

Length of Hospital Stay in Patients with Primary Liver Cancer Undergoing Surgery: Risk Factors and Predictive Model Development

Bin Sun¹, Xiaobo Li², Xiuying He¹, Na Zhang³

¹Department of the International Special Needs Ward, First Hospital of China Medical University, Shenyang, Liaoning, 110001, People's Republic of China; ²Nursing Department, First Hospital of China Medical University, Shenyang, Liaoning, 110001, People's Republic of China; ³Department of Hepatobiliary Surgery, First Hospital of China Medical University, Shenyang, Liaoning, 110001, People's Republic of China

Correspondence: Xiuying He; Na Zhang, Email hexiuying1234@163.com; lnzhangna@126.com

Aim: To identify preoperative risk factors for prolonged length of hospital stay (LOS) in patients undergoing surgery for primary liver cancer and to develop a predictive nomogram.

Methods: We retrospectively analyzed 702 surgical patients from a single center (2020–2023). LOS was modeled using negative binomial regression based on preoperative factors to construct a nomogram. Model performance was evaluated via internal bootstrap validation (1000 resamples), calibration plots, and decision curve analysis. Prolonged LOS was defined as >17 days (75th percentile) for a secondary logistic regression analysis.

Results: Four independent preoperative factors predicted longer LOS: lower serum cholinesterase, higher fibrinogen, intrahepatic cholangiocarcinoma (vs hepatocellular carcinoma), and female sex (all $p < 0.05$). The nomogram showed moderate discriminative ability (apparent AUC ~0.67) with good calibration. The mean absolute error for LOS prediction was ~4.6 days. For predicting prolonged LOS (>17 days), the logistic model achieved an AUC of ~0.67.

Conclusion: We developed an internally validated nomogram using routine preoperative data to estimate the risk of extended hospitalization after liver cancer surgery. This tool may help identify high-risk patients for targeted interventions, although its predictive accuracy is modest, and external validation is required before clinical application.

Keywords: length of stay, liver cancer, nomogram, nursing, surgery

Introduction

Primary liver cancer (PLC), also known as hepatocellular carcinoma, is one of the most prevalent malignancies worldwide and the third leading cause of cancer-related mortality.¹ In 2020, approximately 905,700 new PLC cases were diagnosed globally, leading to about 830,200 deaths.² Asia and Africa bear the highest PLC burden, with China alone accounting for roughly half of all cases. The main PLC subtypes are hepatocellular carcinoma (75–85% of cases) and intrahepatic cholangiocarcinoma (10–15%).³ In China, PLC ranks fourth among malignancies in incidence. There are significant geographic disparities: rural areas have higher incidence and mortality rates than urban areas, particularly in individuals under 65 years of age.^{4,5}

Surgical resection is the primary curative treatment for liver cancer. An increasing number of patients are undergoing liver cancer surgery, which has led to rising hospitalization costs and, in many cases, longer hospital stays. Extended postoperative stays can heighten the risk of hospital-acquired complications and adversely affect patient prognosis.⁶ Minimizing length of stay (LOS) has therefore become a priority in modern hospital management to improve outcomes and reduce costs.⁷ Enhanced recovery after surgery (ERAS) programs exemplify this effort, using multidisciplinary approaches to accelerate rehabilitation, reduce complications, and hasten discharge.⁸

Despite these advances, the factors that determine how long liver cancer patients remain hospitalized after surgery are not fully understood. In general, LOS correlates with disease severity and the quality of care.^{9,10} However, even among

patients with similar clinical profiles, hospitalization duration can vary widely due to various factors. Globally, research on predicting LOS in liver cancer patients is limited, and existing studies show heterogeneous results influenced by differences in social structure, medical systems, and study design.^{11,12} No well-validated predictive model currently exists for this specific context, highlighting a gap in clinical practice. Commonly suspected contributors to prolonged LOS include tumor burden, comorbidities, liver function status, and postoperative complications, but these have not been definitively established in this population.

In light of this gap, our study aimed to identify key risk factors associated with extended hospital stay in patients with primary liver cancer undergoing surgery, and to develop a predictive model to stratify patients by risk of prolonged hospitalization. We hypothesized that routine clinical indicators available at admission could help predict which patients are likely to have longer-than-expected stays. By building a nomogram-based prediction model and validating it internally, we sought to lay the groundwork for early identification of high-risk patients, enabling targeted interventions and improved perioperative management and nursing care for those patients.

Methods

Study Design and Setting

We conducted a single-center retrospective cohort study to investigate factors associated with hospital length of stay (LOS) and to develop a predictive model for LOS among patients undergoing surgery for primary liver cancer. The study was performed in the Department of Hepatobiliary Surgery at the First Hospital of China Medical University (Shenyang, Liaoning, China). All consecutive eligible surgical admissions between January 1, 2020, and October 31, 2023, were included. The dataset was accessed for analysis on December 12, 2023.

Participants

The source dataset comprised 966 hospital admissions for liver cancer surgery during the study period. After applying inclusion and exclusion criteria, 702 admissions qualified for the final analysis.

Inclusion Criteria

Patients were included if they (1) had a pathological diagnosis confirming primary liver cancer and underwent surgical resection during the admission, and (2) were hospitalized for at least one day and discharged after completion of the planned surgical treatment.

Exclusion Criteria

Admissions were excluded if they (1) did not involve primary liver cancer (eg, the final diagnosis was a benign lesion or a metastatic liver tumor rather than PLC), (2) had missing information on the tumor's pathological subtype that could not be retrieved from medical records, or (3) met other pre-specified exclusion conditions (such as incomplete medical records or repeat hospitalization of the same patient within the study window for continuation of treatment). These criteria were designed to yield a homogeneous cohort of first-time surgical admissions for primary liver cancer with complete baseline data.

Pathological Subtype Definition

Pathological subtype was treated as a binary predictor with two categories: hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (ICC). We mapped the pathology report of each case to one of these categories using a standardized rule: diagnoses containing "Hepatocellular carcinoma" (including synonyms like "HCC") were classified as HCC, while those indicating "Cholangiocarcinoma" were classified as ICC if the context confirmed an intrahepatic primary tumor. Rare primary liver tumor types (eg, mixed HCC-CCA or other uncommon histologies) were excluded from the primary analysis. Cases with missing or ambiguous subtype information were not imputed; they were excluded as described above to avoid misclassification.

Outcome Measure (LOS)

The primary outcome was hospital length of stay, defined as the number of days from the date of admission to the date of discharge for the index surgical hospitalization. For patients transferred to other facilities or who died during hospitalization (if any in our cohort), LOS was calculated up to the point of transfer or death. LOS was treated as a continuous count variable (in days) for analysis.

Predictor Variables and Measurements

We considered a broad range of candidate predictors available from the hospital records, focusing on preoperative factors. All laboratory measurements and vital signs were those obtained during the initial assessment upon admission (ie, preoperative baseline values). The candidate predictors included:

Demographics: Sex and age.

Medical history: Presence of underlying viral hepatitis (hepatitis B and/or C) and lifestyle factors such as smoking status and alcohol use history.

Vital signs: Blood pressure at admission (systolic blood pressure, SBP; diastolic blood pressure, DBP).

Anthropometrics: Height, weight, and body mass index (BMI).

Blood type: ABO blood group.

Baseline laboratory tests: Liver function and metabolic panel – total bilirubin (TBIL), direct bilirubin (DBIL), albumin (ALB), aspartate aminotransferase (AST), alanine aminotransferase (ALT), and cholinesterase (ChE); renal function tests – serum creatinine (Crea) and blood urea; coagulation profile – prothrombin time (PT) and international normalized ratio (PT-INR), fibrinogen (Fbg), and activated partial thromboplastin time (APTT); complete blood count – neutrophil count (NE), white blood cell count (WBC), red blood cell count (RBC), hemoglobin (Hb), and platelet count (PLT).

Tumor markers: Alpha-fetoprotein (AFP) and prothrombin induced by vitamin K absence-II (PIVKA-II), when available.

Laboratory values were recorded in standard units: for example, TBIL/DBIL in $\mu\text{mol/L}$, ALB in g/L , AST/ALT/ChE in U/L , Crea in $\mu\text{mol/L}$, Urea in mmol/L , AFP in U/mL , PIVKA-II in mAU/mL , fibrinogen in g/L , PT/APTT in seconds, blood counts in $10^9/\text{L}$ for WBC/NE/PLT (RBC in $10^{12}/\text{L}$, Hb in g/L), and blood pressure in mmHg .

It should be noted that due to the constraints of the retrospective data retrieval platform, detailed intraoperative variables (eg, whether surgery was laparoscopic or open, extent of hepatic resection) and postoperative events (such as specific complications or their severity) were not available for extraction. Consequently, such factors were not included in our candidate predictors and are recognized as potential unmeasured confounders in this study.

Data Quality Control and Preprocessing

We undertook several steps to ensure data quality and consistency before analysis. Categorical variables (like sex and smoking) were standardized to a uniform coding scheme, and obvious data entry errors (eg, misspelled categories or out-of-range values) were corrected when identified. Continuous variables were checked for implausible values; any extreme outliers suggestive of measurement error were flagged and cross-checked with source documents when possible. We verified that LOS values were reasonable (for instance, we excluded any case with $\text{LOS} < 1$ day, since our inclusion criteria required at least one overnight stay). In addition, for biomarkers with highly skewed distributions such as AFP and PIVKA-II, we considered logarithmic transformation in exploratory analyses to assess whether skewness materially affected the results (these were used in sensitivity analyses described later).

Missing Data Handling

We examined the dataset for missing values across all candidate predictors. The extent and pattern of missingness for each variable were analyzed and visualized. Our approach to missing data was as follows:

For pathological subtype, as noted, cases with missing subtype were removed from the primary analysis; we did not attempt to impute this, since subtype defines our cohort and imputation could introduce significant bias if, for example, missingness was associated with outcome.

For other predictors with incomplete data, we assumed that data were missing at random (MAR) conditional on observed information. We employed multiple imputation by chained equations (MICE) to handle missing values for these variables.¹³ The imputation model included the outcome (LOS), all other predictors, and auxiliary variables that might predict missingness. We generated 20 imputed datasets to ensure stability of the imputation estimates and combined results using Rubin's rules.

An exception was made for AFP and PIVKA-II. These tumor markers had higher rates of missingness (because in clinical practice they were not measured in all patients, often being ordered selectively). In the primary analysis, we did not impute absent AFP or PIVKA-II values for patients in whom these tests were not done. Instead, we created binary indicator variables for "AFP measured" and "PIVKA-II measured" to account for the fact of whether the test was performed. The actual AFP and PIVKA-II values were then included in a secondary analysis restricted to the subset of patients who had those tests, rather than being imputed for all. This approach avoids making unverifiable assumptions about unmeasured tumor marker levels.

We also conducted a complete-case analysis (using only those admissions with no missing values in any predictor) as a sensitivity check to compare with the main (imputed) analysis results.

Collinearity Diagnostics

Before fitting the multivariable model, we assessed potential multicollinearity among the continuous predictors. We examined Spearman correlation matrices to identify highly correlated pairs of variables that might introduce redundancy (for example, we anticipated high correlation between PT and PT-INR, and between TBIL and DBIL, as these measure related aspects of coagulation and bilirubin levels, respectively). In instances of strong collinearity, we decided a priori to retain the variable more commonly used in clinical practice or more directly interpretable and to exclude the other. Specifically, PT-INR was retained over the raw PT value, and TBIL was retained over DBIL, since PT-INR and TBIL capture the needed information with potentially less variability across laboratories. We also calculated variance inflation factors (VIFs) for the full set of candidate predictors after these exclusions. No remaining predictor had a VIF >5 (indeed, all were well below this threshold) in the final model, indicating that multicollinearity was not a significant concern.

Model Development (Primary Analysis)

Our primary analysis modeled LOS as a continuous outcome. Given that LOS (in days) is a count variable often exhibiting over-dispersion (variance larger than mean), we chose a generalized linear model with a negative binomial distribution and a log link. This approach can accommodate the right-skewed distribution and over-dispersion commonly observed in LOS data. We prespecified a core set of clinically relevant predictors to include in the multivariable model based on subject-matter knowledge and the results of univariable checks, rather than relying solely on automated selection procedures. The core model predictors were as follows: pathological subtype (HCC vs ICC), sex, cholinesterase, fibrinogen, neutrophil count, diastolic BP, and APTT. These were chosen because they represented distinct domains (tumor characteristics, patient demographics, liver function/nutrition, coagulation/inflammation, hemodynamics) and had shown at least a suggestive association with LOS in initial analyses or prior literature. Continuous predictors were entered linearly in the model; we examined potential non-linear effects (eg, by adding quadratic or spline terms) in exploratory analyses and found no strong deviations from linearity for the main predictors, so the simpler linear terms were used in the final model. The outcome (LOS) was modeled on the log scale as per the negative binomial link function, and results are reported as incidence rate ratios (IRRs) which represent the multiplicative effect on the expected LOS per unit increase in the predictor.

Internal Validation and Performance Assessment

We evaluated the model's performance in terms of discrimination, calibration, and overall prediction error. For internal validation, we used bootstrap resampling. We generated 1000 bootstrap samples from the original dataset, refit the entire modeling procedure on each bootstrap sample (including the imputation and model fitting steps), and evaluated the model predictions on the original dataset to estimate optimism. This allowed us to calculate optimism-corrected performance metrics. We summarized predictive accuracy for continuous LOS using measures such as mean absolute error (MAE) and

root mean squared error (RMSE) between predicted and observed LOS. Calibration of the model was assessed by comparing predicted vs observed LOS across deciles of predicted risk; we plotted a calibration curve and computed the calibration intercept and slope (ideal values being 0 and 1, respectively). Because our outcome is continuous, traditional discrimination metrics like AUC are not directly applicable; however, for a general sense of how well the model separates patients with longer vs shorter stays, we also conducted a secondary analysis dichotomizing LOS (described below) which allowed calculation of an AUC.

Nomogram and Secondary (Prolonged LOS) Analysis

We created a nomogram based on the final negative binomial model to provide an accessible tool for clinicians. The nomogram assigns points for each predictor value and yields an estimated LOS when the points are summed, thereby translating the regression model into a simple graphical scoring system. In addition to predicting actual LOS, we performed a secondary analysis focusing on the prediction of “prolonged” stay. We defined prolonged LOS as a hospital stay longer than 17 days, which corresponds to the 75th percentile of LOS in our cohort. Using this binary outcome (LOS >17 days vs ≤17 days), we fit a logistic regression model with the same set of core predictors. We evaluated the discrimination of this model using the area under the receiver operating characteristic curve (AUC). We also performed decision curve analysis (DCA) to assess the clinical utility of the logistic model across a range of threshold probabilities for prolonged stay. DCA examines the net benefit of using the model to guide interventions (eg, identifying high-risk patients for special management) compared to default strategies of treating all patients as high-risk or none as high-risk, for various risk threshold preferences. This helps determine if the model’s predictions would do more good than harm in a practical decision-making scenario.

Formal Sample Size Assessment

We assessed the adequacy of our sample size for developing a prediction model using the framework proposed by Riley et al. Specifically, we used the `pmsampsize` package in R to estimate the minimum sample size required for a desired level of model performance and stability.¹⁴ We input parameters including the number of candidate predictors (~7 in the core model), an anticipated Cox-Snell R^2 for the model on the log LOS scale, a target global shrinkage factor of ≥ 0.90 (to limit overfitting), and a 5% tolerance for the calibration error. The calculation suggested that our available sample of 702 was sufficient to support the planned model complexity, indicating that the study was adequately powered to detect moderate effect sizes without severe overfitting.

Statistical Software

All analyses were conducted in R version 4.2 (R Foundation for Statistical Computing). Multiple imputation was carried out using the `mice` package. Negative binomial regression modeling was done with the `MASS` package. The nomogram and calibration plot were generated using the `rms` package. ROC analysis was performed with `pROC`, and decision curve analysis utilized the `rmda` package. Our reporting of the study adheres to the STROBE guidelines for observational studies and the TRIPOD statement for reporting of multivariable prediction models.

Ethical Considerations

This study was approved by the Ethics Committee of the First Affiliated Hospital of China Medical University (Approval No. AF-SOP-07-1.2–01, dated December 4, 2023). The requirement for informed consent was waived due to the retrospective nature of the research and use of de-identified patient data. All procedures were carried out in accordance with the Declaration of Helsinki and relevant local regulations. Patient confidentiality was strictly maintained; data were accessed in anonymized form and stored securely.

Results

Study Cohort

Out of 966 surgical admissions initially identified, 702 met the eligibility criteria and were included in the analysis. We excluded 245 admissions that lacked pathological subtype information and 19 admissions that involved diagnoses other than primary liver cancer or extremely rare tumor types (Figure 1). The final cohort consisted of 651 HCC cases and 51 ICC cases. Baseline characteristics of the cohort are summarized in Table 1. In brief, the patient population was predominantly male (80.2%), with a median age of 62 years. Approximately 66.0% of patients had underlying viral hepatitis (HBV or HCV infection). Other baseline features (Table 1) indicate that most patients had preserved liver function (median albumin in normal range, etc.), and the distribution of comorbid conditions was typical for a liver cancer surgical group, although detailed comorbidity data are not shown in the table.

Distribution of LOS

The length of stay for the 702 patients ranged from 1 day to 56 days. The median LOS was 13 days, with an interquartile range (IQR) of 10–17 days. The distribution was right-skewed – most patients stayed around one to two weeks, but a minority had substantially longer hospitalizations extending several weeks (Figure 2). We noted that 25% of patients had LOS greater than 17 days. We chose this value (17 days, the 75th percentile) as a cutoff to define “prolonged LOS” in a secondary analysis, as it represents notably extended hospitalization relative to the typical postoperative course.

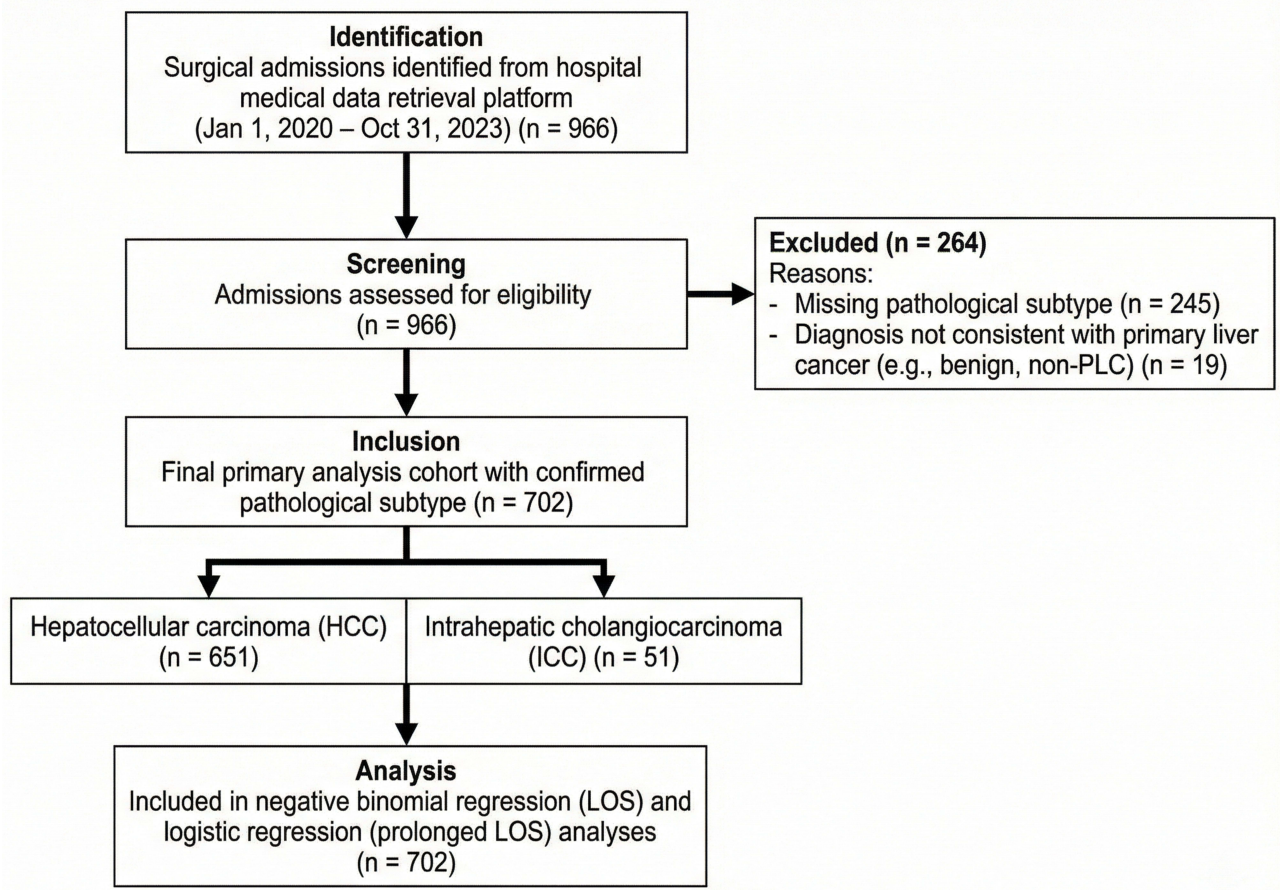


Figure 1 STROBE flow diagram. From 966 surgical admissions (Jan 2020–Oct 2023) identified from the hospital data platform, 245 were excluded due to missing pathological subtype and 19 for non-primary liver cancer or rare diagnoses. The final cohort comprised 702 patients (HCC: 651; ICC: 51), all included in the primary (negative binomial) and secondary (logistic) regression analyses.

Table 1 Baseline Characteristics of the Analysis Cohort (n=702), Overall and by Pathological Subtype

Variable	Overall	HCC	ICC
N	702	651	51
Male sex, n (%)	563 (80.2%)	527 (81.0%)	36 (70.6%)
Viral hepatitis, n (%)	463 (66.0%)	441 (67.7%)	22 (43.1%)
Smoking, n (%)	214 (30.5%)	202 (31.0%)	12 (23.5%)
Alcohol use, n (%)	135 (19.2%)	127 (19.5%)	8 (15.7%)
ABO blood type, n (%)			
A	157 (22.4%)	149 (22.9%)	8 (15.7%)
B	189 (26.9%)	172 (26.4%)	17 (33.3%)
AB	51 (7.3%)	47 (7.2%)	4 (7.8%)
O	228 (32.5%)	209 (32.1%)	19 (37.3%)
Other	0 (0.0%)	0 (0.0%)	0 (0.0%)
Missing	77 (11.0%)	74 (11.4%)	3 (5.9%)
Age (years), median (IQR)	62.00 (55.00–68.00)	62.00 (55.00–68.00)	63.00 (55.50–67.50)
BMI (kg/m ²), median (IQR)	23.90 (21.70–26.10)	23.85 (21.70–26.10)	24.60 (21.65–27.10)
SBP (mmHg), median (IQR)	130.00 (120.00–142.00)	130.00 (120.00–143.00)	130.00 (120.25–137.00)
DBP (mmHg), median (IQR)	83.00 (76.00–90.00)	83.00 (76.00–90.00)	83.00 (78.25–90.00)
Total bilirubin (μmol/L), median (IQR)	14.10 (10.90–18.40)	14.10 (10.83–18.10)	14.70 (11.60–22.25)
Albumin (g/L), median (IQR)	38.80 (36.30–41.10)	38.75 (36.23–41.10)	38.90 (37.20–40.55)
Creatinine (μmol/L), median (IQR)	65.00 (56.00–74.00)	65.00 (57.00–74.00)	60.00 (51.50–70.00)
AST (U/L), median (IQR)	28.00 (21.00–44.00)	28.00 (21.00–44.50)	24.50 (19.00–40.50)
ALT (U/L), median (IQR)	27.00 (18.00–44.00)	27.00 (18.00–44.75)	25.00 (15.00–41.50)
Cholinesterase (U/L), median (IQR)	6635.00 (5575.00–7729.00)	6628.00 (5557.00–7660.50)	7034.00 (5814.75–8432.00)
Urea (mmol/L), median (IQR)	5.70 (4.64–6.88)	5.75 (4.65–6.90)	5.02 (4.40–6.25)
PT-INR, median (IQR)	1.03 (1.00–1.09)	1.03 (1.00–1.09)	1.00 (0.98–1.08)
Fibrinogen (g/L), median (IQR)	2.88 (2.46–3.42)	2.86 (2.45–3.38)	3.09 (2.64–4.03)
APTT (s), median (IQR)	36.90 (34.77–39.60)	37.00 (34.80–39.70)	35.90 (34.25–38.45)
Neutrophils (10 ⁹ /L), median (IQR)	3.01 (2.24–4.07)	2.98 (2.20–4.02)	3.59 (2.96–4.58)
WBC (10 ⁹ /L), median (IQR)	5.29 (4.25–6.72)	5.21 (4.19–6.62)	6.06 (5.14–7.20)
RBC (10 ¹² /L), median (IQR)	4.56 (4.21–4.88)	4.56 (4.21–4.88)	4.52 (4.17–4.88)
Hemoglobin (g/L), median (IQR)	142.00 (130.00–153.00)	142.00 (130.00–153.00)	143.00 (128.50–150.00)
Platelets (10 ⁹ /L), median (IQR)	157.00 (114.00–195.00)	156.00 (114.00–192.75)	167.00 (123.00–234.00)
Length of stay (days), median (IQR)	13.00 (10.00–17.00)	13.00 (10.00–17.00)	14.00 (12.00–17.50)
AFP (U/mL), median (IQR); measured n	13.10 (3.41–236.50); measured n=571	14.00 (3.62–248.00); measured n=529	3.42 (2.58–15.70); measured n=42
PIVKA-II (mAU/mL), median (IQR); measured n	218.25 (40.38–1537.45); measured n=436	258.71 (49.81–1626.22); measured n=403	29.54 (21.12–66.63); measured n=33

Notes: Continuous variables are presented as median (IQR). Categorical variables are presented as n (%). AFP and PIVKA-II summaries are among tested patients only; the number tested is shown after the summary.

Abbreviations: HCC, hepatocellular carcinoma; ICC, intrahepatic cholangiocarcinoma.

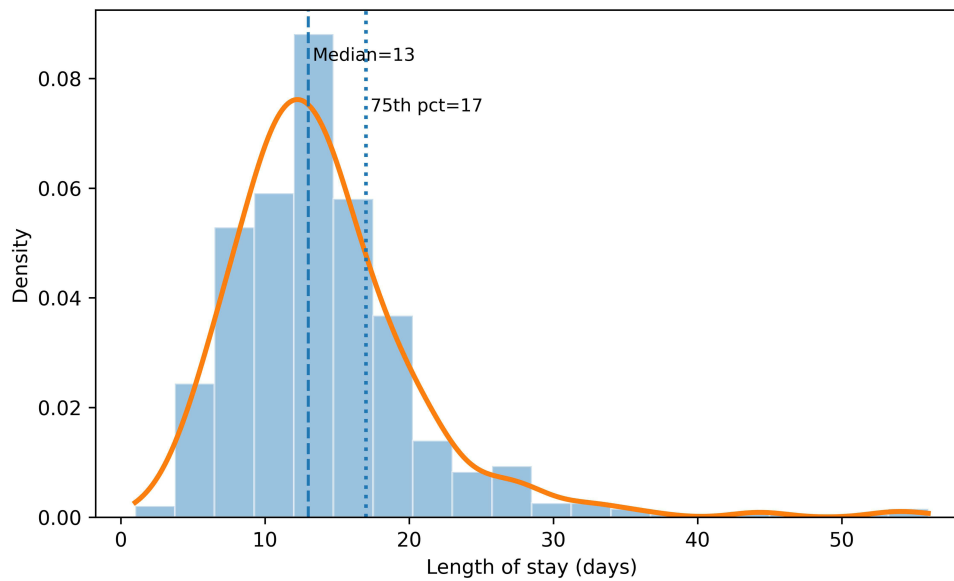


Figure 2 Distribution of LOS. Histogram (density-scaled) overlaid with a kernel density estimate. The dashed line indicates the median LOS; the dotted line indicates the 75th percentile used to define prolonged LOS in the secondary analysis.

Missing Data

Missing data patterns for key variables are presented in Table 2. Overall, the extent of missing data was low for most baseline variables. The majority of laboratory measurements and vital signs were available for >95% of patients. Two

Table 2 Missing Data Summary (Analysis Cohort, n=702)

Variable	Missing_n	Missing_%
PIVKA-II (mAU/mL)	266	37.9
AFP (U/mL)	131	18.7
Blood_type	77	11.0
DBP (mmHg)	25	3.6
SBP (mmHg)	25	3.6
DBIL (umol/L)	25	3.6
AST (U/L)	13	1.9
ChE (U/L)	13	1.9
Fbg (g/L)	10	1.4
BMI	10	1.4
APTT (s)	10	1.4
PT (s)	10	1.4
PT-INR	10	1.4
ALB (g/L)	9	1.3

(Continued)

Table 2 (Continued).

Variable	Missing_n	Missing_%
NE (10 ⁹ /L)	9	1.3
RBC (10 ¹² /L)	9	1.3
Hb (g/L)	9	1.3
WBC (10 ⁹ /L)	9	1.3
PLT (10 ⁹ /L)	9	1.3
Height (cm)	9	1.3
ALT (U/L)	9	1.3
TBIL (umol/L)	9	1.3
Weight (kg)	8	1.1
Urea (mmol/L)	6	0.9
Crea (umol/L)	6	0.9

Notes: Missingness is calculated within the primary analysis cohort (HCC or ICC). Pathological subtype is complete by design because records with missing subtype were excluded from the primary analysis cohort.

notable exceptions were the tumor markers: AFP was not measured in 18.7% of the cohort, and PIVKA-II was not measured in 37.9%. This reflects the selective ordering of these tests in clinical practice (they were likely omitted in cases judged low-risk for certain tumor markers). In our primary modeling, we did not impute these unmeasured values. Instead, as described, we incorporated indicator variables denoting whether AFP or PIVKA-II was tested for each patient. Analyses using the actual AFP and PIVKA-II values were then conducted on the subset of patients who had those results, as a form of sensitivity analysis. Aside from these markers, other variables had minimal missingness. A few laboratory values (eg, some instances of missing AST or ALT) were imputed using MICE, but since the rates were very low, this had negligible impact on overall results. The distribution of missing data did not appear to be associated with the outcome (for example, the proportion of missing AFP was similar among those with short vs long stays), supporting the use of the missing-at-random assumption for imputation.

Collinearity Diagnostics

Before fitting the multivariable model, we addressed collinearity among predictors. As expected, we found very high correlation between prothrombin time (PT) and its INR (International Normalized Ratio) (Table 3 shows that both convey

Table 3 Collinearity Diagnostics: Variance Inflation Factors (VIF) for the Prespecified Full Predictor Set

Predictor	VIF
RBC (10 ¹² /L)	4.140278192243204
Hb (g/L)	4.095773434657155
AST (U/L)	3.2731818905935954
ALT (U/L)	2.969145554696655

(Continued)

Table 3 (Continued).

Predictor	VIF
NE (10 ⁹ /L)	2.373557202501001
WBC (10 ⁹ /L)	2.319627088417671
ChE (U/L)	2.0801284085757623
ALB (g/L)	1.994257367378758
ABO_O	1.7578658404074363
SBP (mmHg)	1.7560291210530727
DBP (mmHg)	1.7339526499261892
ABO_B	1.686240928934406
Fbg (g/L)	1.6773685135020708
PLT (10 ⁹ /L)	1.6726954098073537
Smoke_Yes	1.635105028971763
Crea (umol/L)	1.621636436858601
Drink_Yes	1.591068193871093
Sex_Male	1.555216623522095
PT-INR	1.5507321525339444
Urea (mmol/L)	1.4954214477650005
APTT (s)	1.3674549753763734
Age (years)	1.322449042589236
ABO_AB	1.248043852990718
BMI	1.222757067667572
PIVKA_measured	1.1803745459344674
TBIL (umol/L)	1.1632448111847904
Viral_hepatitis_Yes	1.132296672916383
AFP_measured	1.1263621945794935
Path_subtype_ICC	1.1112032369231009

Notes: VIF values >5 were considered noteworthy and >10 severe. No predictor exceeded VIF=5 after removing PT(s) (in favor of PT-INR) and DBIL (in favor of TBIL).

similar information). We retained PT-INR in the model and dropped the raw PT value. Similarly, total bilirubin (TBIL) and direct bilirubin (DBIL) were strongly correlated (both reflect bilirubin levels); we kept TBIL as the more inclusive measure and omitted DBIL. AST and ALT were also correlated (both are liver enzymes indicating hepatic injury), but we included only one of them (we chose AST; however, ultimately neither AST nor ALT was among the core predictors in the final model based on our prespecified set). After removing these redundant variables, the remaining predictors showed no concerning multicollinearity – all variance inflation factors were below 5 (Table 3). Therefore, multicollinearity was effectively mitigated, and each predictor in the model contributed unique information.

Factors Associated with LOS (Negative Binomial Model)

We modeled LOS as a continuous count outcome using a negative binomial regression with a log link. The model included the core set of seven predictors described in Methods. The regression results (Table 4) identified four predictors with statistically significant associations with LOS, while the others did not reach significance. The significant associations were as follows:

Cholinesterase (ChE): Lower baseline ChE was strongly associated with longer hospital stay. Quantitatively, for each 1000 U/L increase in the ChE level, the expected LOS decreased by about 6.2% (incidence rate ratio [IRR] 0.938 per +1000 U/L, 95% confidence interval [CI] 0.920–0.957, $p < 0.001$). In other words, patients with low cholinesterase (indicative of poorer liver synthetic function or nutritional status) tended to have substantially longer postoperative stays compared to patients with higher cholinesterase levels. This makes clinical sense, as cholinesterase is produced by the liver and low levels may reflect impaired liver function or cachexia, which could slow recovery.

Pathological subtype: Patients with intrahepatic cholangiocarcinoma (ICC) had significantly longer LOS than those with hepatocellular carcinoma (HCC). Having ICC (versus HCC) was associated with approximately a 17.6% increase in expected LOS (IRR 1.176, 95% CI 1.038–1.333, $p = 0.011$). For example, if an HCC patient had an expected stay of 10 days (based on other factors), a similar ICC patient might be expected to stay roughly 11.8 days, all else being equal. This suggests that the nature of the tumor (and possibly the complexity of the surgery or perioperative course) differs between ICC and HCC in a way that affects recovery time.

Fibrinogen (Fbg): Higher fibrinogen levels were associated with longer LOS. Each increase of 1 g/L in fibrinogen predicted about a 3.6% longer stay (IRR 1.036 per +1 g/L, 95% CI 1.001–1.072, $p = 0.043$). For instance, a patient with fibrinogen of 4 g/L would have an expected LOS roughly 7% longer than a patient with fibrinogen of 2 g/L, holding other factors constant. Fibrinogen is an acute phase reactant and high levels can indicate systemic inflammation or a prothrombotic state; this result implies that patients with more inflammation (or a tendency to clot) preoperatively may recover more slowly or experience more complications, thus extending their hospital stay.

Sex: Sex emerged as a significant factor, with male sex associated with a shorter LOS compared to female sex. In the model, male patients had about an 8.9% shorter expected hospital stay than female patients (IRR 0.911, 95% CI 0.838–0.991, $p = 0.030$). To illustrate, if a female patient's predicted LOS is 14 days, a comparable male patient might be predicted around 12.8 days. This finding indicates that, in our cohort, women tended to have slightly longer hospitalizations than men. The reasons for this difference are not immediately clear from our data; it could be due to any number of factors (eg, differences in comorbidity profiles, complication rates, or perhaps socio-cultural factors affecting discharge), and it warrants further investigation.

Table 4 Multivariable Negative Binomial Regression for LOS (Core Clinical Model)

Predictor	IRR (95% CI)	p
Pathology: ICC vs HCC	1.176 (1.038–1.333)	0.011
Cholinesterase (per 1000 U/L)	0.938 (0.920–0.957)	<0.001
Fibrinogen (per 1 g/L)	1.036 (1.001–1.072)	0.043
Neutrophils (per 1×10^9 /L)	1.006 (0.991–1.022)	0.418
Diastolic BP (per 10 mmHg)	1.023 (0.993–1.054)	0.127
APTT (per 5s)	1.031 (0.989–1.074)	0.155
Sex: male vs female	0.911 (0.838–0.991)	0.030

Notes: Continuous predictors are scaled as shown in the predictor label to aid clinical interpretation. An IRR >1 indicates longer expected LOS; IRR <1 indicates shorter expected LOS, holding other variables constant.

Abbreviation: IRR: incidence rate ratio for LOS.

The other predictors in the model – namely baseline neutrophil count, diastolic blood pressure (DBP), and APTT – did not show statistically significant independent associations with LOS (each had $p > 0.05$). For example, although we included neutrophils as a marker of inflammation or infection risk, the model did not find a significant effect of higher neutrophil count on LOS when controlling for the other variables. It's possible that baseline neutrophil count at admission is not a strong predictor of prolonged stay (whereas an increase in neutrophils after surgery might signal a complication). Similarly, DBP and APTT did not significantly influence LOS in the multivariate context; their contributions may have been overshadowed by stronger predictors or they simply may not be major determinants of recovery length within normal ranges. These non-significant variables were retained in the core model to maintain our prespecified set, but one could consider a simplified model without them, which we explored and found the significant factors and performance remained essentially the same. Figure 3 provides a forest plot visualizing the IRRs and 95% CIs for all predictors in the model, highlighting the ones that were significant.

Model Performance and Calibration

The predictive performance of our core model for LOS is summarized in Table 5. The model's apparent (training data) mean absolute error (MAE) was 4.527 days, and the root mean squared error (RMSE) was 6.466 days. This means that on average, the model's LOS predictions were about 4.5 days off from the actual LOS, and there was considerable variability in error for individual predictions (as indicated by the RMSE). After adjusting for optimism using 1000 bootstrap resamples, the MAE increased slightly to 4.574 days and RMSE to 6.547 days, suggesting only minor overfitting. These error rates reflect the inherent difficulty of predicting exact LOS, which can be affected by unpredictable postoperative events.

We assessed model calibration by comparing predicted vs observed LOS. The calibration curve (Figure 4) indicates that the model is well calibrated across deciles of risk. The calibration intercept was -0.062 , very close to 0, indicating little systematic bias (the model slightly under-predicts LOS by a small constant, which is negligible). The calibration slope was 0.988, close to 1, indicating that the model's predictions maintain the correct spread and are neither over- nor under-dispersed relative to actual outcomes. In Figure 4, the points (observed mean LOS in each decile of predicted LOS) lie near the 45-degree line (perfect calibration), which visually confirms good agreement. We did not observe evidence of miscalibration in any particular range of predicted LOS – for example, for patients with very high predicted LOS, the observed LOS was also high, on average, with no systematic deviation.

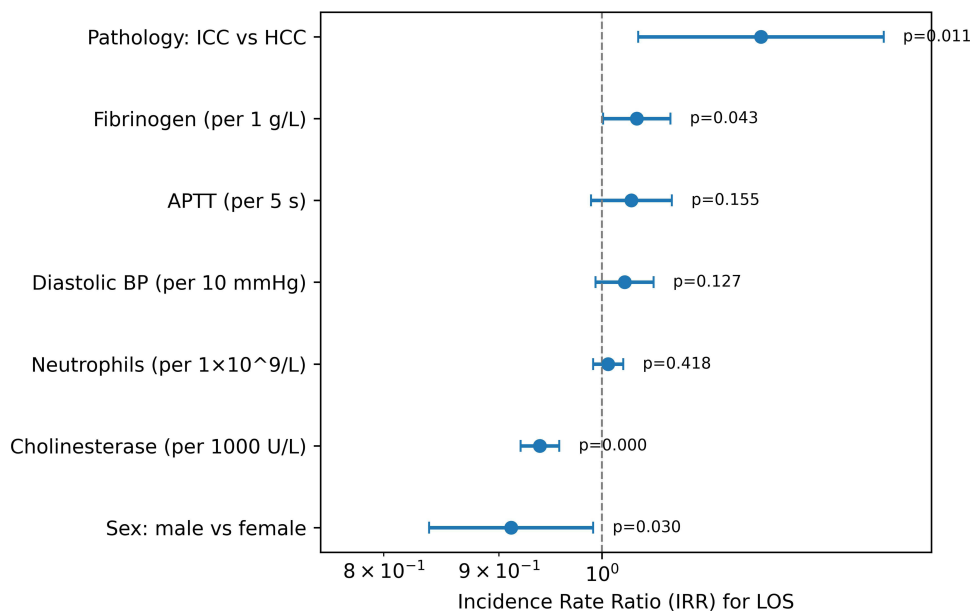


Figure 3 Forest plot of factors associated with LOS. Incidence rate ratios (IRRs) and 95% confidence intervals from the negative binomial regression core model are shown. Continuous predictors are scaled as labeled. $IRR > 1$ indicates a longer expected LOS; $IRR < 1$ indicates a shorter expected LOS.

Table 5 Model Performance and Internal Validation (Core Clinical Model)

Metric	Apparent	Optimism	Optimism-Corrected
MAE (days)	4.527	-0.047	4.574
RMSE (days)	6.466	-0.080	6.547
Calibration intercept ^a	-0.062	-	-0.062
Calibration slope ^a	0.988	-	0.988
Pseudo-R ² (McFadden)	0.016	-	0.016
Overdispersion alpha ^b	0.127	-	0.127
Bootstrap resamples	1000	-	1000

Notes: ^aCalibration intercept and slope were estimated by regressing log(LOS) on log(predicted mean LOS). ^bOverdispersion alpha was estimated from an intercept-only negative binomial model and used as a fixed dispersion parameter in GLM fitting.

Nomogram and Decision Curve Analysis

To facilitate practical use of the model, we constructed a nomogram (Figure 5) based on the core negative binomial model. The nomogram allows a clinician to easily calculate a patient's expected LOS by assigning points for each predictor (ChE level, fibrinogen level, subtype, sex, etc.) and summing them to get a total score, which corresponds to a predicted LOS on the bottom scale. For instance, a patient with very low ChE, high fibrinogen, ICC subtype, and female sex would score more points, translating to a higher predicted LOS, whereas a patient with high ChE, low fibrinogen, HCC, and male sex would score fewer points (predicting a shorter stay). This kind of tool can be useful at the bedside or in preoperative planning meetings to quickly estimate risk.

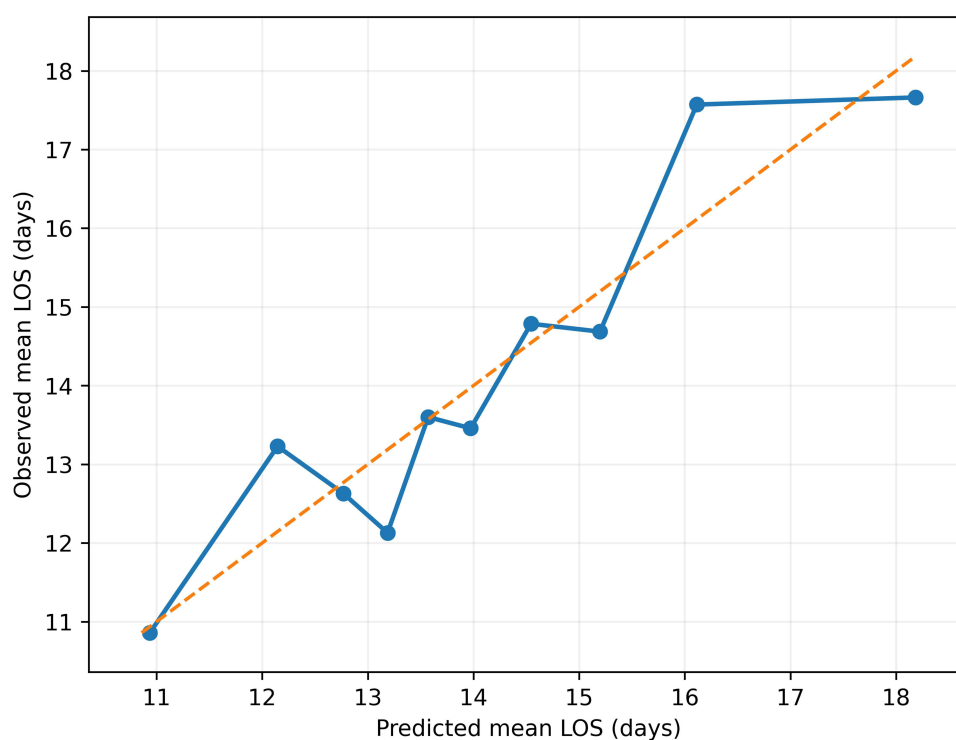


Figure 4 Calibration plot for continuous LOS prediction. Patients were grouped into deciles of predicted mean LOS. Points and the solid line represent the mean observed LOS versus mean predicted LOS in each decile; the dashed line is the line of identity (perfect calibration).

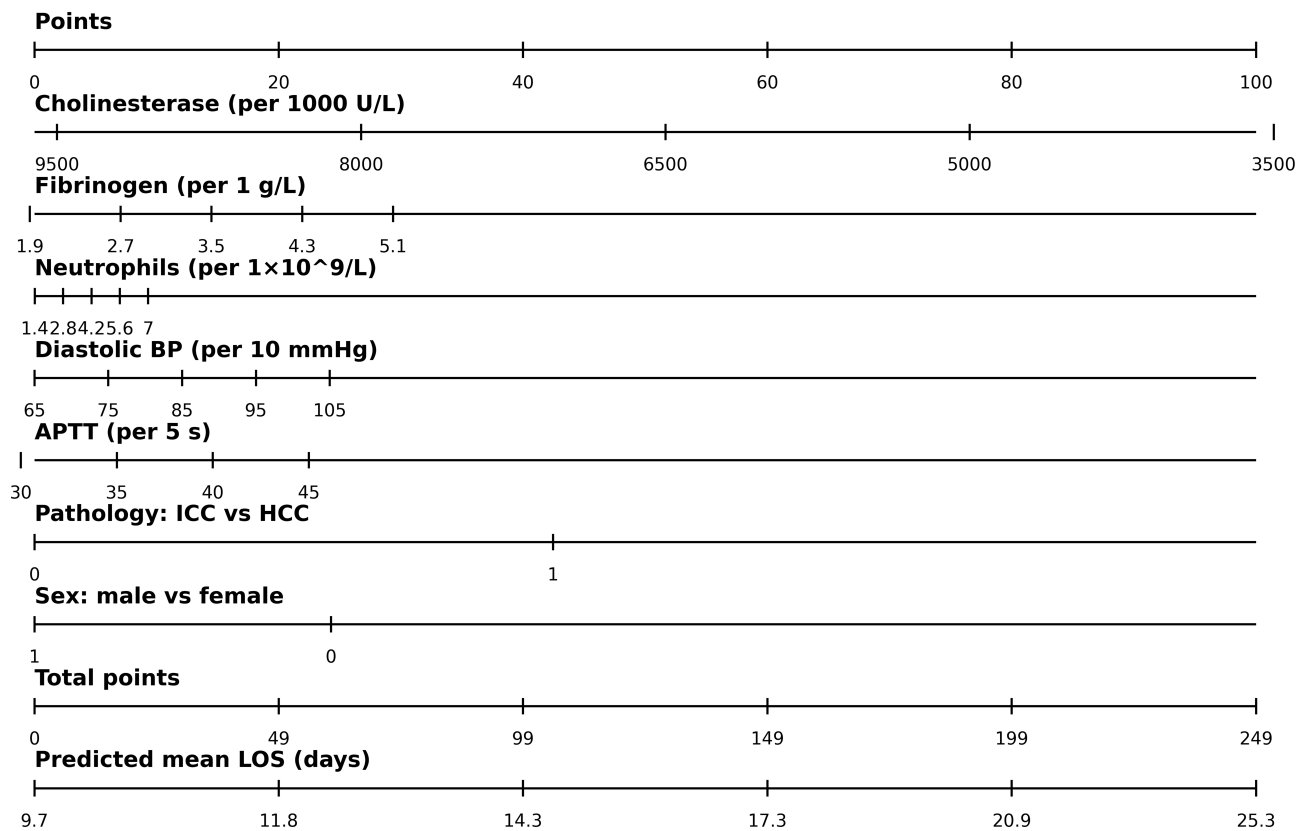


Figure 5 Nomogram to estimate expected LOS. The nomogram converts each predictor value into a point score and maps the total points to predicted mean LOS (in days) based on the core negative binomial model.

In our secondary analysis, we focused on the binary outcome of prolonged LOS (>17 days). We refit a logistic regression using the same core predictors to predict the probability of a prolonged stay for each patient. The performance of this classification model is reported in Table 6. The model's apparent AUC (on the training data) was 0.671, indicating modest discrimination. An AUC of 0.671 means that if we take two patients at random, one who had a prolonged stay and one who did not, the model would correctly assign a higher risk score to the one with prolonged stay about 67.1% of the time. This is only somewhat better than chance (50%), highlighting that distinguishing exactly who will have an

Table 6 Secondary Analysis: Logistic Regression for Prolonged LOS (>17 Days)

Predictor	OR (95% CI)	p
Pathology: ICC vs HCC	1.208 (0.605–2.411)	0.592
Cholinesterase (per 1000 U/L)	0.746 (0.666–0.836)	<0.001
Fibrinogen (per 1 g/L)	1.110 (0.933–1.322)	0.240
Neutrophils (per $1 \times 10^9/L$)	1.056 (0.980–1.137)	0.153
Diastolic BP (per 10 mmHg)	1.044 (0.890–1.224)	0.596
APTT (per 5s)	1.033 (0.826–1.293)	0.773
Sex: male vs female	0.517 (0.332–0.806)	0.004

Notes: Outcome was defined as LOS >17 days (75th percentile). Continuous predictors are scaled as shown in the predictor label. The model AUC was 0.671 in the analysis cohort (apparent).

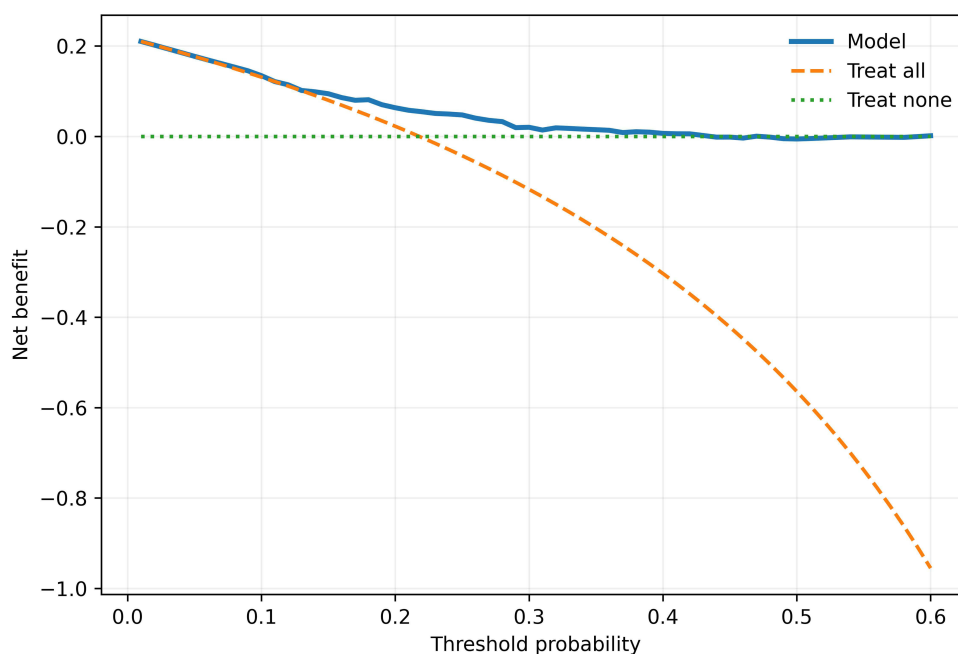


Figure 6 Decision curve analysis (DCA) for prolonged LOS (>17 days). The net benefit of using the model across a range of threshold probabilities is compared with treat-all and treat-none strategies.

extended stay is challenging. We did not have a separate external validation cohort for this logistic model, but the bootstrap validation for the continuous model suggests a slight drop could be expected if externally validated.

We also evaluated the clinical utility of the prolonged-stay model using decision curve analysis (Figure 6). The decision curve shows the net benefit of using our model to guide decisions (eg, an intervention to mitigate prolonged stay) across a range of threshold probabilities from about 0.04 (4%) to 0.60 (60%). The net benefit curve for the model is above those for the default strategies of “treat all” or “treat none” within this range, which implies that there are ranges of risk threshold where using the model to decide, for instance, which patients should receive extra intervention, would do better than simply intervening on everyone or no one. For example, if a clinician would consider a patient high-risk and worthy of intervention when their predicted probability of prolonged stay is, say, 20%, our model provides a positive net benefit in that scenario. In contrast, at very low thresholds (<4%) or very high thresholds (>60%), the model does not add value (which is typical: at extremely low thresholds, one might as well treat everyone; at extremely high thresholds, one might treat no one or only the obvious cases). Overall, the decision curve suggests that the model has potential clinical usefulness in identifying patients at risk for extended hospitalization, as long as the intervention threshold is set in a plausible range.

Discussion

In this retrospective cohort study of 702 primary liver cancer surgery patients, we developed and internally validated a risk prediction model for hospital length of stay. Our nomogram-based model, which utilizes routine preoperative clinical factors, identified several key predictors of prolonged hospitalization: low serum cholinesterase, high fibrinogen, intrahepatic cholangiocarcinoma subtype, and female sex. These findings provide insights into which patients may experience extended recovery periods after liver cancer surgery. To our knowledge, this work represents one of the novel attempts to create a predictive nomogram for LOS in primary liver cancer patients.

The significant predictors in our model largely reflect intrinsic patient and disease characteristics. Cholinesterase and fibrinogen are laboratory markers related to liver function, nutrition, and inflammation; their association with LOS highlights the importance of a patient’s baseline physiological reserve and systemic inflammatory state in recovery. Pathological subtype (ICC vs HCC) being a determinant suggests that the nature of the tumor and the surgery (eg,

anatomic location and complexity of resection) can influence postoperative course. The observation that female sex was linked to longer stays is intriguing and warrants further exploration. Meanwhile, other factors that one might expect to influence LOS (such as baseline neutrophil count as an inflammation marker) did not show independent effects, possibly because their impact manifests only when postoperative complications occur (which our preemptive model does not directly capture).

Our approach in this study complements and extends prior research on postoperative LOS. Many contemporary studies have focused on interventions to reduce LOS, such as enhanced recovery pathways, optimized pain management, and early mobilization. For example, efforts like adjusting perioperative medication regimens to prevent adverse events or implementing exercise programs during cancer treatment have been shown to improve outcomes and potentially shorten hospital stays. By contrast, our study zeroes in on the patient's initial risk profile. We asked: given the patient's condition upon admission, who is predisposed to a longer stay? Identifying these individuals early could allow clinicians to apply existing interventions (like ERAS protocols) more intensively or tailor additional support where needed.

Cholinesterase (ChE): We found that a lower ChE level is a strong predictor of prolonged LOS. Cholinesterase is synthesized in the liver and is a sensitive marker of hepatic synthetic function and overall nutritional status. A low ChE often indicates significant liver dysfunction or malnutrition/cachexia. In our cohort, patients with low ChE stayed considerably longer in the hospital. This aligns with clinical expectations: patients with compromised liver function might have impaired healing, greater risk of complications (like infections or ascites), and slower recovery, all of which can prolong hospitalization. ChE has long been used as a liver function test, and our findings reinforce its relevance in surgical outcomes. This suggests that preoperative nutritional and liver function optimization (for instance, improving protein status, which would reflect in higher albumin and potentially ChE levels) might be important for reducing LOS.

Fibrinogen: Elevated fibrinogen was associated with longer LOS in our study. Fibrinogen is a coagulation factor and an acute-phase reactant that rises in response to inflammation. A high fibrinogen level can indicate that a patient has an underlying inflammatory state or a tendency towards thrombosis. This connection with prolonged LOS is biologically plausible: systemic inflammation can impair recovery and wound healing, and a hypercoagulable state could lead to complications like thrombosis. Supporting this, there is evidence that inflammation and coagulation biomarkers correlate with outcomes in cancer patients. For instance, hypoalbuminemia (a sign of inflammation and poor nutrition) has been linked to higher risk of venous thromboembolism and mortality in cancer populations. In the context of liver cancer, a composite index like the albumin-to-fibrinogen ratio (AFR) has been shown to predict LOS – Li et al reported that a lower AFR (meaning lower albumin, higher fibrinogen) was associated with longer postoperative stays. This is consistent with our finding that high fibrinogen (and by implication, often lower albumin in the same patients) contributes to extended hospitalization. Pathophysiologically, fibrinogen plays a multifaceted role in tissue injury and repair; it is not only critical for clot formation but also interacts with inflammatory pathways. High fibrinogen levels often reflect an ongoing acute-phase response, which can mean the patient's body is already in a stressed or reactive state before surgery. Such patients might have less reserve to handle surgical stress or might be more prone to complications like infections or thromboses, resulting in longer recovery times.^{15,16} Our results underscore the importance of preoperative inflammatory status: it might be beneficial to address and mitigate systemic inflammation (when possible) before surgery, and certainly to closely monitor patients who enter surgery with elevated inflammatory markers.

Pathological subtype (ICC vs HCC): We found that patients with intrahepatic cholangiocarcinoma have significantly longer hospital stays than those with hepatocellular carcinoma, even after accounting for other factors. There are a few possible explanations for this. Surgically, ICC resections can be more complex; they may require more extensive liver resection or additional procedures like bile duct reconstruction, leading to longer operative times and potentially more postoperative complications (such as bile leaks or abscesses). ICC patients might also have different baseline characteristics – for instance, some ICCs are associated with underlying primary sclerosing cholangitis or other conditions that could complicate recovery. Additionally, differences in tumor biology and patient profiles (ICC patients in our cohort were fewer and might have been treated more aggressively or at a more advanced stage on average) could contribute. While our data cannot pinpoint the exact reason, the message is that having ICC as opposed to HCC is a flag for a potentially longer hospital course. Clinicians should be aware that ICC patients may need more resources or longer convalescence. This is in line with clinical intuition and some reports that cholangiocarcinoma surgeries have distinct

challenges compared to HCC surgeries. Our model quantified this difference, which had not been explicitly done in many prior studies.

Sex differences: An interesting finding was that female patients had somewhat longer LOS than male patients, on average. The effect size was about a 9% longer stay, which is modest but notable. The reasons for this are not immediately evident. It could be a coincidental finding, or it might hint at underlying differences not captured in our data. For example, if female patients were, on average, older or had more comorbidities (in our cohort, the proportion of females was smaller, so even a few outliers could influence the result), that might explain a longer recovery. It's also possible that there are sex-based physiological differences in recovery or pain management that influenced discharge timing. Another consideration is that social factors sometimes play a role in discharge planning – it is purely speculative, but differences in family support or caregiver availability could potentially lead to differing discharge readiness between male and female patients. Given the limitations of our retrospective data, we can only acknowledge this finding without overinterpreting it. We have flagged it as an area for further study. Clinicians might take it into account in a broad sense by ensuring that all patients, regardless of sex, receive adequate support. The result serves as a reminder that individual patient factors (some of which may correlate with sex, such as body composition and hormonal influences) can subtly affect recovery trajectories.

Neutrophils and other non-significant factors: We included neutrophil count as a candidate predictor because an elevated neutrophil count can indicate systemic inflammation or infection risk. Interestingly, baseline neutrophil count was not a significant predictor of LOS in our model. This might be because an initial neutrophil count within normal range does not necessarily predict who will get an infection after surgery. It is the occurrence of postoperative complications (like infections) that really drives LOS up, and those events are not captured by a preoperative neutrophil count. So while a patient's neutrophil count on admission might not foretell a prolonged stay, a spike in neutrophils a few days after surgery would, but that is outside the scope of our preoperative model. Similarly, diastolic blood pressure and APTT were not significant predictors in the multivariable context. They may not vary enough or have a strong enough relationship with LOS in our cohort, especially once other factors are accounted for. It is worth noting that our model's focus was on preoperative data; factors like intraoperative blood loss, surgery duration, and postoperative events likely play a huge role in determining LOS, but we could not include those here. These limitations are discussed below.

Model performance: Overall, our predictive model demonstrated moderate performance. The discrimination was fair (AUC ~0.67 for classifying prolonged stays), indicating the model can identify high-risk patients better than chance, but there is considerable overlap in predicted risk between those who did and did not have prolonged stays. This level of performance is not unexpected given the complexity of LOS as an outcome. Many factors influencing LOS are unmeasured or inherently unpredictable (for instance, an unforeseen complication can extend a stay dramatically). In fact, a systematic review of hospital LOS prediction methods found that many models achieve only modest accuracy, often in the 0.60–0.75 AUC range at best. Our model falls into that range. We should also interpret the mean absolute error of ~4.5 days in context: if the typical LOS is about 13 days, an error of 4–5 days is not negligible. It means for an individual patient, the prediction might be off by nearly a week in some cases. This underscores that our model is better suited for risk stratification (eg, identifying a subset of patients likely to have significantly longer stays) rather than for pinpointing an exact discharge date.

On a positive note, the model's calibration was very good, which means it estimates the probability or expected value of LOS reasonably well on a population level. For example, if the model predicts an average LOS of 15 days for a group of patients, that group's actual average LOS was indeed around 15 days. This reliability is important for practical use, because it means we could trust the model's risk stratification to not be systematically biased. The decision curve analysis further suggested that using the model has potential net benefit over certain ranges of decision thresholds – in practical terms, if we set a policy like “provide an extra intervention to any patient predicted to have >X% chance of prolonged stay”, our model would help make that policy more efficient than a blanket or no intervention approach for a reasonable range of X values. For instance, suppose a hospital can offer a special postoperative rehabilitation program to a limited number of patients; using our model to select those patients (say, those with predicted risk above 30%) would likely identify a group where that program averts more prolonged stays than it wastes resources (net benefit analysis quantitatively supports this kind of targeted approach).

Limitations

This study has several important limitations. First, it is a retrospective analysis from a single tertiary hospital, which may limit the generalizability of the findings. Healthcare delivery processes and patient populations can differ between hospitals (and countries), and our model is currently validated only on internal data. External validation is needed to verify its performance elsewhere. Second, as mentioned earlier, we did not have data on certain key factors that undoubtedly affect LOS, such as the details of the surgery (eg, whether the surgery was performed laparoscopically or via open incision, the extent of hepatic resection, concurrent procedures) and postoperative complications. These variables could not be extracted from our electronic system in a reliable manner. Their absence likely reduced the model's predictive power – in essence, our model had to rely only on preoperative factors, some of which serve as proxies but none of which can perfectly predict an event like a complication. For example, an open surgery or a major complication like an infection can add many days to LOS, but our model would not know if a given patient will encounter those. Future efforts should incorporate such variables, perhaps by moving to a dynamic model that updates predictions based on intra- and post-operative events. Third, we did not explicitly include a measure of comorbidity or general health status (such as ASA score or Charlson Comorbidity Index) in our model. Comorbid conditions and overall frailty surely influence recovery – for instance, a patient with severe COPD or heart failure might have a longer stay due to slower mobilization or higher risk of complications. While some of this may be indirectly captured through laboratory values or age, it is not directly accounted for. The Global Burden of Disease study (GBD 2017) for China, for instance, lists hypertension and smoking, as top contributors to health risks; patients with such risk factors might also have prolonged recoveries. We did not have granular data on comorbidities beyond what was in lab results. This is a limitation of our dataset and focusing on surgical admissions. Fourth, our study period spanned nearly four years, including the COVID-19 pandemic era. Changes in hospital policy (eg, infection control measures and resource allocation) during COVID-19 might have affected LOS. We did not adjust for temporal trends or the potential impact of COVID-19 waves on hospital stay (for example, whether patients stayed longer due to quarantine policies at certain times). This could introduce some bias or noise. However, our data did not show a noticeable trend in LOS over time, and fortunately, none of the patients in our cohort had COVID-19 during their admission (as elective surgeries were done in COVID-negative patients). Lastly, while we used bootstrap validation to correct for optimism, no independent external validation was performed. The AUC of 0.67 is the apparent performance; an external test might yield a lower value.

Implications and Future Directions

Despite its limitations, our study provides a framework for anticipating prolonged hospital stays in liver cancer patients. Clinically, if we can identify high-risk patients early (for example, those with low ChE and high fibrinogen), healthcare teams can take proactive steps: these might include optimizing the patient's condition pre-surgery (nutritional support and prehabilitation exercises), planning for closer monitoring or ICU care immediately after surgery, involving multidisciplinary teams (physiotherapy, nutrition, and social work) early in the postoperative phase, and managing expectations by informing patients and families that a longer stay might be needed. From a health system perspective, predictive models like ours could be integrated into electronic health records to trigger alerts or allocate resources (like assigning more experienced nursing staff or prioritizing certain patients for enhanced recovery pathways). Our decision curve analysis implies that such targeted intervention strategies could be beneficial.

However, given the model's modest discrimination, it should not be used in isolation to make clinical decisions. Rather, it can complement clinical judgment. For instance, a surgeon might intuitively identify some high-risk patients (very advanced age, obvious frailty). The model might highlight some less obvious ones (perhaps a patient with normal appearance but very low ChE and high fibrinogen, indicating hidden risk). Combining model output with clinician insight will likely be the way to maximize accuracy. We also suggest that future models incorporate more variables for better accuracy. Machine learning approaches could be explored as well, especially if larger multi-center datasets become available, but one must ensure interpretability and validation.

Our conclusion from this work is that a patient's baseline liver function and inflammatory status significantly influence their recovery trajectory after liver surgery. Interventions aiming to improve these aspects preoperatively (for example, nutritional supplements to raise albumin and ChE, or treating underlying infections/inflammations to lower

fibrinogen) might translate into shorter stays, although that needs to be tested in prospective studies. Additionally, our findings justify closer surveillance for patients identified as high risk; for example, a patient with very low ChE could be monitored more aggressively for complications or kept in step-down units longer, possibly preventing adverse events that would prolong LOS.

Conclusion

We developed a nomogram model based on preoperative factors (serum cholinesterase, fibrinogen, pathological subtype, sex, etc.) to predict the risk of extended hospital stay in primary liver cancer patients undergoing surgery. This predictive tool showed good calibration and modest discriminatory ability. It can potentially aid clinicians and care teams in identifying patients who are likely to require prolonged hospitalization, allowing for early planning and targeted interventions to improve postoperative recovery. For instance, patients predicted to be high-risk could be candidates for more intensive ERAS protocols or additional supportive care, which may help mitigate length of stay. Our work underscores the impact of a patient's baseline liver function and inflammatory state on surgical outcomes, aligning with the principles of fast-track surgery that emphasize optimizing patient condition and anticipating needs.

However, it is important to note that the model's accuracy is limited. It should be considered a preliminary risk stratification tool rather than a definitive predictor. The model's predictions should complement, not replace, clinical judgment. Before this model (or a similar predictive approach) can be implemented broadly, it requires external validation in other cohorts and settings to ensure its generalizability. Additionally, incorporating intra- and postoperative factors in future models could enhance predictive performance. Ultimately, with further refinement and validation, such predictive models could contribute to personalized perioperative care – by flagging high-risk patients, guiding resource allocation, and informing clinical decisions – thereby improving outcomes and efficiency in the management of primary liver cancer surgery patients.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, the authors utilized ChatGPT 3.5 to assist with language polishing. After drafting content, the authors employed the AI tool to improve clarity and grammar. All AI-generated suggestions were carefully reviewed and edited by the authors to ensure accuracy and consistency with our data. The authors take full responsibility for the content of this manuscript.

Acknowledgment

A preprint of this study was posted on ResearchSquare (DOI: 10.21203/rs-4419695/v1) prior to peer-reviewed publication.

Funding

This work was supported by the National Key R&D Program of China under Grant No. 2022YFB2703304.

Disclosure

The authors declare that there are no conflicts of interest relevant to this work.

References

1. Konyn P, Ahmed A, Kim D. Current epidemiology in hepatocellular carcinoma. *Expert Rev Gastroenterol Hepatol*. 2021;15(11):1295–1307. doi:10.1080/17474124.2021.1991792
2. Rungay H, Arnold M, Ferlay J, et al. Global burden of primary liver cancer in 2020 and predictions to 2040. *J Hepatol*. 2022;77(6):1598–1606. doi:10.1016/j.jhep.2022.08.021
3. Massarweh NN, El-Serag HB. Epidemiology of hepatocellular carcinoma and intrahepatic cholangiocarcinoma. *Cancer Control*. 2017;24(3):1073274817729245. doi:10.1177/1073274817729245
4. Rungay H, Ferlay J, de Martel C, et al. Global, regional and national burden of primary liver cancer by subtype. *European J Cancer*. 2022;161:108–118. doi:10.1016/j.ejca.2021.11.023

5. Shi JF, Cao M, Wang Y, et al. Is it possible to halve the incidence of liver cancer in China by 2050? *Int J Cancer*. 2021;148(5):1051–1065. doi:10.1002/ijc.33313
6. Padilla RM, Mayo AM. Patient survival and length of stay associated with delayed rapid response system activation. *Crit Care Nurs Quart*. 2019;42(3):235–245. doi:10.1097/cnq.0000000000000264
7. Peters GM, Kooij L, Lenferink A, van Harten WH, Doggen CJM. The effect of telehealth on hospital services use: systematic review and meta-analysis. *J Med Internet Res*. 2021;23(9):e25195. doi:10.2196/25195
8. Smith TW, Wang X, Singer MA, Godellas CV, Vaince FT. Enhanced recovery after surgery: a clinical review of implementation across multiple surgical subspecialties. *Am J Surg*. 2020;219(3):530–534. doi:10.1016/j.amjsurg.2019.11.009
9. Lee SY, Lee SH, Tan JHH, et al. Factors associated with prolonged length of stay for elective hepatobiliary and neurosurgery patients: a retrospective medical record review. *BMC Health Serv Res*. 2018;18(1):5. doi:10.1186/s12913-017-2817-8
10. Han TS, Murray P, Robin J, Wilkinson P, Fluck D, Fry CH. Evaluation of the association of length of stay in hospital and outcomes. *Int J Qual Health Care*. 2022;34(2). doi:10.1093/intqhc/mzab160
11. Lequertier V, Wang T, Fondrevelle J, Augusto V, Duclos A. Hospital length of stay prediction methods: a systematic review. *Med Care*. 2021;59(10):929–938. doi:10.1097/mlr.0000000000001596
12. Gokhale S, Taylor D, Gill J, et al. Hospital length of stay prediction tools for all hospital admissions and general medicine populations: systematic review and meta-analysis. *Front Med*. 2023;10:1192969. doi:10.3389/fmed.2023.1192969
13. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatric Res*. 2011;20(1):40–49. doi:10.1002/mpr.329
14. Riley RD, Ensor J, Snell KIE, et al. Calculating the sample size required for developing a clinical prediction model. *BMJ*. 2020;368:m441. doi:10.1136/bmj.m441
15. Luyendyk JP, Schoenecker JG, Flick MJ. The multifaceted role of fibrinogen in tissue injury and inflammation. *Blood*. 2019;133(6):511–520. doi:10.1182/blood-2018-07-818211
16. Donkin R, Fung YL, Singh I. Fibrinogen, coagulation, and ageing. *Sub-Cellular Biochemistry*. 2023;102:313–342. doi:10.1007/978-3-031-21410-3_12

Journal of Hepatocellular Carcinoma

Publish your work in this journal

The Journal of Hepatocellular Carcinoma is an international, peer-reviewed, open access journal that offers a platform for the dissemination and study of clinical, translational and basic research findings in this rapidly developing field. Development in areas including, but not limited to, epidemiology, vaccination, hepatitis therapy, pathology and molecular tumor classification and prognostication are all considered for publication. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/journal-of-hepatocellular-carcinoma-journal>

Dovepress
Taylor & Francis Group