-wide association study, OR odds ratio, CI confidence interval.

Cross-species proteomic profiling identifies conserved serum protein signatures for postoperative delirium

Given that postoperative delirium (POD) represents the most common subtype of delirium, we conducted perioperative serum proteomic profiling in a Chinese surgical cohort (18 non-POD vs. 18 POD) and parallel proteomic analyses in aged-matched control mice and POD model mice, aiming to identify core conserved protein drivers of delirium and POD across species (**Figure 5A**). Baseline and perioperative clinical characteristics of the human discovery cohort are summarized in **Figure 5B**. No significant differences were observed between the two groups in terms of sex, age, preoperative MMSE, ASA classification, or surgery type (all $P > 0.05$). However, POD patients had significantly lower education levels ($P = 0.048$), lower BMI ($P = 0.040$), longer duration of surgery ($P = 0.038$), and substantially longer hospitalization days ($P < 0.001$).

Compared to non-POD patients, POD patients exhibited 136 upregulated and 122 downregulated proteins in serum collected 1 h after surgery (**Table S13**). Similarly, POD mice showed 263 differentially expressed proteins relative to controls (90 upregulated, 173 downregulated; **Table S14**). Cross-species comparison identified a conserved signature of 18 upregulated proteins, namely C2, SPON2, ITIH4, CRP, SAA2, SAA1, ALDOA, TKT, PGK1, HSPA8, ACTR3, SPP1, VWF, ENO1, LBP, PZP, ITIH3, and S100A8, and 2 downregulated proteins (APLP1 and CFP) (**Figure 5C-E**). Notably, 12 of these upregulated POD-associated proteins were already significantly elevated in patients 1 h before surgery (**Figure 5F** and **Table S15**). Furthermore, the postoperative-to-preoperative ratios of six proteins (CFP, ACTR3, HSPA8, ENO1, PGK1, and S100A8) showed significant differences between non-POD and POD groups (**Figure 5G** and **Table S16**). A series of models based on the 20 postoperative

proteins yielded excellent predictive performance for POD, with AUCs ranging from 0.88 to 0.90 (**Figure S2A-C** and **Table S17**). Similarly, models containing the six proteins with significant postoperative-to-preoperative ratios also showed good discriminative ability, achieving AUCs of 0.75-0.81 (**Figure S2D-F** and **Table S18**).

Importantly, three plasma proteins, APLP1, C2, and SPON2, were consistently validated across multiple independent datasets as key drivers of delirium risk (**Figure 5H**). Specifically, APLP1 levels were significantly decreased in postoperative serum from POD patients compared to non-POD patients, a finding that was recapitulated in POD model mice and further corroborated in the UK Biobank delirium cohort, where higher APLP1 concentrations were associated with significantly reduced delirium risk (**Figure 5I-K**). In contrast, serum C2 and SPON2 levels were markedly elevated in POD patients both preoperatively and postoperatively, consistently increased in POD mice, and similarly upregulated in the UK Biobank delirium cohort. Importantly, higher concentrations of C2 or SPON2 were associated with significantly increased delirium risk (**Figure 5L-N**).

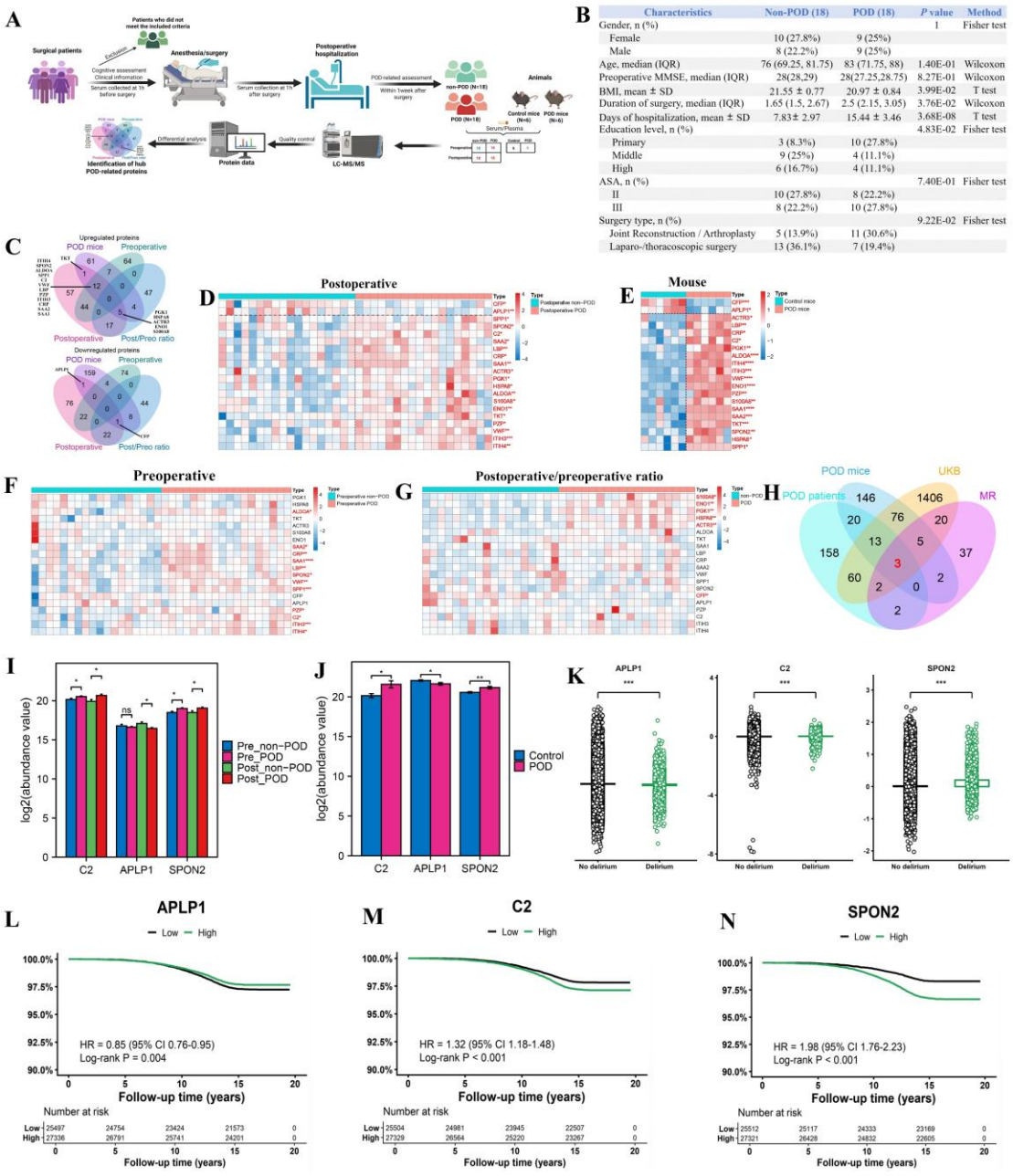


Figure 5 Cross-species proteomic profiling identifies conserved serum protein signatures for postoperative delirium (POD). **A** Schematic overview of the study design, including human perioperative serum proteomics (18 non-POD vs. 18 POD) and parallel mouse models of POD. **B** Baseline and perioperative clinical characteristics of the human discovery cohort stratified by POD status. **C** Venn diagram showing the overlap of differentially expressed proteins between human (POD vs. non-POD) and mouse (POD vs. control) blood samples. **D, E** Heatmap of the 20 conserved proteins (18 upregulated, 2 downregulated) across human (D) and mouse (E). **F** Preoperative serum levels of 12 upregulated POD-associated proteins that were already significantly elevated in POD

patients 1 h before surgery. **G** Postoperative-to-preoperative ratios of six proteins showing significant differences between non-POD and POD groups. **H** Summary of the three core proteins (APLP1, C2, and SPON2) consistently validated across human POD, mouse POD, MR, and UK Biobank cohorts. **I-K** Concentration changes of blood APLP1, C2 and SPON2 in POD patients (I), POD mice (J), and UK Biobank participants with delirium relative to the corresponding control samples (K). **L-N** Associations of APLP1, C2 and SPON2 changes with delirium risk in the UK Biobank. AUC, area under the curve, POD, postoperative delirium.

To trace the potential tissue origins of C2, SPON2, and APLP1, we mapped the expression of their encoding genes across human organs using GTEx transcriptomic data. We found that APLP1 was widely expressed in both male and female brains, and its peripheral expression confined predominantly to the pituitary (**Figure S3**). In contrast, C2 showed negligible expression across all brain regions in both sexes, with predominant expression in the liver and modest expression in other peripheral tissues including adipose tissue, colon, lung, and spleen. SPON2 exhibited low-level expression in the cerebellum and cerebellar hemisphere, and broad expression in peripheral organs such as liver (**Figure S4**).

To further investigate whether transcriptional changes of these genes in brain cells contribute to POD pathogenesis, we examined their expression status at the single-cell level using snRNA-seq data of the brain from non-POD and POD patients (**Figure S5A**). APLP1 was mainly expressed in oligodendrocytes, excitatory neurons, inhibitory neurons. SPON2 showed a sparse expression in excitatory neurons. The expression level of C2 was almost negligible across all cell types (**Figure S5B**). Notably, no significant differences in transcriptional levels of C2, SPON2, or APLP1 were observed across any brain cell type between non-POD and POD groups (**Figure S5C**), suggesting that cerebral expression of these genes is unlikely to play a major role in POD development. Collectively, these findings indicate that peripherally derived C2, SPON2,

and APLP1, rather than their brain-originating transcripts, are likely the key mediators modulating delirium risk.

Targeting peripheral C2, SPON2, and APLP1 confers therapeutic benefit against postoperative delirium in mice

To determine whether systemic changes in C2, SPON2, and APLP1 causally influence susceptibility to POD, we performed cognitive function assessments and histopathological examinations in POD mice following recombinant human APLP1 protein supplementation or liver-specific knockdown of C2 or SPON2 (**Figure 6A**). Immunofluorescence analysis confirmed selective hepatic transduction of rAAV2/8-mCherry-shRNA vectors targeting C2 or SPON2, with positive signals detected exclusively in the liver and absent in other examined organs at 3 weeks post-injection (**Figure S6A**). ELISA further verified successful manipulation of systemic protein levels: APLP1 supplementation significantly elevated its circulating concentrations, whereas liver-specific knockdown of C2 or SPON2 led to marked reductions in their respective serum levels (**Figure S6B**). These findings collectively confirm effective modulation of peripheral APLP1, C2, and SPON2 levels in the POD model.

The open field test revealed that altering circulating levels of these proteins did not affect general locomotor activity or anxiety-like behavior (**Figure S6C-F**). However, the Barnes maze and fear conditioning tests demonstrated that both APLP1 supplementation and liver-specific knockdown of C2 or SPON2 significantly rescued the impaired learning and memory performance in POD mice (**Figure 6B-G**). At the histopathological level, these interventions exerted protective effects within the hippocampus. Specifically, APLP1 supplementation and knockdown of C2 or SPON2 markedly reduced the number of activated glial cells and concurrently decreased the expression of GFAP and IBA1 (**Figure 6H**), and increased the expression of synaptic proteins PSD-95 and synaptophysin (SYN) in the hippocampus (**Figure 6I**). Moreover, these treatments significantly decreased hippocampal levels of pro-inflammatory

cytokines, including TNF-α, IL-6, and IL-1β (**Figure 6J-L**), indicating that targeting these proteins attenuates glial activation, synaptic damage, and neuroinflammation.

Notably, H&E staining of heart, liver, spleen, lung, and kidney revealed no overt histopathological abnormalities (e.g., structural disorganization, inflammatory infiltration, edema, or necrosis) following either APLP1 supplementation or liver-specific knockdown of C2 or SPON2 (**Figure S7**), suggesting favorable safety profiles for these interventions.

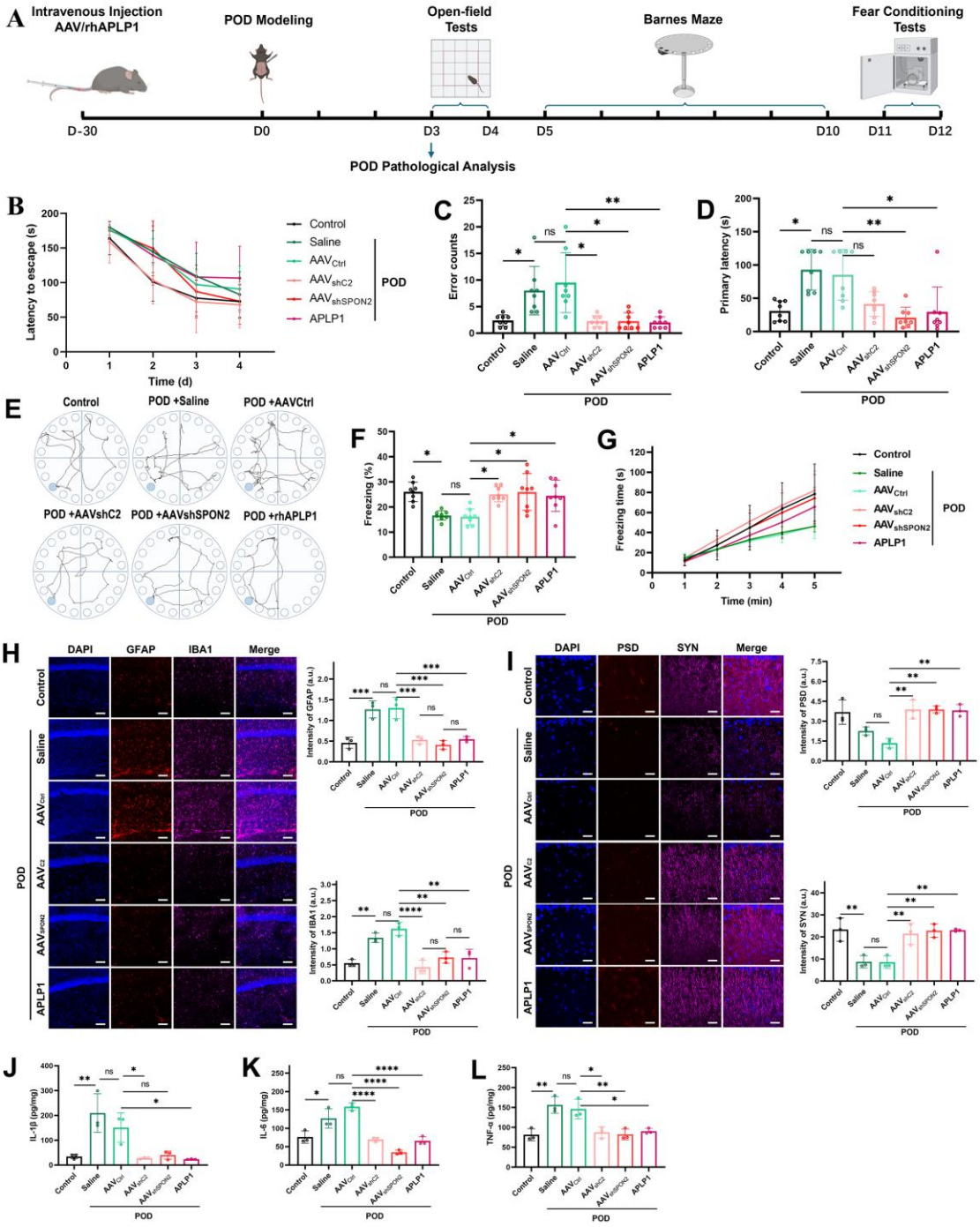


Figure 6 In vivo validation of causal roles for APLP1, C2, and SPON2 in postoperative delirium (POD) pathogenesis. **A** Schematic timeline of the experimental design. POD mice received recombinant human APLP1 supplementation or liver-specific knockdown of C2 or SPON2 via rAAV2/8-shRNA, followed by cognitive behavioral tests and histopathological analyses. **B-G** Barnes maze (B-E) and fear conditioning tests (F, G) showing that APLP1 supplementation and knockdown of C2 or SPON2 significantly rescued learning and memory deficits in POD mice. n =

8 mice per group. **H, I** Representative images and quantification of hippocampal glial activation (IBA-1 and GFAP) and synaptic protein expression (PSD-95 and synaptophysin), demonstrating reduced gliosis and preserved synaptic integrity following interventions. n = 3 mice per group. Scale bar = 100 μm. **J-L** Hippocampal levels of pro-inflammatory cytokines IL-1β, IL-6, and TNF-α showing significant reductions after APLP1 supplementation or C2/SPON2 knockdown. n = 3 mice per group. Data are presented as mean ± SEM. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. OFT open field test, FC fear conditioning; IBA-1, ionized calcium-binding adapter molecule 1; GFAP, glial fibrillary acidic protein; PSD-95 postsynaptic density protein 95; SYN, synaptophysin; TNF-α, tumor necrosis factor-alpha; IL-6, interleukin-6; IL-1β: interleukin-1 beta.

Discussion

Delirium is an acute disturbance of attention, awareness, and cognition that is typically triggered by illness, surgery, medication, or other physiological stress in a vulnerable individual.[12-14]. POD is more than temporary confusion; it is associated with longer hospital stays, functional decline, and persistent cognitive deterioration after surgery.[13-15]. Yet its biological heterogeneity has impeded biomarker and therapeutic development. As a results, we followed the same sequential strategy used in the Results, moving from prospective population discovery to prediction, genetic causal prioritization, cross-species validation, tissue-source inference and experimental intervention. This approach identified a broad proteomic outcomes of delirium and ultimately converged on C2 and SPON2 as risk-associated candidates and APLP1 as a potentially protective candidate. The central contribution is therefore not simply the number of associated proteins, but the triangulation of complementary evidence types that address different questions: association defines the risk associations, prediction tests prognostic information, Mendelian randomization (MR) assesses causal compatibility, perioperative and cross-species profiling establishes acute relevance, and perturbation tests biological modifiability.

The first stage revealed some measured proteins associated with incident delirium among 52,936 UK Biobank participants, including 401 proteins that remained

significant across all temporal strata. UK Biobank enables linkage of deeply phenotyped participants to long-term health outcomes[16], while population-scale plasma proteomics has shown that thousands of circulating proteins capture genetic influences, intermediate traits and future incident delirium risk[17]. The enrichment of immune, inflammatory and chemotactic pathways in our analysis is consistent with neuroinflammatory and immunovascular models in which systemic inflammation perturbs endothelial and blood–brain barrier function, activates glia and disrupts synaptic homeostasis[18-20]. Smaller perioperative studies provided convergent clinical context: C-reactive protein has been associated with POD incidence, duration and severity[21], proteome-wide profiling identified chitinase-3-like protein 1 as both a preoperative risk and postoperative delirium marker[22], and serum proteomics after esophagectomy implicated inflammation, coagulation and C2[23].

The prediction analysis, addressed a distinct translational question. The age-plus-20-protein model achieved areas under the receiver-operating-characteristic curve of 0.77–0.82 across three internal splits, suggesting reproducible discrimination within UK Biobank. However, age was the dominant predictor, only 56 delirium events occurred within five years of baseline sampling, and most events were recorded more than a decade later. The model may thus identify long-term susceptibility to hospitalization, frailty or neurodegeneration rather than imminent POD. It should not yet be regarded as a clinical test. TRIPOD+AI statement emphasizes transparent reporting of participant selection, missing data, model specification and validation[24]. Discrimination must also be accompanied by calibration[25], and clinical value should be evaluated through decision-curve analysis at meaningful thresholds rather than inferred from the AUC alone[26]. Small event numbers, data-driven feature selection and solely internal validation are recognized sources of prediction-model bias in PROBAST[27]. External validation should therefore use cohorts with repeated bedside delirium assessments and compare the protein panel against a strong clinical model incorporating baseline cognition, frailty, comorbidity, medication exposure and surgical characteristics.

MR then narrowed the observational signals by testing whether genetically proxied protein concentrations were compatible with causal effects. Large pQTL resources demonstrate the feasibility of connecting genetically regulated proteins to disease across the phenome[28,29], and proteome-wide MR has prioritized potential therapeutic targets for complex disorders[30,31]. In our analysis, 71 proteins showed genetic evidence of an effect and 15 were directionally concordant with the prospective associations, materially strengthening their prioritization. This concordance is more compelling than either analysis alone because conventional epidemiology and germline genetic inference have different biases. It does not, however, prove that short-term perioperative modulation will reproduce a lifelong genetically mediated difference. Horizontal pleiotropy, linkage disequilibrium, weak instruments and variants that alter assay epitope binding remain possible explanations; even cis-focused drug-target MR reduces rather than eliminates these concerns[32]. Colocalization, independent cis-pQTL replication and explicit adherence to STROBE-MR reporting standards would further test whether the protein and delirium signals share a causal variant[33]. MR should consequently be viewed here as a prioritization layer, with experimental perturbation providing the next, mechanistically different test.

The paired human and mouse perioperative proteomic analyses next asked whether population-prioritized signals were engaged during an acute surgical insult. The convergence of the human POD comparison, aged-mouse surgery model and upstream discovery analyses on C2, SPON2 and APLP1 markedly reduced the candidate space. This cross-species step is important because remote baseline associations may reflect vulnerability, whereas perioperative changes capture the biological response around delirium onset. It also extends earlier human proteomic observations of inflammatory and complement-related changes [22, 23]. Nevertheless, our human analysis included only 18 POD cases and 18 controls, and serum collected one hour after surgery preceded the three-day period during which delirium was assessed. These proteins may therefore represent an early response, a preclinical mediator or a correlate of surgical stress.

Tissue-source analysis provided the next link in this sequence. GTEx offers a systematic atlas of tissue gene regulation[34], and annotation of the human secretome helps distinguish proteins plausibly released into the circulation[35]. Together with brain single-nucleus RNA sequencing, these resources supported a predominantly peripheral origin for C2 and SPON2 and a neuronal association for APLP1. Blood-derived brain extracellular-vesicle studies also show, however, that assigning circulating proteins to the central nervous system is technically difficult and highly dependent on enrichment strategy[36]. Transcript abundance does not equal protein secretion, and neither GTEx nor single-nucleus data capture perioperative flux, proteolytic processing or tissue-specific clearance. Thus, the absence of a strong POD-related transcriptional signal in the examined brain cells argues against, but does not exclude, a central source. Stable-isotope tracing, arteriovenous sampling, tissue-specific rescue and measurement of distinct molecular species are required to establish origin and direction of transport. The consistency between peripheral expression and liver-directed knockdown nevertheless provides functional support for a peripheral-to-central axis.

The intervention experiments followed logically from that source inference. Liver-directed knockdown of C2 or SPON2 and systemic recombinant APLP1 improved postoperative behavioural performance while reducing hippocampal gliosis, cytokine expression and loss of PSD-95 and synaptophysin. These convergent molecular and behavioural effects support modifiability rather than association alone. They should still be framed as attenuation of postoperative cognitive and inflammatory phenotypes, not direct reversal of human delirium. Consensus recommendations note that no single rodent task reproduces fluctuating attention and awareness, and advocate convergent behavioural, physiological and inflammatory measures[37]. Automated phenotyping can better capture acute delirium-like behavioural dynamics[38], whereas Barnes maze and fear conditioning primarily measure learning and memory over days. Future experiments should therefore combine dynamic attention/arousal measures with the current synaptic and inflammatory endpoints and test aged females, comorbid animals and clinically relevant treatment windows.

C2 offers a plausible mechanistic connection to complement-mediated synaptic injury. C2 participates in classical and lectin pathway C3 convertases, thereby influencing downstream C3a and C3b generation. Complement has physiological roles in host defence and developmental synaptic refinement: C1q and C3 mediate developmental synapse elimination[39], while complement-dependent microglial activity contributes to early synapse loss in Alzheimer's disease models[40]. Because complement effects depend on cellular source, activation fragment, timing and disease stage[41], our observation that hepatic C2 knockdown reduced gliosis and synaptic-marker loss is compatible with, but does not yet prove, a liver–C2/C3–brain pathway. It extends evidence that liver-derived C3 contributes to anaesthesia/surgery-induced cognitive impairment and microglial synaptic phagocytosis[42]. The apparently lower postoperative-to-preoperative C2 ratio previously reported in POD[23] need not contradict a risk effect of higher antecedent C2: acute complement activation may consume C2 after surgery. Serial measurement of intact C2, C3, activation fragments and terminal pathway activity, together with downstream rescue experiments, would distinguish predisposition from consumption and establish pathway order. Any therapeutic strategy must also consider infection risk from suppressing host defence.

SPON2 points to a potentially distinct peripheral immune–neurovascular route. In other biological systems, SPON2 regulates macrophage recruitment through integrin–Rho GTPase signalling[43] and can activate NF-κB with increased MMP2 and MMP9 expression[44]. These cancer-derived observations establish plausibility, not a delirium mechanism. Their relevance lies in downstream pathways already implicated in postoperative cognitive dysfunction: MMP9 inhibition can restore aquaporin-4 polarity and glymphatic function[45], while caveolin-1-dependent MMP2/9 activity has been linked to postoperative blood–brain barrier disruption[46]. Our finding that liver-directed SPON2 knockdown reduced hippocampal inflammation is therefore consistent with altered circulating immune-cell recruitment or neurovascular signalling. We did not establish that SPON2 crosses the blood–brain barrier, activates MMPs in this model or signals through a specific integrin. Labelled-protein trafficking, endothelial

permeability assays, MMP activity measurements and cell-specific receptor disruption will be needed to separate direct brain entry from peripheral signalling.

APLP1 showed the opposite direction and may represent synaptic resilience. APLP1 is a neuronal amyloid precursor protein-family member that acts as a synaptic adhesion molecule supporting dendritic-spine maintenance and basal transmission[47]; APP-family proteins more broadly contribute to neuronal development, plasticity and repair[48]. Recombinant APLP1 preserved synaptic markers and improved postoperative behaviour in our model, but the active species and receptor remain unknown. Soluble recombinant protein may differ from membrane-bound, cleaved or extracellular-vesicle-associated APLP1. This distinction matters because APLP1 can interact with LAG3 and facilitate pathological α-synuclein transmission[49], whereas APLP1-positive extracellular vesicles have been proposed as early neurological disease biomarkers[50]. Dose response, pharmacokinetics, tissue distribution and long-term safety must therefore precede therapeutic interpretation. More generally, our observational analyses were predominantly European, the perioperative cohort was small and restricted to older Chinese patients with femoral fracture, and only aged male mice were studied. Routine-record delirium definitions likely missed brief or hypoactive episodes. Finally, large trials show that altering anaesthetic technique alone does not reliably prevent POD[51,52], underscoring its multifactorial nature. Overall, the ordered evidence chain prioritizes C2, SPON2 and APLP1 as experimentally supported mediators, while prospective replication, serial biosampling, source tracing and mechanism-specific rescue remain essential before biomarker or therapeutic use.

Methods

UKB Study

Ethics. Ethical approval for the UK Biobank (reference no. 11/NW/0382) was granted by the North West Multicenter Research Ethics Committee. Under UK Biobank application number 854614, this study adhered to the Declaration of Helsinki principles.

Study participants. Participants with available baseline plasma proteomic data and linked clinical records were selected from the UK Biobank[53]. Plasma proteins were measured using the Olink proximity extension assay platform. Proteins with more than 20% missing values were excluded; missing values for the remaining proteins were imputed using the median, and protein measurements were standardized before analysis. In association analyses, each protein was entered as a continuous exposure and scaled to z-scores (per SD) for comparability of effect sizes. A total of 2,911 proteins were retained. Clinical variables included age, sex, smoking status, alcohol intake, diabetes, cardiovascular disease, and educational attainment. Incident delirium was identified from linked hospital inpatient records, International Classification of Diseases codes, and self-reported diagnoses. When multiple records were available, the earliest recorded date was defined as the diagnosis date. Participants with delirium diagnosed at or before baseline were excluded. Follow-up was calculated from the baseline assessment until incident delirium, death, and loss to follow-up. The final analytical cohort included 52,936 participants, of whom 1,269 developed incident delirium during follow-up.

Identification of delirium-related proteins. We visualized standardized levels of 2,911 proteins between delirium groups using boxplots. For all proteins, we plotted Kaplan-Meier curves and used log-rank tests to compare delirium-free survival after splitting participants at the median protein level (high vs. low). Cox models estimated HRs and 95% CIs. To test protein-delirium associations across different time points, we repeated Cox models in 4 blocks: 0-5 years (52,880 non-delirium vs 56 delirium), 5-10 years (51,240 non-delirium vs 399 delirium), 10-15 years (48,420 non-delirium vs 814 delirium) years, and full follow-up time (51,667 non-delirium vs 1,269 delirium). Bonferroni-adjusted P value < 0.05 was considered statistically significant.

Feature selection and optimization. LASSO regression reduced 401 candidates to 89 non-zero features. Random forest with five-fold cross-validation ranked features by importance; cumulative AUC plateaued at 21 features. SHAP analysis quantified each

feature's contribution[54]. Integrated selection identified 20 proteins and age as the final panel.

Prediction model development and evaluation. The 21-feature panel was used to develop five prediction models: Elastic Net, neural network (nnet), LDA, naïve Bayes, and logistic regression. Models were trained under three stratified splits (60:40, 70:30, 80:20) with five-fold cross-validation for parameter tuning. Discrimination was assessed by ROC-AUC with 95% CIs; optimal thresholds were determined by Youden index. Performance metrics included accuracy, sensitivity, specificity.

Surgical participants and POD assessment

Ethics. This study was approved by the Ethics Committee of Shanghai Fourth People's Hospital Affiliated to Tongji University School of Medicine (Approval No. 2024177-001) in July 2024 and registered in the Chinese Clinical Trial Registry (No. ChiCTR2400091413). Written informed consent was obtained from all participants, and all procedures were conducted in accordance with the principles outlined in the Declaration of Helsinki.

Inclusion and exclusion criteria

This prospective study enrolled patients undergoing open reduction and intramedullary nail fixation of femoral fractures under general anesthesia between November 2024 and December 2025. Inclusion and exclusion criteria were established with reference to the methodology of prior clinical investigations on postoperative delirium (POD)[55,56]. The inclusion criteria were as follows: (1) patients of Chinese Han ethnicity; (2) age ≥ 60 years; (3) scheduled for surgery under general anesthesia; (4) classified as American Society of Anesthesiologists (ASA) physical status I–III; (5) normal cognitive function at enrollment, defined as a Mini-Mental State Examination (MMSE) score > 27; (6) native Mandarin Chinese speakers; (7) ability to provide written informed consent, with sufficient verbal and written communication skills. Exclusion criteria, determined through medical record review, included: (1) presence of preoperative delirium; (2) documented history of neurological disorders, significant hepatic or renal impairment,

or metabolic diseases; (3) prior diagnosis of psychiatric conditions (e.g., major depressive disorder in an acute phase, schizophrenia, intellectual disability) or hereditary familial diseases; (4) significant visual or hearing impairment that could interfere with study assessments; and (5) refusal to participate.

Diagnosis of postoperative delirium (POD). Delirium was identified based on the International Statistical Classification of Diseases (ICD) codes, specifically ICD-10 (F05) and ICD-9 (293.0). All participating surgical patients were confirmed to have intact cognitive function preoperatively, as assessed by a MMSE score greater than 27 one day prior to surgery[55]. Diagnosis of POD was established within the first three postoperative days using the 3-Minute Diagnostic Interview for CAM (3D-CAM) instrument. Cases were confirmed as POD if they met the positive criteria of the CAM during this period, consistent with established diagnostic protocols[55,56]. A total of 36 elderly surgical patients were enrolled in the study, comprising 18 POD cases and 18 non-POD controls.

Serum collection

Fasting morning blood were collected at two time points: one hour prior to surgery (as baseline) and one hour after surgery. Serum samples were prepared from the blood and stored at −80 °C for subsequent molecular analyses, following established protocols as previously detailed[57].

Mice

C57BL/6J male mice at 18 months of age were utilized for the establishment of a delirium model. The animals were maintained in a controlled environment with standard conditions, including a 12-hour light/dark cycle, temperature of 24 ± 1°C, and relative humidity of 55 ± 5%, and were provided with ad libitum access to food and water. All experimental procedures were carried out in compliance with the guidelines of the Chinese Council on Animal Care and were approved by the Animal Ethics Committee of Tongji University (Approval No. TJBH00921101).

Viral vector preparation and administration

Recombinant adeno-associated virus vectors [rAAV-TBG-mCherry-shRNA(C2)-WPREs and rAAV-TBG-mCherry-shRNA (SPON2)-WPREs] along with a scrambled control virus [rAAV2/8-TBG-mCherry-shRNA (Scramble)-WPREs] were designed and produced by BrainVTA Technology (Wuhan, China). The mechanism underlying liver-specific modulation of target gene expression has been described previously[58]. In accordance with established protocols[58], each mouse received a tail-vein injection of 200 μL rAAV2/8 viral solution (1×10^{12} genomic copies per mouse), and interventions were maintained for three weeks.

Administration of recombinant APLP1 protein

Recombinant human APLP1 protein (HY-P7502, MCE) was prepared in physiological buffer, aliquoted, and stored at −80 °C for subsequent use. For each injection, a dose of 1 mg kg^{-1} was administered. 18-month-old control and delirium mice were administrated with either 0.9% saline (vehicle) or the recombinant human APLP1 via intravenous injection 1 h before surgery, with repeated administration at 6 h and 24 h postoperatively. Specific treatment schedules and experimental timelines are provided in the corresponding figure legends.

Delirium model

This procedure followed an established protocol[59]. Animals were anesthetized with 2 ± 0.2% isoflurane in 90% oxygen via a transparent acrylic chamber. A midline abdominal incision was made, the intestines were externalized, and the superior mesenteric artery was isolated and briefly clamped for 15 seconds. This clamping procedure was repeated three times over a total surgical duration of 10 min. The peritoneal layer and skin were closed with sterile 4-0 chromic gut sutures. After surgery, mice were kept under 2 ± 0.2% isoflurane anesthesia in a transparent chamber for 2 hours. Upon anesthetic discontinuation, a topical analgesic cream (2.5% lidocaine + 2.5% prilocaine) was applied to the incision site. A second application was

given 3 hours post-surgery. Mice were returned to their home cages once full mobility was restored.

Behavioural testing

Open field test. Spontaneous locomotor activity and anxiety-like behavior were assessed using the open field test[60]. After a 30-minute acclimatization period in the testing room on the third postoperative day, each mouse was placed in the center of a gray polyvinyl chloride arena and allowed to explore freely for 5 minutes. Motor function and anxiety levels were evaluated based on the average speed and time spent in the central zone, respectively, using Smart Video Tracking Software (Panlab; Harvard Apparatus).

Barnes maze test. Spatial learning and memory were evaluated using the Barnes maze, as previously described[57]. After four consecutive days of training (3-4 trials per day, days 5–9 post-surgery), the escape box was removed on the test day (day 10). Mice were allowed to search for the target location for 2 minutes. Performance was video-recorded and analyzed blindly using Smart Video Tracking Software (Panlab; Harvard Apparatus). The primary latency to enter the escape compartment and the percentage of time spent in the target quadrant were recorded.

Fear conditioning test. Contextual fear memory was examined following a previously established protocol[57]. On the habituation day, mice freely explored the conditioning chamber for 10 minutes. During training (day 11 post-surgery), baseline freezing was recorded for 2 minutes before delivery of five foot-shocks. Memory retrieval was tested 24 hours later (day 12) by returning mice to the same chamber for 5 minutes. Freezing, defined as the absence of movement for at least 2 seconds, was quantified using ANY-maze software (Stoelting Co.).

Sample collection of mice

Following deep anesthesia, blood was obtained from mice via retro-orbital bleeding, and serum was subsequently separated. Mice were euthanized under deep anesthesia, after which the brain, heart, liver, spleen, kidney, and lung tissues were harvested. All

collected samples were processed and stored for downstream molecular analyses according to previously described protocols[57].

Proteomics analysis of the serum from perioperative patients and mice

Protein Analysis. Serum samples were collected from 18 surgical patients without POD and 18 with POD both preoperatively and postoperatively, as well as from 6 control mice and 6 POD mice, for proteomic analysis. Proteomic profiling was performed on a Q Exactive HF-X mass spectrometer equipped with a Thermo UltiMate 3000 RSLCnano chromatography system. Peptides were separated using a trap column followed by an analytical column with a 125-min linear gradient from 1% to 90% mobile phase B at a flow rate of 400 nL/min. The mass spectrometer was operated in data-independent acquisition (DIA) mode.

Data analysis. DIA raw files were processed with DIA-NN (version 1.8.0) against a predicted human or mouse proteome spectral library. Peptide- and protein-level false discovery rates (FDR) were controlled at <1%. Protein abundances were quantified based on the top three intense peptides, normalized, and subjected to quality filtering.

Postoperative/preoperative protein abundance ratio processing. To identify proteins affected by surgical stimulation during the perioperative period in surgical patients, we calculated the ratio of postoperative to preoperative (Post/Preo ratio) protein profiles in 18 non-POD and 18 POD individuals. Differences in these protein profile ratios were then compared between the non-POD and POD groups.

Differential protein expression. Normalized intensities underwent $\log_2$ transformation before statistical analysis. The limma package (v3.50.3) in R (4.1.3) was used to perform differential expression, with significance thresholds set at $|\log_2 FC| > 0.25$ and Benjamini-Hochberg adjusted $P < 0.05$. For the Post/Pre ratio dataset, differential analysis was conducted between the non-POD and POD groups using Wilcoxon test, and proteins meeting $|\log_2 FC| > 0.1$ with $P < 0.05$ were considered significant.

Model development and evaluation for POD classification. Four machine-learning algorithms including Random Forest, XGBoost, CatBoost, and Extra Trees, were employed to construct classifiers distinguishing 18 non-POD from 18 POD patients based on two feature sets: (1) the 20 postoperative hub proteins, and (2) the Post/Pre ratios of six time-shared proteins. For each feature set, data were split 2:1 into training and test sets, and model training was repeated across five random seeds (42, 100, 200, 300, and 400) to ensure robustness. SHAP dependency plots were used to identify the most discriminative features. Model performance was evaluated on test sets using AUC, accuracy, sensitivity, and specificity. The optimal model was determined by the highest average AUC.

Causal inference analysis between plasma proteomics and delirium

Genome-Wide Association Study (GWAS) data source. Protein quantitative trait loci (pQTLs) for 4,907 circulating proteins were derived from the deCODE study (N=35,559)[61]. Delirium GWAS summary statistics were obtained from FinnGen R12 (4,381 cases, 469,981 controls; https://r12.finngen.fi/), with cases defined by ICD-10 F05 or ICD-9 2930[62]. Two-sample MR used pQTLs as instrumental variables.

Instrumental variable selection. Protein pQTLs were selected as IVs according to MR assumptions[60]: (1) significant association with exposure ($P < 5\times10^{-8}$); (2) independence (LD r^2 <0.1 within 1 Mb, 1000 Genomes European reference); (3) no direct association with delirium ($P > 0.05$); (4) MHC region (chr6:26–34 Mb) excluded; (5) F-statistic >10. PhenoScanner (phenoscanner R package v. 1.0) was used to remove variants linked to known delirium risk factors (dementia, CVD, stroke, sleep disorders, Parkinson's; $P < 5\times10^{-8}$).

MR estimation and sensitivity analyses. Wald ratio (single SNP) or IVW (multiple SNPs) estimated causal effects; A P value < 0.05 was considered statistically significant for all MR tests. As previously established[60], exposures with more than three IVs were considered potentially susceptible to heterogeneity. Several sensitivity analyses were conducted to assess the robustness of the MR findings: Cochran's Q for heterogeneity

($P_{Q\text{-}stat} > 0.05$), MR-PRESSO global test for outliers ($P_{Global\ Test} > 0.05$), MR-Egger intercept for directional pleiotropy ($P_{Egger\text{-}Intercept} > 0.05$), leave-one-out for influential SNPs, and Steiger test for reverse causation ($P_{Steiger\ test} < 0.05$ indicating correct causal direction).

Origin identification of proteins

To trace the potential tissue origins of delirium-associated circulating proteins, we retrieved bulk RNA-seq TPM data from the GTEx portal (GTEx_Analysis_2017-06-05_v8_RNASeQCv1.1.9_gene_tpm.gct), covering both brain regions and peripheral organs across 11,584 male and 5,878 female samples. After $\log_2$(TPM+1) transformation, mean multi-organ mRNA expression levels of genes encoding each circulatory protein were calculated for cerebral and peripheral tissues. Expression patterns across organs were visualized using the pheatmap (v.1.0.12) and gganatogram (v.1.1.1) packages.

To integrate protein-level evidence with transcriptional data, we referenced The Human Protein Atlas (HPA; https://www.proteinatlas.org/) for proteins previously reported to enter the circulation and exert systemic effects in humans[63]. An integrated evaluation at both transcriptomic and proteomic levels was performed to further infer the tissue origins of these circulatory proteins.

Single-nucleus RNA-seq for the brain between non-POD and POD patients

Data processing and cell type annotation. The snRNA-Seq data (GSE291019) was obtained from GEO database[64]. UMI counts processed in Seurat v5.0.2. Nuclei filtered (200-6000 genes, 500-25,000 UMIs, <5% mitochondrial RNA), log-normalized (scale 10,000), and 2000 variable genes selected. Batch correction by IntegrateData; data scaled for age/sex/mitochondrial content. Clusters were identified at a resolution of 0.4 and annotated by canonical marker genes into the following cell types: astrocytes, endothelial cells, excitatory neurons, inhibitory neurons, fibroblasts, microglia, oligodendrocytes, and oligodendrocyte precursor cells (OPCs)[64].

Differential expression analysis of target genes. For C2, SPON2, and APLP1, log-normalised expression per cell type was compared between non-POD and POD using two-sided Wilcoxon tests ($P < 0.05$). Results shown as mean ± SEM bar plots with significance asterisks.

Function analyses

To explore underlying pathways, we performed GSEA on the 2,911 UKB proteins using clusterProfiler (v4.6.0), and significance was set at q-value < 0.05 after 5,000 permutations. Additionally, KEGG and GO enrichment analyses were conducted on the 20 POD-related proteins using the same package, with a q-value < 0.05 as threshold.

Immunofluorescence

Immunofluorescence staining was conducted following our previously established protocol[57]. Briefly, brain and liver tissues were fixed via transcardial perfusion with 4% paraformaldehyde (PFA) for 24 hours. After fixation, tissues were cryoprotected in 30% sucrose solution for 72 hours and sectioned at a thickness of 40 μm using a freezing microtome. Sections were washed three times (10 minutes each) in 0.1 M phosphate-buffered saline (PBS), then blocked for 1 hour at room temperature in a solution containing 5% donkey serum and 0.5% Triton X-100. Subsequently, sections were incubated overnight at 4°C with the following primary antibodies: rabbit anti-Iba1 (1:1000, 17198S, CST), rabbit anti-PSD95 (1:1000, ET1602-20, HUABIO), mouse anti-synaptophysin (1:1000, 67864-1-Ig, Proteintech), mouse anti-GFAP (1:1000, 3670T, CST).

Following primary antibody incubation, sections were counterstained with DAPI and incubated with the appropriate species-specific secondary antibodies: goat anti-rabbit Alexa Fluor 647 (1:500, ab150079, Abcam), goat anti-rabbit Alexa Fluor 594 (1:500, ab150080, Abcam), goat anti-mouse Alexa Fluor 647 (1:500, ab150115, Abcam) goat anti-mouse Alexa Fluor 594 (1:500, 150116, Abcam). Finally, sections were mounted and imaged using an Olympus FV3000 confocal microscope with FV31S-SW acquisition software.

Hematoxylin and eosin (H&E) staining

Five major organs (heart, lung, liver, spleen, kidney) were fixed in neutral buffered formalin and processed for paraffin embedding. Three-micron sections were stained with H&E using conventional methods[57].

Statistical analysis

All statistical analyses were performed using R 4.1.3, and figures were generated with ggplot2 3.4.4. Baseline characteristics were compared between delirium groups in UKB and POD cohorts. Categorical variables as n (%) and compared by chi-square or Fisher's exact test. Continuous variables were expressed as mean (SD or SEM) or median (IQR). Two-group comparisons used t-test or Wilcoxon rank-sum test, while multi-group comparisons used ANOVA with Tukey's HSD or Kruskal-Wallis with Dunn's test, depending on distribution. Two-way repeated-measures ANOVA with Bonferroni correction was applied for repeated measures. Statistical significance was set at two-sided $P < 0.05$.

Acknowledgement

The authors acknowledge the Department of Anesthesiology and Perioperative Medicine at Shanghai Fourth People's Hospital for facilitating patient enrollment and providing help with biospecimen and clinical information acquisition. We are also thankful to all the patients who consented to participate in this study.

Funding

This work was supported by Major Program of National Natural Science Foundation of China (82293640 and 82293643 to L.X), National Natural Science Foundation of China (82501531 to H.Z), Scientific Research Founds of Shanghai Fourth People’s Hospital (sykyqd08701 to X. H; sykyqd07401 to Y.Z; SY-XRZT-2024-1002, sykyqd09101and 2026-0674-YB-01 to J. W), Postdoctoral Fellowship Program of CPSF (GZC20261227 to J. W), and Supported by the Postdoctoral Fellowship Program of CPSF under Grant Number GZC20261227.

Author contributions

Xinwei Huang, Xinhua Wei, Jing Wang, and Lize Xiong designed the study. Xinwei Huang, Jingxin Yan and Yinggang Zheng performed the analysis. Xinwei Huang, Xinhua Wei, Jing Wang, and Lize Xiong supervised the study. Peilin Cong, Hui Zhang, and Jing Wang collected clinical information and samples. Jing Wang, Qingyuan Miao, Liuqing Yang, and Ban Feng performed experiment. Xinwei Huang, Jingxin Yan and Yinggang Zheng, and Jing Wang wrote the manuscript. Xinwei Huang formed the key revision group. Xinwei Huang and Lize Xiong provided study directions. All the authors contributed to the final revision of the manuscript.

Data availability

The data that support the findings of this study are not publicly available due to ethical restrictions imposed by the relevant institutional review board and legal provisions concerning participant privacy. De-identified data may be made available by the corresponding author, Prof. Xinwei Huang (huanggenetics@tongji.edu.cn), upon reasonable request and subject to a formal data access agreement. All other public datasets utilized in this manuscript are available from their respective sources as shown in the Methods.

Declarations

Ethics approval

This research was conducted under UK Biobank application number 854614. For surgical participants, this study was approved by the Ethics Committee of Shanghai Fourth People's Hospital Affiliated to Tongji University School of Medicine (Approval No. 2024177-001) in July 2024 and registered in the Chinese Clinical Trial Registry (No. ChiCTR2400091413). For animal experiment, all procedures were carried out in compliance with the guidelines of the Chinese Council on Animal Care and were approved by the Animal Ethics Committee of Tongji University (Approval No. TJBH00921101).

AI use statement

During manuscript preparation, the authors used ChatGPT 5.2 for language editing and text refinement, including improving the clarity and structure of selected sections and figure/table descriptions. AI tools were not used to perform autonomous data analysis, interpret results, or generate study findings. All content was reviewed and verified by the authors, who take full responsibility for the manuscript.

Conflicts of interest

The authors declare no competing interests.

References

1. Lander HL, Dick AW, Joynt Maddox KE, et al. Postoperative Delirium in Older Adults Undergoing Noncardiac Surgery. *JAMA network open* 2025; **8**(7): e2519467.
2. Heffernan Á B, Steinruecke M, Dempsey G, et al. Role of glia in delirium: proposed mechanisms and translational implications. *Molecular psychiatry* 2025; **30**(3): 1138-47.
3. Wu CR, Chang KM, Tranyor V, Chiu HY. Global incidence and prevalence of delirium and its risk factors in medically hospitalized older patients: A systematic review and meta-analysis. *International journal of nursing studies* 2025; **162**: 104959.
4. Cheng SL, Hsu TW, Kao YC, et al. Haloperidol in treating delirium, reducing mortality, and preventing delirium occurrence: Bayesian and frequentist meta-analyses. *Critical care (London, England)* 2025; **29**(1): 126.
5. Mattison MLP. Delirium. *Annals of internal medicine* 2020; **173**(7): Itc49-itc64.
6. Lim XM, Lim ZHT, Ignacio J. Nurses' experiences in the management of delirium among older persons in acute care ward settings: A qualitative systematic review and meta-aggregation. *International journal of nursing studies* 2022; **127**: 104157.
7. Cheng C, Wan H, Cong P, et al. Targeting neuroinflammation as a preventive and therapeutic approach for perioperative neurocognitive disorders. *Journal of neuroinflammation* 2022; **19**(1): 297.
8. Cong P, Huang X, Zhang Q, et al. DNMT3a Deficiency Contributes to Anesthesia/Surgery-Induced Synaptic Dysfunction and Cognitive Impairment in Aged Mice. *Aging cell* 2025; **24**(4): e14458.
9. Vasunilashorn SM, Lunardi N, Newman JC, et al. Preclinical and translational models for delirium: Recommendations for future research from the NIDUS delirium network. *Alzheimer's & dementia : the journal of the Alzheimer's Association* 2023; **19**(5): 2150-74.
10. Wu Q, Cong P, Li Z, et al. Inhibiting Liver-Derived C3 Protein Rescues Anesthesia/Surgery-Induced Cognitive Impairment, Synaptic Disorders, and Microglial Phagocytosis. *Advanced science (Weinheim, Baden-Wurttemberg, Germany)* 2026; 13(2): e02034.
11. Sun BB, Chiou J, Traylor M, et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* 2023; 622(7982): 329-38.
12. Wilson, J. E. et al. Delirium. Nat Rev Dis Primers. 6, 90 (2020).

13. Inouye, S. K., Westendorp, R. G. J. & Saczynski, J. S. Delirium in elderly people. Lancet 383, 911–922 (2014).

14. Marcantonio, E. R. Delirium in hospitalized older adults. N. Engl. J. Med. 377, 1456–1466 (2017).

15. Saczynski, J. S. et al. Cognitive trajectories after postoperative delirium. N. Engl. J. Med. 367, 30–39 (2012).

16. Sudlow, C. et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 12, e1001779 (2015).

17. Sun, B. B. et al. Plasma proteomic associations with genetics and health in the UK Biobank. Nature 622, 329–338 (2023).

18. Cerejeira, J., Firmino, H., Vaz-Serra, A. & Mukaetova-Ladinska, E. B. The neuroinflammatory hypothesis of delirium. Acta Neuropathol. 119, 737–754 (2010).

19. Wang, P. et al. Neurovascular and immune mechanisms that regulate postoperative delirium superimposed on dementia. Alzheimers Dement. 16, 734–749 (2020).

20. Heffernan, Á. B. et al. Role of glia in delirium: proposed mechanisms and translational implications. Mol. Psychiatry 30, 1138–1147 (2025).

21. Dillon, S. T. et al. Higher C-reactive protein levels predict postoperative delirium in older patients undergoing major elective surgery: a longitudinal nested case-control study. Biol. Psychiatry 81, 145–153 (2017).

22. Vasunilashorn, S. M. et al. Proteome-wide analysis using SOMAscan identifies and validates chitinase-3-like protein 1 as a risk and disease marker of delirium among older adults undergoing major elective surgery. J. Gerontol. A Biol. Sci. Med. Sci. 77, 484–493 (2022).

23. Khan, S. H. et al. Serum proteomic analysis in esophagectomy patients with postoperative delirium: a case-control study. Heart Lung 63, 35–41 (2024).

24. Collins, G. S. et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ 385, e078378 (2024).

25. Steyerberg, E. W. et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 21, 128–138 (2010).

26. Vickers, A. J., van Calster, B. & Steyerberg, E. W. Decision curve analysis to evaluate the clinical benefit of prediction models. Spine J.21(10):1643-1648 (2021).

27. Wolff, R. F. et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann. Intern. Med. 170, 51–58 (2019).

28. Ferkingstad, E. et al. Large-scale integration of the plasma proteome with genetics and disease. Nat. Genet. 53, 1712–1721 (2021).

29. Kurki, M. I. et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature 613, 508–518 (2023).

30. Zheng, J. et al. Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. Nat. Genet. 52, 1122–1131 (2020).

31. Chen, L. et al. Systematic Mendelian randomization using the human plasma proteome to discover potential therapeutic targets for stroke. Nat. Commun. 13, 6143 (2022).

32. Schmidt, A. F. et al. Genetic drug target validation using Mendelian randomisation. Nat. Commun. 11, 3255 (2020).

33. Skrivankova, V. W. et al. Strengthening the reporting of observational studies in epidemiology using Mendelian randomization: the STROBE-MR statement. JAMA 326, 1614–1621 (2021).

34. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. Science 369, 1318–1330 (2020).

35. Uhlén, M. et al. The human secretome. Sci. Signal. 12, eaaz0274 (2019).

36. Badhwar, A. P. et al. Assessment of brain-derived extracellular vesicle enrichment for blood biomarker analysis in age-related neurodegenerative diseases. Alzheimers Dement. 20, 4411–4422 (2024).

37. Vasunilashorn, S. M. et al. Preclinical and translational models for delirium: recommendations for future research from the NIDUS delirium network. Alzheimers Dement. 19, 2150–2174 (2023).

38. Cao, S. et al. Automated phenotyping of postoperative delirium-like behaviour in mice reveals the therapeutic efficacy of dexmedetomidine. Commun. Biol. 6, 807 (2023).

39. Stevens, B. et al. The classical complement cascade mediates CNS synapse elimination. Cell 131, 1164–1178 (2007).

40. Hong, S. et al. Complement and microglia mediate early synapse loss in Alzheimer mouse models. Science. 352, 712–716 (2016).

41. Gomez-Arboledas, A., Acharya, M. M. & Tenner, A. J. The role of complement in synaptic pruning and neurodegeneration. Immunotargets Ther. 10, 373–386 (2021).

42. Wu, Q. et al. Inhibiting liver-derived C3 protein rescues anesthesia/surgery-induced cognitive impairment, synaptic disorders, and microglial phagocytosis. Adv. Sci. 13, e02034 (2026).

43. Zhang, Y. L. et al. SPON2 promotes M1-like macrophage recruitment and inhibits hepatocellular carcinoma metastasis by distinct integrin–Rho GTPase–Hippo pathways. Cancer Res. 78, 2305–2317 (2018).

44. Wu, M., Kong, D. & Zhang, Y. SPON2 promotes the bone metastasis of lung adenocarcinoma via activation of the NF-κB signaling pathway. Bone 167, 116630 (2023).

45. Zhu, B. et al. MMP-9 inhibition alleviates postoperative cognitive dysfunction by improving glymphatic function via regulating AQP4 polarity. Int. Immunopharmacol. 126, 111215 (2024).

46. Li, X. et al. Caveolin-1-mediated blood–brain barrier disruption via MMP2/9 contributes to postoperative cognitive dysfunction. Neurochem. Res. 50, 210 (2025).

47. Schilling, S. et al. APLP1 is a synaptic cell adhesion molecule, supporting maintenance of dendritic spines and basal synaptic transmission. J. Neurosci. 37, 5345–5365 (2017).

48. Müller, U. C., Deller, T. & Korte, M. Not just amyloid: physiological functions of the amyloid precursor protein family. Nat. Rev. Neurosci. 18, 281–298 (2017).

49. Mao, X. et al. Aplp1 interacts with Lag3 to facilitate transmission of pathologic α-synuclein. Nat. Commun. 15, 4663 (2024).

50. Choi, Y. et al. Blood-derived APLP1-positive extracellular vesicles are potential biomarkers for the early diagnosis of brain diseases. Sci. Adv. 11, eado6894 (2025).

51. Neuman, M. D. et al. Spinal anesthesia or general anesthesia for hip surgery in older adults. N. Engl. J. Med. 385, 2025–2035 (2021).

52. Li, T. et al. Effect of regional vs general anesthesia on incidence of postoperative delirium in older patients undergoing hip fracture surgery: the RAGA randomized trial. JAMA 327, 50–58 (2022).

53. Sudlow, C. et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 12, e1001779 (2015).

54. Lundberg, S. M. et al. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2, 56–67 (2020).

55. Zhao, Y. et al. Postoperative delirium risk in patients with hyperlipidemia: a prospective cohort study. J. Clin. Anesth. 98, 111573 (2024).

56. Li, T. et al. Effect of regional vs general anesthesia on incidence of postoperative delirium in older patients undergoing hip fracture surgery: the RAGA randomized trial. JAMA 327, 50–58 (2022).

57. Wu, Q. et al. Inhibiting liver-derived C3 protein rescues anesthesia/surgery-induced cognitive impairment, synaptic disorders, and microglial phagocytosis. Adv. Sci. 13, e02034 (2026).

58. Wu, Y. et al. Hepatic soluble epoxide hydrolase activity regulates cerebral Aβ metabolism and the pathogenesis of Alzheimer's disease in mice. Neuron 111, 2847–2862.e10 (2023).

59. Cao, S. et al. Automated phenotyping of postoperative delirium-like behaviour in mice reveals the therapeutic efficacy of dexmedetomidine. Commun. Biol. 6, 807 (2023).

60. Chen, L. et al. Systematic Mendelian randomization using the human plasma proteome to discover potential therapeutic targets for stroke. Nat. Commun. 13, 6143 (2022).

61. Ferkingstad, E. et al. Large-scale integration of the plasma proteome with genetics and disease. Nat. Genet. 53, 1712–1721 (2021).

62. Kurki, M. I. et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature 613, 508–518 (2023).

63. Uhlén, M. et al. The human secretome. Sci. Signal. 12, eaaz0274 (2019).

64. Ishii, T. et al. Glial contribution to the pathogenesis of post-operative delirium revealed by multi-omic analysis of brain tissue from neurosurgery patients. bioRxiv https://doi.org/10.1101/2025.03.13.643155 (2025).

Supplementary Files

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

- SupplementalFigures.docx
- SupplementaryTable.xlsx