

Cerebrospinal Fluid Chloride as an Independent Prognostic Biomarker in Cryptococcal Meningitis: A Dual-Center Externally Validated Machine Learning Study with SHAP-Based Interpretability

Qihang Zou

Shanghai Changzheng Hospital

Keyao Yu

Shanghai Changzheng Hospital

Dongzhen Liu

Shanghai Changzheng Hospital

Xinlin Zhu

Shanghai Changzheng Hospital

Songting Li

Shanghai Changzheng Hospital

Chao Zhang

Shanghai Changzheng Hospital

Xiaopan Wang

Shanghai Changzheng Hospital

Wanqing Liao

Shanghai Changzheng Hospital

Wenting Lin

Shanghai Changzheng Hospital

Keming Zhang

Shanghai Changzheng Hospital

Weihua Pan

panweihua9@sina.com

Shanghai Changzheng Hospital

Research Article

Keywords: Cryptococcal meningitis, cerebrospinal fluid chloride, prognosis prediction, CatBoost, incremental predictive value, interpretable machine learning

Posted Date: August 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10261941/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Cryptococcal meningitis (CM) carries a high 10-week mortality rate, necessitating accurate early prognostic tools. While cerebrospinal fluid chloride (CSF-Cl) is a routine test, its independent prognostic value in CM remains ill-defined. Furthermore, existing machine learning models lack external validation and clinical interpretability. This study aimed to validate CSF-Cl as an independent biomarker and to develop an externally validated, interpretable prognostic model for CM.

Methods

We enrolled 253 treatment-naïve CM patients from two tertiary hospitals. Univariate and multivariate logistic regression identified independent prognostic risk factors across the full cohort. Patients were stratified into a training cohort (The Fourth People's Hospital of Nanning, $n = 194$) and an independent external test cohort (Jiangxi Provincial Chest Hospital, $n = 59$). In the training set, nine machine learning classifiers were developed following feature selection via LASSO-recursive feature elimination (LASSO-RFE) and sample balancing via synthetic minority oversampling technique (SMOTE). Paired models with and without CSF-Cl were compared to quantify the incremental value of CSF-Cl using Δ AUC, net reclassification index (NRI), integrated discrimination improvement (IDI), and decision curve analysis (DCA). SHapley Additive exPlanations (SHAP) was deployed to dissect dose-response relationships and feature interactions, with subgroup analyses to define risk thresholds.

Results

Multivariate regression confirmed CSF-Cl as an independent risk factor for unfavorable CM outcomes (OR = 1.059, 95% CI 1.011–1.110, $P = 0.017$). CatBoost demonstrated the best generalizability among the nine models evaluated, yielding an AUC of 0.712 and precision of 0.722 in the external test set. Incorporation of CSF-Cl increased the external AUC from 0.651 to 0.712 (Δ AUC = 0.061), with an NRI of 0.132 and IDI of 0.040. SHAP analysis uncovered a positive dose-response association between CSF-Cl and adverse prognosis, alongside synergistic interactions with intracranial pressure (ICP) and HIV status. The model retained robust performance in HIV-positive patients and those aged < 50 years (AUC = 0.752 and 0.759, respectively). A CSF-Cl threshold of ≥ 115 mmol/L for risk stratification was derived, and an online interpretable prediction tool was implemented.

Conclusions

Elevated baseline CSF-Cl is independently linked to 10-week unfavorable prognosis in CM and substantially augments the predictive power of conventional clinical markers. Our dual-center externally validated CatBoost model with SHAP-enabled interpretability supports early, precise risk stratification for CM patients.

1. Introduction

Cryptococcal meningitis (CM) is the most prevalent fungal central nervous system infection in both HIV-infected and non-HIV immunocompromised patients, accounting for 15–20% of all AIDS-related deaths globally [1]. In 2021, approximately 194,000 new CM cases were reported worldwide, with 147,000 associated deaths, corresponding to an overall case fatality rate of 75.8% [2]. Even with standard-of-care antifungal therapy, 10-week mortality remains as high as 50% in low- and middle-income countries [1, 3]. Despite the relative standardization of antifungal treatment regimens, clinical outcomes vary substantially across patients, and early, accurate identification of high-risk individuals is pivotal to improving survival.

Currently, prognostic evaluation in clinical practice relies primarily on age, mental status, intracranial pressure, and routine cerebrospinal fluid (CSF) biochemical parameters (e.g., protein, glucose) [4–8]. However, these conventional markers exhibit limited predictive performance, and indices like CSF glucose are substantially confounded by systemic blood glucose fluctuations [9].

Cerebrospinal fluid chloride (CSF-Cl) is a routine CSF biochemical assay that has long been considered a non-specific marker. In classic bacterial meningitis, elevated blood-brain barrier permeability combined with the Gibbs–Donnan effect results in decreased CSF-Cl concentrations [10, 11]. However, the pathophysiological features of CM differ fundamentally from those of acute bacterial infection: cryptococcal capsular polysaccharides induce a hyperviscous CSF state [12], which may give rise to distinct CSF-Cl alteration patterns. Previous studies have documented reduced pre-treatment CSF-Cl levels in CM patients [5, 13], but emerging clinical observations suggest a more nuanced association with prognosis [14, 15]. To date, no multicenter cohort has systematically validated the independent predictive value and incremental contribution of CSF-Cl in CM.

Furthermore, CM prognosis is governed by the intertwined effects of host immunity, pathogen burden, and treatment response, with intricate interactions among these variables [16, 25]. Traditional logistic regression is insufficient to capture nonlinear relationships in small, imbalanced datasets. Notably, currently available machine learning models for CM prognosis share three critical limitations: lack of independent external validation, absence of standardized interpretability frameworks, and scarcity of clinically translatable tools, all of which hinder their real-world clinical implementation [17].

To address these critical research gaps, this study pursues two core objectives. First, based on a dual-center retrospective cohort, we validate whether CSF-Cl serves as an independent risk factor for 10-week unfavorable prognosis in patients with CM, and quantify its incremental predictive value relative to conventional clinical indicators using the net reclassification index (NRI) and integrated discrimination improvement (IDI). Second, we construct an independently externally validated machine learning prediction model with SHapley Additive exPlanations (SHAP)-enabled interpretability, delineate the dose-response relationship of CSF-Cl and its key interactions, and develop a streamlined online prediction tool for early clinical risk stratification.

2. Patients and Methods

2.1 Study Population and Design

This multicenter retrospective observational cohort study enrolled patients with confirmed CM from two tertiary hospitals: the Fourth People's Hospital of Nanning (December 2014 to June 2019) and Jiangxi Provincial Chest Hospital (2012 to 2018).

Inclusion criteria were as follows: (1) clinically confirmed CM, defined by positive CSF India ink staining or positive cryptococcal culture; (2) first-onset disease with no prior standard antifungal therapy before enrollment; (3) complete baseline CSF chloride data and other clinical and laboratory records; (4) age ≥ 14 years.

Exclusion criteria included: (1) concurrent central nervous system infections such as tuberculous, bacterial, or viral meningitis; (2) missing baseline CSF-Cl data; (3) age < 14 years; (4) incomplete clinical data precluding outcome assessment.

A total of 253 patients with CM were ultimately included. The study protocol was approved by the institutional review boards of both hospitals, and written informed consent was obtained from all patients or their legal guardians. Cohort stratification was performed as follows: the full cohort of 253 patients was used for univariate and multivariable logistic regression analysis of prognostic risk factors. The cohort was split by enrollment center into two datasets: a training set (Fourth People's Hospital of Nanning, $n = 194$) used for feature selection, sample balancing, model training, and

hyperparameter optimization; and an independent external test set (Jiangxi Provincial Chest Hospital, n = 59), which remained blinded throughout all modeling procedures and was used exclusively to validate the external generalizability of the final model.

2.2 Data Collection and Outcome Definition

Baseline patient data were collected, including demographic characteristics (age, sex), clinical symptoms (fever, nausea and vomiting, seizures, altered mental status, fatigue), cranial imaging findings (parenchymal lesions, hydrocephalus, cerebral infarction, meningeal enhancement), and laboratory parameters (ICP, white blood cell count, CSF protein, CSF glucose, CSF-CI, HIV infection status, and disseminated infection status).

The primary outcome was assessed 10 weeks after diagnosis or initiation of antifungal therapy [26], and was categorized as favorable or unfavorable. Favorable prognosis was defined as survival with negative CSF culture at both 2 and 10 weeks, or stable clinical symptoms with complete symptom resolution at 10 weeks. Unfavorable prognosis was defined as death; progressive disease (worsening consciousness, persistent fever or headache, and persistently positive CSF culture or smear); or hospital transfer due to treatment failure or disease recurrence.

2.3 Missing Data Handling

The overall missing rate of baseline data across the 253 patients was 4.8%, with missing rates for all individual variables below 10% (the highest missing rate was 7.1% for CSF glucose), indicating low-level missing at random. To maximize sample size and minimize single-imputation bias, we applied K-nearest neighbor (KNN) imputation (k = 5, Euclidean distance). All imputation parameters were trained and applied exclusively within the training set; the internal validation set and external test set used the imputation model built on the training set to avoid data leakage. After imputation, the full cohort had 100% data completeness (**Supplementary Material 1**).

2.4 Conventional Statistical Analysis

Statistical analyses were performed using SPSS 26.0 (IBM Corp., Armonk, NY, USA). For continuous variables, normality was tested first; normally distributed variables are presented as mean $\pm$ standard deviation, and between-group comparisons were performed using independent samples t-tests. Non-normally distributed variables are presented as median (interquartile range, IQR), and between-group comparisons were conducted using the Mann-Whitney U test. Categorical variables are expressed as n (%), and between-group differences were assessed using the chi-square test.

Variables with a P-value < 0.1 in univariate analysis were entered into a multivariable binary logistic regression model to identify independent prognostic risk factors after adjusting for confounders. Odds ratios (OR) and 95% confidence intervals (CI) were calculated. A two-tailed α level of 0.05 was set as the threshold for statistical significance.

2.5 Machine Learning Modeling Workflow

Modeling was implemented in Python 3.8 (Python Software Foundation, Wilmington, DE, USA). Random seeds (42, 2026) were fixed throughout the procedure to ensure experimental reproducibility.

2.5.1 Feature selection: Within the training set, least absolute shrinkage and selection operator (LASSO) regression (optimal $\lambda = 0.0203$ determined via 10-fold cross-validation) was combined with recursive feature elimination (RFE) to remove redundant variables with no predictive value, retaining the optimal feature subset [27].

2.5.2 Sample balancing: Given the low proportion of unfavorable prognosis cases in the training set, the synthetic minority oversampling technique (SMOTE) was applied to synthesize minority class samples [28], balancing the positive-to-negative sample ratio to 1:1. The external test set retained its original real-world distribution without any balancing processing.

2.5.3 Model construction: Nine classical classification models were developed: logistic regression (LR), KNN, support vector machine (SVM), random forest (RF), extremely randomized trees (ET), gradient boosting machine (GBM), XGBoost,

LightGBM, and CatBoost. Hyperparameter tuning was performed via 5-fold stratified cross-validation combined with grid search, using the internal validation set area under the receiver operating characteristic curve (AUC) as the optimization target. For CatBoost, an ordered boosting strategy was adopted to avoid target leakage of categorical features and reduce the risk of overfitting in small samples.

2.5.4 Performance evaluation: Evaluation metrics included AUC, accuracy, precision, recall, and F1-score for the training set, internal validation set, and external test set. External test set AUC was used as the core criterion for assessing generalizability.

2.6 Quantification of Incremental Predictive Value of CSF-CI

Two parallel control CatBoost models were constructed, with identical dataset partitioning, feature selection, oversampling, and hyperparameter settings except for the inclusion of CSF-CI: (1) Model A (full model): included all 10 features selected via LASSO-RFE, including CSF-CI; (2) Model B (baseline model): excluded CSF-CI, with all remaining features identical to Model A.

Evaluation metrics included: (1) Δ AUC of receiver operating characteristic curves between the two models compared via DeLong's test; (2) clinical reclassification metrics: NRI and IDI, with a risk classification threshold of 0.5; (3) clinical net benefit compared via decision curve analysis (DCA); (4) calibration curves to assess the concordance between model-predicted probabilities and actual event rates. NRI > 0 and IDI > 0 indicate that the added marker significantly improves model discrimination and risk stratification performance.

2.7 Model Interpretability and Robustness Validation

SHAP interpretability analysis

SHAP values were calculated based on the optimal CatBoost model. Visualizations included a SHAP summary plot (global feature importance ranking), a CSF-CI single-feature dependence plot (dose-response relationship), and a feature interaction heatmap (interaction strength between CSF-CI and ICP, and between CSF-CI and HIV status).

Subgroup robustness analysis

Stratification was performed in the independent external test set by HIV status, age, altered mental status, ICP, CSF glucose, and sex. AUC, sensitivity, specificity, F1-score, and calibration error were calculated for each subgroup to verify the cross-population stability of the model.

3. Results

3.1 Baseline Characteristics and Univariate Analysis

Of the 253 enrolled patients with CM, 185 (73.1%) had a favorable 10-week prognosis and 68 (26.9%) had an unfavorable prognosis. Univariate analysis (Table 1) showed that patients in the unfavorable prognosis group were older (53.00 ± 12.53 vs. 47.00 ± 13.28 years, $P = 0.020$) and had a higher HIV infection rate (60.29% vs. 32.43%, $P < 0.001$). Regarding clinical manifestations, the rates of altered mental status (38.24% vs. 13.51%, $P < 0.001$), fatigue (86.76% vs. 62.70%, $P < 0.001$), seizures (19.12% vs. 8.11%, $P = 0.019$), and nausea and vomiting (63.24% vs. 47.03%, $P = 0.032$) were significantly higher in the unfavorable prognosis group.

Table 1
Baseline characteristics and univariate analysis of the total cohort

Clinical / Baseline Characteristics	Poor Prognosis (N = 68)	Good Prognosis (N = 185)	p-value
Sex, n (%)			0.981
Male	46 (67.65%)	123 (66.49%)	
Female	22 (32.35%)	62 (33.51%)	
Age, years, Median [IQR]	53.00 [39.00–67.00]	47.00 [35.00–61.00]	0.020
HIV Infection Status, n (%)			< 0.001
Negative	27 (39.71%)	125 (67.57%)	
Positive	41 (60.29%)	60 (32.43%)	
Brain Parenchymal Lesions, n (%)			0.612
Absent	57 (83.82%)	148 (80.00%)	
Present	11 (16.18%)	37 (20.00%)	
Hydrocephalus, n (%)			0.787
Absent	64 (94.12%)	171 (92.43%)	
Present	4 (5.88%)	14 (7.57%)	
Cerebral Infarction, n (%)			0.765
Absent	65 (95.59%)	173 (93.51%)	
Present	3 (4.41%)	12 (6.49%)	
Meningeal Enhancement, n (%)			0.140
Absent	62 (91.18%)	153 (82.70%)	
Present	6 (8.82%)	32 (17.30%)	
Fever, n (%)			0.415
Absent	16 (23.53%)	55 (29.73%)	
Present	52 (76.47%)	130 (70.27%)	
Altered Mental Status, n (%)			< 0.001
Absent	42 (61.76%)	160 (86.49%)	
Present	26 (38.24%)	25 (13.51%)	
Headache, n (%)			0.981
Absent	7 (10.29%)	17 (9.19%)	
Present	61 (89.71%)	168 (90.81%)	
Nausea and Vomiting, n (%)			0.032

Clinical / Baseline Characteristics	Poor Prognosis (N = 68)	Good Prognosis (N = 185)	p-value
Sex, n (%)			0.981
Male	46 (67.65%)	123 (66.49%)	
Absent	25 (36.76%)	98 (52.97%)	
Present	43 (63.24%)	87 (47.03%)	
Seizures, n (%)			0.019
Absent	55 (80.88%)	170 (91.89%)	
Present	13 (19.12%)	15 (8.11%)	
Fatigue, n (%)			< 0.001
Absent	9 (13.24%)	69 (37.30%)	
Present	59 (86.76%)	116 (62.70%)	
Site of Infection, n (%)			0.030
Disseminated	36 (52.94%)	127 (68.65%)	
Localized / Non-disseminated	32 (47.06%)	58 (31.35%)	
Intracranial Pressure (mmHg), n (%)			0.077
(0, 80)	16 (23.53%)	65 (35.14%)	
[80, 203]	7 (10.29%)	6 (3.24%)	
[204, 272]	11 (16.18%)	27 (14.59%)	
[273, 544]	34 (50.00%)	87 (47.03%)	
CSF Protein (mg/L), Median [IQR]	629.00 [334.50–1489.45]	693.00 [399.00–1100.20]	0.978
CSF Glucose (mmol/L), Median [IQR]	1.94 [1.20–2.84]	2.29 [1.47–3.00]	0.110
CSF Chloride (mmol/L), Mean ± SD	118.18 ± 8.04	114.20 ± 7.79	< 0.001
WBC Count (×10 ⁶ /L), Median [IQR]	12.00 [11.00–103.00]	12.00 [11.00–76.00]	0.403
Time from Symptom Onset to Diagnosis (weeks), Median [IQR]	2.43 [1.43–4.29]	3.00 [1.43–4.29]	0.068

Table 1 footnotes: IQR, interquartile range; CSF, cerebrospinal fluid; WBC, white blood cell. Data are presented as median [IQR], mean ± SD, or n (%). P-values were derived from independent samples t-tests, Mann-Whitney U tests, or chi-square tests, as appropriate.

Unexpectedly, CSF-Cl concentrations were significantly higher in the unfavorable prognosis group than in the favorable group (118.18 vs. 114.20 mmol/L, $P < 0.001$). This pattern contrasts sharply with the hypochloride profile observed in pyogenic meningitis, suggesting the presence of unique ion transport dysregulation in CM. No significant differences were observed in CSF glucose or protein levels between the two groups (all $P > 0.05$).

3.2 Multivariable Logistic Regression Analysis

Variables with $P < 0.1$ in univariate analysis were entered into the multivariable adjusted model. As shown in Table 2, HIV infection was the strongest risk factor for unfavorable prognosis (OR = 5.947, 95% CI: 2.730–12.953, $P < 0.001$); altered mental status (OR = 3.961, 95% CI: 1.826–8.592, $P < 0.001$) and fatigue (OR = 3.626, 95% CI: 1.507–8.728, $P = 0.004$) were also strong independent risk factors. CSF glucose was identified as a protective factor (OR = 0.745, $P = 0.049$).

After adjustment for the above factors, CSF-CI remained independently associated with prognosis (OR = 1.059, 95% CI: 1.011–1.110, $P = 0.017$), with each 1 mmol/L increase in CSF-CI corresponding to a 5.9% elevated risk of unfavorable prognosis. This robust association provides a solid statistical basis for its inclusion in the machine learning feature set.

Table 2
Results of multivariable logistic regression analysis for the prognostic model

Variable	B	OR (95% CI)	P value
HIV Status	1.783	5.947(2.730,12.953)	< 0.001
Level of Consciousness	1.376	3.961(1.826,8.592)	< 0.001
Fatigue	1.288	3.626(1.507,8.728)	0.004
Age	0.028	1.029(1.006,1.052)	0.012
Cerebrospinal Fluid Chloride	0.057	1.059(1.011,1.110)	0.017
Cerebrospinal Fluid Glucose	-0.294	0.745(0.556,0.999)	0.049
Nausea/Vomiting	0.524	1.689(0.858,3.325)	0.129
Site of Infection	0.181	1.199(0.589,2.438)	0.617

Table 2 footnotes: B, regression coefficient; OR, odds ratio; CI, confidence interval.

3.3 Feature Selection and Model Construction

The optimal penalty coefficient λ for LASSO regression in the training set was 0.0203. Combined with RFE to iteratively remove redundant variables, 10 core predictive features were ultimately retained: age, HIV infection status, meningeal enhancement, fever, altered mental status, nausea and vomiting, fatigue, ICP, CSF-CI, and CSF white blood cell count. The LASSO coefficient path showed that the regression coefficient of CSF-CI was not shrunk to zero, further confirming its independent predictive value and justifying its inclusion in the final modeling feature set (Fig. 1).

3.4 Generalization Performance of Nine Machine Learning Models

On the external blind test set, the nine models were ranked by AUC as follows: CatBoost (0.712) > RF (0.707) > ET (0.702) > LightGBM (0.679) > XGBoost (0.642) > LR (0.577). Distance-based models (KNN and SVM) both achieved discrimination performance below 0.55.

CatBoost achieved the highest external test set precision (0.722) among all models, meaning that 72.2% of patients classified as high-risk by the model ultimately experienced unfavorable outcomes. This low false-positive rate for high-risk identification helps avoid excessive medical intervention in low-risk patients. Although the training set showed excellent fit (AUC = 0.995), no abrupt drop in performance was observed between internal validation and external testing, indicating that the ordered boosting structure effectively mitigates overfitting in small samples. Full performance metrics for all models are presented in Table 3.

Table 3
Performance comparison of nine machine learning models

Model	Training AUC	Training Acc	Validation AUC	Validation Acc	External AUC	External Acc	External R	External P	External F1
LR	0.857	0.724	0.673	0.667	0.577	0.610	0.393	0.579	0.468
KNN	1.000	1.000	0.778	0.667	0.531	0.542	0.536	0.441	0.484
SVM	1.000	0.988	0.786	0.846	0.533	0.593	0.393	0.611	0.478
RF	1.000	0.972	0.738	0.821	0.707	0.661	0.500	0.737	0.596
ET	0.931	0.854	0.710	0.718	0.702	0.678	0.536	0.714	0.612
GBM	0.987	0.943	0.754	0.821	0.614	0.576	0.500	0.560	0.528
XGBoost	0.999	0.984	0.766	0.795	0.642	0.559	0.429	0.545	0.480
LightGBM	0.976	0.919	0.790	0.872	0.679	0.627	0.464	0.650	0.542
CatBoost	0.995	0.955	0.770	0.821	0.712	0.661	0.464	0.722	0.565

Table 3 footnotes: AUC, area under the receiver operating characteristic curve; Acc, accuracy; R, recall; P, precision; F1, F1-score.

3.5 Validation of Incremental Predictive Value of CSF Chloride

To quantify the independent contribution of CSF-CI, we constructed two CatBoost models for comparison: one including CSF-CI (Model A) and one excluding CSF-CI (Model B). Results from the external test set (**Fig. 2A**) showed that Model A achieved an AUC of 0.712, higher than the 0.651 of Model B, with an upward trend in AUC ($\Delta\text{AUC} = +0.061$, DeLong's test $P = 0.064$). Model A also outperformed Model B in accuracy (0.661), precision (0.722), and F1-score (0.565). Although DeLong's test did not reach conventional statistical significance ($P = 0.064$), this may be attributed to the limited sample size and event count of the dual-center cohort, and highlights that the practical clinical value of added markers should also be prioritized.

Further evaluation using the gold-standard clinical prediction metrics NRI and IDI showed an NRI of 0.132 (95% CI: 0.011–0.253), indicating that 13.2% of patients were correctly reclassified into appropriate risk strata after adding CSF-CI. Specifically, 17.9% of patients with unfavorable prognosis were correctly up-classified to high risk, and 7.7% of patients with favorable prognosis were correctly down-classified to low risk. The IDI was 0.040, with a 0.061 increase in mean predicted probability for the event group and a 0.021 decrease for the non-event group, confirming that CSF-CI significantly improves the overall discriminative ability of the model.

DCA(**Fig. 2B**) showed that within the clinically common risk threshold range of 0.1–0.6, the net benefit curve of Model A (with CSF-CI) consistently lay above that of Model B (without CSF-CI), as well as above the treat-all and treat-none strategies. For calibration performance (**Fig. 2C**), both models showed good goodness of fit, but Model A had a Brier score of 0.231, superior to 0.265 for Model B, indicating more accurate probability prediction by Model A. Collectively, the inclusion of CSF-CI aligns with the pathophysiological mechanisms of CM, and provides clinically actionable decision support with greater net benefit while maintaining good calibration

Figure 2 Performance comparison of prediction models with and without CSF-CI. (A) Receiver operating characteristic curves with 95% confidence intervals: solid blue line represents Model A (with CSF-CI, AUC = 0.712), dashed red line represents Model B (without CSF-CI, AUC = 0.651), DeLong's test $P = 0.064$. (B) Decision curves: Model A (green line) yields higher clinical net benefit than Model B (red line) and the treat-all strategy across a wide range of thresholds. (C)

Calibration curves: dashed gray line represents ideal calibration; Model A (blue line, Brier score = 0.231) shows better calibration accuracy than Model B (red line, Brier score = 0.265).

3.6 Model Interpretability Analysis (SHAP)

To elucidate the internal decision-making mechanism of the prediction model and verify its clinical plausibility, we performed global and local interpretability analyses based on SHAP values (Fig. 3).

Global feature importance ranking (Fig. 3A) revealed that HIV infection status, ICP, CSF-CI, and altered mental status were the core driving variables influencing model predictions. Elevated feature values (e.g., HIV positivity, increased ICP, altered mental status, corresponding to red scatter points) were significantly associated with positive SHAP values, indicating increased risk of unfavorable prognosis. Conversely, normal or negative feature values (corresponding to blue scatter points) were associated with negative SHAP values, indicating reduced risk.

SHAP dependence plot analysis (Fig. 3B) further revealed a nonlinear association between CSF-CI and prognosis. When standardized CSF-CI was below 0, SHAP values rose rapidly with increasing CSF-CI concentrations, but plateaued after standardized values exceeded 0.5. Meanwhile, color stratification of scatter points indicated a potential interaction effect between CSF-CI and ICP: low ICP status (blue) was predominantly observed in the low CSF-CI range.

Additionally, the feature interaction heatmap based on SHAP interaction values quantified the synergistic effects between features (Fig. 3C). In patients with “low CSF-CI and low ICP”, SHAP interaction values showed a strongly negative contribution (-0.74), reflecting pronounced protective synergy. In contrast, the “high CSF-CI and high ICP” state showed a strongly positive interactive contribution (up to 0.81), indicating additive high risk to patient prognosis when both risk factors coexist. Similarly, the lethal synergistic effect of high CSF-CI was significantly stronger in HIV-positive patients than in HIV-negative patients (interaction value 0.90 vs. 0.76). These findings reinforce the pathophysiological relevance of interactions between key clinical indicators.

3.7 Model Generalizability and Subgroup Robustness

In the external test set ($n = 59$), the CatBoost model achieved the following overall performance: AUC = 0.712, accuracy = 0.661, recall = 0.464, precision = 0.722, F1-score = 0.565. The precision of 0.722 was the highest among the nine models, indicating that over 70% of patients flagged as high-risk by the model indeed experienced unfavorable outcomes.

Subgroup analysis (Table 4) showed favorable model performance in HIV-positive patients (AUC = 0.752), patients aged < 50 years (AUC = 0.759), male patients (AUC = 0.755), and patients with decreased CSF glucose (AUC = 0.722). The AUC in the normal ICP subgroup reached 0.938, but this subgroup included only 16 patients with a sensitivity of 0.250, so results should be interpreted with caution. Calibration errors for all subgroups were within clinically acceptable ranges.

Table 4
Performance metrics of subgroup analysis in the external test set

Stratification Factor	Subgroup	n	Events	Event Rate (%)	AUC	Sensitivity	Specificity	F1 Score	Calibration Error
HIV Status	HIV-negative	10	4	40.00%	0.583	0	0.833	0	0.243
HIV Status	HIV-positive	49	24	49.00%	0.752	0.583	0.760	0.636	0.053
Age	< 50 years	28	11	39.30%	0.759	0.545	0.765	0.571	0.008
Age	≥ 50 years	31	17	54.80%	0.693	0.471	0.786	0.571	0.17
Consciousness Impairment	No impairment	51	20	39.20%	0.700	0.500	0.774	0.541	0.027
Intracranial Pressure (ICP)	Normal ICP	16	8	50.00%	0.938	0.250	1.000	0.400	0.248
Intracranial Pressure (ICP)	Elevated ICP	43	20	46.50%	0.667	0.600	0.696	0.615	0.025
CSF Glucose	Decreased (< 2.5 mmol/L)	34	16	47.10%	0.722	0.500	0.778	0.571	0.108
CSF Glucose	Normal (≥ 2.5 mmol/L)	25	12	48.00%	0.705	0.500	0.769	0.571	0.055
Sex	Male	40	20	50.00%	0.755	0.500	0.800	0.588	0.111

Table 4 footnotes: Data are stratified by clinically relevant subgroups. Calibration error reflects the model's goodness-of-fit within each stratum.

3.8 Dose-Response Relationship and Threshold Analysis of CSF Chloride

To provide a quantifiable early warning threshold for clinical practice, we identified the nonlinear inflection point of CSF-CI for prognostic risk based on the LOESS fitted curve from the SHAP dependence plot (Fig. 3B). Using the mean and standard deviation of CSF-CI in the model training set (115.37 ± 8.19 mmol/L), we calculated via reverse standardization that when CSF-CI reached approximately 115.4 mmol/L, the SHAP value just crossed the zero point, marking the transition of CSF-CI from a normal state to a risk factor for unfavorable prognosis. When CSF-CI further increased to approximately 121.9 mmol/L, the SHAP dependence curve entered a high-level plateau phase, indicating that the risk of unfavorable prognosis had approached its maximum.

These nonlinear thresholds derived from the machine learning model (approximately 115 mmol/L and 122 mmol/L) are highly consistent with the lower bound and mid-high level of the widely accepted normal reference range for CSF chloride (115–130 mmol/L) in domestic and international literature. Accordingly, we propose 115 mmol/L as the high-risk warning threshold for early frontline clinical screening, and 122 mmol/L as the confirmatory threshold for very-high-risk stratification.

Additionally, based on the optimal CatBoost model, we developed an online interactive prognostic prediction tool (**Supplementary Material 2**). By inputting patient parameters including age, HIV status, mental status, ICP, and CSF

chloride, the tool outputs the 10-week unfavorable prognosis probability and risk level, with visualized SHAP waterfall plots to explain the basis of the prediction.

4. Discussion

4.1 Summary of Key Findings

To our knowledge, this is the first dual-center retrospective cohort study to systematically demonstrate and quantify the independent incremental prognostic value of CSF-Cl in CM. Our core findings are fourfold. First, a baseline CSF-Cl level ≥ 115 mmol/L serves as an independent risk factor for 10-week unfavorable prognosis in patients with CM (OR = 1.059), with each 1 mmol/L elevation associated with an approximately 5.9% increase in the risk of death or treatment failure. Second, incorporation of CSF-Cl improved the external AUC of the machine learning model by 0.061, with an NRI of 0.132 and IDI of 0.040, representing a meaningful enhancement in patient risk stratification. Third, CSF-Cl exhibits a dose-response relationship with unfavorable prognosis and exerts synergistic predictive effects with intracranial hypertension and HIV infection. Finally, the CatBoost model demonstrated stable generalizability following independent external validation, and when combined with the SHAP interpretability framework and online prediction tool, shows high feasibility for clinical translation.

4.2 Pathophysiological Hypotheses for Elevated CSF-Cl

A particularly noteworthy finding of this study is that elevated, rather than reduced, baseline CSF-Cl is an independent risk factor for unfavorable CM outcomes. This observation appears paradoxical at first glance given the classic hypochloride profile seen in bacterial meningitis. However, when considered in the context of cryptococcal biology and the unique pathophysiology of CM, several plausible mechanistic hypotheses can be proposed.

First, capsular polysaccharide-mediated impairment of CSF outflow leads to local chloride accumulation. *Cryptococcus neoformans* continuously secretes large quantities of poorly degradable capsular polysaccharides, which deposit in arachnoid granulations and meningeal lymphatic vessels and block CSF reabsorption pathways, creating localized stagnant dead spaces within the CSF compartment. Prior studies have demonstrated that chloride concentrations in ventricular CSF are substantially higher than those in lumbar CSF [18, 19]. When CSF circulation is obstructed, chloride cannot be cleared via normal reabsorption, resulting in accumulative elevation of CSF-Cl measured by lumbar puncture. Our SHAP interaction analysis revealed a strong positive synergistic effect between high CSF-Cl and elevated ICP (interaction value = 0.81). This association likely reflects parallel increases in ICP and CSF-Cl retention driven by a shared underlying cause—high capsular polysaccharide burden. Together, these two markers reflect the severity of CSF circulatory obstruction and thus portend particularly poor prognosis.

Second, dysfunction of choroid plexus NKCC1 transporters may enhance active chloride secretion into the CSF. CSF production depends on $\text{Na}^+/\text{K}^+/\text{2Cl}^-$ cotransporters (NKCC1) expressed on choroid plexus epithelial cells [20]. Chronic meningeal inflammation triggered by cryptococcal infection may upregulate NKCC1 expression via proinflammatory cytokines [21], promoting active chloride secretion into the CSF. When CSF circulation is intact, excess chloride can be cleared via reabsorption; however, when arachnoid granulations are obstructed, chloride accumulates continuously, ultimately manifesting as elevated CSF-Cl. In contrast to bacterial meningitis, where reduced CSF-Cl is driven primarily by passive chloride leakage across a disrupted blood-brain barrier, CM involves a dual mechanism of increased active secretion combined with impaired outflow. This fundamental pathophysiological distinction likely explains the opposing patterns of CSF-Cl changes observed in these two forms of meningitis.

Furthermore, fungal chloride channels may contribute to virulence regulation. *Cryptococcus neoformans* expresses the CLC-type chloride/proton antiporter Clc1, which is involved in virulence-related processes including calcium homeostasis maintenance, cell wall integrity regulation, and high-temperature tolerance [22, 29]. Elevated CSF chloride concentrations

may activate fungal Clc1 channels, enhancing cryptococcal survival and proliferation within the central nervous system microenvironment. This in turn increases fungal burden and capsular polysaccharide release, creating a vicious cycle of elevated CSF-Cl → enhanced fungal virulence → worsened obstruction → further CSF-Cl elevation that accelerates disease progression.

Additionally, HIV-associated immunodeficiency may amplify the prognostic impact of these ion disturbances. SHAP interaction analysis showed a higher interaction value between CSF-Cl and HIV positivity (0.90) than with HIV negativity (0.76). In HIV-infected patients, CD4⁺ T cell depletion impairs effective clearance of cryptococci and capsular polysaccharides, sustaining CSF inflammation and exacerbating ion transport dysregulation [23]. HIV itself can also directly damage choroid plexus epithelial cells [24], further disrupting chloride homeostasis. Consequently, the prognostic warning value of elevated CSF-Cl is amplified in immunocompromised populations, consistent with our finding of a higher AUC in the HIV-positive subgroup (0.752).

Importantly, although these mechanistic hypotheses are supported by indirect evidence from prior studies, our SHAP analysis can only demonstrate statistical associations and cannot establish direct causality. The underlying molecular mechanisms require further validation through targeted basic experiments and prospective pathophysiological investigations.

From a clinical perspective, our rigorous controlled analysis demonstrates that incorporation of CSF-Cl improves net reclassification by 13.2% (NRI = 0.132). This finding carries substantial clinical decision-making relevance: in routine practice, clinicians are typically alert to extremely low CSF-Cl values (e.g., < 110 mmol/L) but often overlook the "gray zone" of 115–120 mmol/L. Our model indicates that adverse prognostic effects begin to emerge once CSF-Cl exceeds 115 mmol/L, which may represent a pathophysiological tipping point from compensatory ion regulation to decompensated circulatory obstruction.

4.3 Model Selection and Interpretability

Among the nine machine learning algorithms compared in this study, CatBoost achieved the best overall generalizability. Its core advantages stem from its ordered boosting mechanism, which effectively mitigates target leakage in small, imbalanced cohorts and thereby reduces overfitting. Additionally, CatBoost offers native support for categorical variables without the need for complex dummy variable transformation, making it well suited to the heterogeneous nature of real-world clinical data.

Most current medical machine learning models operate as "black boxes", outputting only predicted probabilities without explaining how individual features contribute to outcomes. This lack of transparency contributes to low clinician acceptance and hinders real-world implementation. To address this core barrier to clinical translation, we integrated a SHAP-based global and local interpretability framework that clearly delineates the independent contribution of CSF-Cl, its dose-response profile, and its key interactions with other clinical features. This approach not only improves the clinical acceptability of the model but also generates testable statistical hypotheses to guide future mechanistic investigations.

4.4 Study Limitations

Several limitations of this study should be acknowledged when interpreting the results.

First, the retrospective cohort design carries inherent selection bias and cannot fully balance all baseline confounders. Additionally, because data were extracted from existing medical records, potentially prognostic variables such as baseline antifungal drug concentrations, cryptococcal strain genotypes, and immune cytokine profiles were not available, introducing variable omission bias.

Second, the total sample size of 253 patients, with only 59 in the independent external validation set and recruitment limited to two centers, means that the generalizability of the model to broader populations requires confirmation in larger

multicenter prospective cohorts.

Third, our model was built using only a single pre-treatment CSF-CI measurement, without incorporating dynamic changes in CSF-CI during treatment. Therefore, the prognostic value of temporal evolution in electrolyte levels cannot be assessed.

Fourth, the model achieved moderate discriminative performance with an external AUC of 0.712, leaving room for further optimization. Integration of CSF inflammatory cytokine profiles and quantitative neuroimaging markers may improve predictive accuracy in future work.

Lastly, DeLong's test showed that the difference in AUC between models with and without CSF-CI did not reach conventional statistical significance ($P = 0.064$). This is likely attributable to reduced statistical power caused by the limited sample size ($n = 59$) and event count in the external validation set. Nevertheless, clinically oriented metrics including NRI, IDI, and DCA consistently demonstrated clear incremental value, supporting the clinical utility of CSF-CI incorporation.

4.5 Clinical Implications and Future Directions

The nonlinear dose-response pattern and dual-threshold characteristics of CSF-CI identified in our model carry important clinical guidance value.

Clinicians typically focus only on extremely low or high CSF-CI values. Our machine learning model demonstrates that pathophysiological harm begins to emerge once CSF-CI exceeds 115 mmol/L, and that adverse prognostic effects approach saturation above 122 mmol/L. These findings indicate that attention should not be limited to extreme values; patients with CSF-CI in the rising range of 115–122 mmol/L also warrant early intervention and close monitoring. When patients also have HIV infection, altered mental status, or markedly elevated ICP, the prognostic warning value of high CSF-CI is further amplified. In such cases, early initiation of comprehensive interventions—including intensified intracranial pressure reduction, extended induction-phase antifungal therapy, and immunological support—should be considered.

Future studies can advance this work in several directions. First, prospective multicenter cohorts should be conducted to further validate the predictive performance of our model. Second, simultaneous dynamic monitoring of serum and CSF electrolytes will help elucidate the molecular transport mechanisms underlying CSF-CI elevation. Third, integration of metagenomic sequencing and CSF transcriptomic data will enable deeper exploration of regulatory pathways associated with CSF-CI dysregulation. Finally, dynamic time-series prediction models incorporating serial biochemical changes during treatment will enable real-time, longitudinal prognostic assessment throughout the clinical course.

5. Conclusions

CSF-CI is a robust independent prognostic factor for 10-week unfavorable outcomes in CM, delivering demonstrable incremental discriminative value beyond conventional clinical markers for risk stratification (NRI = 0.132; IDI = 0.040). The prognostic effect of CSF-CI exhibits a nonlinear dose-dependent pattern, with pronounced synergistic interactions with ICP and HIV serostatus. Our CatBoost prediction model, independently validated across two clinical centers, demonstrates robust and generalizable performance. When paired with SHAP-enabled interpretability analysis and a dedicated online prediction tool, it forms an integrated translational framework bridging mechanistic biological insight to rapid, point-of-care risk evaluation. We recommend that a baseline CSF-CI threshold of ≥ 115 mmol/L be adopted as a standard early warning criterion for poor CM prognosis, and integrated into routine clinical CSF biochemical assessment workflows.

Declarations

Ethics approval and consent to participate

This study was performed in accordance with the Declaration of Helsinki. Ethical approval was obtained from the Institutional Review Boards of Nanning Fourth People's Hospital and Jiangxi Chest Hospital. All participants provided written informed consent prior to study participation.

Consent for publication

Written informed consent was obtained from participants interviewed in person, and verbal informed consent was obtained and documented for those interviewed by telephone. All participants agreed to the publication of anonymized data.

Availability of data and materials

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

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China (Grant Nos. 82402652 and 82472304).

Authors' contributions

ZQH and YKY contributed to the conception, design, writing, and critical review of the work. LDZ contributed to data collection, processing, analysis, interpretation, and writing. ZXL and LST contributed to supervision and literature review. ZC, WXP and LWQ contributed to data collection, processing, analysis, and interpretation. PWH and ZKM contributed to supervision, funding acquisition, and critical review. LWT contributed to supervision, data collection, processing, analysis, and interpretation. All authors have read and approved the final manuscript to be published and agree to be accountable for all aspects of the work.

Acknowledgements

Not applicable.

References

1. Tugume L, Okirwoth M, Abassi M, et al. Cryptococcal meningitis. *Nat Rev Dis Primers*. 2023;9(1):62.
2. Denning DW. Global incidence and mortality of severe fungal disease. *Lancet Infect Dis*. 2024;24(7):e428–38.
3. Rajasingham R, Govender NP, Boulware DR, et al. The global burden of HIV-associated cryptococcal meningitis in 2020: a modelling analysis. *Lancet Infect Dis*. 2022;22(12):1748–55.
4. Zhang C, et al. The clinic-based predictive modeling for prognosis of patients with cryptococcal meningitis. *BMC Infect Dis*. 2023;23(1):352.
5. Chen Z, et al. Clinical characteristics and cerebro-spinal fluid cytokine changes in patients with acquired immunodeficiency syndrome and central nervous system infection. *Exp Ther Med*. 2019;18(1):523–30.
6. Liu J et al. Clinical characteristics, treatment, and outcome of low-risk non-HIV-associated cryptococcal meningitis: a retrospective cohort study. *Med Mycol*. 2023;61(12).
7. de Oliveira L, et al. Early clinical and microbiological predictors of outcome in hospitalized patients with cryptococcal meningitis. *BMC Infect Dis*. 2022;22(1):138.

8. Liang Y, et al. Analysis of clinical characteristics and prognostic factors in AIDS complicated by cryptococcal meningitis. *Eur J Clin Microbiol Infect Dis*. 2026;45(1):261–9.
9. Zheng G, et al. Development and verification of a discriminate algorithm for diagnosing post-neurosurgical bacterial meningitis: a multicenter observational study. *J Clin Lab Anal*. 2020;34(2):e23069.
10. He X, et al. Immune cell and biochemical marker levels in the cerebrospinal fluid and blood of patients with brucella meningitis and their clinical significance. *Am J Transl Res*. 2024;16(9):5097–105.
11. Hickman PE, et al. Use of chloride gradient for rapid identification of CSF leaks. *Pathology*. 2024;56(4):579–81.
12. Fries BC, et al. Phenotypic switching of *Cryptococcus neoformans* can produce variants that elicit increased intracranial pressure in a rat model of cryptococcal meningoencephalitis. *Infect Immun*. 2005;73(3):1779–87.
13. Huang Y et al. Prediction of hospital discharge outcome from changes in cerebrospinal fluid/serum albumin quotient and cerebrospinal fluid lactate dehydrogenase in patients with cryptococcal meningitis. *Future Microbiol*. 2021.
14. Qin J, et al. A retrospective analysis of central and peripheral metabolic characteristics in patients with cryptococcal meningitis. *Neurol Ther*. 2024;13(3):763–84.
15. Jiang YK, et al. Cerebrospinal fluid cytokine and chemokine patterns correlate with prognosis of HIV-uninfected cryptococcal meningitis: a prospective observational study. *Front Immunol*. 2022;13:993495.
16. Okwir M, et al. High burden of cryptococcal meningitis among antiretroviral therapy-experienced human immunodeficiency virus-infected patients in northern Uganda in the era of test and treat: implications for cryptococcal screening programs. *Open Forum Infect Dis*. 2022;9(2):ofac004.
17. Kelly CJ, Karthikesalingam A, Suleyman M, et al. Key challenges for delivering clinical impact with artificial intelligence. *BMC Med*. 2019;17(1):195.
18. Dong Q, Huang Z, Yu P, Song E, Chen Z, Qin F. Ventricular and lumbar cerebrospinal fluid analysis in 77 HIV-negative patients with cryptococcal meningitis who received a ventriculoperitoneal shunt. *Sci Rep*. 2022;12(1):21366.
19. Van Wieren A, Casadevall A. Pathophysiology of intracranial hypertension in cryptococcal meningoencephalitis. *mBio*. 2026;17(5):e0031226.
20. Yang X, Jin L, Shi C, Yue H, He H, Zhu W, et al. The clinical characteristics and treatment of patients with autoimmune glial fibrillary acidic protein astrocytopathy (GFAP-A): a retrospective study of 29 patients. *Brain Behav*. 2026;16(1):e71159.
21. Xu H, He J, Du H, Jing X, Liu X. Evaluation of the choroid plexus epithelium inflammation TLR4/NF- κ B/NKCC1 signal pathway activation in the development of hydrocephalus. *CNS Neurosci Ther*. 2024;30(10):e70085.
22. Li D, Zhang X, Li Z, Yang J, Pan J, Zhu X. *Cryptococcus neoformans* Ca(2+) homeostasis requires a chloride channel/antiporter Clc1 in JEC21, but not in H99. *FEMS Yeast Res*. 2012;12(1):69–77.
23. Meya DB, et al. The immunopathogenesis of cryptococcal immune reconstitution inflammatory syndrome. *Curr Opin Infect Dis*. 2016;29(1):10–22.
24. Okurut S, et al. Landmark clinical observations and immunopathogenesis pathways linked to HIV and *Cryptococcus* fatal central nervous system co-infection. *Mycoses*. 2020;63(8):840–53.
25. Xu L, Chen Z, Liu J. The impact of comorbidities on clinical efficacy of cryptococcal meningitis treatment. *Zhonghua Yi Yuan Gan Ran Xue Za Zhi [Chinese J Nosocomiology]*. 2019;29.
26. Liu J, et al. Development and validation of a machine learning model to predict prognosis in HIV-negative cryptococcal meningitis patients: a multicenter study. *Eur J Clin Microbiol Infect Dis*. 2023;42(10):1183–94.
27. Takefuji Y. Mitigating biases in feature selection and importance assessments in predictive models using LASSO regression. *Oral Oncol*. 2024;159:107090.
28. Mustafa S, et al. Tomek links-based SMOTE method for class imbalance in blood cell classification with dual path sliding window attention model. *Talanta*. 2026;303:129473.

29. Idnurm A, Reedy JL, Nussbaum JC, Heitman J. *Cryptococcus neoformans* virulence gene discovery through insertional mutagenesis. Eukaryot Cell. 2004;3(2):420–9.

Figures

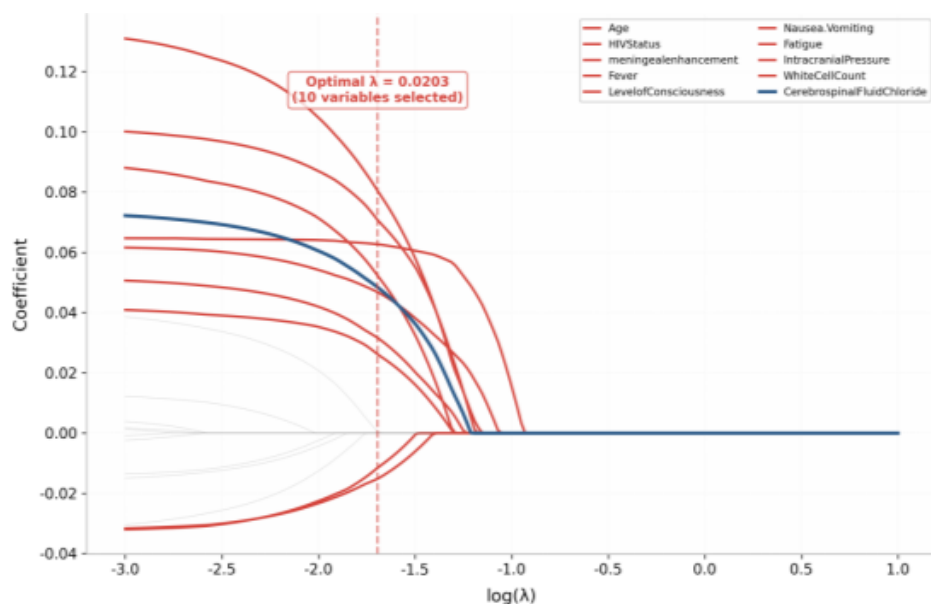


Figure 1

LASSO regression feature selection path. The x-axis represents the log-transformed penalty parameter $\log(\lambda)$, and the y-axis represents the regression coefficient of each variable. As $\log(\lambda)$ increases, coefficients of non-informative variables are shrunk to zero (gray lines). The red vertical dashed line indicates the optimal regularization parameter ($\lambda = 0.0203$) selected via 10-fold cross-validation, at which point 10 clinically informative variables are retained in the model.

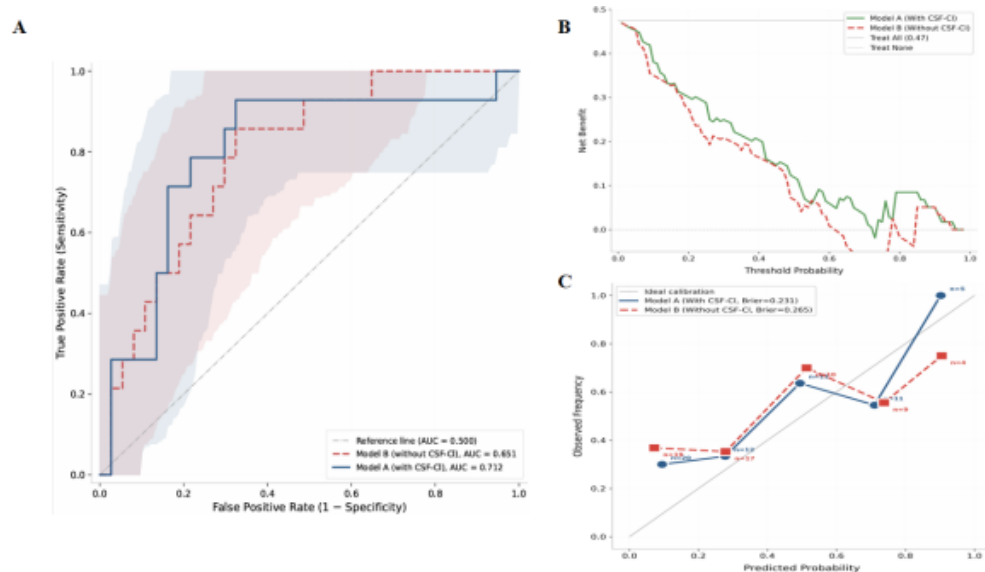


Figure 2

Performance comparison of prediction models with and without CSF-CI. (A) Receiver operating characteristic curves with 95% confidence intervals: solid blue line represents Model A (with CSF-CI, AUC = 0.712), dashed red line represents Model B (without CSF-CI, AUC = 0.651), DeLong's test $P = 0.064$. (B) Decision curves: Model A (green line) yields higher clinical net

benefit than Model B (red line) and the treat-all strategy across a wide range of thresholds. (C) Calibration curves: dashed gray line represents ideal calibration; Model A (blue line, Brier score = 0.231) shows better calibration accuracy than Model B (red line, Brier score = 0.265).

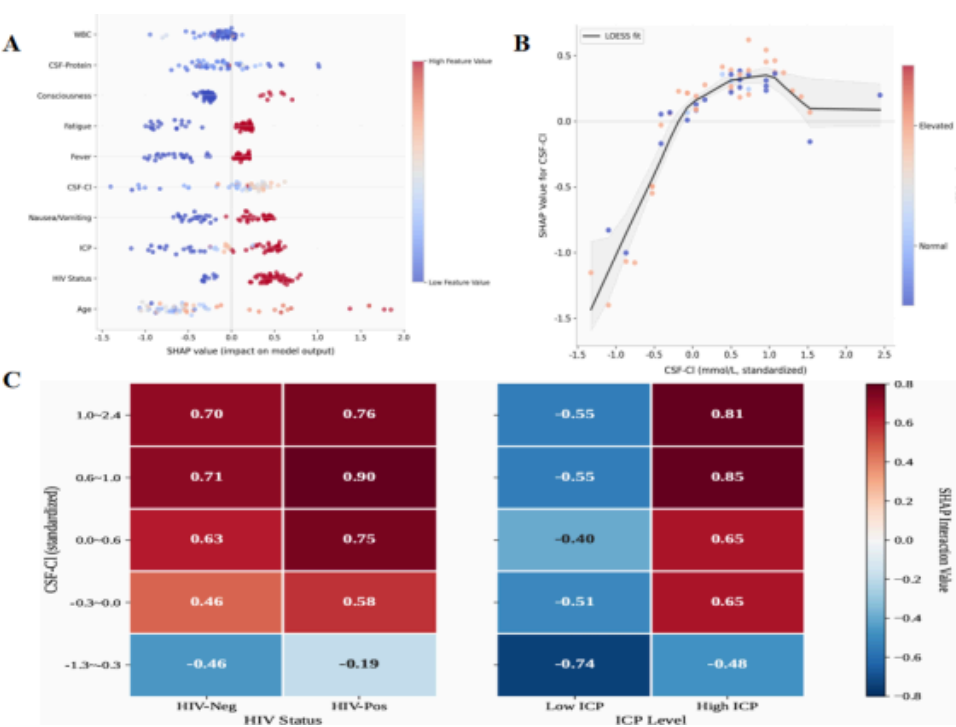


Figure 3

SHAP-based interpretability analysis of the CM prognostic prediction model. (A) Global feature contribution summary plot: the y-axis lists clinical features ranked by importance; the x-axis represents SHAP values (positive values indicate increased risk of unfavorable prognosis; negative values indicate reduced risk). Color gradient from blue to red represents increasing feature values. (B) SHAP dependence plot for CSF-CI: shows the nonlinear relationship between standardized CSF-CI values and SHAP values. The black solid line is the LOESS fitted trend line (gray area represents 95% confidence interval), and scatter points are colored by ICP level. (C) Feature interaction heatmaps: left panel shows the interaction between CSF-CI and HIV status; right panel shows the interaction between CSF-CI and ICP. Values within grids represent SHAP interaction values (positive/negative values indicate additional effects of synergistically increasing or decreasing risk, respectively).

Supplementary Files

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

- [SupplementaryMaterial.docx](#)