Am J Transl Res 2026;18(5):3816-3827
www.ajtr.org /ISSN:1943-8141/AJTR0169783

Original Article

Prediction of hospital length of stay in decompensated cirrhosis using hematologic and inflammatory indices

Longjiao Qiao[1*], Yanqiu Zhang[2*], Li Xiao[1*], Shixiu Ma[1], Mingyan Li[1], Yuebin Luo[1], Xinhui Xu[1], Ling Zhao[1]

[1]Department of Infectious Diseases, Funan County People's Hospital, Northeast Intersection of Wanghua Road and Longquan Road, Lucheng Town, Funan County, Fuyang 236300, Anhui, China; [2]School of Pharmacy, Anhui Medical University; Key Laboratory of Anti-inflammatory and Immune Medicine, Ministry of Education; Institute of Clinical Pharmacology, Anhui Medical University, Hefei 230032, Anhui, China. [*]Equal contributors.

Received November 4, 2025; Accepted March 25, 2026; Epub May 15, 2026; Published May 30, 2026

Abstract: Objective: To identify predictors of prolonged hospital length of stay (LOS) in decompensated cirrhosis (DC) patients and to develop a prediction model based on hematologic and inflammatory indices. Methods: We retrospectively studied 338 hospitalized patients with DC. Admission blood indices reflecting hematologic status (hemoglobin, hematocrit, red blood cell (RBC) count, platelets, mean corpuscular volume, red cell distribution width [RDW]) and inflammation (monocyte-to-lymphocyte ratio [MLR], neutrophil-monocyte-lymphocyte ratio [NMLR], systemic inflammation response index [SIRI]) were compared between patients with LOS ≥ 7 days vs < 7 days. Multivariable logistic regression and machine learning were used to identify independent predictors of LOS ≥ 7 days, with model performance assessed by area under the receiver operating characteristic curve (AUC) on a held-out test set. Results: Of 338 patients (median age 58 years, 33% female), 52% had LOS ≥ 7 days. Prolonged LOS was associated with more severe cytopenias (lower hemoglobin, hematocrit, RBC count, platelets) and heightened inflammation (higher RDW and inflammatory ratios; all $P < 0.05$). Nonlinear thresholds were observed. An admission RBC count < 5.2×10^{12}/L and hemoglobin > 117 g/L were linked to reduced risk of prolonged stay. A predictive model using these 11 markers achieved moderate discrimination (test AUC 0.72; sensitivity 72%, specificity 58%). Conclusions: Admission hematologic and inflammatory markers are independent predictors of prolonged LOS in DC. A machine learning model derived from routinely available low-cost tests provided moderate risk stratification for identifying patients at increased risk and supporting early clinical management.

Keywords: Decompensated cirrhosis, hematologic, inflammatory, hospital length

Introduction

Cirrhosis is a major cause of morbidity and mortality worldwide and it causes about approximately one million deaths each year [1-4]. In the United States it is the eighth leading cause of death, with a heavy burden on the healthcare system, high hospitalization rates and rising medical costs [5]. When cirrhosis progresses to the decompensated stage patients often develop ascites, variceal hemorrhage, hepatic encephalopathy, and hepatorenal syndrome [6-8]. Patients will experience frequent hospitalizations, a high short-term mortality rate, and frequent early readmissions [9]. Decompensated cirrhosis (DC) is estimated to induce billions of dollars in annual healthcare costs (approximately $4.5 billion in the US), as its treatment requires long-term and resource-intensive inpatient care [9]. For these patients with severe liver disease, effective management is complex and often requires multidisciplinary collaboration to address complications in multiple organ systems simultaneously [5].

Hospital length of stay (LOS) is an important intermediate outcome in DC [2]. It reflects disease severity and the efficiency of inpatient management. Prolonged LOS indicates a more complicated clinical course such as refractory ascites or poorly controlled infections. It also increases the use of healthcare resources and raises medical costs. LOS is clinically important, and it has received relatively less atten-

https://doi.org/10.62347/IWMZ6721

tion in cirrhosis research than endpoints such as mortality or 30-day readmission. Most prognostic studies on DC have focused on predicting death or early rehospitalization. Existing risk models, including MELD-Na-based scores, often have limited accuracy in capturing the full complexity of the inpatient course [9]. The factors influencing LOS in DC patients are multifactorial and not fully elucidated. These include the extent of liver and kidney dysfunction, number and severity of complications, co-existing infections, nutritional status, and even system-level practices. For example, patients presenting with acute hepatic encephalopathy or other severe complications typically have significantly longer hospital stays [9, 10]. There is also growing evidence that laboratory markers of hematologic and inflammatory derangements carry prognostic value in advanced liver disease. More than 70% of patients hospitalized with acute decompensation are anemic. Severe anemia is associated with significantly higher short-term mortality [11]. Similarly, elevated inflammation-based indices (such as neutrophil-to-lymphocyte ratio [NLR] or related composite scores) are associated with more advanced cirrhosis features and poorer outcomes [12]. These observations suggest that a patient's hematologic and immune-inflammatory status at admission might correspond to their ability to recover quickly or, conversely, indicating potentially prolonged hospital stays.

More systematic research of LOS in DC is needed. Utilizing modern data-driven analytical methods can help distinguish the impact of various clinical and laboratory factors on LOS. These methods can reveal nonlinear relationships or threshold effects that traditional linear models may miss. This study analyzed a large cohort of hospitalized DC patients to identify key predictors of prolonged LOS. We focused on readily available hematologic and inflammatory parameters (reflecting anemia severity, erythrocyte morphology, and systemic inflammation) as candidate risk factors. This selection was based on their biological relevance to advanced cirrhosis and preliminary associations with outcomes noted in prior studies. We developed and validated an interpretable prediction model for prolonged hospitalization in DC. Early identification of high-risk patients may help clinicians plan treatment and allocate medical resources more effectively. It may also support targeted management such as infection control or nutritional support and improve overall care efficiency. The overall study design is shown in **Figure 1**.

Methods

Study population

This retrospective cohort study included 338 adult patients with DC who were admitted to a tertiary hospital between January 2022 and December 2024. The Medicine Ethics Committee of Funan County People's Hospital reviewed and approved this research (FNLL2025-1027654). All patient information was anonymized in accordance with the Declaration of Helsinki. The diagnosis of DC was established based on clinical, laboratory, and imaging criteria, including the presence of at least one major complication such as ascites, variceal bleeding, hepatic encephalopathy, or hepatorenal syndrome. Patients under 18 years of age, with acute liver failure, hematologic malignancies, active gastrointestinal bleeding at admission, or incomplete clinical data were excluded. All patient demographic, laboratory, and hospitalization information were collected from the hospital electronic medical record system.

Definition of outcomes

The primary outcome was hospital LOS defined as the days from admission to discharge. Patients were divided into the short-stay (< 7 days) and the prolonged-stay (≥ 7 days) based on the median LOS of the cohort. This threshold was chosen to reflect significant clinical differences in hospital burden and to be consistent with the previous research results on the care of hospitalized liver cirrhosis.

Laboratory measurements

Routine hematologic and biochemical parameters were tested within 24 hours of admission using standardized automated analyzers in the hospital's central laboratory. The hematologic indices including hematocrit (HCT), hemoglobin (Hb), red blood cell (RBC) count, mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), red cell distribution width (RDW, reported as coefficient of variation [RDW-CV] and standard deviation [RDW-SD]), and platelet (PLT) count.

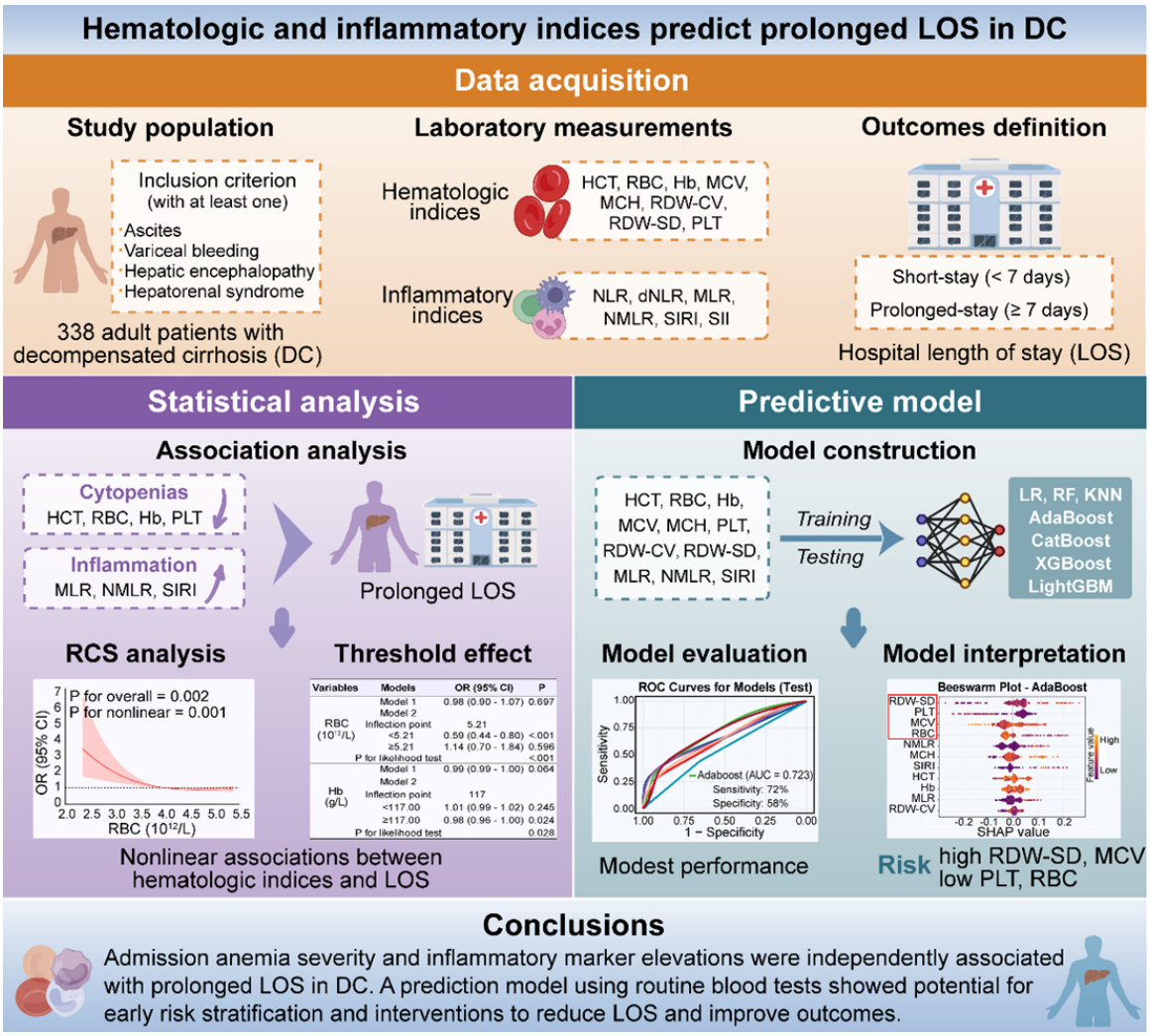

Figure 1. Flowchart of the study. This study investigates factors influencing hospital length of stay (LOS) in patients with decompensated cirrhosis (DC), focusing on admission hematologic and inflammatory markers. Data from 338 adult hospitalized DC patients were analyzed to identify predictors of prolonged hospitalization (≥ 7 days). Severe anemia with lower HCT, RBC, Hb and increased inflammation reflected by higher MLR were independently associated with longer LOS. Nonlinear associations were observed for RBC count and Hb level. A predictive model based on 11 blood markers showed modest performance (AUC 0.72) to identify high-risk patients. This study suggests that these blood indices can help clinicians stratify risk early and guide targeted interventions to reduce hospitalization duration and improve outcomes. HCT, hematocrit; RBC, red blood cell; Hb, hemoglobin; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; RDW-CV, red cell distribution width-coefficient of variation; RDW-SD, red cell distribution width-standard deviation; PLT, platelet; NLR, neutrophil-to-lymphocyte ratio; dNLR, derived neutrophil-to-lymphocyte ratio; MLR, monocyte-to-lymphocyte ratio; NMLR, neutrophil-to-monocyte-to-lymphocyte ratio; SIRI, systemic inflammation response index; SII, systemic immune-inflammation index.

The following comprehensive indicators based on inflammation were derived: monocyte-to-lymphocyte ratio (MLR), neutrophil-monocyte-lymphocyte ratio (NMLR), and systemic inflammation response index (SIRI = neutrophil count × monocyte count/lymphocyte count). These indices were chosen to reflect erythrocyte morphology and systemic immune activation.

Statistical analysis

Sample size was estimated using the empirical rule of at least 10 events per predictor variable (EPV) in logistic regression. A total of 175 events with LOS ≥ 7 days and 11 candidate predictors produced an EPV of 15.9, which supported the stability of the multivariable analy-

sis. All laboratory indices had a missing rate below 2%. Missing values were replaced by the median of each variable.

Distribution of continuous variables was tested using the Kolmogorov-Smirnov test. Continuous data are presented as median and interquartile range [IQR]. Categorical data are reported as counts (percentages). Group differences were analyzed by the Mann-Whitney U test for continuous data and the χ^2 test for categorical data.

Univariate analyses identified hematologic variables associated with LOS at $P < 0.05$. Clinically relevant variables were included in multivariable logistic regression to determine independent predictors of prolonged hospitalization. Restricted cubic spline and segmented regression analyses were used to assess nonlinear threshold effects of Hb and RBC count on LOS [13, 14].

Predictive models were developed and internally validated using 5-fold cross-validation. Model performance was evaluated by the area under the receiver operating characteristic (ROC) curve (AUC), accuracy, sensitivity, specificity, and F1 score. Shapley additive explanation (SHAP) analysis quantified the contribution and direction of each variable and improved model interpretation.

All analyses were conducted using R software (version 4.4.1). A two-sided *P* value < 0.05 was considered statistically significant.

Results

Hematologic and inflammatory profiles associated with length of hospital stay

Among 338 patients with DC, 163 (48.2%) had a hospital stay of less than 7 days, and 175 (51.8%) remained hospitalized for more than 7 days. The median age of the overall cohort was 58.0 years (IQR, 53.0-67.5), and 32.8% were female. No significant differences were observed in age, sex distribution, body mass index (BMI), white blood cells (WBC), monocytes (MONO), neutrophils (NEUT) or lymphocytes (LYMPH) between the two groups (all $P > 0.05$) (**Table 1**).

Patients with longer hospitalization showed significantly lower HCT (median, 0.35 [IQR, 0.31-0.39] vs 0.38 [0.32-0.42]; P = 0.003), RBC count (3.78 [3.06-4.34] vs 4.16 [3.50-4.60]; P < 0.001), Hb level (118.0 [102.5-135.0] vs 127.0 [104.0-142.0]; P = 0.020), and PLT count (72.0 [50.0-109.5] vs 87.0 [56.0-139.8]; P = 0.015). In contrast, patients hospitalized for more than 7 days exhibited significantly higher red cell distribution width (RDW-CV: 0.14 [0.13-0.16] vs 0.14 [0.13-0.16]; P = 0.006; RDW-SD: 50.4 [46.1-55.5] vs 47.8 [44.5-51.9]; P < 0.001) and MCV (94.1 [88.3-98.1] vs 91.8 [86.4-97.2]; P = 0.035) (**Table 1**).

Inflammation-based indices also differed between groups. The MLR (0.35 [0.23-0.51] vs 0.31 [0.21-0.41]; P = 0.008), NMLR (3.08 [2.16-4.95] vs 2.77 [1.99-3.88]; P = 0.047), and SIRI (0.78 [0.44-1.47] vs 0.65 [0.39-1.06]; P = 0.021) were all significantly elevated among patients with prolonged hospitalization. No significant differences were found for NLR, systemic immune-inflammation index (SII). Collectively, anemia, thrombocytopenia, and heightened systemic inflammatory responses were correlated with longer hospital stays in patients with DC (**Table 1**).

Threshold effects of red blood cell count and hemoglobin on length of hospital stay

Restricted cubic spline and segmented regression analyses revealed nonlinear associations between hematologic parameters and hospitalization duration among patients with DC (**Figure 2A**, **2B**).

For RBC count, a significant threshold effect was observed (P for likelihood test < 0.001). Below the inflection point of 5.21×10^{12}/L, lower RBC count was significantly associated with a higher likelihood of prolonged hospitalization (odds ratio [OR], 0.59; 95% confidence interval [CI], 0.44-0.80; P < 0.001). When RBC count exceeded 5.21×10^{12}/L, the association was no longer significant (OR, 1.14; 95% CI, 0.70-1.84; P = 0.596) (**Table 2**).

For Hb concentration, a nonlinear relationship was identified (P for likelihood test = 0.028). Hb levels below 117 g/L showed no significant association with hospital stay (OR, 1.01; 95% CI, 0.99-1.02; P = 0.245). While Hb concentrations above 117 g/L were inversely correlated with prolonged hospitalization (OR, 0.98; 95% CI, 0.96-1.00; P = 0.024) (**Table 2**).

Table 1. Baseline hematologic and inflammatory characteristics of patients with decompensated cirrhosis according to length of hospital stay

Variables	Total (n = 338)	LOS < 7 days (n = 163)	LOS ≥ 7 days (n = 175)	Statistic	P
Age, years, M (Q_1, Q_3)	58.00 (53.00, 67.50)	56.50 (51.75, 67.00)	59.00 (53.50, 69.00)	Z = -1.07	0.286
BMI, kg/m^2, M (Q_1, Q_3)	24.22 (21.88, 26.87)	24.61 (21.96, 27.12)	24.11 (21.86, 26.37)	Z = -1.22	0.224
Sex, n (%)				χ^2 = 0.02	0.902
Male	227 (67.16)	110 (67.48)	117 (66.86)		
Female	111 (32.84)	53 (32.52)	58 (33.14)		
WBC, 10^9/L, M (Q_1, Q_3)	3.83 (2.80, 5.21)	3.77 (2.79, 4.86)	3.83 (2.81, 5.50)	Z = -1.06	0.288
MONO, 10^9/L, M (Q_1, Q_3)	0.28 (0.19, 0.42)	0.27 (0.18, 0.41)	0.30 (0.20, 0.43)	Z = -1.48	0.140
MONO, % (decimal), M (Q_1, Q_3)	0.07 (0.06, 0.10)	0.07 (0.06, 0.09)	0.08 (0.06, 0.10)	Z = -1.34	0.180
NEUT, 10^9/L, M (Q_1, Q_3)	2.29 (1.59, 3.35)	2.23 (1.61, 3.21)	2.46 (1.58, 3.54)	Z = -1.12	0.262
NEUT, % (decimal), M (Q_1, Q_3)	0.63 (0.56, 0.72)	0.62 (0.56, 0.71)	0.66 (0.55, 0.73)	Z = -1.30	0.194
LYMPH, 10^9/L, M (Q_1, Q_3)	0.89 (0.57, 1.43)	0.95 (0.58, 1.50)	0.83 (0.52, 1.27)	Z = -1.52	0.128
LYMPH, % (decimal), M (Q_1, Q_3)	0.24 (0.17, 0.32)	0.26 (0.20, 0.32)	0.24 (0.16, 0.31)	Z = -1.88	0.060
HCT, L/L, M (Q_1, Q_3)	0.36 (0.31, 0.41)	0.38 (0.32, 0.42)	0.35 (0.31, 0.39)	Z = -2.97	0.003
RBC, 10^{12}/L, M (Q_1, Q_3)	3.98 (3.35, 4.47)	4.16 (3.50, 4.60)	3.78 (3.06, 4.34)	Z = -3.66	< 0.001
Hb, g/L, M (Q_1, Q_3)	121.00 (103.00, 138.00)	127.00 (104.00, 142.00)	118.00 (102.50, 135.00)	Z = -2.33	0.020
PLT, 10^9/L, M (Q_1, Q_3)	79.00 (52.00, 125.00)	87.00 (56.00, 139.75)	72.00 (50.00, 109.50)	Z = -2.43	0.015
MCV, fL, M (Q_1, Q_3)	92.75 (87.43, 97.57)	91.80 (86.35, 97.15)	94.10 (88.30, 98.10)	Z = -2.11	0.035
MCH, pg, M (Q_1, Q_3)	31.20 (29.13, 33.18)	30.90 (28.80, 32.55)	31.50 (29.45, 33.45)	Z = -2.12	0.034
MCHC, g/L, M (Q_1, Q_3)	333.00 (321.25, 346.00)	332.00 (319.50, 343.00)	334.00 (323.00, 347.00)	Z = -1.60	0.109
RDW-CV, % (decimal), M (Q_1, Q_3)	0.14 (0.13, 0.16)	0.14 (0.13, 0.16)	0.14 (0.13, 0.16)	Z = -2.76	0.006
RDW-SD, fL, M (Q_1, Q_3)	49.20 (45.50, 53.98)	47.80 (44.45, 51.90)	50.40 (46.05, 55.45)	Z = -3.88	< 0.001
NLR, M (Q_1, Q_3)	2.62 (1.78, 4.07)	2.44 (1.71, 3.46)	2.75 (1.84, 4.44)	Z = -1.79	0.073
dNLR, M (Q_1, Q_3)	0.87 (0.81, 0.89)	0.87 (0.82, 0.89)	0.86 (0.80, 0.89)	Z = -0.17	0.866
MLR, M (Q_1, Q_3)	0.33 (0.22, 0.47)	0.31 (0.21, 0.41)	0.35 (0.23, 0.51)	Z = -2.63	0.008
NMLR, M (Q_1, Q_3)	2.93 (2.03, 4.57)	2.77 (1.99, 3.88)	3.08 (2.16, 4.95)	Z = -1.99	0.047
SIRI, M (Q_1, Q_3)	0.70 (0.40, 1.26)	0.65 (0.39, 1.06)	0.78 (0.44, 1.47)	Z = -2.32	0.021
SII, M (Q_1, Q_3)	203.23 (117.90, 365.50)	197.18 (127.59, 368.40)	207.94 (112.07, 360.37)	Z = -0.28	0.779

Abbreviations: M, median; Q_1, first quartile; Q_3, third quartile; Z, Mann-Whitney U test statistic; χ^2, chi-square test; LOS, length of stay; BMI, body mass index; WBC, white blood cell; MONO, monocyte; NEUT, neutrophil; LYMPH, lymphocyte; HCT, hematocrit; RBC, red blood cell; Hb, hemoglobin; PLT, platelet; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RDW-CV, red cell distribution width-coefficient of variation; RDW-SD, red cell distribution width-standard deviation; NLR, neutrophil-to-lymphocyte ratio; dNLR, derived neutrophil-to-lymphocyte ratio; MLR, monocyte-to-lymphocyte ratio; NMLR, neutrophil-to-monocyte-to-lymphocyte ratio; SIRI, systemic inflammation response index; SII, systemic immune-inflammation index.

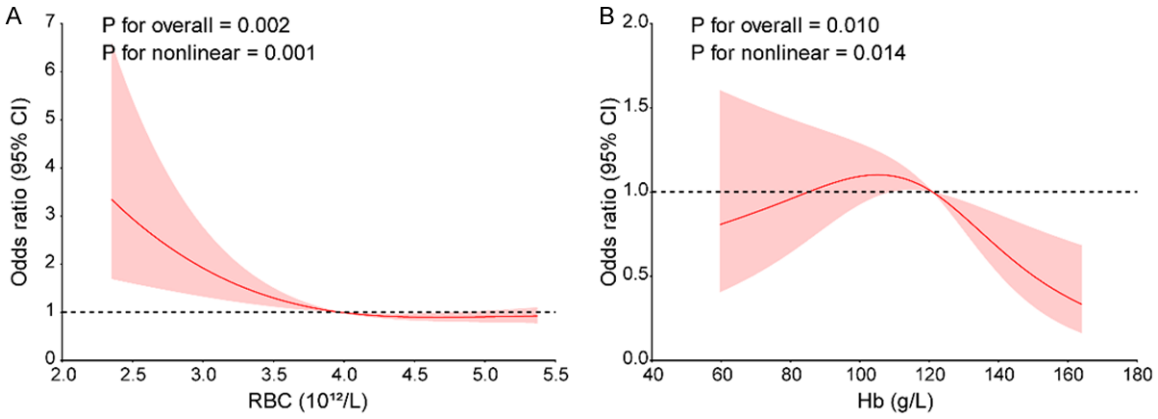

Figure 2. The nonlinear associations between hematologic parameters and hospitalization duration among patients with DC. Restricted cubic spline analyses of the associations of (A) RBC count and (B) Hb level with hospitalization duration in patients with DC.

Table 2. Threshold effects of red blood cell count and hemoglobin on length of hospital stay

Variables	Models	OR (95% CI)	P
RBC (10^{12}/L)	Model 1 Fitting model by standard linear regression	0.98 (0.90-1.07)	0.697
	Model 2 Fitting model by two-piecewise linear regression		
	Inflection point	5.21	
	< 5.21	0.59 (0.44-0.80)	< 0.001
	≥ 5.21	1.14 (0.70-1.84)	0.596
	P for likelihood test		< 0.001
Hb (g/L)	Model 1 Fitting model by standard linear regression	0.99 (0.99-1.00)	0.064
	Model 2 Fitting model by two-piecewise linear regression		
	Inflection point	117.00	
	< 117.00	1.01 (0.99-1.02)	0.245
	≥ 117.00	0.98 (0.96-1.00)	0.024
	P for likelihood test		0.028

Abbreviations: OR, odds ratio; CI, confidence interval; RBC, red blood cell; Hb, hemoglobin.

Anemia severity showed a nonlinear threshold effect on hospitalization risk. Hb levels above approximately 117 g/L and RBC counts around 5.2 × 10^{12}/L may be associated with a shorter hospital stay in patients with DC.

Model performance and validation

A multivariable machine learning model was constructed to predict prolonged hospitalization (≥ 7 days) in patients with DC. Eleven hematologic and inflammatory parameters were selected as input features based on their univariate associations and clinical interpretability: HCT, RBC, Hb, PLT, MCV, MCH, RDW-CV, RDW-SD, MLR, NMLR, and SIRI. These indicators reflect anemia severity, erythrocyte morphology, and systemic inflammatory activation. These key factors influence hepatic decompensation and clinical recovery.

Seven supervised learning algorithms including logistic regression, random forest, adaptive boosting (AdaBoost), k-nearest neighbor (KNN), extreme gradient boosting (XGBoost), light gradient boosting machine (LightGBM), and categorical boosting (CatBoost) were developed. Model performance was evaluated using 5-fold cross-validation in the training set and an independent testing set.

AdaBoost achieved an AUC of 0.872 in the training set (**Figure 3A**). The confusion matrix showed the best generalization performance in the testing set with an accuracy of 0.65, F1 score of 0.67, recall of 0.72, and AUC of 0.723. Sensitivity was 0.72 and specificity was 0.58. The model showed good discriminative ability without clear overfitting (**Figure 3A-C**; **Table 3**). Calibration analysis showed good agreement between predicted probability and observed events and the Hosmer-Lemeshow test was not significant (P > 0.05), indicating no significantly poor fit. CatBoost (AUC = 0.720) and logistic regression (AUC = 0.674) showed comparable but slightly lower predictive accuracy. Random forest and KNN exhibited overfitting, with perfect training results but substantially reduced test AUCs (0.64 and 0.55, respectively) (**Figure 3A**, **3B**; **Table 3**).

Hematologic and inflammatory parameters captured important biological differences among DC patients. Variables related to erythrocyte indices and immune activation showed stable predictive value. AdaBoost model showed the most balanced and interpretable prediction performance and may support early risk assessment and clinical decision making in advanced liver disease.

Feature contribution and model interpretation

Model interpretation was assessed using SHAP analysis to measure the contribution of each predictor in AdaBoost model. Eleven hematologic and inflammatory variables were included: HCT, RBC, Hb, PLT, MCV, MCH, RDW-CV, RDW-SD, MLR, NMLR, and SIRI.

The global SHAP summary plot showed that RDW-SD, PLT, MCV, and RBC were the most

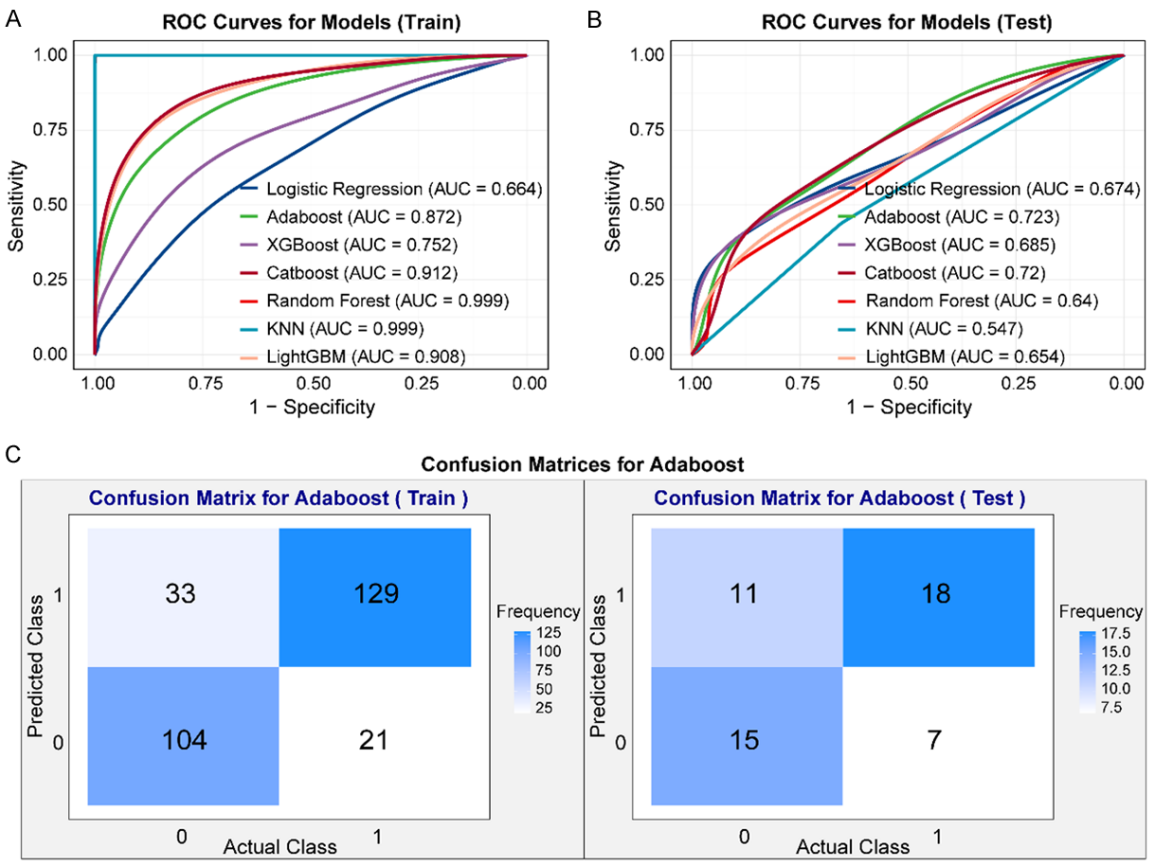

Figure 3. Machine learning algorithms for predicting the risk of prolonged hospitalization in DC. A. The ROC curves of the prolonged hospitalization prediction models in the training set. B. The ROC curves of the prolonged hospitalization prediction models in the testing set. C. Confusion matrix analysis of AdaBoost prediction model for prolonged hospitalization in DC.

Table 3. Performance of machine learning models for predicting prolonged hospitalization in patients with decompensated cirrhosis

Model	Dataset	AUC	Accuracy	Recall	F1 score	Sensitivity	Specificity
Logistic Regression	Train	0.664	0.634	0.560	0.615	0.560	0.715
	Test	0.674	0.647	0.560	0.609	0.560	0.731
Random Forest	Train	0.999	0.999	0.999	0.999	0.999	0.999
	Test	0.640	0.569	0.440	0.500	0.440	0.692
Adaboost	Train	0.872	0.812	0.860	0.827	0.860	0.759
	Test	0.723	0.647	0.720	0.667	0.720	0.577
KNN	Train	0.999	0.999	0.999	0.999	0.999	0.999
	Test	0.547	0.549	0.440	0.489	0.440	0.654
XGBoost	Train	0.752	0.707	0.653	0.700	0.653	0.766
	Test	0.685	0.627	0.520	0.578	0.520	0.731
LightGBM	Train	0.908	0.836	0.793	0.835	0.793	0.883
	Test	0.654	0.627	0.480	0.558	0.480	0.769
Catboost	Train	0.912	0.847	0.833	0.850	0.833	0.861
	Test	0.720	0.647	0.520	0.591	0.520	0.769

Abbreviations: AdaBoost, adaptive boosting; KNN, k-nearest neighbor; XGBoost, extreme gradient boosting; LightGBM, light gradient boosting machine; CatBoost, categorical boosting.

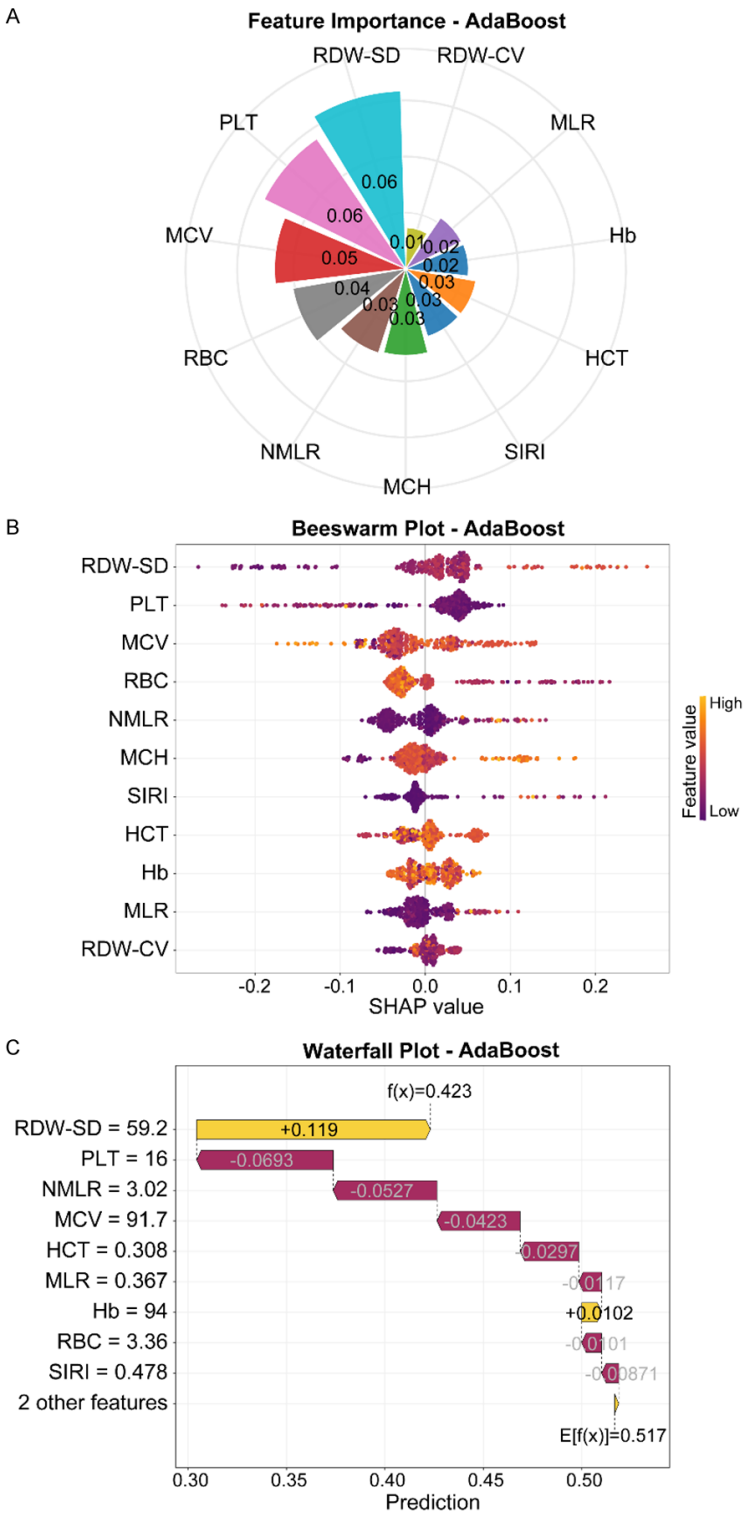

Figure 4. Feature importance for predicting prolonged hospitalization based on SHAP analysis. A. Global feature importance of the eleven hematologic and inflammatory markers for the AdaBoost prolonged hospitalization prediction model. B. The SHAP beeswarm plot shows the effect of each feature on the AdaBoost prolonged hospitalization prediction model. C. Waterfall plot demonstrates the impact of specific hematologic and inflammatory variables on the predicted risk of prolonged hospitalization.

influential predictors of the AdaBoost model (**Figure 4A**, **4B**). Higher RDW-SD and MCV increased predicted risk, while lower PLT and RBC were associated with higher risk.

The SHAP waterfall plot showed individual prediction patterns. Patients with low RBC count, Hb, and PLT levels together with elevated inflammatory indices such as MLR, NMLR, and SIRI showed higher predicted probability of extended hospital stay (predicted risk = 0.52 vs baseline 0.42) (**Figure 4C**). These features matched the biological mechanisms of DC, where anemia, systemic inflammation, and impaired hematopoiesis reflect advanced disease severity and organ dysfunction.

Feature interpretation showed that the AdaBoost model identified meaningful relationships between hematologic abnormalities, inflammation, and prolonged hospitalization. This interpretable model supports the use of hematologic and inflammatory markers as noninvasive predictors of hospitalization burden in DC.

Discussion

In this single-center Chinese cohort of patients with DC, routine hematologic and inflammatory indices abnormalities were strongly associated with prolonged hospitalization. Patients with hospital stay ≥ 7 days exhibited more severe anemia (lower Hb, HCT, RBC count) and thrombocytopenia, and elevated markers of systemic inflammation (e.g. higher RDW and MLR) compared to those discharged within one week. Reduced oxygen-carrying capacity and increased inflammatory

burden reflected more advanced cirrhosis and were associated with slower recovery and longer hospitalization. RBC count and Hb showed nonlinear threshold relationships with hospital stay. When RBC count fell below ~5.2 × 10^{12}/L, the risk of prolonged hospitalization rose sharply. While values above these level showed limited additional benefit. Using eleven such hematologic/inflammatory parameters, we developed a machine learning model (AdaBoost) that achieved modest but significant discrimination of patients at risk for longer admission (test-set AUC ~0.72, sensitivity 72%, specificity 58%). SHAP analysis confirmed that features reflecting red cell abnormalities and innate immune activation (including RDW-SD, PLT, MCV, and RBC, etc.) were the dominant contributors to prediction. Anemia, disordered hematopoiesis, and systemic inflammation are closely related to disease progression and hospitalization burden in DC.

Our observations align with and extend prior evidence on the prognostic significance of hematologic derangements in cirrhosis. Anemia is exceedingly common in advanced liver disease - affecting over half of patients with decompensation [15] - due to multifactorial causes such as portal hypertensive hypersplenism, nutritional deficiencies, bone marrow suppression, and gastrointestinal blood loss [15, 16]. Correspondingly, anemia has been linked to worse clinical outcomes. For example, Paternostro et al. reported that anemia (Hb ≤ 12 g/dL) was present in ~62% of decompensated cirrhotics (versus 19% of compensated) and independently predicted a higher risk of hepatic decompensation events and mortality during follow-up [15]. Ren et al. recently demonstrated in a large Chinese multi-center study that severe anemia (Hb < 8 g/dL) was associated with a ~1.6-fold increased hazard of 90-day and 1-year mortality in hospitalized cirrhotic patients [11]. Our data build on these reports by focusing on LOS as an outcome: we found even moderate anemia correlated with prolonged hospitalization, and that maintaining Hb above ~117 g/L was associated with shorter stays. The Hb threshold of about 11.7 g/dL is of clinical significance. This value is higher than the 7 g/dL transfusion threshold used in acute variceal bleeding, where restrictive transfusion improved survival [17]. Hemoglobin that remains mildly low outside acute bleeding may slow recovery in cirrhosis patients. Anemia in patients with cirrhosis will aggravate fatigue and weakness, reduce liver and kidney perfusion, and damage wound healing or tissue regeneration. These changes may prolong hospitalization. Previous studies have shown that low hemoglobin at discharge are associated with a high rate of re-hospitalization within 30 days in cirrhosis patients. This indicates anemia is a modifiable risk factor for post-hospital outcomes [18]. Maintaining adequate oxygen-carrying capacity through appropriate transfusion or hematinic treatment may promote earlier recovery and discharge.

The association between inflammatory cell indices and LOS was also clear. DC is now recognized as a state of systemic inflammation and immune dysregulation [19, 20]. Bacterial translocation from the gut and frequent infections (spontaneous peritonitis, pneumonia, etc.) can activate persistent inflammation and lead to organ dysfunction such as kidney injury and hepatic encephalopathy and can aggravate portal hypertension [19]. Markers such as the NLR have been shown to correlate with cirrhosis severity and prognosis. A recent meta-analysis of 27 studies found that elevated NLR independently predicts higher mortality in end-stage liver disease and the association was stronger in Asian patient cohorts [21]. NLR showed no significant difference between short and long hospitalization groups in our study. WBC counts were often normal or reduced in both groups and this pattern may relate to leukopenia caused by hypersplenism. Patients with prolonged hospitalization showed median MLR values about 13% higher and SIRI values about 20% higher than patients with shorter stays (both $P < 0.05$). These indices reflected the innate immune activation in advanced cirrhosis and indicated pro-inflammatory monocytes/macrophages are over-represented relative to lymphocytes [19]. In support of this, prior work has identified a low lymphocyte-to-monocyte ratio (LMR, the inverse of MLR) as a poor prognostic marker in cirrhosis. Qi et al. reported that cirrhotic patients with HBV-related decompensation and an LMR below ~2.0 had significantly higher 30-day and 90-day mortality, and LMR was an independent predictor of survival in multivariate analysis [22]. Our results are concordant in that a high monocyte count (and high MLR) portended worse inpatient courses.

Monocytes drive production of cytokines like TNF-α, IL-6, and IL-1β, which can induce vasodilation, hypotension, and further immune dysfunction (the so-called "cirrhosis-associated immune syndrome") [19]. This inflammatory milieu likely contributed to complications necessitating longer hospitalization. It is worth noting that one inflammation-based score that did not differ significantly between our groups was the SII (SII = platelet count × neutrophil count/lymphocyte count). While SII has prognostic value in some chronic liver disease contexts (e.g. fatty liver disease and hepatocellular carcinoma), it may be less relevant in DC where platelet counts are often depressed. In our cohort, median platelets were low in both short-stay and long-stay patients (87 vs 72 × 10^9/L). This may reduce the ability of SII to distinguish risk. Thrombocytopenia is common in cirrhosis and occurs in about 70-80% of compensated patients [16]. Low platelet levels reflect portal hypertension severity and reduced thrombopoietin synthesis. Patients with longer hospitalization showed lower platelet counts. Platelet levels usually change slowly during admission and recovery is uncommon without intervention. Indices such as RDW and MLR reflect ongoing physiological stress and may be more useful for short-term risk stratification.

Elevated RDW was one of the strongest predictors of prolonged stay in our analysis. This finding is similar to previous studies in critical illness and liver disease. RDW reflects variation in RBC size. It often elevates in nutritional anemias (iron, vitamin B12, folate deficiencies) and bone marrow dysfunction [23]. In cirrhosis, RDW may rise due to iron-restricted erythropoiesis caused by chronic bleeding or anemia of inflammation. Folate deficiency, especially in alcohol-associated liver disease may also contribute. High RDW has been associated with increased mortality in acute DC. Turcato et al. found that nonsurvivors had a significantly higher RDW than survivors (median 17.4% vs 15.5%), and each 1% RDW increase conferred ~20% higher odds of 30-day mortality [24]. This study corroborates RDW as a marker of poor clinical trajectory. Patients hospitalized ≥ 7 days had an RDW roughly 5% higher than those with shorter admissions (50 vs 48 fL, $P < 0.001$). This likely indicates greater heterogeneity of the erythrocyte population due to deficient erythropoiesis or hemolysis in the longer-stay group. RDW is also correlate with systemic inflammatory markers (e.g. C-reactive protein and IL-6) [25], linking it to the inflammation-frailty axis in chronic disease. RDW can be regarded as a marker of poor nutritional status, chronic inflammation and bone marrow stress and these conditions are linked to slower recovery in cirrhosis. SHAP analysis of the AdaBoost model showed that RDW especially RDW-SD contributed most to prediction. RDW-CV is a relative index because it is influenced by MCV while RDW-SD measures the absolute distribution width of erythrocyte volume in fL. MCV is often elevated in patients with cirrhosis. This change is usually related to vitamin deficiency or alcohol exposure. RDW-SD may better reflect bone marrow stress and ineffective erythropoiesis. The model's decision-making aligns with clinical plausibility. Patients with high RDW, MLR (monocytosis), and low Hb, HCT were classified as high risk. These features were common in patients with advanced decompensation. Patients with preserved hematocrit and lower inflammatory ratios were usually predicted as short-stay cases. These patients often had more stable cirrhosis or responded faster to therapy.

Routine hematologic and inflammatory markers provide significant prognostic information for the hospital course of DC patients. Anemia, especially elevated RDW and reduced Hb, RBC count was independently associated with prolonged LOS and more severe disease. Systemic inflammation such as high MLR, NMLR, and SIRI also showed independent associations with longer LOS. Our machine learning model based on these features showed moderate ability to identify patients at risk for longer hospital stay. The prediction pattern was consistent with clinical observations. Combining hematologic and inflammatory indicators with traditional liver scores may help identify high-risk cirrhotic patients early during admission. This stratification could guide targeted clinical management. Patients with high RDW may benefit from nutritional supplementation including iron and vitamin B12 and folate because elevated RDW often reflects nutritional deficiency. More liberal transfusion thresholds (e.g., maintaining Hb > ~12 g/dL) may enhance tissue oxygenation and functional status. A high inflammatory burden may prompt a more diligent search for subclinical infections and consi-

der appropriate immunomodulatory strategies. Clinical management of cirrhosis may improve by addressing nutrition, anemia and infection control at the same time.

This study has several limitations. It was a retrospective analysis conducted in a single tertiary center in China. Selection bias may exist and may not fully apply to other populations and healthcare settings. The predominant etiology of cirrhosis in our cohort is likely related viral hepatitis. While alcoholic and non-alcoholic fatty liver disease are more prevalent in Western countries. Differences in disease etiology, patient demographics, and healthcare delivery systems may influence the predictive accuracy and clinical applicability in different regions. The model was designed to assess prolonged hospitalization risk based on admission parameters. Patients who developed complications such as infections or variceal bleeding during hospitalization were not excluded. These events are common in decompensated cirrhosis and substantially affect length of stay. Future studies should validate our predictive model in larger, multicenter cohorts to confirm stability and applicability. Future models may combine hematologic indices with established clinical predictors such as MELD score, infection status and ascites severity to improve prediction accuracy. This approach may improve prediction accuracy. A more comprehensive model may help clinicians identify high-risk patients earlier and support more efficient clinical management.

Disclosure of conflict of interest

None.

Abbreviations

AdaBoost, Adaptive Boosting; AUC, Area Under the Curve; BMI, Body Mass Index; CatBoost, Categorical Boosting; DC, Decompensated Cirrhosis; Hb, Hemoglobin; HCT, Hematocrit; KNN, K-Nearest Neighbor; LightGBM, Light Gradient Boosting Machine; LOS, Length of Stay; MCH, Mean Corpuscular Hemoglobin; MCV, Mean Corpuscular Volume; MLR, Monocyte-to-Lymphocyte Ratio; NLR, Neutrophil-to-Lymphocyte Ratio; NMLR, Neutrophil-Monocyte-Lymphocyte Ratio; PLT, Platelets; RBC, Red Blood Cell; RDW-CV, Red Cell Distribution Width-Coefficient of Variation; RDW-SD, Red Cell Distribution Width-Standard Deviation; SHAP, Shapley Additive Explanation; SII, Systemic Immune-Inflammation Index; SIRI, Systemic Inflammation Response Index; XGBoost, Extreme Gradient Boosting.

Address correspondence to: Drs. Xinhui Xu and Ling Zhao, Department of Infectious Diseases, Funan County People's Hospital, Northeast Intersection of Wanghua Road and Longquan Road, Lucheng Town, Funan County, Fuyang 236300, Anhui, China. E-mail: xxh0052@126.com (XHX); zhaoling769@163.com (LZ)

References

[1] Li J, Pang S, Huang H, Lu Y, Tang T, Wu J and Chen M. Association between red cell distribution width-to-albumin ratio and all-cause mortality in critically ill cirrhotic patients with sepsis: a retrospective analysis of the MIMIC-IV database. Front Med (Lausanne) 2025; 12: 1610726.

[2] Kim WR, Raghunathan K, Martin GS, Davis EA, Sindhwani NS, Telang S and Lodaya K. Timely albumin infusion may improve resource utilization in patients with cirrhosis and spontaneous bacterial peritonitis. Biomed Res Int 2024; 2024: 6673823.

[3] Sun L, Ding P, Mao W and Wu J. D-dimer-to-albumin ratio: a novel indicator to predict poor outcomes in patients with hbv-associated decompensated cirrhosis. Biomed Res Int 2022; 2022: 9062383.

[4] Garbuzenko DV. Therapeutic possibilities of gut microbiota modulation in acute decompensation of liver cirrhosis. World J Hepatol 2023; 15: 525-537.

[5] Celdir MG, Wehby G, Prakash S and Tanaka T. July effect in hospitalized cirrhosis patients: a US nationwide study using difference-in-differences analysis. PLoS One 2025; 20: e0316445.

[6] Beyoğlu D, Simillion C, Storni F, De Gottardi A and Idle JR. A metabolomic analysis of cirrhotic ascites. Molecules 2022; 27: 3935.

[7] Chang YC, Fang YT, Chen HC, Lin CY, Chang YP, Tsai YH, Chen YM, Huang KT, Chang HC, Wang CC, Lin MC and Fang WF. The survival of septic patients with compensated liver cirrhosis is not inferior to that of septic patients without liver cirrhosis: a propensity score matching analysis. J Clin Med 2022; 11: 1629.

[8] Shi Y, Shen W, Xu G, Wang X and Ning B. Hepatic venous pressure gradient and rebleeding risk of patients with nonalcoholic steatohepatitis cirrhosis after variceal bleeding. Front Med (Lausanne) 2023; 10: 1224506.

[9] Ifrah A, Fromer R, Gayner AH and Yeung HM. Discharge outcomes of hospitalized patients with new onset decompensated cirrhosis. Dig Dis Sci 2024; 69: 3220-3225.
[10] Jeong J, Tanaka M and Iwakiri Y. Hepatic lymphatic vascular system in health and disease. J Hepatol 2022; 77: 206-218.
[11] Ren H, Li H, Deng G, Wang X, Zheng X, Huang Y, Chen J, Meng Z, Gao Y, Qian Z, Liu F, Lu X, Shang J, Wang S, Yin S, Tan W, Hou Y, Xiong S, Long L, Li B, Luo S, Zhang W and Shi Y. Severe anemia is associated with increased short-term and long-term mortality in patients hospitalized with cirrhosis. Ann Hepatol 2023; 28: 101147.
[12] Dejanović B, Barak O, Čolović P, Janjić N, Savić Ž, Gvozdanović N and Ružić M. Hospital mortality in acute decompensation of alcoholic liver cirrhosis: can novel survival markers outperform traditional ones? J Clin Med 2024; 13: 6208.
[13] Qiao L, Zhang Y, Wang J, Wu X, Zhang C, Xue Y, Cao J, Jin J, Wang T, Li J and Liang Y. Machine learning-based prediction of preeclampsia using first-trimester inflammatory markers and red blood cell indices. BMC Pregnancy Childbirth 2025; 25: 1083.
[14] Zhu Y, Zhang Y, Hu W, Cao J, Feng B, Jin J, Zhang S, Tang Q, Qiao L and Liang Y. First-trimester pan-immune-inflammation value predicts preeclampsia in a dose-dependent linear pattern. BMC Pregnancy Childbirth 2025; 25: 1250.
[15] Paternostro R, Kapzan L, Mandorfer M, Schwarzer R, Benedikt S, Viveiros A, Bauer D, Ferlitsch M, Zoller H, Trauner M and Ferlitsch A. Anemia and iron deficiency in compensated and decompensated cirrhosis: prevalence and impact on clinical outcomes. J Gastroenterol Hepatol 2020; 35: 1619-1627.
[16] Fierro-Angulo OM, González-Regueiro JA, Pereira-García A, Ruiz-Margáin A, Solis-Huerta F and Macías-Rodríguez RU. Hematological abnormalities in liver cirrhosis. World J Hepatol 2024; 16: 1229-1244.
[17] Villanueva C, Colomo A, Bosch A, Concepción M, Hernandez-Gea V, Aracil C, Graupera I, Poca M, Alvarez-Urturi C, Gordillo J, Guarner-Argente C, Santaló M, Muñiz E and Guarner C. Transfusion strategies for acute upper gastrointestinal bleeding. N Engl J Med 2013; 368: 11-21.
[18] Pompili E, Baldassarre M, Zaccherini G, Tufoni M, Iannone G, Pratelli D, Palmese F, Vizioli L, Faggiano C, Bedogni G, Domenicali M and Caraceni P. Low haemoglobin level predicts early hospital readmission in patients with cirrhosis and acute decompensation. JHEP Rep 2023; 5: 100698.
[19] Cheng J, Ju H, Wang G, He C and Wang W. Association of systemic inflammation response index with short-term all-cause mortality in decompensated liver cirrhosis patients. J Inflamm Res 2024; 17: 8985-8995.
[20] Diaz-Ruiz R, Poca M, Roman E, Panadero-Gomez R, Cuyàs B, Bañares I, Morales A, Puerto M, Lopez-Esteban R, Blazquez E, Fernández-Castillo M, Correa-Rocha R, Rapado-Castro M, Breton I, Bañares R, Soriano G and Garcia-Martinez R. Vitamin D supplementation is associated with inflammation amelioration and cognitive improvement in decompensated patients with cirrhosis. Nutrients 2025; 17: 226.
[21] Cai XH, Tang YM, Chen SR, Pang JH, Chong YT, Cao H and Li XH. Prognostic value of neutrophil-to-lymphocyte ratio in end-stage liver disease: a meta-analysis. World J Hepatol 2024; 16: 477-489.
[22] Qi X. Peripheral blood lymphocyte-to-monocyte ratio predicts mortality in patients with HBV-related decompensated cirrhosis. Clin Lab 2019; 65: [Epub ahead of print].
[23] Seyit M, Avci E, Yilmaz A, Kemanci A, Caliskan A, Ozen M, Oskay A, Aybek H and Türkcuer İ. Evaluation of the diagnostic value of hematologic parameters and ratios in SARS-CoV-2 VOC-202012/01 mutant population. Cureus 2022; 14: e28285.
[24] Turcato G, Campagnaro T, Bonora A, Vignola N, Salvagno GL, Cervellin G, Ricci G, Maccagnani A and Lippi G. Red blood cell distribution width independently predicts 1-month mortality in acute decompensation of cirrhotic patients admitted to emergency department. Eur J Gastroenterol Hepatol 2018; 30: 33-38.
[25] Song Q, Wu X, Taheri FA, Meng L, Wang W and Mo X. Association of red blood cell distribution width with short- and long-term all-cause mortality in patients with acute pancreatitis and sepsis. BMC Gastroenterol 2025; 25: 539.