Updates in Surgery
https://doi.org/10.1007/s13304-025-02476-8

ORIGINAL ARTICLE

Development and validation of machine learning tools for predicting postoperative complications in acute calculous cholecystitis

Ottavia Cicerone[1,2] · Simone Frassini[1,2] · Anna Gallotti[4] · Alessandro Vanoli[3] · Luca Ansaloni[1,2] · Marcello Maestri[1,2] · Paola Fugazzola[2] · S.P.Ri.M.A.C.C.[2] · Collaborative Group[2]

Received: 8 July 2025 / Accepted: 7 November 2025

Abstract

Background Acute calculous cholecystitis (ACC) is a common gastrointestinal emergency, with early laparoscopic cholecystectomy (ELC) as the standard of care. However, the risk of postoperative complications remains significant. This study developed and validated two machine learning models—CholeSurgRisk I (preoperative) and CholeSurgRisk II (comprehensive)—to predict major postoperative complications in ACC.

Methods A prospectively collected, multicenter database of 1,253 patients was retrospectively analyzed. Lasso regression identified key predictive variables among demographic, clinical, and perioperative factors. Three machine learning algorithms (Random Forest, XGBoost, Decision Tree) were trained and tested, comparing their performance via AUC-ROC.

Results CholeSurgRisk I achieved an AUC-ROC of 0.8456, while incorporating intraoperative variables (CholeSurgRisk II) improved performance to 0.8903. To facilitate clinical use, a web-based tool - "CholeSurgRisk I" - was developed based on the preoperative model, providing real-time, patient-specific risk estimations.

Conclusion Machine learning enhances perioperative risk stratification in ACC. CholeSurgRisk I facilitates early preoperative assessment, whereas CholeSurgRisk II refines predictions by integrating intraoperative factors. The user-friendly application "CholeSurgRisk I" offers individualized complication risk forecasts, potentially aiding clinical decision-making and optimizing outcomes.

Keywords Acute cholecystitis · Laparoscopic cholecystectomy · Postoperative complications; Machine learning; Risk prediction.

Ethical considerations.

The study protocol was approved by the medical Ethics Board of the trial coordinating center at the IRCCS Policlinico San Matteo, Pavia (Italy). Secondary approvals were obtained from all local ethics committees in the participating centers. Patients gave orally and written informed consent prior to inclusion. The S.P.Ri.M.A.C.C. trial was conducted in accordance with the declaration of Helsinki.

All authors have read and agreed to the final version of the manuscript.

Collaborators:

S.P.Ri.M.A.C.C. Collaborative Group:

Goran Augustin [1], Trpimir Morić [1], Selmy Awad [2], Azzah M Alzahrani [2], Mohamed Elbahnasawy [3], Damien Massalou [4], Belinda De Simone [5], Zaza Demetrashvili [6], Athina-Despoina Kimpizi [7], Dimitrios Schizas [8], Dimitrios Balalis [9], Nikolaos Tasis [10], Maria Papadoliopoulou [11], Petrakis Georgios [12], Konstantinos Lasithiotakis [13], Orestis Ioannidis [14], Lovenish Bains [15], Matteo Magnoli [16], Pasquale Cianci [17], Nunzia Ivana Conversano [17], Alessandro Pasculli [18], Jacopo Andreuccetti [19], Elisa Arici [19], Giusto Pignata [19], Guido A.M. Tiberio [20], Mauro Podda [21], Cristina Murru [22], Massimiliano Veroux [23],Costanza Distefano

✉ Marcello Maestri
m.maestri@smatteo.pv.it

1 Dipartimento di scienze clinico-chirurgiche, diagnostiche e pediatriche, University of Pavia, Pavia, Italy

2 Fondazione IRCCS Policlinico San Matteo, SC Chirurgia Generale 1, P.le Golgi 19, Pavia 27100, Italy

3 Fondazione IRCCS Policlinico San Matteo, SC Anatomia Patologica, Pavia, Italy

4 Fondazione IRCCS Policlinico San Matteo, SC Radiologia, Pavia, Italy

[23], Danilo Centonze [23], Fausto Catena [24], Francesco Favi [24], Raffaele Bova [24], Girolamo Convertini [24], Vanni Agnoletti [24], Andrea Balla [25], Diego Sasia [26], Giorgio Giraudo [26], Anania Gabriele [27], Nicola Tartaglia [28], Giovanna Pavone [28], Fabrizio D'Acapito [29], Nicolò Fabbri (30), Francesco Ferrara (31), Stefania Cimbanassi (32), Luca Ferrario (32), Stefano Cioffi (32), Marco Ceresoli (33), Chiara Fumagalli (33), Luca Degrate (33), Maurizio Degiuli (34), Silvia Sofia (34), Leo Licari (35), Andrea Anderloni (36), Ahmed Ghaly (36), Andrea Dagnoni (36), Claudio Tassi (36), Tommaso Dominioni (36), Matteo Tomasoni (36), Lorenzo Cobianchi (36), Enrico Cicuttin (36), Jacopo Viganò (36), Camilla Nikita Farè (36), Benedetta Sargenti (36), Mario Improta (36), Valeria Musella (36), Giulia Gambini (36), Alberto Patriti (37),Diego Coletta (37), Luigi Conti (38), Michele Malerba (39), Muratore Andrea (40), Marcello Calabrò (40), Beatrice De Zolt (40), Gabriele Bellio (41), Alessio Giordano (42), Davide Luppi (43), Carlo Corbellini (44),Gianluca Matteo Sampietro (44), Chiara Marafante (45), Stefano Rossi (46), Andrea Mingoli (47), Pierfrancesco Lapolla (47), Pierfranco M Cicerchia (47), Leandro Siragusa (48), Michele Grande (48), Claudio Arcudi (49), Amedeo Antonelli (49), Danilo Vinci (49), Ciro De Martino (50), Mariano Fortunato Armellino (50), Enrica Bisogno (50), Diego Visconti (51), Mauro Santarelli (51), Elena Montanari (51), Alan Biloslavo (52), Paola Germani (52), Claudia Zaghi (53), Naoki Oka (54), Mohd Azem Fathi (55), Daniel Ríos-Cruz (56), Edgard Efren Lozada Hernandez (57), Ibrahim Umar Garzali (58), Liliana Duarte (59), Ionut Negoi (60), Andrey Litvin (61), Sharfuddin Chowdhury (62), Salem M. Alshahrani (62), Silvia Carbonell-Morote (63), Juan J. Rubio-Garcia (63), Claudia Cristina Lopes Moreira (64), Iñigo Augusto Ponce (64), Fernando Mendoza-Moreno (65), Anna Muñoz Campaña (66), Heura Llaquet Bayo (66), Andrea Campos Serra (66), Begoña Estraviz-Mateos (67), Izaskun Markinez-Gordobil (67), Aitor Landaluce-Olavarria (67), Mario Serradilla-Martín (68), Antonio Cano-Paredero (68), Miguel Ángel Dobón-Rascón (68), Hytham Hamid (69), Oussama Baraket (70), Emre Gonullu (71), Sezai Leventoglu (72), Yilmaz Turk (72), Çağrı Büyükkasap (72), Ulaş Aday (73), Yasin Kara (74), Hamit Ahmet Kabuli (75), Semra Demirli Atici (76), Elif Colak (77), Serge Chooklin (78), Serhii Chuklin (78), Elena Martín-Pérez (79), Ángela de la Hoz Rodríguez (79), Marcello Di Martino (80), Salomone Di Saverio (81), Ari Leppäniemi (82), Ernest E. Moore (83), Andrew B. Peitzman (84), Massimo Sartelli (85), Walter Biffl (86).

[1] Department of Surgery, University Hospital Centre Zagreb, Zagreb, Croatia.

[2] Department of General Surgery, Mansoura University Hospitals, Mansoura, Egypt.

[3] University Faculty of Medicine, Tanta, Egypt.

[4] Acute care surgery, CHU de Nice, Nice, France.

[5] Dpt of Emergency, Digestive and Metabolic Minimally invasive.

surgery, Poissy and Saint Germain en Laye Hospitals, Poissy, France.

[6] State Medical University, Tbilisi, Georgia.

[7] Hippocration General Hospital of Athens, Athens, Greece.

[8] National and Kapodistrian University of Athens, Laikon General Hospital, Athens, Greece.

[9] Saint Savvas Cancer Hospital, Athens, Greece.

[10] Naval and Veterans Hospital, Department of Surgery, Athens, Greece.

[11] University General Hospital Attikon-UoA, Chaidari, Greece.

[12] Surgery Department General Hospital of Chania—Saint George, Crete, Greece.

[13] Surgery University Hospital of Heraklion, Medical School of Heraklion, Crete, Greece.

[14] General Hospital George Papanikolaou, Thessaloniki, Greece.

[15] Maulana Azad Medical College, New Delhi, India.

[16] Chirurgia generale, Ospedale Monsignor Galliano, Acquiterme, Italy.

[17] Lorenzo Bonomo Hospital, Andria, Italy.

[18] University of Bari, Bari, Italy.

[19] Chirurgia 2, ASST Spedali Civili of Brescia, Brescia, Italy.

[20] Surgical Clinic, University of Brescia, Brescia, Italy.

[21] Department of Surgical Science, University of Cagliari, Cagliari, Italy.

[22] PO Santissima Trinità, Cagliari, Italy.

[23] Azienda Policlinico Università di Catania, Catania, Italy.

[24] Bufalini Hospital, Cesena, italy.

[25] UOC of General and Minimally Invasive Surgery, San Paolo Hospital, Civitavecchia, Italy.

[26] Azienda Ospedaliera Santa Croce e Carle, Cuneo, Italy.

[27] Azienda Ospedaliero-Universitaria di Ferrara, Arcispedale Sant'Anna, Ferrara, Italy.

[28] University Hospital, Foggia, Italy.

[29] Ospedale Morgagni Pierantoni, Forlì, Italy.

(30) Ospedale del Delta, Lagosanto, Italy.

(31) ASST Santi Paolo e Carlo, Milano, Italy.

(32) Chirurgia generale Trauma team, Ospedale Niguarda, Milano, Italy.

(33) General and Emergency Surgery, School of Medicine and Surgery, Milano-Bicocca University, Monza, Italy.

(34) Surgical oncology and digestive surgery, A.O.U. San Luigi Gonzaga, Orbassano, Italy.

(35) FBF Buccheri La Ferla, Palermo, Italy.
(36) Fondazione IRCCS Policlinico San Matteo, Pavia, Italy.
(37) AO Ospedali Riuniti Marche Nord, Pesaro, Italy.
(38) G, Da Saliceto Hospital, Piacenza, Italy.
(39) ASL2 Savonese, Pietra Ligure, Italy.
(40) Chirurgia generale Ospedale E. Agnelli, Pinerolo, Italy.
(41) Ospedale Immacolata Concezione, Piove di Sacco, Italy.
(42) "S.Stefano" New Hospital, Prato, Italy.
(43) ASMN, Reggio Emilia, Italy.
(44) Ospedale di Rho—ASST Rhodense, Rho, Italy.
(45) Chirurgia Generale Ospedale degli Infermi, Rivoli, Italy.
(46) San Filippo Neri Hospital, Roma, Italy.
(47) Policlinico Umberto I Sapienza, Roma, Italy.
(48)University Tor Vergata, Roma, Italy.
(49) UOC Chirurgia Bariatrica, Ospedale San Carlo di Nancy, Roma, Italy.
(50) University Hospital "S. Giovanni di Dio e Ruggi d'Aragona", Salerno, Italy.
(51) Chirurgia Generale d'Urgenza e PS, AOU Città della Salute e della Scienza, Torino, Italy.
(52) General Surgery Department, Cattinara University Hospital, ASUGI, Trieste, Italy.
(53) Ospedale San Bortolo, Vicenza, Italy.
(54) Kurashiki Central Hospital, Kurashiki, Japan.
(55) Department of Surgery, Hospital Universiti Sains Malaysia, Kota Bharu, Malaysia.
(56) Instituto Mexicano del seguro social, Ciudad de México, Mexico.
(57) Regional Hospital of High Speciality of Bajio, Leon Guanajuato, Mexico.
(58) Aminu Kano Teaching Hospital, Kano, Nigeria.
(59) Centro Hospitalar Tondela-Viseu, Viseu, Portugal.
(60) Emergency Hospital of Bucharest, Bucharest, Romania.
(61) Immanuel Kant Baltic Federal University, Regional Clinic Hospital, Kaliningrad, Russia.
(62) General Surgery, King Saud Medical City, Riyadh, Saudi Arabia.
(63) Department of Surgery and Liver Transplantation. Hospita General Universitario Dr. Balmis, Alicante, Spain.
(64) Hospital universitario Donostia, Donostia-San Sebastian, Spain.
(65) Hospital Universitario Príncipe de Asturias, Madrid, Spain.
(66) Emergency Surgery Unit, Parc Tauli Hospital, Sabadell, Spain.
(67) Urduliz Hospital, Urduliz, Spain.
(68) Department of Surgery, Miguel Servet University Hospital, Zaragoza, Spain.
(69) Kuwaiti Specialized Hospital, Khartoum, Sudan.
(70) Bizerte hospital, Bizerte, Tunis.
(71) Sakarya Training and Research Hospital, Adapazari/Sakarya, Turkey.
(72) Gazi University, School of Medicine, Department of Surgery, Ankara, Turkey.
(73) Department of Gastroenterological Surgery, Dicle University School of Medicine, Diyarbakir, Turkey. (74) Kanuni Sultan Süleyman Training and Research Hospital, Instanbul, Turkey.
(75) Bakirkoy Dr Sadi Konuk Training and Research Hospital, Instanbul, Turkey.
(76) Tepecik Training and Research Hospital, Konak/İzmir, Turkey.
(77) Samsun Training and Research Hospital, Samsun, Turkey.
(78) Regional Clinical hospital, Lviv, Ukraine.
(79) Department of General and Digestive Surgery, Hospital Universitario La Princesa, Instituto de Investigación Sanitaria Princesa (IIS-IP), Universidad Autónoma de Madrid (UAM), Madrid, Spain.
(80) Hepato-Biliary and Liver Transplantation Department, AORN Cardarelli, Napoli.
(81) Department of Surgery, Madonna Del Soccorso Hospital, San Benedetto del Tronto, Italy.
(82) Abdominal Center, Helsinki University Hospital and University of Helsinki, Helsinki, Finland.
(83) Denver Health System—Denver Health Medical Center, Denver, USA.
(84) Department of Surgery, University of Pittsburgh School of Medicine, UPMC-Presbyterian, Pittsburgh, USA.
(85) Department of Surgery, Macerata Hospital, 62,100 Macerata, Italy. (86) Division of Trauma/Acute Care Surgery, Scripps Clinic Medical Group, La Jolla, CA, USA.
.

Introduction

Gallbladder gallstones have an incidence of approximately 10–15% in the general population, with key risk factors including female sex, obesity, pregnancy, physical inactivity, advanced age, and genetic predisposition. Among these patients, 20–40% will develop complications, with acute calculous cholecystitis (ACC) occurring in 10–15% of cases [1, 2].

The most widely adopted guidelines for the management of acute cholecystitis are the Tokyo Guidelines and those issued by the World Society of Emergency Surgery. According to these recommendations, treatment selection is influenced by multiple factors, including the patient's general condition, the timing of clinical presentation, and the rapidity of disease progression [1, 3, 4].

It is well established that early laparoscopic cholecystectomy (ELC) represents the gold standard whenever feasible, as performing surgery at the earliest possible stage is associated with reduced complication rates, shorter hospital stays, and lower healthcare costs. The Tokyo Guidelines define the optimal surgical window as within 72 h from symptom onset [5–7].

Several predictive scores have been developed to assess surgical risk and/or mortality, including the ASA score, POSSUM and APACHE score [8–12]. Multiple studies indicate that ELC remains a superior alternative compared to non-surgical management, offering better survival outcomes even in high-risk patients, provided that careful selection criteria are applied. Current guidelines define non-operable patients as those who either refuse surgery or are deemed unfit for the procedure, without providing a standardized definition. However, recent studies suggest that some patients with high-risk scores may benefit from non-surgical treatments [13–15].

The reported complication rate for acute cholecystitis ranges from 7% to 26%, while mortality rates vary between 0% and 10%. Given the high prevalence of this condition, reducing complications and mortality remains a priority in clinical practice [10, 16, 17].

Therefore, establishing a valid and reliable method to identify high-risk patients for postoperative complications, who could benefit from less invasive treatments, would be of great importance.

This study aims to analyze patients undergoing surgery for acute cholecystitis across multiple centers and develop two complementary machine learning models: a preoperative model, based on clinical and demographic variables available before surgery, and a comprehensive model, incorporating intraoperative factors to refine risk prediction. By identifying key determinants of postoperative complications, this approach seeks to enhance risk stratification, optimize surgical decision-making, and improve patient outcomes while mitigating the burden of acute cholecystitis on healthcare systems [18, 19].

Materials and methods

We retrospectively analyzed the S.P.Ri.M.A.C.C. (Scores for Prediction of RIsk for postoperative major Morbidity after cholecystectomy in Acute Calculous Cholecystitis) study database. The S.P.Ri.M.A.C.C. study, a WSES prospective multicenter observational study, collected 1,253 patients from 79 centers located in 19 different countries from 1st September 2021 to 1st September 2022.

Inclusion criteria

- A confirmed diagnosis of ACC in accordance with the 2018 Tokyo Guidelines.
- Suitability for ELC during the index hospital stay. Notably, opting for a different surgical approach (e.g., open surgery or a bailout technique such as subtotal cholecystectomy) was not considered an intraoperative exclusion criterion.
- Age ≥ 18 years.
- Evaluation for possible common bile duct (CBD) stones, with preoperative ERCP if these stones were confirmed.
- Provision of a signed and dated informed consent form.
- Willingness to adhere to all study-related procedures and availability for the entire study duration.

Exclusion criteria

- Pregnancy or breastfeeding.
- Acute cholecystitis of a non-gallstone etiology.
- Concurrent cholangitis or pancreatitis.
- Intraoperative management of CBD stones.
- Any condition that could pose additional risk to the patient or prevent full participation in or completion of the study.

All patients underwent initial clinical evaluation and imaging studies upon hospital admission. Key demographic and clinical data were collected, including age, sex, comorbid conditions, vital signs, laboratory values, and radiological findings. In addition, data on prior abdominal surgeries and the presence of systemic complications were documented (Table 1).

Dataset Pre-processing

To prepare the dataset for analysis, text entries that contained numerical values represented with commas or unconventional formats were corrected and cast into standard numeric types. Missing values within numerical columns were imputed using the median of the respective variable, whereas missing categorical values were imputed using the most frequent category. These preprocessing steps ensured a clean, consistent dataset suitable for machine learning analysis.

Table 1 Patients baseline Characteristics. This table summarizes key baseline characteristics of 1,253 patients with acute calculous cholecystitis. For each variable, "n/y" indicates the number of patients without and with the specified condition (No/Yes). Continuous variables are shown as mean ± standard deviation. The Clavien-Dindo (CD) scale categorizes postoperative complications according to their severity

N	1253
Age	59.43±17.03
Sex (M/F)	666/587
BMI	27.24±4.81
Perforated gallbladder (n/y)	1122/109
Common bile duct > 7 mm (n/y)	1115/136
Total Bilirubin > 1.8 mg/dl (n/y)	986/265
Associated common bile duct stones (n/y)	1132/71
Hypotension requiring dopamine (n/y)	1209/40
Decreased level of consciousness (n/y)	1206/41
pO2/FiO2 < 300 (n/y)	1199/40
Oliguria (n/y)	1181/69
Creatinine > 2 mg/dl (n/y)	1172/79
INR > 1.5 (n/y)	1165/81
Platelet count < 100.000/mm3 (n/y)	1218/31
White blood cells > 18,000/mm3 (n/y)	1002/244
Palpable tender mass (n/y)	814/434
Duration of complaints > 72 h (n/y)	773/477
Gangrenous cholecystitis (n/y)	915/336
Pericholecystic abscess (n/y)	1092/154
Hepatic abscess (n/y)	1213/32
Biliary peritonitis (n/y)	1190/53
Emphysematous cholecystitis (n/y)	1201/45
History of myocardial infarction (n/y)	1109/86
Congestive heart failure (n/y)	1124/70
Peripheral vascular disease (n/y)	1152/42
Cerebrovascular disease (n/y)	1139/55
Dementia (n/y)	1160/34
Hemiplegia (n/y)	1182/10
Chronic pulmonary disease (n/y)	1086/108
Connective tissue disease (n/y)	1167/23
Peptic ulcer disease (n/y)	1137/54
Mild liver disease (without portal hypertension) (n/y)	1125/66
Moderate or severe liver disease (n/y)	1178/14
Diabetes without end-organ damage (n/y)	982/211
Diabetes with end-organ damage (n/y)	1160/31
Moderate or severe renal disease (n/y)	1134/56
Tumor without metastasis (n/y)	1146/47
Metastatic solid tumor (n/y)	1178/11
Leukemia (n/y)	1183/7
Lymphoma (n/y)	1184/5
AIDS or HIV + (n/y)	1183/2
Previous abdominal interventions (n/y)	895/319
Previous percutaneous cholecystostomy? (n/y)	1184/30
Warfarin therapy (n/y)	1160/40
Raised Jugular venous pressure (n/y)	1184/7
Atrial fibrillation (n/y)	1125/68
Patient functional status before surgery partially or totally dependent (n/y)	1113/120
Temperature	36.867±0.816
Systolic blood pressure	131.996±21.842
Pulse	83.249±14.923
Respiratory rate	18.375±9.234
Operative time (min)	100.681±47.336
Peritoneal Soiling (n/y)	552/701

Table 1 (continued)

N	1253
Peritoneal Soiling (Minor serous fluid) (n/y)	920/333
Peritoneal Soiling (Purulent fluid) (n/y)	1097/156
Peritoneal Soiling (Free bowel content) (n/y)	1248/5
Peritoneal Soiling (Pus or blood) (n/y)	1240/13
Presence of malignancy (n/y)	1184/69
Mode of surgery (Elective) (n/y)	871/382
Mode of surgery (Emergency: resuscitation of > 2 h possible or operation < 24 h after admission) (n/y)	464/789
Mode of surgery (Emergency: immediate surgery <2 h needed)) (n/y)	1219/34
Conversion to open surgery (n/y)	1045/105
Bail out procedure (n/y) ♣ Subtotal (partial) cholecystectomy ♣ Fundus first technique ♣ Cholecystostomy (drainage only)	1141/108
Intraoperative complication (n/y) ♣ Bleeding (> 500 ml) ♣ Biliary tree injury ♣ Bowel perforation ♣ Major vascular injury ♣ General anesthesia respiratory complications ♣ General anesthesia cardiac complications	1200/51
Length of stay (days)	5.466±5.298
Post-operative in-hospital complications (n/y)	1061/188
Days between symptoms and cholecystectomy	8.681±40.775
Clavien Dindo (CD)	• No complications: *n* = 1061 • CD1: *n* = 48 • CD2: *n* = 62 • CD3a: *n* = 38 • CD3b: *n* = 16 • CD4a: *n* = 9 • CD4b: *n* = 4 • CD5: *n* = 11

Feature selection using Lasso regression

The primary outcome variable was defined as post-operative in-hospital complications. Since the dataset contained a relatively large number of variables, we applied a Lasso regression model to identify the most predictive features [20]. Numerical variables were standardized to ensure consistent scaling. Using 5-fold cross-validation, the Lasso method penalized less predictive features by shrinking their coefficients to zero. Variables with non-zero coefficients were retained for subsequent modeling. This approach allowed us to reduce the complexity of the model, focusing only on the variables most strongly associated with the outcome, and minimizing the risk of overfitting.

Binary outcome definition and model training

The Clavien-Dindo classification of complications was recoded into a binary outcome: values of no complications or CD1–CD2 were labeled as 0, while CD3a–CD5 were labeled as 1 [21]. This dichotomization was chosen to simplify the modeling approach and emphasize the distinction between mild and severe complications. Stratified random sampling was used to split the dataset into an 80% training set and a 20% test set, ensuring that the distribution of outcomes was preserved across subsets.

To evaluate both preoperative and intraoperative risk assessment, we developed two separate machine learning models:

1. Preoperative Model: Trained using only variables available before surgery, including patient demographics, comorbidities, and preoperative clinical parameters.
2. Comprehensive Model: Integrated both preoperative and intraoperative variables, incorporating factors such

as operative time, conversion to open surgery, intraoperative complications, and peritoneal contamination.

A Random Forest classifier was then trained using the features selected through Lasso regression. Cross-validation (5 folds) was conducted on the training set to evaluate the model's robustness and estimate the area under the curve (AUC). The final model was applied to the test set, where its performance was assessed by calculating the AUC and generating a receiver operating characteristic (ROC) curve [22].

In addition to Random Forest, we employed two other machine learning algorithms - XGBoost and Decision Tree classifiers - to evaluate the robustness and generalizability of our predictive models. Each algorithm was trained and validated using the same set of features identified by Lasso regression, as well as the same binary outcome (0 for no or mild complications, 1 for severe complications).

For the Decision Tree classifier, we selected a maximum depth of 3 to balance interpretability with performance. The Decision Tree's structure was visualized to provide insight into the decision-making process, and performance was evaluated using a test set ROC AUC.

XGBoost was configured to handle the binary outcome with an evaluation metric of log-loss. Standard hyperparameters were used, and the model's ROC AUC and classification metrics were assessed on the same test set. This approach allowed for direct comparison across different algorithms and between preoperative and intraoperative models to determine the most effective strategy for postoperative complication prediction.

Feature importance analysis

To assess the relative contribution of each predictor, feature importance scores were extracted from the trained Random Forest model. These scores indicate the degree to which each variable influenced the model's predictions. The feature importance values were normalized and plotted as a horizontal bar chart, allowing a visual comparison of the predictors' relative contributions. This analysis highlights the most influential features driving the model's performance, providing insights into which variables are most strongly associated with post-operative complications.

Web application

To facilitate the clinical usability of the preoperative model (CholeSurgRisk I), a web-based application was developed.

To enhance clinical interpretability, the application does not rely on an arbitrary threshold but instead uses the Youden Index to determine the optimal cutoff, ensuring a balance between sensitivity and specificity for classifying patients as high-risk or low risk [23, 24]. This threshold balances sensitivity and specificity, optimizing the model's performance for real-world application (eFigure 1 in the Supplement).

The application processes the entered clinical parameters through the trained model and returns the risk for major postoperative complications.

Results

Postoperative complications observed

The eTable 1 in the Supplemental file summarizes the postoperative complications observed in the study cohort. Notably, some patients experienced multiple complications, and the Clavien-Dindo classification was assigned based on the most severe complication recorded for each individual. The most frequently reported complications included respiratory complications ($N=45$), deep infections such as intra-abdominal collections and abscesses ($N=35$), and postoperative bleeding ($N=32$). Other complications involved cardiac, renal, and thromboembolic events, as well as biliary tree injuries and wound infections.

LASSO regression analysis and feature selection

To identify the most relevant predictors of postoperative complications, we applied Lasso regression with cross-validation for feature selection: the optimal alpha parameter selected by cross-validation was 0.019320 for the preoperative model and 0.017316 for the comprehensive model, indicating the level of regularization applied to the model.

After performing Lasso regression, we retained two sets of predictive features, distinguishing between:

- Preoperative-only model: Using variables available before surgery.
- Comprehensive model: Incorporating both preoperative and intraoperative factors to enhance predictive accuracy.

In the preoperative model, 14 predictive variables were retained as their coefficients remained nonzero, indicating their relevance in predicting complications. The most significant predictors included age, pulse, perforated gallbladder, pO2/FiO2 < 300, creatinine > 2 mg/dl, INR > 1.5, pericholecystic abscess, hepatic abscess, biliary peritonitis, emphysematous cholecystitis, congestive heart failure, warfarin therapy, gangrenous cholecystitis, and sex. (eTable 2 in the Supplement).

Table 2 Model Comparison. The table shows, for each model, the AUC on the independent test set, the 5-fold cross-validation AUC, and the precision for class 1 (high-risk patients)

Model	5-Fold CV AUC	Test Set ROC AUC	Precision (Class 1)
Preoperative RF	0.6687 (±0.0243)	0.8456	1.00
Comprehensive RF	0.7914 (±0.0419)	0.8903	0.67
Preoperative XGBoost	-	0.8074	0.62
Preoperative Decision Tree	-	0.8634	0.62
Comprehensive XGBoost	-	0.7614	0.50
Comprehensive Decision Tree	-	0.7216	0.60

In contrast, the comprehensive model integrated intraoperative variables, identifying additional predictors such as operative time, conversion to open surgery, intraoperative complications, peritoneal soiling, and bailout procedures as determinants of postoperative complications (eTable 3 in the Supplement).

Models performance and validation

Using the selected features, Random Forest (RF), XGBoost, and Decision Tree classifiers were constructed and evaluated.

Among the models, Random Forest (RF) demonstrated the best predictive performance, achieving a test set AUC of 0.8903 in the comprehensive model, compared to 0.8456 in the preoperative-only model. While XGBoost and Decision Tree showed competitive performance, RF outperformed in terms of recall and overall discrimination, particularly in the preoperative model where it provided the highest precision for Class 1, which was of primary interest (Table 2).

ROC curve of both preoperative and comprehensive models

Among the three tested classifiers (Decision Tree, XGBoost, and Random Forest), Random Forest yielded the highest ROC AUC and the most consistent predictive performance. This superior performance, combined with its ability to handle complex interactions between variables and provide robust feature importance rankings, led to its selection as the final model (Fig. 1) (eFigs. 2 and 3 in the Supplement).

Feature importance of the random forest model

The feature importance analysis of the preoperative model showed that:

- Age, pulse, perforated gallbladder, sex and creatinine > 2 mg/dL were the strongest predictors.
- Markers of systemic inflammation and hemodynamic instability, such as pericholecystic or hepatic abscess, INR > 1.5 and Gangrenous cholecystitis, also had substantial contributions.

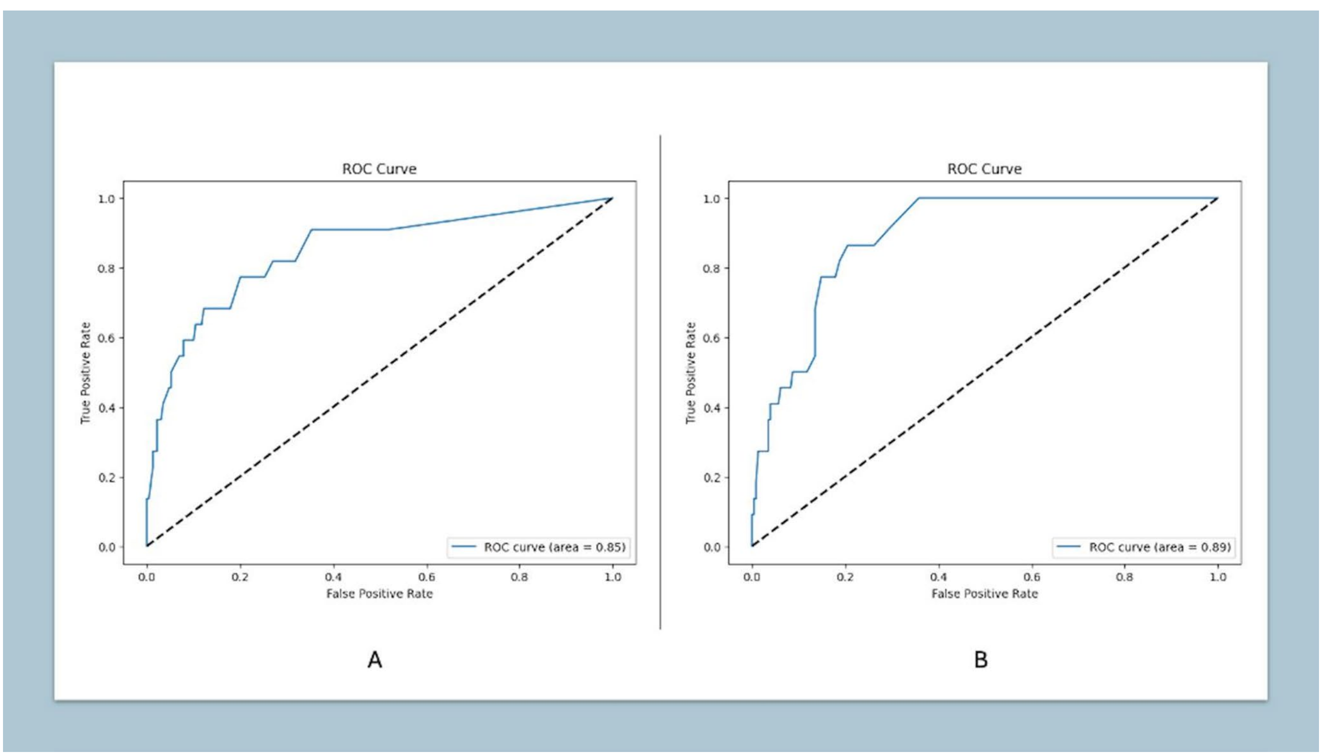

Fig. 1 ROC Curve of the Random Forest Model. On the left the preoperative model (A), on the right the comprehensive one (B)

When intraoperative variables were included, the comprehensive model emphasized additional predictors:

- Conversion to open surgery, bailout procedure, operative time and intraoperative complications were strong predictors.
- Age and pulse remined influential even in the comprehensive model.
- Peritoneal soiling (local pus) also emerged as a key predictor, underscoring the impact of intraoperative findings on postoperative risk.

The feature importance distributions for both models are presented in Fig. 2 (Preoperative model) and Fig. 3 (Comprehensive model). These findings reinforce the role of both preoperative patient status and intraoperative complexity in determining postoperative complications.

Web Application: "CholeSurgRisk I" [25].

The web-based prediction tool was successfully implemented and tested using real-world clinical data and provides early-stage risk estimation.

This application provides an intuitive and accessible tool for clinicians to estimate complication risks in real-time. By integrating this model into clinical practice, surgical decision-making can be optimized, and resource allocation can be managed more efficiently in acute cholecystitis management.

Discussion

Predicting postoperative complications in ACC is a critical aspect of perioperative decision-making, as these adverse events significantly impact patient morbidity, hospital costs, and overall surgical outcomes [10, 26]. In this study, we developed and validated two distinct machine learning models to predict the risk of major postoperative complications in patients undergoing cholecystectomy for ACC. The first model focused on preoperative risk assessment, utilizing only variables available before surgery, whereas the second model incorporated intraoperative factors, providing a more comprehensive risk evaluation.

By leveraging a prospective dataset of 1,253 patients, we applied Lasso regression for feature selection and trained three different machine learning models (Random Forest, XGBoost, and Decision Tree) to compare their predictive performance. Among these, the Random Forest model demonstrated the highest discrimination ability in both approaches (preoperative: ROC AUC 0.8456, intraoperative: ROC AUC 0.8903) and was ultimately selected for deployment in a web-based clinical application.

Postoperative complications following cholecystectomy range from minor adverse events to moderate and severe complications, including biliary leaks, intra-abdominal abscesses, and, in the most severe cases, septic shock [10, 26, 27].

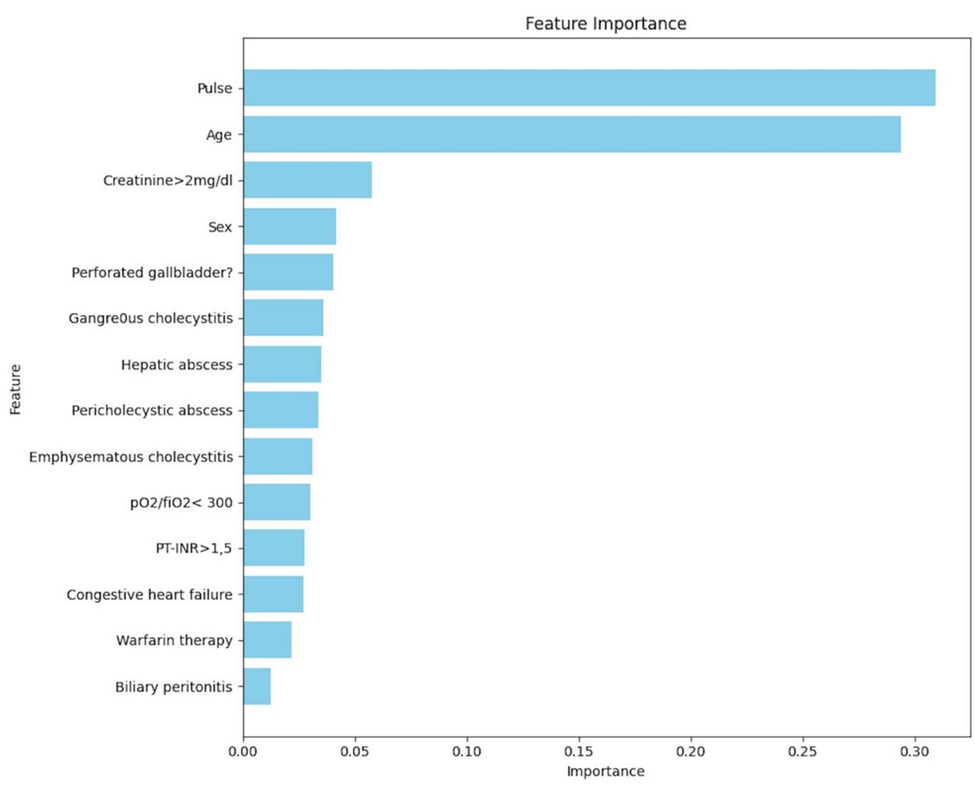

Fig. 2 Feature Importance Analysis of the Preoperative Model. The bar chart illustrates the relative importance of each predictor variable in the preoperative model. A higher importance score indicates a stronger influence on the model's prediction

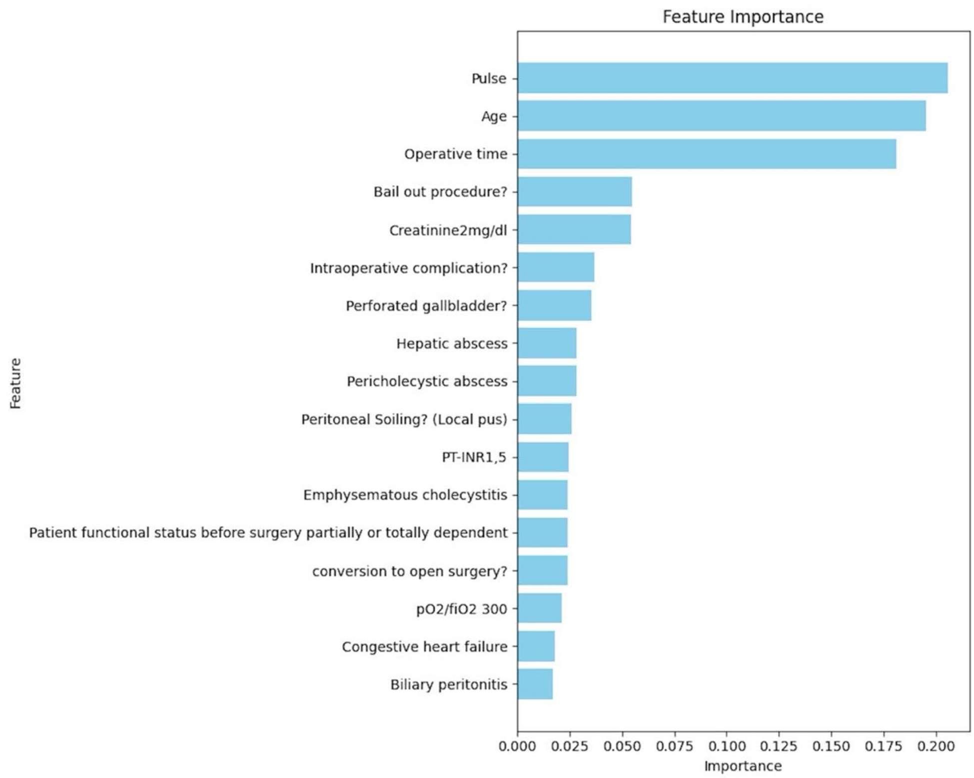

Fig. 3 Feature Importance Analysis of the Comprehensive Model. This bar chart shows the relative importance of each predictor variable in the comprehensive model, incorporating both preoperative and intraoperative factors. A higher importance score indicates a stronger influence on the model's prediction

Although ELC is widely recognized as the gold standard for treating ACC, certain high-risk patients may benefit from alternative treatment strategies, such as endoscopic drainage or perioperative optimization [7, 15, 28].

Several risk assessment scores, including CCI, ASA score, POSSUM score, and APACHE II, have been proposed to estimate perioperative morbidity [1, 6, 9, 11–13]. However, these traditional scoring systems rely on predefined weighting of variables and linear assumptions, which may oversimplify complex interactions among patient-specific factors. As a result, they often fail to integrate dynamic and non-linear relationships that can better capture the multifactorial nature of surgical risk.

Unlike static models, machine learning-based approaches provide an adaptive, data-driven framework capable of identifying intricate patterns across clinical, laboratory, and intraoperative variables. By leveraging these advantages, our study advances the field by implementing a more precise and individualized risk prediction model, demonstrating high accuracy in forecasting postoperative complications.

Through Lasso regression, we identified 14 key predictors in the preoperative model, reflecting patient demographics, comorbidity burden, systemic inflammatory response, and hemodynamic status. These features align with previous literature emphasizing the role of patient frailty and inflammatory markers in surgical risk stratification [8, 16, 29]. In contrast, the intraoperative model retained 17 predictive variables, incorporating additional procedural risk factors such as operative time, conversion to open surgery, intraoperative complications, and peritoneal soiling. These findings highlight the dynamic nature of surgical risk, where intraoperative events significantly contribute to postoperative outcomes. The superior performance of the intraoperative model (ROC AUC 0.8903 vs. 0.8456 preoperatively) underscores the importance of incorporating real-time surgical findings into risk prediction models.

To bridge the gap between predictive modeling and clinical application, we developed a web-based risk assessment tool: "CholeSurgRisk I" [25]. The app is freely accessible using the link: https://mlcolecisti-43pfetc6hyq6m6kzvyeat6.streamlit.app .This user-friendly application allows healthcare providers to input relevant clinical parameters and obtain an individualized risk prediction for major postoperative complications in real time. The "CholeSurgRisk I" app focuses on preoperative risk assessment, assisting clinicians in identifying high-risk patients before surgery and guiding preoperative counseling, optimization and decision- making.

The proposed machine learning models enhance personalized risk stratification, offering an individualized and objective estimation of major complications following cholecystectomy. This predictive capability supports shared decision-making, enabling both surgeons and patients to make informed choices regarding treatment options and expectations based on a data-driven assessment of surgical risk. Second, preoperative risk assessment plays a crucial role in patient optimization. Early identification of high-risk patients may also prompt adjustments in perioperative management to mitigate complications and improve postoperative recovery. Moreover, an accurate risk prediction model can support surgical decision-making, assisting clinicians in selecting the most appropriate treatment strategy tailored to each patient's risk profile.

Limitations

Despite the promising results, this study is based on a retrospective analysis of a prospectively collected dataset, which may introduce selection bias. Further studies are needed to assess the feasibility of implementing this model in daily surgical practice and to evaluate its impact on clinical decision-making and patient outcomes.

Conclusion and key points

This study highlights the potential of machine learning to enhance perioperative risk stratification in ACC. We developed and validated two complementary Random Forest models: CholeSurgRisk I, which provides preoperative risk assessment, and CholeSurgRisk II, which integrates intraoperative findings for a more refined evaluation.

The preoperative model was successfully deployed as web-based risk assessment tool, offering clinicians a user-friendly, real-time solution to predict the probability of major postoperative complications. By facilitating personalized risk estimation, this tool may optimize surgical decision-making and aid in preoperative patient optimization.

Key points

- This study developed two machine learning-based models for predicting major postoperative complications: a preoperative model (CholeSurgRisk I) and a comprehensive intraoperative model (CholeSurgRisk II).
- The intraoperative model improved prediction accuracy (ROC AUC 0.8903 vs. 0.8456 preoperatively), highlighting the importance of intraoperative variables in risk assessment.
- CholeSurgRisk I, our user-friendly web application, enables real-time, individualized risk stratification, supporting clinicians in making personalized, data-driven surgical decisions.
- The study's limitations include its retrospective analysis of prospectively collected data, which may introduce selection bias, and the need for further validation to assess the real-world implementation and clinical impact of the predictive model.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s13304-025-02476-8.

Author contributions Maestri M. and Fugazzola P. are co-last authors. Concept and design: Cicerone O., Maestri M.; Writing – Original draft: Cicerone O., Fugazzola P.; Data curation: S.P.Ri.M.A.C.C. Collaborative Group; Methodology: all authors; Machine learning model development: Cicerone O., Maestri M.; References screening: Frassini S., Gallotti A., Vanoli A.; Writing – Review and editing: Ansaloni L., Maestri M.; Project administration: Maestri M.

Funding This research received no external funding.

Data availability Data available upon reasonable request.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

1. Pisano M, Allievi N, Gurusamy K, Borzellino G, Cimbanassi S, Boerna D et al (2020) 2020 world society of emergency surgery updated guidelines for the diagnosis and treatment of acute calculus cholecystitis. World J Emerg Surg 15(1):61
2. Lammert F, Gurusamy K, Ko CW, Miquel JF, Méndez-Sánchez N, Portincasa P et al (2016) Gallstones Nat Rev Dis Primers 2:16024
3. Miura F, Takada T, Kawarada Y, Nimura Y, Wada K, Hirota M et al (2007) Flowcharts for the diagnosis and treatment of acute cholangitis and cholecystitis: Tokyo guidelines. J Hepatobiliary Pancreat Surg 14(1):27–34
4. Yokoe M, Hata J, Takada T, Strasberg SM, Asbun HJ, Wakabayashi G et al (2018) Tokyo guidelines 2018: diagnostic criteria and severity grading of acute cholecystitis (with videos). J Hepatobiliary Pancreat Sci 25(1):41–54
5. Fugazzola P, Abu-Zidan FM, Cobianchi L, Dal Mas F, Ceresoli M, Coccolini F, et al. Timing of Early Cholecystectomy for AcuteCalculous Cholecystitis: A Multicentric Prospective Observational Study. Healthcare (Basel). 2023; 11(20)
6. Keus F, de Jong JA, Gooszen HG, van Laarhoven CJ. Laparoscopic versus open cholecystectomy for patients with symptomaticcholecystolithiasis. Cochrane Database Syst Rev. 2006(4):CD006231
7. Altieri MS, Brunt LM, Yang J, Zhu C, Talamini MA, Pryor AD (2020) Early cholecystectomy (< 72 h) is associated with lower rate of complications and bile duct injury: a study of 109,862 cholecystectomies in the state of new York. Surg Endosc 34(7):3051–3056
8. Sato M, Endo K, Harada A, Shijo M. Risk Factors of Postoperative Complications in Laparoscopic Cholecystectomy for AcuteCholecystitis. JSLS. 2020; 24(4)
9. Copeland GP, Jones D, Walters M (1991) POSSUM: a scoring system for surgical audit. Br J Surg 78(3):355–360
10. Tufo A, Pisano M, Ansaloni L, de Reuver P, van Laarhoven K, Davidson B et al (2021) Risk prediction in acute calculous cholecystitis: A systematic review and Meta-Analysis of prognostic factors and predictive models. J Laparoendosc Adv Surg Tech A 31(1):41–53
11. Saklad M (1941) Grading of patients for surgical procedure. Anesthesiology 2:281–284

12. Knaus WA, Draper EA, Wagner DP, Zimmerman JE (1985) APACHE II: a severity of disease classification system. Crit Care Med 13(10):818–829
13. Kivivuori A, Salminen P, Ukkonen M, Ilves I, Vihervaara H, Zalevskaja K et al (2023) Laparoscopic cholecystectomy versus antibiotic therapy for acute cholecystitis in patients over 75 years: randomized clinical trial and retrospective cohort study. Scand J Surg 112(4):219–226
14. Tartaglia D, Coccolini F, Cremonini C, Strambi S, Musetti S, Cicuttin E et al (2022) Laparoscopic cholecystectomy for acute calculous cholecystitis in elderly. More complex but equally safe and effective. Ann Ital Chir 93:550–556
15. Loozen CS, van Santvoort HC, van Duijvendijk P, Besselink MG, Gouma DJ, Nieuwenhuijzen GA et al (2018) Laparoscopic cholecystectomy versus percutaneous catheter drainage for acute cholecystitis in high risk patients (CHOCOLATE): multicentre randomised clinical trial. BMJ 363:k3965
16. Ausania F, Guzman Suarez S, Alvarez Garcia H, Senra del Rio P (2015) Casal Nuñez E. Gallbladder perforation: morbidity, mortality and preoperative risk prediction. Surg Endosc 29(4):955–960
17. Ma Y, Luo M, Guan G, Liu X, Cui X, Luo F (2025) An explainable predictive machine learning model of gangrenous cholecystitis based on clinical data: a retrospective single center study. World J Emerg Surg 20(1):1
18. Bi Q, Goodman KE, Kaminsky J, Lessler J (2019) What is machine learning? A primer for the epidemiologist. Am J Epidemiol 188(12):2222–2239
19. Jiang T, Gradus JL, Rosellini AJ (2020) Supervised machine learning: A brief primer. Behav Ther 51(5):675–687
20. Xi LJ, Guo ZY, Yang XK, Ping ZG (2023) [Application of LASSO and its extended method in variable selection of regression analysis]. Zhonghua Yu Fang Yi Xue Za Zhi 57(1):107–111
21. Dindo D, Demartines N, Clavien PA (2004) Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg 240(2):205–213
22. Metz CE (1978) Basic principles of ROC analysis. Semin Nucl Med 8(4):283–298
23. Fluss R, Faraggi D, Reiser B (2005) Estimation of the Youden index and its associated cutoff point. Biom J 47(4):458–472
24. Lai CY, Tian L, Schisterman EF (2012) Exact confidence interval Estimation for the Youden index and its corresponding optimal cut-point. Comput Stat Data Anal 56(5):1103–1114
25. Cicerone O, CholeSurgRisk I 2025 [Available from: https://mlcolecisti-43pfetc6hyq6m6kzvyeat6.streamlit.app
26. Gallaher JR, Charles A (2022) Acute cholecystitis: A review. JAMA 327(10):965–975
27. Machado NO (2011) Biliary complications postlaparoscopic cholecystectomy: mechanism, preventive measures, and approach to management: a review. Diagn Ther Endosc 2011:967017
28. Jang WS, Lim JU, Joo KR, Cha JM, Shin HP, Joo SH (2015) Outcome of Conservative percutaneous cholecystostomy in high-risk patients with acute cholecystitis and risk factors leading to surgery. Surg Endosc 29(8):2359–2364
29. Cillara N, Podda M, Cicalò E, Sotgiu G, Provenzano M, Fransvea P et al (2023) A prospective cohort analysis of the prevalence and predictive factors of delayed discharge after laparoscopic cholecystectomy in italy: the DeDiLaCo study. Surg Laparosc Endosc Percutan Tech 33(5):463–473