Research Square

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

Predicting recurrence of biliary colic after conservative management in the emergency department: a prospective cohort study with a parametric Weibull survival model

David Alejandro Benalcázar Villavicencio
david.benalcazar@ucuenca.edu.ec

University of Cuenca
Diana Gabriela Ibadango Ávila
University of Cuenca

Article

Keywords: Biliary colic, Cholelithiasis, Recurrence, Emergency department, Conservative management, Cholecystectomy, Survival analysis, Prognostic model, Weibull distribution

Posted Date: November 1st, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11297016/v1

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

Additional Declarations: No competing interests reported.

Abstract

Background

Biliary colic is usually managed conservatively in the emergency department (ED), yet recurrence is frequent and its timing is poorly characterized, particularly in Latin American settings with limited access to elective cholecystectomy. We aimed to develop and internally validate a survival-based model predicting recurrence of biliary colic after conservative ED management.

Methods

Single-center prospective cohort of adults with uncomplicated biliary colic and ultrasound-confirmed gallstones discharged from the ED of a tertiary hospital in Cuenca, Ecuador, followed for up to 365 days. The primary outcome was symptomatic recurrence requiring an ED visit; cholecystectomy and biliary complications before recurrence were censored. Predictors were screened with Cox regression; multivariable Cox and parametric Weibull models were compared by Harrell's C-index and the Akaike information criterion. Internal validation used 1000 bootstrap resamples; calibration was assessed at 7, 15, 30 and 60 days overall and by predicted-risk group. A sensitivity analysis used Fine–Gray competing-risks regression.

Results

Of 507 patients (median age 38 years; 76.5% women), 223 (44.0%) recurred; 25% of recurrences occurred within 7 days of discharge and recurrence-free survival was 67.8% at 30 days and 43.1% at 365 days. Biliary complications developed in 10.7% and only 7.3% underwent cholecystectomy before recurrence. In the final Weibull model, a positive Murphy's sign (hazard ratio [HR] 7.09; 95% CI 4.73–10.61) and diabetes mellitus (HR 2.18; 95% CI 1.33–3.58) were independent predictors. The shape parameter ($\gamma = 0.419$; 95% CI 0.376–0.468) indicated a monotonically decreasing baseline hazard. Discrimination was acceptable (C-index 0.692; bootstrap-corrected 0.692, 95% CI 0.665–0.723) with good calibration from 30 days onwards and underestimation of very early risk in the highest-risk group. Fine–Gray estimates were consistent (subdistribution HR 5.73 and 2.06, respectively).

Conclusions

Recurrence after conservative management of biliary colic is frequent and concentrated in the first days after discharge. Two bedside variables, Murphy's sign and diabetes mellitus, identify patients at high risk of early recurrence who may benefit from prioritized surgical evaluation and structured early follow-up. External validation in multicenter cohorts is required.

Trial registration:

Not applicable (observational study).

Background

Gallstone disease is among the most prevalent digestive disorders worldwide, affecting 10–15% of adults in the United States [1] and up to 26.7% of Hispanic women [2]. Its global burden rose steadily between 1990 and 2019, from 127 to 193 million prevalent cases [3]. In Ecuador, cholelithiasis was the leading cause of hospital discharge in 2024, with 51 781 discharges, a 41.6% increase over seven years [4]. Uncomplicated symptomatic cholelithiasis, or biliary colic, is the most common clinical presentation and accounts for an estimated 335 000 emergency department (ED) visits per year in the United States alone [5].

Laparoscopic cholecystectomy is the definitive treatment, but in the absence of complications most patients are discharged from the ED with analgesia and referral for elective surgery [6]. Fewer than half of them undergo cholecystectomy within one year, and a substantial proportion re-present with recurrent pain or biliary complications in the interval [5, 7, 8]. The reported recurrence risk is heterogeneous, ranging from 18.6–37.1% in population-based series to nearly 90% at 10 years in older cohorts, with most events clustering in the first two years [8–10]. In the United States alone, gallstone disease generates approximately 1.1 million ED visits, 615 000 hospital discharges and 2000 deaths each year [11]. Delaying surgery therefore exposes patients to recurrence, cholecystitis,

choledocholithiasis, cholangitis or biliary pancreatitis [12], while systematic early cholecystectomy may lead to unnecessary operations; the 5-year follow-up of the SECURE trial showed that a restrictive surgical strategy did not increase biliary complications, but that only two-thirds of operated patients became pain-free [13]. Recent randomized evidence also suggests that conservative management is cheaper over 24 months without a significant quality-of-life penalty [14], leaving the decision to individual surgeon judgment and producing wide variation in practice [15].

Emergency surgeons and emergency physicians therefore lack objective criteria to decide, at the index visit, which patients with uncomplicated biliary colic can safely wait and which should be prioritized for early cholecystectomy. The RELAPSTONE multicenter study recently described recurrence patterns after a first admission for symptomatic gallstone disease using Cox regression, and identified biliary colic as the presentation with the highest relapse risk [16]. However, semiparametric models leave the shape of the baseline hazard unspecified and cannot establish whether recurrence risk is highest immediately after discharge or accumulates steadily over time, information that directly informs the timing of surgical evaluation. Moreover, most available evidence comes from Europe, the United States and Japan, and may not transfer to Latin American health systems, where access to elective surgery is constrained and follow-up is less structured.

We conducted a prospective cohort study of patients discharged from the ED with uncomplicated biliary colic in a tertiary public hospital in Ecuador. Our aims were to describe the incidence and timing of symptomatic recurrence under conservative management, to identify independent predictors available at the index visit, and to develop and internally validate a parametric survival model that explicitly characterizes the temporal dynamics of recurrence risk.

Methods

Study design and setting

We performed a single-center, prospective, observational cohort study in the ED of Hospital Vicente Corral Moscoso, a tertiary public teaching hospital in Cuenca, Ecuador, which attended 1178 patients with ICD-10 codes K80, K80.2 and K80.8 in 2024. Patients were enrolled consecutively between 1 January and 31 December 2025 and followed until recurrence, cholecystectomy, biliary complication or the end of the observation period on 31 December 2025 (maximum follow-up 365 days). The study was reported according to the TRIPOD statement for multivariable prediction models [17] and the STROBE recommendations for cohort studies [18] (Additional file 2).

Participants

Eligible patients were aged over 16 years, had ultrasonographic evidence of gallstones or biliary sludge at the index visit, and presented with a typical episode of biliary pain fulfilling Rome IV criteria (constant epigastric or right upper quadrant pain of moderate to severe intensity lasting at least 30 minutes) [19]. All participants provided written informed consent. We excluded patients with clinical, biochemical or imaging evidence of complicated gallstone disease at the index visit (acute cholecystitis, choledocholithiasis, cholangitis, pancreatitis or biliary fistula), suspected functional biliary disorders, suspected or confirmed biliary or pancreatic malignancy, previous cholecystectomy, patients scheduled for urgent surgery at first contact, and records with more than 20% missing data in the variables critical for survival analysis.

The minimum sample size was calculated for an expected proportion of 50%, 5% precision and 95% confidence, corrected for a finite population of 1178 (n = 290) and inflated by 10% for losses to follow-up (n = 319). A total of 507 patients were ultimately enrolled. With 223 recurrence events and two predictors in the final model, the events-per-variable ratio was well above conventional thresholds for model development [20].

Data collection and follow-up

A single investigator (DABV) collected all data prospectively on a structured case report form at the index visit, through direct patient interview, standardized clinical examination with calibrated instruments, and review of the official laboratory and ultrasound reports. Candidate predictors included demographic and anthropometric variables; comorbidities (diabetes mellitus, hypertension); clinical features (symptom duration, vital signs, nausea, fever, Murphy's sign, first-ever episode, pharmacological treatment at first contact);

laboratory values (leukocyte count, neutrophil percentage, gamma-glutamyl transferase, alkaline phosphatase, aspartate and alanine aminotransferases, total, direct and indirect bilirubin); and ultrasound findings (number and size of stones, sludge, impacted stone in the gallbladder neck, wall thickening ≥ 4 mm, hydrops, sclero-atrophic gallbladder and wall–echo–shadow sign). Murphy's sign was assessed clinically as inspiratory arrest on palpation of the right upper quadrant. All predictors were measured before the outcome and without knowledge of it.

Follow-up was performed by systematic review of the institutional electronic records, including ED visits, outpatient consultations, operative records and pathology reports. No active telephone follow-up was undertaken. Data were entered into Microsoft Excel and analyzed in Stata 17.0 (StataCorp, College Station, TX).

Outcomes

The primary outcome was symptomatic recurrence, defined as a new episode of biliary pain prompting an ED visit after discharge from the index episode. The secondary outcome was a biliary complication (acute cholecystitis or acute-on-chronic cholecystitis with clinical and laboratory decompensation, biliary pancreatitis, choledocholithiasis, cholangitis, Mirizzi syndrome or gallstone ileus). Follow-up time was measured from ED discharge to the first of recurrence, cholecystectomy, biliary complication or end of observation.

Cholecystectomy before recurrence was treated as censoring rather than as a competing risk because it is a treatment decision driven by the patient's own clinical course and affected only 7.3% of the cohort. Biliary complication before recurrence (5.9%) is a true competing event; it was nevertheless censored because the research question was cause-specific, namely the temporal mechanism of recurrence among patients who remain at risk, which is the quantity estimated by the Weibull hazard. The robustness of this decision was examined in a competing-risks sensitivity analysis.

Statistical analysis

Continuous variables were tested for normality with the Kolmogorov–Smirnov test and summarized as mean ± standard deviation or median (interquartile range, IQR); categorical variables as counts and percentages. Recurrence-free survival was estimated with the Kaplan–Meier method, and the temporal pattern of events was described with a smoothed hazard function and a kernel density of time to recurrence.

Each candidate predictor was first examined in a univariable Cox model; complementary univariable Weibull models yielded virtually identical estimates and are not shown. The proportional hazards (PH) assumption was tested with scaled Schoenfeld residuals [21]. Because Murphy's sign, the strongest predictor, violated the PH assumption, univariable screening was repeated in Cox models stratified by Murphy's sign, which allows each stratum its own baseline hazard without imposing a functional form on the effect of Murphy's sign. Some continuous variables were dichotomized at clinically interpretable thresholds (age ≥ 30 years; leukocytes ≥ 18 × 10^3/μL) to favor bedside applicability.

Multivariable modeling started from the covariates that were clinically relevant according to univariable screening and the literature (diabetes mellitus, biliary sludge, impacted stone and Murphy's sign). Four specifications were estimated: Cox and Weibull models, each with Murphy's sign either as a covariate or as a stratification variable. Since parametric models do not admit stratification in the strict sense, the "stratified" Weibull specification retained Murphy's sign as a covariate and additionally allowed it to modify the shape parameter. Non-significant covariates were removed by backward elimination if their removal did not worsen the Bayesian information criterion. Models were compared on discrimination (Harrell's C-index and Somers' D, computed on the linear predictor with correction for censoring) and, within the same likelihood family, on the Akaike information criterion (AIC); AIC was not compared across the Cox (partial likelihood) and Weibull (full likelihood) families. The Weibull model was chosen a priori as the parametric candidate because the smoothed hazard was monotonically decreasing.

The final model was checked for the plausibility of the Weibull form and proportionality with log[− log S(t)] versus log(t) plots by Murphy strata, and for global fit with Cox–Snell residuals. The sensitivity of the Schoenfeld test to the time scale (untransformed, Kaplan–Meier, rank and log time) was examined. Discrimination was reported as Harrell's C with 95% confidence intervals (CI). Internal validation followed Harrell's optimism-correction procedure with 1000 bootstrap resamples [22]. Patients were grouped into low, intermediate and high predicted-risk tertiles of the linear predictor and Kaplan–Meier curves were compared across groups.

Calibration was assessed by comparing model-predicted with Kaplan–Meier observed survival at 7, 15, 30 and 60 days, overall and within each risk group. As a sensitivity analysis, cumulative incidence functions accounting for competing biliary complications were compared with Kaplan–Meier estimates, and a Fine–Gray subdistribution hazards model [23] including a time-dependent term for Murphy's sign was fitted. Two-sided $p < 0.05$ was considered statistically significant.

Results

Cohort characteristics

Of 508 enrolled patients, one with a follow-up time of zero (enrolled on the last day of the observation period) was excluded, leaving 507 patients for analysis. Baseline characteristics are summarized in Table 1. The median age was 38 years (IQR 28–52), 76.5% were women and the median body mass index was 28.0 kg/m^2. Diabetes mellitus was present in 4.9% and hypertension in 10.0%. The index episode was the first-ever biliary colic in 52.1% of patients, Murphy's sign was positive in 63.1%, and 43.6% received pharmacological treatment at first contact. Laboratory values were largely within normal limits. Multiple gallstones were the predominant ultrasound finding (70.6%), with sludge in 12.6%, an impacted stone in 4.5% and gallbladder wall thickening in 10.5%.

Table 1
Baseline characteristics at the index emergency department visit (N = 507)

Characteristic	Value
Demographic and anthropometric	
Age, years, median (IQR)	38 (28–52)
Female sex, n (%)	388 (76.5)
Height, cm, median (IQR)	155 (150–162)
Weight, kg, median (IQR)	69 (62–78)
Body mass index, kg/m², median (IQR)	28.0 (25.4–31.5)
Comorbidities	
Diabetes mellitus, n (%)	25 (4.9)
Hypertension, n (%)	51 (10.0)
Clinical features	
First-ever episode of biliary colic, n (%)	264 (52.1)
Duration of pain before presentation, hours, median (IQR)	12 (4–24)
Positive Murphy's sign, n (%)	320 (63.1)
Nausea, n (%)	378 (74.6)
Fever (≥ 37.5°C), n (%)	35 (6.9)
Pharmacological treatment at first contact, n (%)	221 (43.6)
Heart rate, beats/min, mean ± SD	79.3 ± 14.9
Systolic blood pressure, mmHg, mean ± SD	124.3 ± 16.9
Diastolic blood pressure, mmHg, mean ± SD	74.2 ± 10.9
Laboratory	
Leukocytes, ×10³/μL, median (IQR)	8.9 (7.2–11.2)
Neutrophils, %, mean ± SD	66.8 ± 13.1
GGT, U/L, median (IQR)	41.2 (22.5–79.5)
Alkaline phosphatase, U/L, median (IQR)	105.0 (85.5–124.0)
AST, U/L, median (IQR)	28.4 (21.3–46.7)
ALT, U/L, median (IQR)	28.5 (18.1–54.3)
Total bilirubin, mg/dL, median (IQR)	0.50 (0.40–0.73)
Direct bilirubin, mg/dL, median (IQR)	0.20 (0.20–0.32)
Indirect bilirubin, mg/dL, median (IQR)	0.30 (0.20–0.40)
Ultrasound findings	
Sludge only, n (%)	21 (4.1)
Single stone, n (%)	128 (25.2)
Multiple stones, n (%)	358 (70.6)

IQR interquartile range, SD standard deviation, GGT gamma-glutamyl transferase, AST aspartate aminotransferase, ALT alanine aminotransferase. Continuous variables are presented as median (IQR) when non-normally distributed (Kolmogorov–Smirnov test) and as mean ± SD otherwise. Stone diameter was measured in the 486 patients with measurable stones.

Characteristic	Value
Demographic and anthropometric	
Largest stone diameter, mm, median (IQR) (n = 486)	10 (6–17)
Biliary sludge (isolated or with stones), n (%)	64 (12.6)
Impacted stone in gallbladder neck, n (%)	23 (4.5)
Wall–echo–shadow sign, n (%)	31 (6.1)
Gallbladder wall thickening (≥ 4 mm), n (%)	53 (10.5)
Gallbladder hydrops, n (%)	5 (1.0)
Sclero-atrophic gallbladder, n (%)	2 (0.4)

IQR interquartile range, SD standard deviation, GGT gamma-glutamyl transferase, AST aspartate aminotransferase, ALT alanine aminotransferase. Continuous variables are presented as median (IQR) when non-normally distributed (Kolmogorov–Smirnov test) and as mean ± SD otherwise. Stone diameter was measured in the 486 patients with measurable stones.

Recurrence, complications and surgery during follow-up

During follow-up, 223 patients (44.0%) experienced symptomatic recurrence, 217 (42.8%) remained event-free at the end of observation, 37 (7.3%) were censored at cholecystectomy and 30 (5.9%) at a biliary complication (Table 2). Recurrence was strikingly early: 25% of patients recurred within 7 days of discharge (95% CI 4–13), and the median time to recurrence was 256 days (95% CI 159–not estimable). Recurrence-free survival was 67.8% (95% CI 63.5–71.8) at 30 days, 59.6% (55.0–63.9) at 90 days, 53.7% (48.8–58.4) at 180 days and 43.1% (34.3–51.6) at 365 days; the Kaplan–Meier cumulative incidence of recurrence at one year was therefore 56.9%, higher than the crude proportion because it accounts for censoring. The Kaplan–Meier curve showed an abrupt initial decline followed by a progressively flatter slope, and the smoothed hazard peaked in the first weeks and then declined monotonically (Fig. 1). The density of time to recurrence was markedly right-skewed (Additional file 1: Figure S3A).

Table 2
Outcomes during follow-up and recurrence-free survival (N = 507)

Outcome	Value
Status at end of follow-up	
Symptomatic recurrence (event), n (%)	223 (44.0)
Event-free at end of observation (administrative censoring), n (%)	217 (42.8)
Censored at cholecystectomy before recurrence, n (%)	37 (7.3)
Censored at biliary complication before recurrence, n (%)	30 (5.9)
Time to recurrence (Kaplan–Meier)	
25th percentile, days (95% CI)	7 (4–13)
Median, days (95% CI)	256 (159–NE)
75th percentile, days	NE
Recurrence-free survival, % (95% CI)	
30 days	67.8 (63.5–71.8)
90 days	59.6 (55.0–63.9)
180 days	53.7 (48.8–58.4)
365 days	43.1 (34.3–51.6)
Biliary complications and surgery (whole follow-up)	
Any biliary complication, n (%)	54 (10.7)
Cholecystectomy before any event, n (%)	37 (7.3)
Cholecystectomy after recurrence or complication, n (%)	120 (23.7)
Cholecystectomy at any time, n (%)	157 (31.0)

CI confidence interval, NE not estimable. The 75th percentile could not be estimated because recurrence-free survival did not fall below 25% during follow-up. Of the 54 patients with a biliary complication, 30 developed it before any recurrence (censored) and 24 after recurrence. The breakdown of complications is given in Additional file 1: Table S5.

A biliary complication developed in 54 patients (10.7%; 61 events, seven patients with two concurrent complications): acute-on-chronic cholecystitis in 4.7%, acute cholecystitis in 3.6%, biliary pancreatitis in 1.8%, choledocholithiasis in 1.0%, Mirizzi syndrome in 0.4%, cholangitis in 0.4% and gallstone ileus in 0.2%. Thirty of these complications preceded any recurrence and 24 followed it. Overall, 157 patients (31.0%) underwent cholecystectomy at some point: 37 (7.3%) before any event and 120 (23.7%) after recurrence (n = 90) or complication (n = 30); thus 40.4% of patients who recurred were eventually operated on. Histopathology was available for all 157 resected gallbladders: 94.9% showed at least one pathological finding, predominantly chronic cholecystitis (65.6%), and two specimens (1.3%) harboured incidental adenocarcinoma (Additional file 1: Table S4).

Univariable analysis

The full univariable screening of all candidate predictors is presented in Additional file 1: Table S1, and the variables reaching statistical significance or clinical relevance are summarized in Table 3. Murphy's sign was by far the strongest predictor (HR 6.41; 95% CI 4.28–9.61; $p < 0.001$), but it violated the PH assumption (Schoenfeld $p = 0.005$). Diabetes mellitus (HR 1.98; 95% CI 1.20–3.25), an impacted stone (HR 1.72; 1.02–2.91), marked leukocytosis ≥ 18 × 10^3/μL (HR 3.41; 1.40–8.31; based on only five patients, all of whom recurred) and direct bilirubin per mg/dL (HR 1.40; 1.08–1.81) were associated with a higher hazard of recurrence, whereas age ≥ 30 years was associated with a lower hazard (HR 0.72; 0.55–0.95). Sex, body mass index, vital signs, symptom duration, first-ever episode, fever, nausea, liver enzymes, stone number and size, wall thickening and sludge (HR 1.32; 0.90–1.93) were not significantly associated with recurrence. After stratification by Murphy's sign, only diabetes mellitus retained an independent

association (HR 2.00; 95% CI 1.22–3.29; p = 0.006), whereas age, leukocytosis, direct bilirubin and impacted stone lost statistical significance. Kaplan–Meier curves for these factors are shown in Fig. 2.

Table 3
Univariable Cox regression for symptomatic recurrence: selected candidate predictors

Variable	Events/exposed	HR (95% CI)	p	p, stratified by Murphy's sign
Positive Murphy's sign	196/320	6.41 (4.28–9.61)	< 0.001	— (stratification variable)
Diabetes mellitus	17/25	1.98 (1.20–3.25)	0.007	0.006
Impacted stone in gallbladder neck	15/23	1.72 (1.02–2.91)	0.043	0.154
Leukocytosis ≥ 18 × 10^3/μL	5/5	3.41 (1.40–8.31)	0.007	0.099
Direct bilirubin (per mg/dL)	—	1.40 (1.08–1.81)	0.011	0.120
Age ≥ 30 years	142/348	0.72 (0.55–0.95)	0.021	0.551
Biliary sludge	31/64	1.32 (0.90–1.93)	0.155	0.106

HR hazard ratio, CI confidence interval. HRs are from unstratified univariable Cox models; the last column shows the p value after repeating the analysis in a Cox model stratified by Murphy's sign, which therefore has no estimate for that variable. Murphy's sign violated the proportional hazards assumption (Schoenfeld residuals p = 0.005). The complete screening of all candidate predictors is provided in Additional file 1: Table S1.

Multivariable models and model selection

In the multivariable Cox model stratified by Murphy's sign, diabetes mellitus remained the only significant covariate (HR 2.07; 95% CI 1.26–3.41), while sludge (HR 1.38; 0.94–2.02) and impacted stone (HR 1.53; 0.90–2.60) did not; their removal improved the BIC (2312.9 vs 2320.5) without loss of fit (Table 4). The PH assumption was satisfied in the stratified models (global Schoenfeld p = 0.308 and 0.249). However, stratification excluded Murphy's sign from the linear predictor and reduced discrimination to a C-index of 0.518, whereas including it as a covariate in the Cox model restored discrimination (C 0.690) at the cost of a PH violation (Murphy's sign p = 0.006; global p = 0.012) (Additional file 1: Table S2).

Table 4
Multivariable models for symptomatic recurrence of biliary colic

Variable	Cox, stratified by Murphy's sign — full model HR (95% CI); p	Cox, stratified by Murphy's sign — final model HR (95% CI); p	Weibull, Murphy's sign as covariate — final model HR (95% CI); p
Diabetes mellitus	2.07 (1.26–3.41); 0.004	2.00 (1.22–3.29); 0.006	2.18 (1.33–3.58); 0.002
Positive Murphy's sign	Stratification variable	Stratification variable	7.09 (4.73–10.61); <0.001
Biliary sludge	1.38 (0.94–2.02); 0.098	—	—
Impacted stone	1.53 (0.90–2.60); 0.113	—	—
Shape parameter γ (95% CI)	—	—	0.419 (0.376–0.468)
Scale parameter λ (95% CI)	—	—	0.0183 (0.0117–0.0286)
Likelihood-ratio χ^2 (df); p	10.98 (3); 0.012	6.20 (1); 0.013	145.45 (2); <0.001
AIC / BIC	2307.9 / 2320.5	2308.6 / 2312.9	1461.6 / —
Harrell's C-index	—	0.518	0.692
Global Schoenfeld test, p	0.308	0.249	Not applicable

HR hazard ratio, CI confidence interval, df degrees of freedom, AIC Akaike information criterion, BIC Bayesian information criterion. N = 507; events = 223. AIC values are not comparable between the Cox (partial likelihood) and Weibull (full likelihood) families. The final Weibull model is $h(t \mid x) = \gamma\lambda t^{(\gamma-1)} \cdot \exp(0.779 \times \text{diabetes} + 1.959 \times \text{Murphy's sign})$, with t in days; ln γ = −0.869 (p < 0.001) and ln λ = −4.002 (intercept, p < 0.001). Recurrence–free survival is $S(t \mid x) = \exp[-\lambda\, t^{\gamma} \exp(0.779 \times \text{diabetes} + 1.959 \times \text{Murphy's sign})]$. A Cox model with Murphy's sign as a covariate achieved a C-index of 0.690 but violated the proportional hazards assumption (Additional file 1: Table S2).

The Weibull models resolved this tension. In the specification with Murphy's sign as a covariate, both diabetes mellitus (HR 2.18; 95% CI 1.33–3.58; p = 0.002) and Murphy's sign (HR 7.09; 95% CI 4.73–10.61; p < 0.001) were independent predictors, and the shape parameter γ = 0.419 (95% CI 0.376–0.468; ln γ = −0.869; p < 0.001) confirmed a monotonically decreasing baseline hazard (Table 4). Allowing the shape parameter to depend on Murphy's sign did not improve discrimination (C 0.692 in both) or fit (AIC 1463.3 vs 1461.6), and the interaction term was not significant (p = 0.582), indicating that the decreasing temporal pattern was shared by patients with and without Murphy's sign (Additional file 1: Table S3). The Weibull model with Murphy's sign as a covariate, which had the lowest AIC within its family and discrimination equivalent to the unstratified Cox model, was therefore selected as the final model:

$$h(t \mid x) = \gamma\, \lambda\, t^{\gamma-1} \cdot \exp(0.779 \times \text{DM} + 1.959 \times \text{MURPHY})$$

Final model and diagnostics

Under the final model the hazard of recurrence at time t (days) is $h(t \mid x) = \gamma\, \lambda\, t^{\gamma-1} \cdot \exp(0.779 \times \text{DM} + 1.959 \times \text{MURPHY})$, where DM and MURPHY are indicator variables for diabetes mellitus and a positive Murphy's sign, γ = 0.419 and the scale parameter λ = 0.0183 (95% CI 0.0117–0.0286; ln λ = −4.002). The corresponding recurrence-free survival is $S(t \mid x) = \exp[-0.0183 \times t^{0.419} \times \exp(0.779 \times \text{DM} + 1.959 \times \text{MURPHY})]$, which allows individual predictions at any horizon: at 30 days the predicted probability of remaining recurrence-free is 0.927 for a patient with neither predictor, 0.847 with diabetes only, 0.583 with a positive Murphy's sign only and 0.309 with both (0.886, 0.769, 0.425 and 0.155, respectively, at 90 days). The model-based survival and hazard functions and the hazard curves for the lowest- and highest-risk profiles are shown in Additional file 1: Figures S1 and S3B. The log[− log S(t)] versus log(t) plots by Murphy strata were approximately linear and parallel, supporting the Weibull form, and the Cox–Snell residual plot followed the 45° reference line closely with divergence only in the sparsely populated right tail (Additional file 1: Figure S2). The significant Schoenfeld test for Murphy's sign obtained with untransformed time (p = 0.006) did not hold with Kaplan–Meier (p = 0.448), rank (p = 0.466) or log (p = 0.219) time transformations, suggesting a gradual late attenuation of its effect rather than a structural violation.

Predictive performance and internal validation

Somers' D for the final model was 0.384, corresponding to a Harrell's C-index of 0.692 (95% CI 0.663–0.721). Bootstrap validation with 1000 resamples yielded a mean optimism of − 0.0005 and an optimism-corrected C-index of 0.692 (95% CI 0.665–0.723), indicating stable coefficients and negligible overfitting. Because the linear predictor takes only four distinct values, 47.4% of comparable pairs (36 445 of 76 845) were tied on predicted risk, which bounds the attainable C-index.

Tertiles of the linear predictor produced unequally sized groups: low risk (n = 179; neither predictor), intermediate risk (n = 311) and high risk (n = 17; both predictors). Observed Kaplan–Meier curves separated clearly (Fig. 3A). Recurrence-free survival in the low-risk group was 92.7%, 90.1% and 87.2% at 30, 90 and 180 days, with confidence intervals that did not overlap those of the other groups; in the intermediate-risk group it was 55.0% at 30 days and 35.9% at 180 days; and in the high-risk group it fell to 28.2% at 30 days and 14.1% between 90 and 180 days, although this estimate relies on very few patients.

Overall calibration was acceptable: predicted versus observed survival was 0.817 vs 0.750 at 7 days, 0.761 vs 0.698 at 15 days, 0.699 vs 0.678 at 30 days and 0.629 vs 0.627 at 60 days. By risk group (Fig. 3B), the low-risk group lay on the identity line at every horizon, whereas at 7 and 15 days the intermediate- and especially the high-risk groups fell below it (observed 0.35 vs predicted 0.53 at day 7 in the high-risk group), i.e. the model underestimated very early risk in patients with a positive Murphy's sign. From day 30 onwards all three groups approached the identity line with confidence intervals including it.

Sensitivity analyses for competing risks

Kaplan–Meier estimates and cumulative incidence functions treating biliary complication as a competing event differed by less than three percentage points at every horizon (e.g. 56.9% vs 54.3% cumulative recurrence at 365 days). In the Fine–Gray model, diabetes mellitus (subdistribution HR 2.06; 95% CI 1.41–3.02) and Murphy's sign (subdistribution HR 5.73; 95% CI 3.86–8.51) retained strong effects, with a significant attenuation of the effect of Murphy's sign over time (time interaction p = 0.007), whereas sludge lost significance (subdistribution HR 1.19; 95% CI 0.81–1.72).

Discussion

In this prospective cohort of patients discharged from a Latin American ED with uncomplicated biliary colic, symptomatic recurrence affected 44.0% of patients within a maximum of one year, one in four recurrences occurred within a week of discharge, and 10.7% of patients developed a biliary complication while under conservative management. A parametric Weibull model showed that the hazard of recurrence is highest immediately after discharge and declines monotonically thereafter, and identified two bedside variables, a positive Murphy's sign and diabetes mellitus, as independent predictors. The model discriminated acceptably, showed no optimism on bootstrap validation, and was well calibrated beyond the first month, with underestimation of very early risk confined to patients with a positive Murphy's sign.

Comparison with existing evidence. The most direct comparator is RELAPSTONE, which followed 3016 patients after a first admission for symptomatic gallstone disease in 18 Spanish and Mexican centers and reported recurrence-free survival of 79%, 71% and 63% at 3, 6 and 12 months with a median time to relapse of 2.1 months [16]. Our figures were considerably lower (59.6%, 53.7% and 43.1%), which is methodologically expected: RELAPSTONE pooled all forms of symptomatic gallstone disease, and biliary colic was precisely the presentation with the highest relapse hazard in that study (HR 1.70; 95% CI 1.31–2.21), whereas our cohort consisted exclusively of uncomplicated biliary colic. The coexistence in our cohort of a 25th percentile of 7 days with a median of 256 days reflects a bimodal risk: a subgroup with highly active lithiasis that relapses almost immediately and a subgroup with an indolent course. The concentration of early events is conceptually similar to the pattern reported by Helenius et al. after non-operative management of acute cholecystitis, where 71% of relapses occurred within 6 months [24], and to the high short-term complication risk described by Kousgaard et al. in symptomatic uncomplicated gallstone disease, with a median of 5 months to the first event [25]. Randomized evidence pooled by Ghose et al. confirms that conservative management carries a higher risk of recurrent biliary colic (RR 2.75) and total biliary complications (RR 3.63) than early cholecystectomy, without differences in mortality [26], and the 5-year SECURE data place the complication rate of a restrictive strategy at 14.9% [13]. Our 10.7% complication rate accrued over a much shorter horizon, suggesting that with longer observation the burden in our setting would match or exceed that of European series.

The predictors partly converge and partly diverge with RELAPSTONE, which identified multiple stones and elevated ALT as risk factors and older age, sphincterotomy and leukocytosis > 11 000/mm^3 as protective [16]. Both studies found a protective effect of age, which in our cohort emerged above 30 years (HR 0.72). Leukocytosis pointed in the opposite direction in our univariable analysis (HR 3.41 for ≥ 18 × 10^3/μL), but this estimate rests on five patients and must be regarded as exploratory; the discrepancy may reflect the base

population, since in a cohort that includes complicated disease a strong inflammatory response accelerates surgical referral and thereby reduces the opportunity for recurrence. First-ever presentation did not predict recurrence, consistent with Lund's classic series of 526 non-operated patients [27] but in contrast with the National Cooperative Gallstone Study and Japanese cohorts [10, 28]; contemporary ultrasound and clinical detection of active gallbladder inflammation at the first episode may have displaced symptom history as a risk marker.

Interpretation of the predictors. Murphy's sign carried the largest effect (HR 7.09). Classically, uncomplicated biliary colic is defined by transient cystic duct obstruction without established wall inflammation, so a positive Murphy's sign is considered a hallmark of acute cholecystitis. Its presence in 63.1% of our patients, none of whom met criteria for cholecystitis at the index visit, suggests that in a substantial proportion it signals early or subclinical gallbladder inflammation, a transitional state within the continuum from simple colic to established cholecystitis. Under this hypothesis Murphy's sign acts as a low-cost clinical marker of active wall inflammation, which both marks a more severe index episode and predisposes to early recurrence by perpetuating wall irritation and biliary stasis. Consistently, its effect attenuated over time in the Fine–Gray model, as expected for an early inflammatory process, and the very early miscalibration of the model was confined to groups defined by its presence. Diabetes mellitus approximately doubled the hazard of recurrence, in keeping with the impaired gallbladder motility, sphincter of Oddi dysfunction and chronic pro-inflammatory state that characterize these patients [29, 30]. Neither predictor requires laboratory or imaging technology, which favors implementation in resource-limited EDs.

The behavior of the remaining candidate predictors is instructive. Sludge and an impacted stone were associated with events in a preliminary analysis that pooled recurrence and complications into a composite outcome, but this association disappeared when the outcome was restricted to symptomatic recurrence, and sludge lost significance in the Fine–Gray model. These ultrasound features appear to be markers of stone burden and progression whose effect is oriented toward complicated disease rather than toward painful recurrence per se; pooling both outcomes may therefore overestimate their prognostic value. Conversely, leukocytosis and bilirubin lost significance not with the outcome definition but upon adjustment for Murphy's sign, suggesting that their prognostic information is largely subsumed by a stronger clinical marker of the same inflammatory substrate.

Methodological considerations. The choice of a parametric Weibull model over the conventional Cox model was driven by two features of the data. First, the smoothed hazard was clearly non-constant, and the Weibull shape parameter ($\gamma = 0.419$) quantified this decreasing pattern explicitly, something a Cox model cannot provide and that RELAPSTONE, relying solely on Cox regression, could not address [16]. Second, Murphy's sign violated the PH assumption in the Cox framework, forcing stratification and thereby removing the strongest predictor from the linear predictor, with a collapse of discrimination to 0.518. The Weibull model in the PH metric shares the proportionality assumption, but the sensitivity analysis showed the deviation to be modest and dependent on the time scale, and the Weibull model additionally admits an accelerated failure time interpretation that does not rely on it. The single hazard ratio for Murphy's sign should nevertheless be read as a weighted average over follow-up.

Clinical implications. Our data challenge the premise that uncomplicated biliary colic is a uniformly benign condition amenable to open-ended observation. The elective cholecystectomy rate before any event in our cohort (7.3%) was almost four times lower than the 29.5% observed at 24 months in the conservative arm of the C-GALL trial [14], reflecting either different surgical thresholds, barriers to elective surgery, or both. Delay is not innocuous: in a nationwide cohort of 34 925 patients, cholecystectomy for recurrent rather than first-episode cholecystitis was associated with higher odds of bile duct injury (OR 2.44), bowel perforation (OR 2.54) and conversion to open surgery (OR 1.76) [31]. The decreasing hazard identified here implies that the window for preventing early recurrence is narrow and begins at ED discharge. Patients with a positive Murphy's sign or diabetes mellitus should be identified and prioritized for surgical evaluation at the index visit rather than referred to general waiting lists, and ED discharge protocols should include explicit warning about early recurrence and structured follow-up during the first week. Whether such risk-based prioritization reduces re-attendance and progression to complications is a hypothesis that requires prospective evaluation.

Strengths and limitations. Strengths of this study include its prospective design with standardized data collection by a single investigator, a sample size well above the minimum calculated, reporting according to TRIPOD, explicit modeling of the baseline hazard, bootstrap internal validation and competing-risks sensitivity analyses, and the fact that it provides one of the first systematic characterizations of the natural history of biliary colic in a real-world Latin American cohort. Several limitations must be acknowledged. Selection bias is inherent to an ED-based cohort, which excludes patients who do not seek care and may therefore overestimate recurrence in the general population of symptomatic gallstone carriers. In the opposite direction, follow-up relied on institutional records without active patient contact, so recurrences or cholecystectomies at other institutions were not captured; the

reported rates should be interpreted as a floor rather than a ceiling. The study was single-center and the model has not been externally validated. Ultrasound features are subject to inter-observer variability, which was not assessed. Some cut-points were derived from the data and may carry optimism not captured by internal validation; several subgroups, including marked leukocytosis and impacted stone, had few events, and the effect of diabetes was estimated from 25 exposed patients (17 events), so the precision of its hazard ratio is limited. The parsimony of the final model bounds its discrimination: with two binary predictors the linear predictor takes only four values, which explains both the modest C-index and the small high-risk group (n = 17), whose estimates are indicative only. Potential predictors such as the number of previous episodes, dyslipidemia, metabolic syndrome, recent pregnancy, rapid weight loss and family history were not collected. Finally, treating cholecystectomy and complications as censoring may require competing-risks models in settings with higher early surgery rates, although our sensitivity analyses suggest that the impact on the present estimates was small.

Conclusions

Recurrence of biliary colic after conservative management in the ED is frequent, early and accompanied by a non-negligible rate of biliary complications. The hazard of recurrence is highest in the days immediately after discharge and declines thereafter, so the opportunity to prevent early recurrence lies at the index visit. A parsimonious Weibull model based on two bedside variables, Murphy's sign and diabetes mellitus, stratifies patients into groups with markedly different recurrence-free survival, with acceptable discrimination and stable internal validation. Patients with either predictor should be considered for prioritized surgical evaluation and structured early follow-up. External validation in multicenter cohorts, preferably including Latin American populations with different levels of access to elective surgery, and prospective evaluation of risk-based prioritization are the necessary next steps before routine clinical implementation.

Abbreviations

AIC
Akaike information criterion
ALT
Alanine aminotransferase
AST
Aspartate aminotransferase
BIC
Bayesian information criterion
CI
Confidence interval
DM
Diabetes mellitus
ED
Emergency department
GGT
Gamma-glutamyl transferase
HR
Hazard ratio
ICD-10
International Classification of Diseases, 10th revision
IQR
Interquartile range
OR
Odds ratio
PH
Proportional hazards
RR
Risk ratio

SD
Standard deviation
SHR
Subdistribution hazard ratio
STROBE
Strengthening the Reporting of Observational Studies in Epidemiology
TRIPOD
Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis
WES
Wall–echo–shadow sign.

Declarations

Ethics approval and consent to participate

The study protocol was approved by the Human Research Ethics Committee of the University of Cuenca (Comité de Ética de Investigación en Seres Humanos, CEISH-UC), approval code CEISH-UC-2024-002EO-ESP-EYD, decision letter No. CEISH-UC-2024-637, dated 3 December 2024, and authorized by the authorities of Hospital Vicente Corral Moscoso. All participants provided written informed consent before enrolment. The study was conducted in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable. No individual patient data, images or identifying details are reported.

Availability of data and materials

The de-identified dataset and the Stata analysis code supporting the conclusions of this article are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This research received no specific grant from any funding agency. DABV received a postgraduate training scholarship from the Ecuadorian Secretaría de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT), which had no role in the design of the study, data collection, analysis, interpretation or writing of the manuscript.

Authors' contributions

DABV conceived and designed the study, collected the data, performed the statistical analysis, interpreted the results and drafted the manuscript. DGIA supervised the study design and analysis, contributed to the interpretation of the results and critically revised the manuscript for important intellectual content. Both authors read and approved the final manuscript.

Acknowledgements

The authors thank the staff of the Emergency Department of Hospital Vicente Corral Moscoso for their support during patient enrolment. A large language model (Claude, Anthropic) was used to assist with translation from Spanish and language editing of the manuscript; the authors reviewed and edited all content and take full responsibility for it.

Authors' information

DABV is a resident in Emergency and Disaster Medicine at the University of Cuenca and this work derives from his specialist degree thesis. DGIA is a general surgeon at Hospital Vicente Corral Moscoso and a lecturer at the Faculty of Medical Sciences, University of Cuenca, and supervised the thesis.

References

1. Stinton, L. M. & Shaffer, E. A. Epidemiology of gallbladder disease: cholelithiasis and cancer. *Gut Liver*. **6** (2), 172–187. https://doi.org/10.5009/gnl.2012.6.2.172 (2012).
2. Everhart, J. E. & Ruhl, C. E. Burden of digestive diseases in the United States part III: liver, biliary tract, and pancreas. *Gastroenterology* **136** (4), 1134–1144. https://doi.org/10.1053/j.gastro.2009.02.038 (2009).
3. Li, J. et al. Global burden of gallbladder and biliary diseases: a systematic analysis for the Global Burden of Disease Study 2019. *J. Gastroenterol. Hepatol.* **37** (7), 1389–1399. https://doi.org/10.1111/jgh.15859 (2022).
4. Instituto Nacional de Estadística y Censos (INEC). Boletín Técnico N°01-2025-ECEH. Registros Estadísticos de Camas y Egresos Hospitalarios, 2024. Quito: INEC. (2025). https://www.ecuadorencifras.gob.ec/documentos/web-inec/Estadisticas_Sociales/Camas_Egresos_Hospitalarios/2024/Boletin_Tecnico_de_Camas_y_Egresos_Hospitalarios_2024.pdf. Accessed 7 Sep 2026.
5. Makutonin, M., Moghtaderi, A., Newton, S., Ma, Y. & Meltzer, A. C. Biliary colic in the emergency department: a state-wide analysis of one-year costs and clinical outcomes. *Surg. Open. Sci.* **12**, 9–13. https://doi.org/10.1016/j.sopen.2023.02.002 (2023).
6. Gallaher, J. R. & Charles, A. Acute cholecystitis: a review. *JAMA* **327** (10), 965–975. https://doi.org/10.1001/jama.2022.2350 (2022).
7. Altieri, M. S. et al. What happens to biliary colic patients in New York State? 10-year follow-up from emergency department visits. *Surg. Endosc*. **32** (4), 2058–2066. https://doi.org/10.1007/s00464-017-5902-5 (2018).
8. Portincasa, P., Moschetta, A. & Palasciano, G. Cholesterol gallstone disease. *Lancet* **368** (9531), 230–239. https://doi.org/10.1016/S0140-6736(06)69044-2 (2006).
9. Lee, B. J. H., Yap, Q. V., Low, J. K., Chan, Y. H. & Shelat, V. G. Cholecystectomy for asymptomatic gallstones: Markov decision tree analysis. *World J. Clin. Cases*. **10** (29), 10399–10412. https://doi.org/10.12998/wjcc.v10.i29.10399 (2022).
10. Tomida, S. et al. Long-term ursodeoxycholic acid therapy is associated with reduced risk of biliary pain and acute cholecystitis in patients with gallbladder stones: a cohort analysis. *Hepatology* **30** (1), 6–13. https://doi.org/10.1002/hep.510300108 (1999).
11. Unalp-Arida, A. & Ruhl, C. E. Burden of gallstone disease in the United States population: prepandemic rates and trends. *World J. Gastrointest. Surg.* **16** (4), 1130–1148. https://doi.org/10.4240/wjgs.v16.i4.1130 (2024).
12. Strasberg, S. M. Acute calculous cholecystitis. *N Engl. J. Med.* **358** (26), 2804–2811. https://doi.org/10.1056/NEJMcp0800929 (2008).
13. Comes, D. J. et al. Restrictive strategy vs usual care for cholecystectomy in patients with abdominal pain and gallstones: 5-year follow-up of the SECURE randomized clinical trial. *JAMA Surg.* **159** (11), 1235–1243. https://doi.org/10.1001/jamasurg.2024.3080 (2024).
14. Innes, K. et al. Laparoscopic cholecystectomy versus conservative management for adults with uncomplicated symptomatic gallstones: the C-GALL RCT. *Health Technol. Assess.* **28** (26), 1–151. https://doi.org/10.3310/MNBY3104 (2024).
15. Quintana, J. M. et al. Appropriateness variation in cholecystectomy. *Eur. J. Public. Health*. **14** (3), 252–257. https://doi.org/10.1093/eurpub/14.3.252 (2004).
16. Velamazán, R. et al. Symptomatic gallstone disease: recurrence patterns and risk factors for relapse after first admission, the RELAPSTONE study. *United Eur. Gastroenterol. J.* **12** (3), 286–298. https://doi.org/10.1002/ueg2.12544 (2024).
17. Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. M. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): the TRIPOD statement. *BMJ* **350**, g7594. https://doi.org/10.1136/bmj.g7594 (2015).
18. von Elm, E. et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet* **370** (9596), 1453–1457. https://doi.org/10.1016/S0140-6736(07)61602-X (2007).
19. Cotton, P. B., Elta, G. H., Carter, C. R., Pasricha, P. J. & Corazziari, E. S. Gallbladder and sphincter of Oddi disorders. *Gastroenterology* **150** (6), 1420–9e2. https://doi.org/10.1053/j.gastro.2016.02.033 (2016).
20. Riley, R. D. et al. Calculating the sample size required for developing a clinical prediction model. *BMJ* **368**, m441. https://doi.org/10.1136/bmj.m441 (2020).
21. Grambsch, P. M. & Therneau, T. M. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* **81** (3), 515–526. https://doi.org/10.1093/biomet/81.3.515 (1994).

22. Harrell, F. E. Jr, Lee, K. L. & Mark, D. B. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat. Med.* **15** (4), 361–387. https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4%3C361::AID-SIM168%3E3.0.CO;2-4 (1996).
23. Fine, J. P. & Gray, R. J. A proportional hazards model for the subdistribution of a competing risk. *J. Am. Stat. Assoc.* **94** (446), 496–509. https://doi.org/10.1080/01621459.1999.10474144 (1999).
24. Helenius, L., Linder, F. & Osterman, E. Relapse in gallstone disease after non-operative management of acute cholecystitis: a population-based study. *BMJ Open. Gastroenterol.* **12** (1), e001680. https://doi.org/10.1136/bmjgast-2024-001680 (2025).
25. Kousgaard, N., Rasmussen, S. K. M., Möller, S., Koulaouzidis, A. & Mark-Christensen, A. Symptomatic uncomplicated gallstone disease is associated with a high short-term risk of gallstone-related complications: a contemporary cohort study. *Scand. J. Gastroenterol.* **59** (8), 954–960. https://doi.org/10.1080/00365521.2024.2361756 (2024).
26. Ghose, K. et al. Conservative management versus early cholecystectomy for gallstone disease: a meta-analysis of randomized controlled trials. *JGH Open.* **9** (8), e70253. https://doi.org/10.1002/jgh3.70253 (2025).
27. Lund, J. Surgical indications in cholelithiasis: prophylactic cholecystectomy elucidated on the basis of long-term follow up on 526 nonoperated cases. *Ann. Surg.* **151** (2), 153–162. https://doi.org/10.1097/00000658-196002000-00001 (1960).
28. Thistle, J. L., Cleary, P. A., Lachin, J. M., Tyor, M. P. & Hersh, T. The natural history of cholelithiasis: the National Cooperative Gallstone Study. *Ann. Intern. Med.* **101** (2), 171–175. https://doi.org/10.7326/0003-4819-101-2-171 (1984).
29. Figueiredo, J. C. et al. Sex and ethnic/racial-specific risk factors for gallbladder disease. *BMC Gastroenterol.* **17** (1), 153. https://doi.org/10.1186/s12876-017-0678-6 (2017).
30. Fujita, N. et al. Evidence-based clinical practice guidelines for cholelithiasis 2021. *J. Gastroenterol.* **58** (9), 801–833. https://doi.org/10.1007/s00535-023-02014-6 (2023).
31. Edblom, M. et al. Early cholecystectomy for recurrent versus first-time cholecystitis: nationwide population-based study. *BJS Open.* **10** (1), zraf166. https://doi.org/10.1093/bjsopen/zraf166 (2026).

Figures

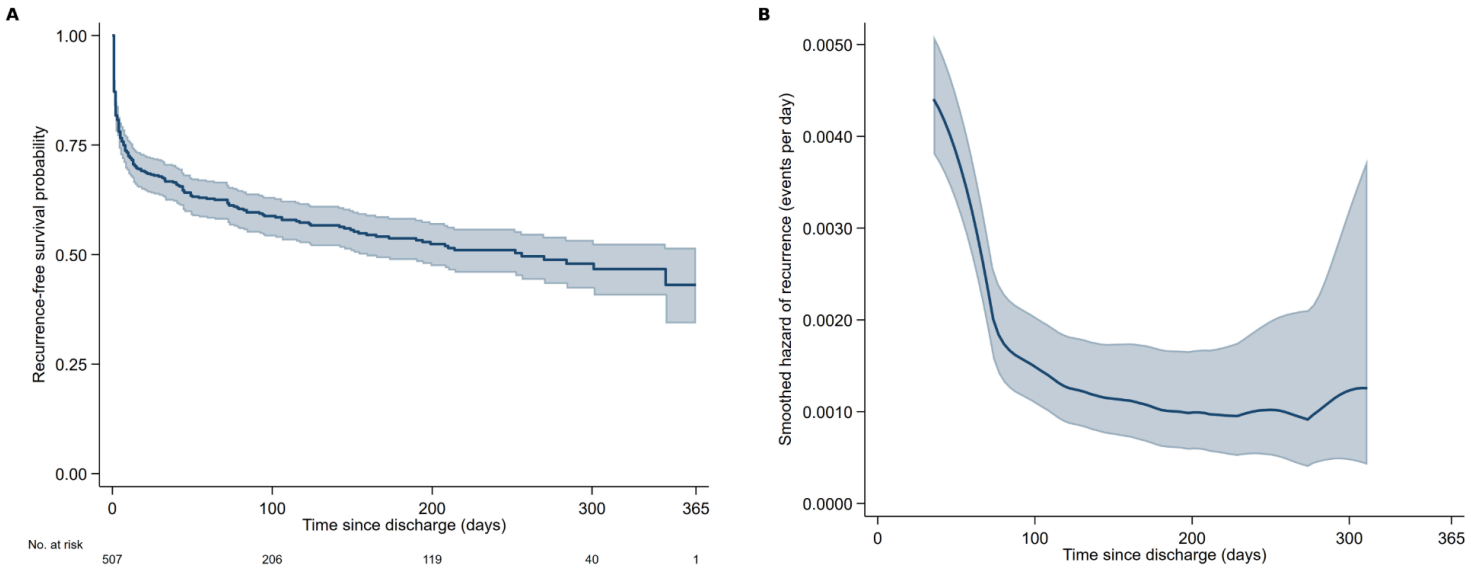


Figure 1

Recurrence-free survival and hazard of recurrence after emergency department discharge.

AKaplan–Meier estimate of recurrence-free survival for the whole cohort (N = 507) with 95% confidence band and number of patients at risk. **B** Smoothed hazard function of symptomatic recurrence with 95% confidence interval, showing a peak in the first weeks after discharge followed by a monotonic decline.

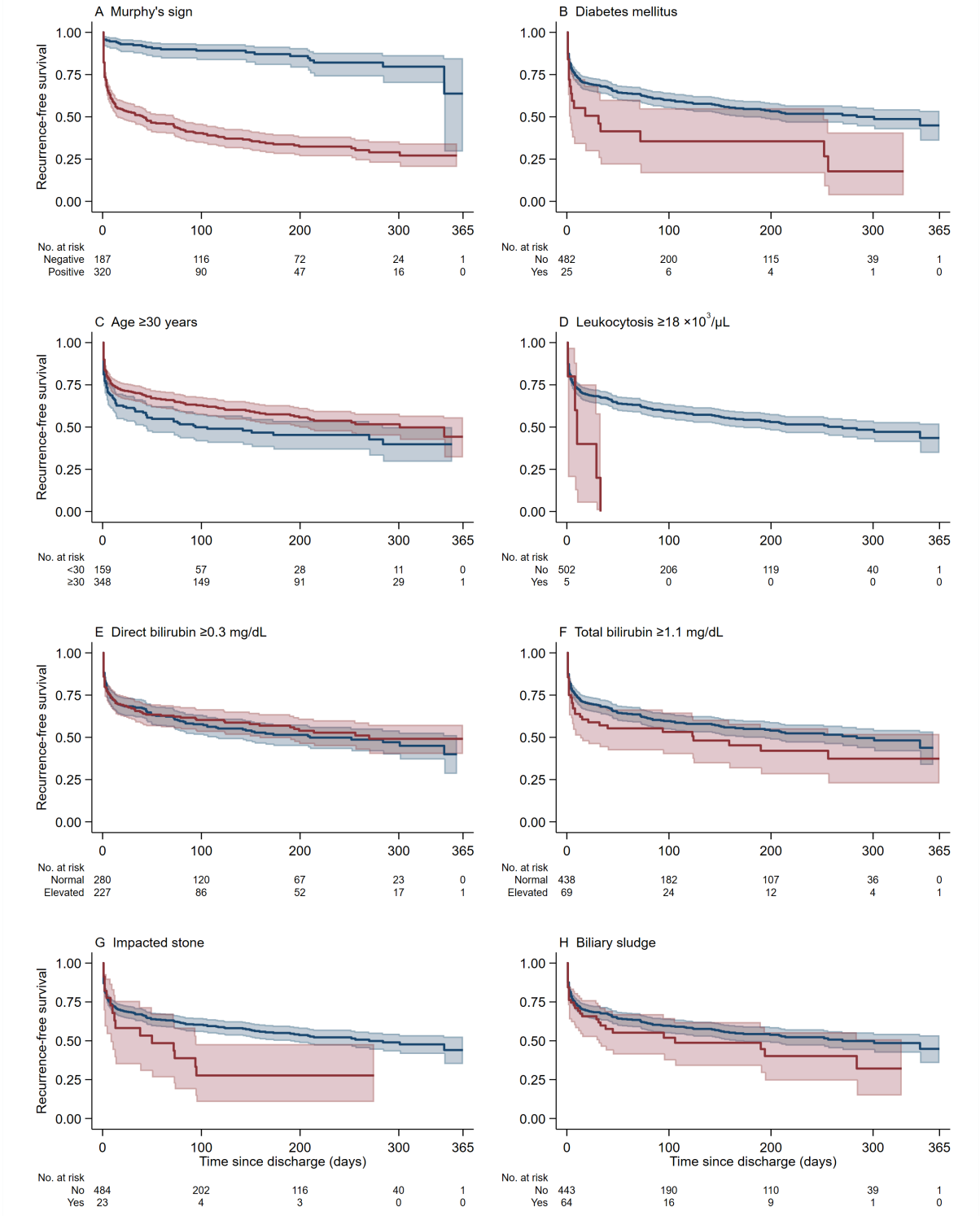


Figure 2

Kaplan–Meier curves of recurrence-free survival according to candidate predictors.

A Murphy's sign, **B** diabetes mellitus, **C** age ≥30 years, **D** leukocytosis ≥18 × 10^3/μL, **E** direct bilirubin ≥0.3 mg/dL, **F** total bilirubin ≥1.1 mg/dL, **G** impacted stone in the gallbladder neck, **H** biliary sludge. Patients exposed to each factor are shown in dark red and reference groups in navy; shaded areas are 95% confidence intervals. Numbers at risk are given below each panel.

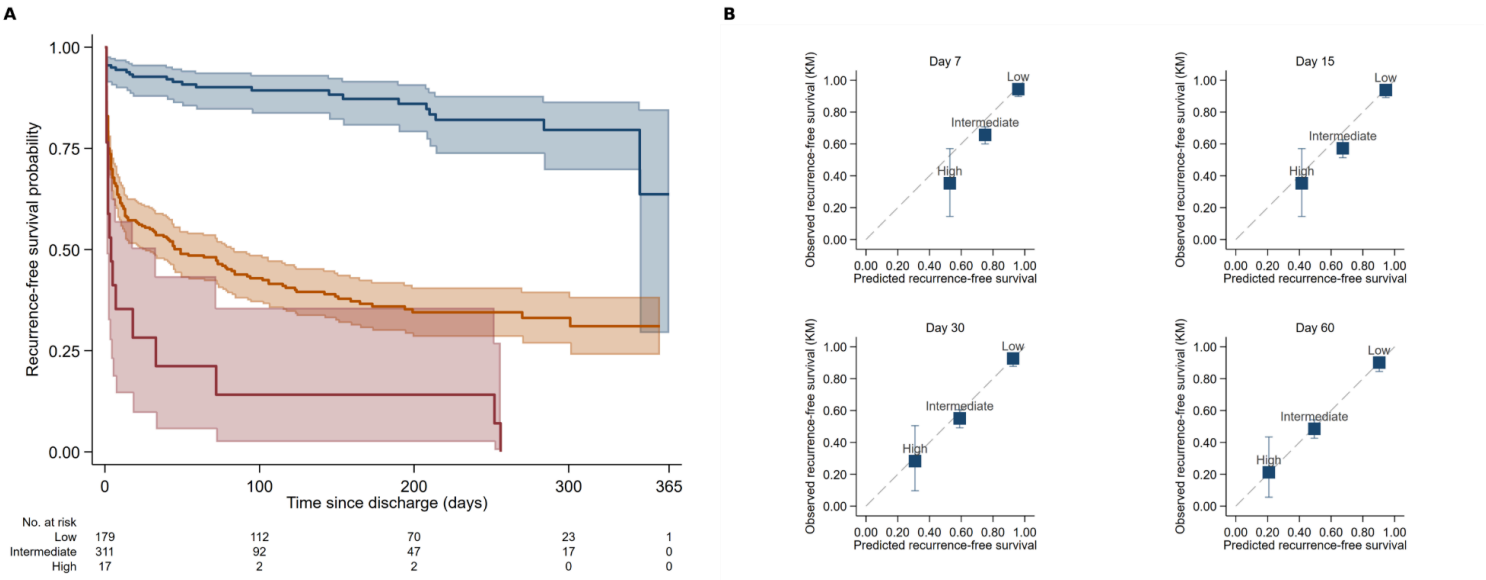


Figure 3

Risk stratification and calibration of the final Weibull model.

AObserved Kaplan–Meier recurrence-free survival by tertile of the model linear predictor: low risk (navy, n = 179), intermediate risk (orange, n = 311) and high risk (dark red, n = 17). **B** Calibration plots at 7, 15, 30 and 60 days comparing model-predicted survival with observed Kaplan–Meier survival (with 95% confidence intervals) in each risk group (low, intermediate, high); the dashed line is the identity line.

Supplementary Files

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

- Additionalfile2TRIPODchecklist.docx
- Additionalfile1Supplementary.docx