World Journal of Emergency Surgery

RESEARCH

Open Access

A radiomics-clinical predictive model for difficult laparoscopic cholecystectomy based on preoperative CT imaging: a retrospective single center study

Rui-Tao Sun[1†], Chang-Lei Li[1†], Yu-Min Jiang[1], Ao-Yun Hao[1], Kui Liu[1], Kun Li[1], Bin Tan[1], Xiao-Nan Yang[2], Jiu-Fa Cui[2], Wen-Ye Bai[1], Wei-Yu Hu[1], Jing-Yu Cao[1*] and Chao Qu[1*]

Abstract

Background Accurately identifying difficult laparoscopic cholecystectomy (DLC) preoperatively remains a clinical challenge. Previous studies utilizing clinical variables or morphological imaging markers have demonstrated suboptimal predictive performance. This study aims to develop an optimal radiomics-clinical model by integrating preoperative CT-based radiomics features with clinical characteristics.

Methods A retrospective analysis was conducted on 2,055 patients who underwent laparoscopic cholecystectomy (LC) for cholecystitis at our center. Preoperative CT images were processed with super-resolution reconstruction to improve consistency, and high-throughput radiomic features were extracted from the gallbladder wall region. A combination of radiomic and clinical features was selected using the Boruta-LASSO algorithm. Predictive models were constructed using six machine learning algorithms and validated, with model performance evaluated based on the AUC, accuracy, Brier score, and DCA to identify the optimal model. Model interpretability was further enhanced using the SHAP method.

Results The Boruta-LASSO algorithm identified 10 key radiomic and clinical features for model construction, including the Rad-Score, gallbladder wall thickness, fibrinogen, C-reactive protein, and low-density lipoprotein cholesterol. Among the six machine learning models developed, the radiomics-clinical model based on the random forest algorithm demonstrated the best predictive performance, with an AUC of 0.938 in the training cohort and 0.874 in the validation cohort. The Brier score, calibration curve, and DCA confirmed the superior predictive capability of this model, significantly outperforming previously published models. The SHAP analysis further visualized the importance of features, enhancing model interpretability.

†Rui-Tao Sun and Chang-Lei Li contributed equally to this work.

*Correspondence:
Jing-Yu Cao
cjy7027@163.com
Chao Qu
2011110479@bjmu.edu.cn

Full list of author information is available at the end of the article

Conclusion This study developed the first radiomics-clinical random forest model for the preoperative prediction of DLC by machine learning algorithms. This predictive model supports safer and individualized surgical planning and treatment strategies.

Keywords Laparoscopic cholecystectomy, Radiomics, Machine learning, Surgical difficulty prediction

Background

Cholecystectomy is the standard treatment for cholecystitis, with laparoscopic cholecystectomy (LC) recognized as the gold standard for gallbladder surgery [1]. According to the Tokyo Guidelines, LC is preferred over open cholecystectomy due to its advantages, including reduced postoperative pain, shorter hospital stays, faster recovery, lower risk of wound infections, and improved quality of life [2, 3]. However, in certain cases, LC can present significant technical challenges, necessitating conversion to open surgery. The complexity of the procedure is often exacerbated by progressive inflammatory and fibrotic changes associated with prolonged disease duration [2]. Even when conversion is not required, difficult LC is associated with increased intraoperative bleeding, prolonged operative time, the need for drainage placement, extended hospitalization, and a higher risk of complications such as bile duct injury [4–8]. Accurately predicting difficult LC holds substantial clinical significance. Risk-based patient stratification allows for complex procedures to be managed by experienced hepatobiliary surgeons or performed under their direct supervision, optimizing both patient outcomes and surgical training programs. Furthermore, it facilitates preoperative counseling, ensuring informed consent and allowing patients and their families to plan accordingly [9–11]. More importantly, in high-risk patients with severe gallbladder inflammation, early or urgent biliary drainage can be recommended as a temporizing measure, followed by elective surgery once the patient's overall condition has improved [2, 12].

Most existing predictive models rely primarily on clinical data and serum biomarkers, yet imaging-based predictive parameters remain underutilized. While some studies have incorporated morphological imaging markers, such as gallbladder wall thickening and pericholecystic fluid collection, their observer-dependent nature limits reliability and generalizability across clinical settings [4, 9]. Radiomics, an advanced imaging analysis technique, extracts high-dimensional quantitative imaging features—encompassing morphological, textural, and intensity-based parameters—from medical imaging data. These radiomic features can quantify subvisual phenotypic characteristics beyond the perceptual capabilities of the human eye [13, 14]. However, to our knowledge, no published studies have investigated the use of radiomics-based predictive models for assessing surgical difficulty in LC.

Therefore, this study aims to develop and validate a radiomics-clinical predictive model for evaluating LC difficulty. By integrating high-throughput CT radiomic feature selection with preoperative clinical variables, we seek to enhance predictive accuracy and provide a robust tool for individualized surgical planning.

Methods

Patients

This study was approved by the Institutional Review Board of The Affiliated Hospital of Qingdao University. We retrospectively identified all patients who underwent cholecystectomy at The Affiliated Hospital of Qingdao University between March, 2018, and March, 2023.

Inclusion criteria: (I) Patients who underwent cholecystectomy for cholecystitis only; (II) Preoperative upper abdominal CT scan performed within seven days before surgery at our institution [15]; (III) Complete and reliable clinical and imaging data.

Exclusion criteria: (I) Preoperative percutaneous transhepatic gallbladder drainage or endoscopic transpapillary gallbladder drainage; (II) Presence of gallbladder polyps, adenomyomatosis, or other space-occupying lesions; (III) Intraoperative frozen section analysis or postoperative pathology indicating malignancy; (IV) Patients undergoing unexpected surgical procedures (e.g., common bile duct incision, ulcer resection); (V) Gallbladder calcification or atrophic gallbladder, which were difficult to visualize on CT.

Clinical data collected for each patient included demographic information (gender, age, BMI), history of abdominal surgery, history of endoscopic retrograde cholangiopancreatography (ERCP), laboratory values (Supplementary E1), ultrasound parameters (long diameter of the gallbladder [LDGB], short diameter of the gallbladder [SDGB], gallbladder wall thickness [GBWT], common bile duct diameter [CBDD]), and cholecystitis type (acute vs. chronic). Surgical data were also collected, including operative time, conversion to open surgery, intraoperative complications (e.g., bile duct injuries, injuries to adjacent organs, need for surgical repair or anastomotic procedures), intraoperative blood loss, requirement for abdominal drainage, and length of hospital stay. Patients were then categorized into the difficult laparoscopic cholecystectomy (DLC) group and non-difficult laparoscopic cholecystectomy (NDLC) group based on predefined scoring criteria.

Image acquisition and preprocessing

CT scans were performed using institutional spiral CT scanners, including SOMATOM Definition Flash and SOMATOM Perspective (Siemens Medical Systems), BRILLIANCE16 (Philips Healthcare), and Optima CT670, CT660, CT620, and Revolution CT (GE Healthcare). Scanning parameters were as follows: automatic tube current and tube voltage 120 kV, detector collimation 64×0.6 or 128×0.625 mm, matrix 512×512, slice thickness 5 mm, slice interval 5 mm.

All patients were scanned in the supine position, covering the upper abdomen. To ensure feature reproducibility and mitigate variability across different scanners, all images were resampled to a standardized voxel size of $1 \times 1 \times 5$ mm^3 using a B-spline interpolation algorithm. Voxel intensity values were discretized using a fixed bin width of 32 units, with the window width and level set to 350 and 40, respectively. Preprocessing procedures adhered strictly to the International Biomarker Standardization Initiative to maintain scientific rigor and reproducibility.

Super-resolution reconstruction

Unlike solid organs, the gallbladder has both a serosal and mucosal surface, making delineation of its wall particularly challenging. Given that normal gallbladder wall thickness ranges from one to three mm, marginal regions constitute a significant proportion of the structure, and precise delineation is crucial for accurate analysis. However, the spatial resolution of medical images is often limited due to hardware constraints, acquisition time, and radiation exposure.

To enhance image clarity, we employed a super-resolution reconstruction technique based on a generative adversarial network framework [16, 17]. Generative adversarial networks consist of two adversarial neural networks: a generator, which synthesizes high-resolution images from low-resolution inputs, and a discriminator, which distinguishes real high-resolution images from generated ones. The generator progressively refines its high-resolution to low-resolution mapping, while the discriminator enhances its capacity to differentiate between synthetic and real images.

We curated a large-scale dataset and applied standardized preprocessing techniques, including noise reduction, artifact correction, and intensity normalization. Paired high-resolution to low-resolution samples were generated by controlled downsampling of original high-resolution images. The composite loss function integrated: (I) Gradient loss, ensuring structural consistency in edge features. (II) L1 loss, minimizing pixel-wise deviations. (III) Perceptual loss, based on deep feature comparisons using a pre-trained neural network, preserving anatomical fidelity [18].

This approach increased spatial resolution by a factor of four, transforming$1 \times 1 \times 5$ mm^3 voxels into $0.25 \times 0.25 \times 5$ mm3, effectively eliminating jagged edges and blurring (Supplementary F1). The improved image resolution enhanced region-of-interest (ROI) delineation and minimized interobserver variability, which is crucial for radiomic analysis. All super-resolution reconstruction tasks were performed using Python (version 3.7).

Image segmentation and radiomics features extraction

ROIs were manually delineated on CT images using 3D Slicer (version 5.6.1, https://www.slicer.org/). For each patient, three consecutive slices where the gallbladder wall was clearly visible were selected, while the top and bottom slices were excluded to minimize partial volume effects. Gallbladder wall regions were carefully contoured, avoiding bile and gallstones. The Clot's triangle region was left unmarked if visualization was challenging. To assess interobserver reproducibility, two radiologists (with eight [The 8th author] and fifteen years [The 9th author] of experience) independently segmented 200 randomly selected cases, blinded to clinical data. The inter-class correlation coefficient was computed for radiomics feature consistency, with inter-class correlation coefficient > 0.80 indicating high reproducibility. The remaining segmentations were completed by the radiologist with eight years of experience.

Radiomic features were extracted using PyRadiomics (version 3.0.1, http://pyradiomics.readthedocs.io/) on Python (version 3.7). Features were categorized into six groups: (I) Geometry, (II) Intensity, (III) Texture (GLCM, GLDM, GLRLM, GLSZM, NGTDM), (IV) Fractal, (V) Hessian, and (VI) Topology (Supplementary E2).

Radiomics and clinical feature selection

Both radiomic features and clinical characteristics must undergo statistical preprocessing and analytical procedures before downstream selection. For radiomic features, to further mitigate potential overfitting, we subjected the filtered features to minimum redundancy maximum relevance (mRMR) selection and retained the 30 most relevant features. The most representative features were selected for the radiomics model establishment through the least absolute shrinkage and selection operator (LASSO) method. The optimal regularization weight (λ) was determined through ten-fold cross-validation utilizing a minimum criteria approach to select the λ value corresponding to the lowest cross-validation error. Finally, the optimal features with corresponding nonzero coefficients were linearly radiomics-clinical to obtain a Rad-Score for classification analysis.

For clinical characteristics, apart from LASSO, we also used Boruta, which had been validated in previous radionics studies as a reliable method to determine the most

relevant parts [19]. The Boruta algorithm utilizes Random Forest-based iterative feature selection by comparing original features against shuffled "shadow attributes" as noise benchmarks. Features are retained only if their permutation importance (Z-score) consistently surpasses the maximum shadow attribute score across iterations, ensuring statistically validated relevance. This process eliminates spurious correlations and enhances high-dimensional data stability, effectively identifying strong and subtle predictive variables. Only characteristics retained in both LASSO and Boruta would be used for downstream modeling.

Data balancing and model construction, evaluation, and explanation

To address class imbalance, the Synthetic Minority Oversampling Technique (SMOTE) was employed before model training [20]. SMOTE identifies the k-nearest neighbors for each minority class instance within the feature space. For every selected instance, synthetic data points are created along the line segments connecting it to its neighbors. This interpolation introduces diversity into the minority class while preserving the intrinsic distribution of the original dataset. Unlike random oversampling, which duplicates existing instances and risks model overfitting, SMOTE expands the minority class's representation by filling gaps in the feature space. The synthetic instances encourage classifiers to form broader, more generalized decision regions, thereby improving robustness against majority class dominance. By balancing class distributions through synthetic generation and strategic majority class reduction, SMOTE mitigates bias and enhances predictive performance on imbalanced datasets.

Three models were constructed, including the clinical model, radiomics model, and radiomics-clinical combined model. Six machine learning algorithms were applied: logistic regression, support vector machine, random forest (RF), decision trees, eXtreme Gradient Boostin, and naïve Baye. The ratio of the training cohort to the validation cohort was 7:3 in our research, and the ten-fold cross-validation was employed to evaluate the performance of our models. Model performance was assessed using area under the receiver operating characteristic curve (AUC), accuracy, and Brier score. Decision curve analysis (DCA) and calibration curves were used to evaluate clinical utility. SHapley Additive exPlanations (SHAP) analysis was employed to rank feature importance and provide individualized model interpretation [21, 22].

Statistical analysis

All radiomic features underwent Z-score normalization to ensure comparability across scales after extraction. The normality of distribution was assessed using the Shapiro-Wilk test. Features conforming to normality were subjected to independent samples t-tests for intergroup comparison, while non-normally distributed features were analyzed via Mann-Whitney U tests. Features with statistically significant differences ($p < 0.05$) were retained. Pairwise Pearson correlation coefficients were computed to mitigate multicollinearity, and feature pairs exhibiting high redundancy ($|r| > 0.9$) were processed through recursive elimination.

For clinical characteristics, continuous variables were initially assessed for normality using the Shapiro-Wilk test. Normally distributed data were presented as mean ± standard deviation, while non-normally distributed variables were summarized as median with interquartile range (IQR). Categorical variables were expressed as frequency percentages. Between-group comparisons utilized the following statistical approaches: Chi-square tests for categorical variables, independent two-sample t-tests for normally distributed continuous variables, and the Kruskal-Wallis rank-sum test for nonparametric distributions. A p-value of ≤ 0.001 was considered to denote statistical significance for clinical characteristics. To account for potential confounding effects during modeling, pairwise correlations between continuous variables were evaluated through Spearman's rank correlation analysis [23]. Feature pairs exhibiting redundancy > 0.80 were manually processed. All statistical analysis, calculations, Boruta, SMOTE, model construction, and SHAP were performed using R software (version 4.3.2).

Results

Definition of difficult laparoscopic cholecystectomy

There are no universally accepted criteria for assessing the intraoperative difficulty of LC. Most studies define surgical complexity using parameters such as operative time, conversion to open surgery, and intraoperative complications [4, 5, 9]. However, these metrics are inherently influenced by surgeon experience, institutional protocols, and individual variations in surgical technique, limiting their generalizability. To partially mitigate operator-dependent bias, Bourgouin et al. proposed defining DLC as procedures requiring 1.5 times the individual base time [24]. While this method improves consistency, it imposes constraints on sample size eligibility and retrospective data standardization. Although intraoperative findings provide valuable insights into surgical complexity [11, 25], retrospective studies often face challenges in capturing nuanced details of surgical difficulty. To address these limitations, we developed a new scoring system based on six parameters: operative time, conversion to open surgery, intraoperative complications, intraoperative blood loss, abdominal drainage

Table 1 The new scoring criteria for DLC classification

Clinical characters		Score
Operative time, (min)		
	(0,60)	0
	[60,70)	1
	[70,80)	2
	[80,90)	3
	[90,+∞)	4
intraoperative blood loss, (ml)		
	(0,5]	0
	(5,10]	1
	(10,20]	2
	(20,+∞)	4
length of hospital stay, (day)		
	[0,2]	0
	3	1
	[4,+∞)	3
need for conversion		
	Yes	4
	No	0
intraoperative complications		
	Yes	4
	No	0
usage of abdominal drainage		
	Yes	4
	No	0

Score 0–3: Normal laparoscopic cholecystectomy

Score ≥ 4: Difficult laparoscopic cholecystectomy

placement and length of hospital stay. A total score of ≤ three was classified as NDLC, whereas ≥ four was classified as DLC (Table 1). These scoring criteria were established by a panel of nine experienced surgeons from our department, including: two surgeons with > 30 years of experience, two surgeons with > 20 years of experience, one surgeon with > 15 years of experience and four surgeons with > eight years of experience.

Given that operative time and conversion to open surgery are widely recognized as key indicators of surgical difficulty [4, 9], these parameters were incorporated into our scoring system. Additionally, acute inflammatory changes and chronic fibrotic remodeling in cholecystitis can distort anatomical structures, increasing the risk of bile duct injury and intraoperative bleeding [15, 26]. Severe adhesions or inflammation often necessitate subhepatic drain placement at our center to facilitate early detection of bile leaks, occult hemorrhage, and fluid collections [27]. Lastly, as the standard postoperative hospital stay is two days, an extended length of stay serves as a proxy for surgical complexity and postoperative complications [6].

Clinical characteristics of patients

The workflow of this study is illustrated in Fig. 1. A total of 2,983 patients who underwent laparoscopic cholecystectomy were retrospectively reviewed. After applying the exclusion criteria, 2,055 patients were included in the final analysis. Based on our DLC scoring system, patients were categorized into two groups: DLC group (646, 31.44%) and the NDLC group (1409, 68.56%) (Fig. 2).

Table 2 presents the baseline clinical characteristics of the study cohort. Compared to the NDLC group, patients in the DLC group were: older (60.90 ± 13.48 vs. 57.27 ± 13.59, $p < 0.001$), predominantly male (55.60% vs. 37.00%, $p < 0.001$), and had higher BMI (25.78 ± 3.33 vs. 25.23 ± 3.55, $p = 0.001$). Patients also showed a higher proportion of ERCP history (4.70% vs. 1.20%,

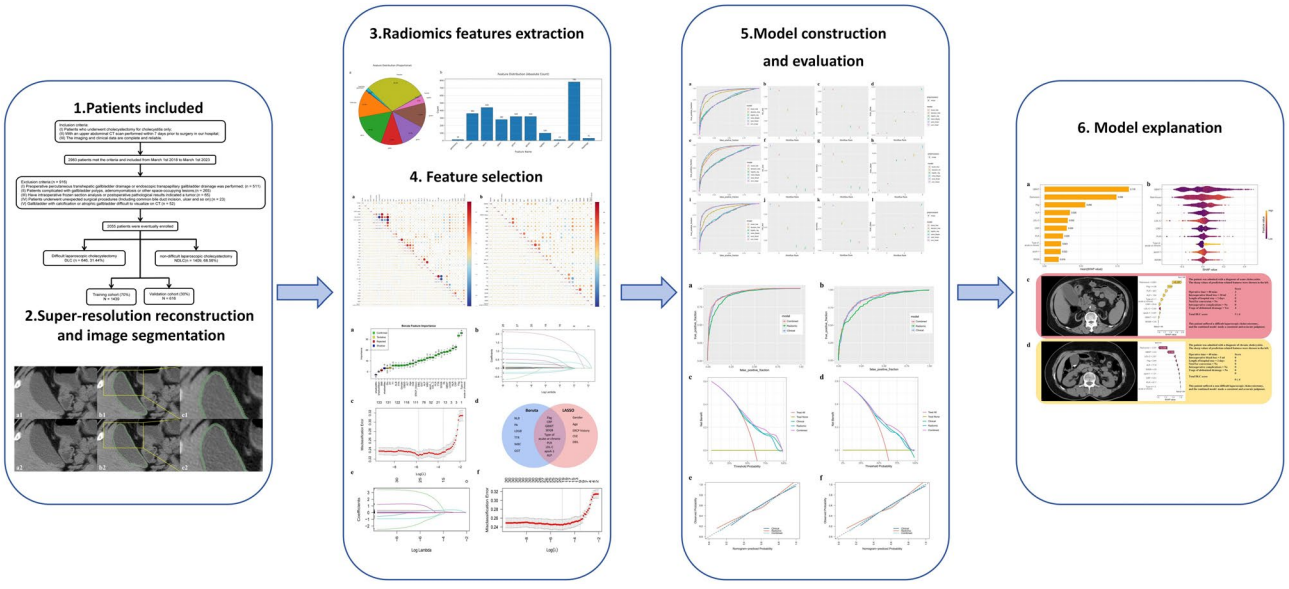

Fig. 1 Workflow of our study

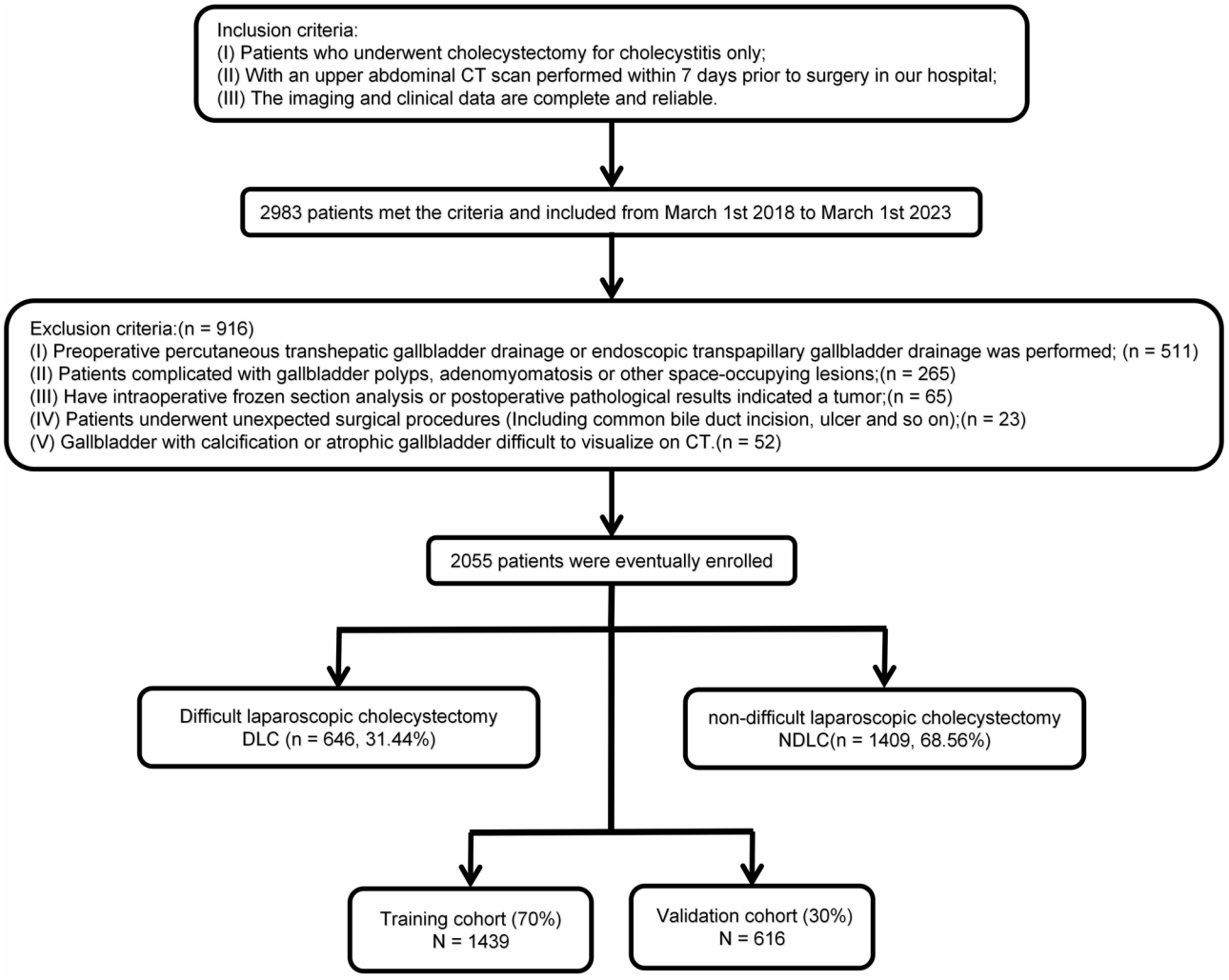

Fig. 2 Flowchart illustrating the patient enrollment process and study design

$p < 0.001$) and acute cholecystitis (33.30% vs. 5.60%, $p < 0.001$). And Laboratory markers of systemic inflammation were significantly elevated in the DLC group, including: C-Reactive Protein (CRP) (17.68 ± 41.94 vs. 2.47 ± 11.07, $p < 0.001$), White Blood Cell Count (WBC) (7.05 ± vs. 6.01 ± 1.89, $p < 0.001$), Lymphocyte Percentage (28.15 ± 12.41 vs. 34.28 ± 9.58, $p < 0.001$), platelet-to-lymphocyte ratio (PLR) (149.30 ± 80.00 vs. 125.74 ± 56.62, $p < 0.001$), Neutrophil-to-Lymphocyte Ratio (NLR) (3.14 ± 3.58 vs. 2.03 ± 2.24, $p < 0.001$). Additionally, biochemical and coagulation parameters were significantly altered, including: lower Low-Density Lipoprotein Cholesterol (LDL-C) (2.64 ± 0.72 vs. 2.89 ± 0.76, $p < 0.001$) and Apolipoprotein A-1 (apoA-1) (1.22 ± 0.28 vs. 1.35 ± 0.26, $p < 0.001$), elevated Alkaline Phosphatase (ALP) (85.88 ± 53.64 vs. 74.08 ± 30.71, $p < 0.001$) and increased fibrinogen (Fbg) (3.67 ± 1.21 vs. 2.98 ± 0.65, $p < 0.001$). Ultrasound parameters were also significantly different between groups: GBWT (0.43 ± 0.21 vs. 0.33 ± 0.14, $p < 0.001$), SDGB (3.11 ± 1.02 vs. 2.69 ± 0.79, $p < 0.001$) and LDGB (7.55 ± 2.00 vs. 6.83 ± 1.48, $p < 0.001$). These findings suggest that a combination of clinical, laboratory, and imaging parameters may enable early prediction of DLC risk.

Feature selection and covariate removal

A correlation analysis was performed on the predictive clinical variables with a p-value threshold of < 0.001 (Fig. 3a). Clinical features with inter-correlation coefficients > 0.80 were manually curated to retain representative features from each correlated cluster, based on clinical relevance and statistical significance. The results revealed a strong correlation between NLR and several other features, including neutrophil percentage, neutrophil count, and lymphocyte percentage. Given the high research value of NLR in multiple studies and its strong correlation with the outcome variable, we decided to retain NLR in the analysis [28–30]. Consequently,

Table 2 The baseline characteristics of the patients

Variable		Difficult laparoscopic cholecystectomy (DLC) *n* = 646	Non-difficult laparoscopic cholecystectomy (NDLC) *n* = 1409	*P* value
Age, (years)		60.90 ± 13.48	57.27 ± 13.59	**< 0.001***
Gender, (n%)				**< 0.001**#
	Male	359 (55.57%)	521 (36.98%)	
	Femle	287 (44.43%)	888 (63.02%)	
BMI, (kg/m2)		25.78 ± 3.33	25.23 ± 3.55	**0.001***
Abdominal operation history (n%)				0.821
	Yes	92 (14.24%)	206 (14.62%)	
	No	554 (85.76%)	1203 (85.38%)	
ERCP history, (n%)				**< 0.001**#
	Yes	31 (4.80%)	17 (1.21%)	
	No	615 (95.20%)	1392 (98.79%)	
Type of acute or chronic, (n%)				**< 0.001**#
	Acute	215 (33.28%)	79 (5.61%)	
	Chronic	431 (66.72%)	1330 (94.39%)	
WBC, ($\times 10^9$/L)		7.05 ± 3.14	6.01 ± 1.89	**< 0.001***
Neut%, (%)		61.98 ± 13.57	55.84 ± 10.32	**< 0.001***
Lymph%, (%)		28.15 ± 12.41	34.28 ± 9.58	**< 0.001***
Mono%, (%)		7.37 ± 2.08	7.11 ± 1.93	0.005
Eos%, (%)		2.02 ± 2.19	2.33 ± 2.09	0.100
Baso%, (%)		0.47 ± 0.31	0.51 ± 0.28	0.004
Neut#, ($\times 10^9$/L)		4.65 ± 3.02	3.54 ± 4.04	**< 0.001***
Lymph#, ($\times 10^9$/L)		1.75 ± 0.67	2.04 ± 2.49	0.003
Mono#, ($\times 10^9$/L)		0.51 ± 0.26	0.65 ± 8.67	0.684
Eos#, ($\times 10^9$/L)		0.12 ± 0.17	0.13 ± 0.14	0.639
Baso#, ($\times 10^9$/L)		0.03 ± 0.02	0.04 ± 0.28	0.481
RBC, ($\times 10^{12}$/L)		4.50 ± 0.51	4.54 ± 0.49	0.114
Hb, (g/L)		136.90 ± 29.50	135.91 ± 21.18	0.392
Hct, (%)		59.13 ± 4.38	40.94 ± 4.35	0.150
MCV, (fL)		90.08 ± 4.70	90.20 ± 4.95	0.597
MCH, (pg)		30.25 ± 1.86	30.18 ± 1.92	0.436
MCHC, (g/L)		335.74 ± 11.75	334.53 ± 11.82	0.031
PLT, ($\times 10^9$/L)		229.48 ± 68.18	227.49 ± 58.17	0.495
MPV, (fL)		10.04 ± 1.00	10.08 ± 1.00	0.485
PDW, (%)		13.38 ± 2.52	13.18 ± 2.51	0.098
PLR		149.30 ± 80.00	125.74 ± 56.62	**< 0.001***
NLR		3.14 ± 3.58	2.03 ± 2.24	**< 0.001***
CRP, (mg/L)		17.68 ± 41.94	2.47 ± 11.07	**< 0.001***
TP, (g/L)		74.28 ± 15.14	73.69 ± 13.36	0.373
ALB, (g/L)		43.32 ± 7.92	44.53 ± 7.25	0.001*
PA, (mg/L)		222.75 ± 74.83	253.61 ± 59.48	**< 0.001***
TBIL, (μmol/L)		20.73 ± 16.72	17.36 ± 10.88	**< 0.001***
DBIL, (μmol/L)		8.43 ± 5.68	10.41 ± 6.09	**< 0.001***
IBIL, (μmol/L)		12.62 ± 8.63	11.90 ± 6.76	0.040
ALT, (U/L)		46.63 ± 72.1	34.58 ± 48.82	**< 0.001***
AST, (U/L)		28.45 ± 42.85	23.30 ± 24.57	**0.001***
GGT, (U/L)		87.69 ± 141.82	57.00 ± 105.61	**< 0.001***
ALP, (U/L)		85.88 ± 53.64	74.08 ± 30.71	**< 0.001***
LDH, U/L)		179.73 ± 72.12	172.69 ± 67.95	0.033
ChE, (kU/L)		8030.95 ± 1670.34	8611.72 ± 1553.57	**< 0.001***
TG, (mmol/L)		1.28 ± 0.68	1.38 ± 1.08	0.040
TC, (mmol/L)		4.01 ± 1.37	4.12 ± 1.62	0.144

Table 2 (continued)

Variable	Difficult laparoscopic cholecystectomy (DLC) *n* = 646	Non-difficult laparoscopic cholecystectomy (NDLC) *n* = 1409	*P* value
HDL-C, (mmol/L)	1.18 ± 0.31	1.25 ± 0.28	**< 0.001***
LDL-C, (mmol/L)	2.64 ± 0.72	2.89 ± 0.76	**< 0.001***
apoA-1, (g/L)	1.22 ± 0.28	1.35 ± 0.26	**< 0.001***
apoB, (g/L)	0.84 ± 0.21	0.87 ± 0.21	**0.001***
Lp(a), (mg/dL)	214.70 ± 178.94	209.91 ± 179.34	0.574
FFA, (mmol/L)	0.51 ± 0.20	0.48 ± 0.19	**< 0.001***
Glu, (mmol/L)	5.60 ± 1.70	5.32 ± 1.27	**< 0.001***
Urea, (mmol/L)	5.12 ± 1.88	5.11 ± 1.26	0.923
PT, (s)	11.36 ± 1.41	11.10 ± 1.22	**< 0.001***
PT-INR	1.00 ± 0.11	0.98 ± 0.09	**< 0.001***
PTA, (%)	103.38 ± 22.94	107.25 ± 22.69	**< 0.001***
APTT, (s)	31.26 ± 3.94	31.26 ± 3.44	0.968
APTTR	1.08 ± 0.12	1.08 ± 0.11	0.220
TTR	1.09 ± 0.11	1.13 ± 0.09	**< 0.001***
TT, (s)	15.76 ± 2.25	16.29 ± 2.09	**< 0.001***
Fbg, (g/L)	3.67 ± 1.21	2.98 ± 0.65	**< 0.001***
LDGB, (cm)	7.55 ± 2.00	6.83 ± 1.48	**< 0.001***
SDGB, (cm)	3.11 ± 1.02	2.69 ± 0.79	**< 0.001***
GBWT, (mm)	0.43 ± 0.21	0.33 ± 0.14	**< 0.001***
CBDD, (mm)	0.55 ± 0.15	0.53 ± 0.12	**0.001***

*statistical significance($p \leq 0.001$) after independent two-sample t-tests;

#statistical significance($p \leq 0.001$) after Chi-square tests

WBC, white blood cell count; Neut%, neutrophil percentage; Lymph%, lymphocyte percentage; Mono%, monocyte percentage; Eos%, eosinophil percentage; Baso%, basophil percentage; Neut#, neutrophil count; Lymph#, lymphocyte count; Mono#, monocyte count; Eos#, eosinophil count; Baso#, basophil count; RBC, red blood cell count; Hb, hemoglobin; Hct, hematocrit; MCV, mean corpuscular volume; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; PLT, platelet count; MPV, mean platelet volume; PDW, platelet distribution width; PLR, platelet-to-lymphocyte ratio; NLR, neutrophil-to-lymphocyte ratio; CRP, C-reactive protein; TP, total protein; ALB, albumin; PA, prealbumin; TBIL, total bilirubin; DBIL, direct bilirubin; IBIL, indirect bilirubin; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gamma-glutamyl transferase; ALP, alkaline phosphatase; LDH, lactate dehydrogenase; ChE, cholinesterase; TG, triglyceride; TC, total cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; apoA-1, apolipoprotein A-1; apoB, apolipoprotein B; Lp(a), lipoprotein(a); FFA, free fatty acid; Glu, glucose; PT, prothrombin time; PT-INR, prothrombin time international normalized ratio; PTA, prothrombin time activity; APTT, activated partial thromboplastin time; APTTR, activated partial thromboplastin time ratio; TTR, thrombin time ratio; TT, thrombin time; Fbg, fibrinogen; LDGB, long diameter of the gallbladder; SDGB, short diameter of the gallbladder; GBWT, gallbladder wall thickness; CBDD, common bile duct diameter

neutrophil percentage, neutrophil count, and lymph percentage were excluded. Additionally, clinical features such as direct bilirubin (DBIL), alanine aminotransferase (ALT), apoA-1, LDL-C, and prothrombin time (PT) were retained, while total bilirubin (TBIL), aspartate aminotransferase (AST), high-density lipoprotein cholesterol (HDL-C), apolipoprotein B (apoB), PT-international normalized ratio (PT-INR), and prothrombin activity (PTA) were excluded. The correlations between the remaining variables were further confirmed using heatmaps (Fig. 3b). Feature selection was then conducted using the Boruta and LASSO algorithms. The results of clinical feature screening using the Boruta algorithm are shown in Fig. 4a. Boruta classified features into three categories based on their relative importance compared to the shuffled "shadow" features (blue area). The box plot depicts the Z-values of each variable during model calculation. Features in the green area (e.g., type of acute or chronic disease, Fbg) were confirmed as statistically significant predictors, exhibiting higher importance scores than the maximum shadow feature (shadowMax). Features in the yellow area (e.g., PT, glucose [Glu]) showed tentative importance, with scores overlapping the shadowMax threshold and requiring further validation. Features in the red area (e.g., BMI) were deemed unimportant, as their importance fell below the shadowMin threshold. This analysis retained 25 confirmed features and two tentative features for subsequent modeling, effectively filtering out noise variables like BMI.

The LASSO regression method was then used in combination with Boruta to select the final clinical features. The coefficient paths (Fig. 4b) demonstrated how feature coefficients shrank toward zero as the regularization parameter (λ) increased, with key features maintaining non-zero coefficients across a wide range of λ values, indicating their robust association with the outcome. Cross-validation (Fig. 4c) was used to identify the optimal λ value that minimized mean squared error, balancing model complexity and predictive accuracy. Ultimately, 14 clinical features were retained for further analysis. The

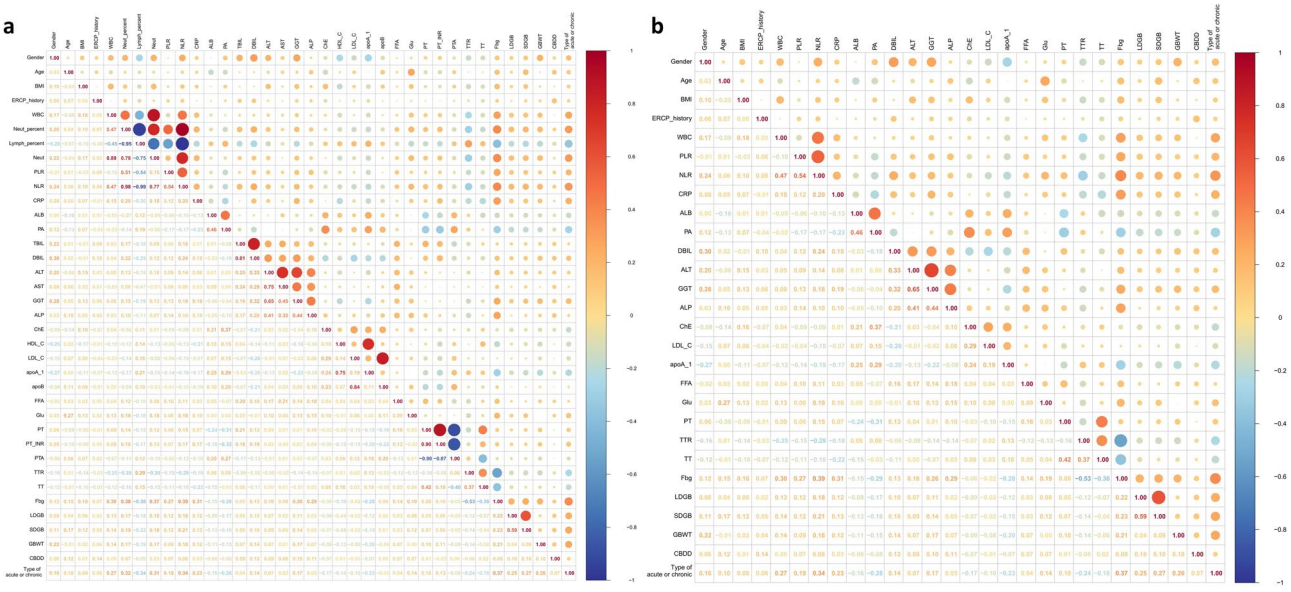

Fig. 3 The correlation heatmap between the clinical variables. **(a)** Correlation Heatmap of all variables. **(b)** Correlation heatmap after removing collinear variables

intersection of the top 15 Boruta-identified features and the 14 LASSO-retained features resulted in the selection of nine clinical features for subsequent modeling: type of acute or chronic disease, Fbg, C-reactive protein, GBWT, SDGB, PLR, apoA-1, ALP, and LDL-C (Fig. 4d).

Radiomics features extraction and selection

In total, 2,658 handcrafted radiomics features were extracted across ten categories for each imaging modality (Supplementary F2). To refine the feature set, a three-step filtering process was applied: first, an interclass correlation coefficient analysis across all features was performed to assess feature reproducibility [with median = 0.88 (range: 0.53–1.00)]; second, statistical tests (t-test or Mann-Whitney U test) were used to identify significant differences between groups; and third, Pearson correlation coefficient analysis was applied to remove highly correlated redundant features. After these steps, 890 features remained for further analysis.

Subsequently, two feature selection methods were used: mRMR and LASSO. The mRMR method was first applied to eliminate redundant and irrelevant features, retaining the 30 most relevant features (Supplementary E3). Then, LASSO regression was conducted to further optimize the feature subset. This process resulted in the selection of nine key radiomics features for model construction (Fig. 4e and f). The final Rad-Score was computed using the following formula:

Rad-Score = -1.59856720711946 + 0.423362681787926 * wavelet_LLH_glszm_GrayLevelNonUniformity + 0.117175965304045 * exponential_ngtdm_Busyness + 0.00201245862173653 * original_shape_MinorAxisLength + 0.000641368517257317*original_topology_SkeletonLength – 1.41953267045166e-05 * wavelet_LLL_glszm_GrayLevelVariance – 5.70803319069496e-05 * wavelet_HLL_ngtdm_Strength – 0.00102732335840969 * logarithm_hessian_AxisZ90Percentile – 0.00207504538088785 * wavelet_HHH_ngtdm_Strength – 0.00465363023203582 * gradient_hessian_AxisZStd.

Model construction and evaluation

Six different machine learning algorithms were evaluated for their predictive performance. Table 3 summarizes the performance metrics of these models. We constructed three types of predictive models: (I) a radiomics model based on Rad-Score alone, (II) a clinical model based on selected clinical features, and (III) a radiomics-clinical model combining both Rad-Score and clinical features. Across all six algorithms, the radiomics-clinical model outperformed the radiomics and clinical models in both the training cohort (Fig. 5) and validation cohort (Supplementary F3). Among the tested algorithms, the RF model demonstrated the highest predictive performance, with an AUC of 0.906 (radiomics), 0.932 (clinical), and 0.938 (radiomics-clinical) in the training cohort (Fig. 6a), and an AUC of 0.840 (radiomics), 0.867 (clinical), and 0.874 (radiomics-clinical) in the validation cohort (Fig. 6b). Given its superior performance, the RF-based radiomics-clinical model was selected as the final predictive model for difficult DLC.

To further validate its effectiveness, we compared our radiomics-clinical model with previously published predictive models (Table 4). The results demonstrated that

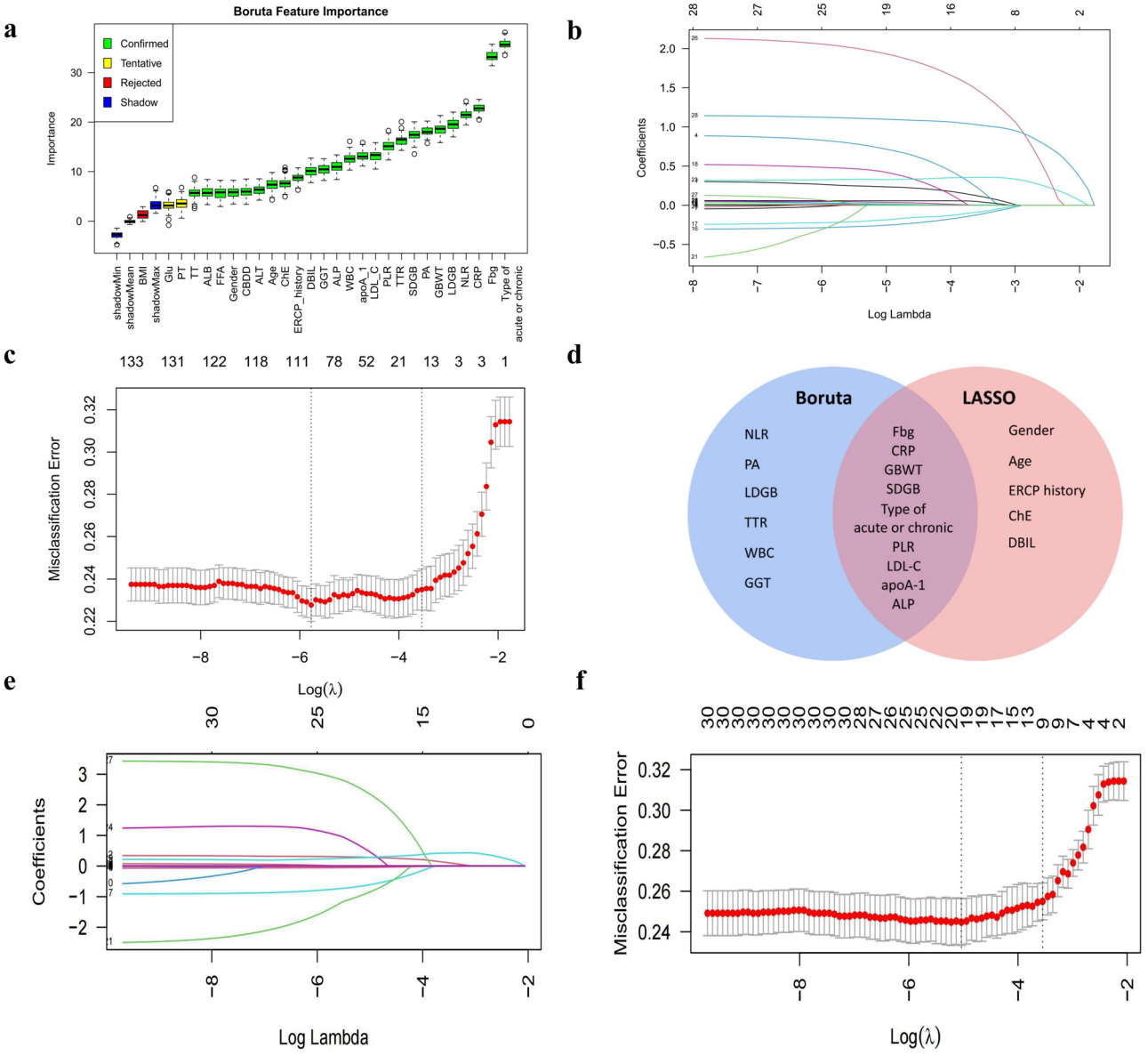

Fig. 4 **(a)** The Boruta algorithm classified features into three distinct categories and features in the green area were confirmed as statistically important predictors. **(b)**, **(c)** The lasso process for clinical feature selection. **(d)** The intersection of the top 15 Boruta-identified features and 14 Lasso-retained features. **(e)**, **(f)**The lasso process for radiomics feature selection

our model achieved higher AUC values than all prior models, indicating superior predictive accuracy and clinical applicability. DCA was conducted to assess the clinical utility of the predictive models. In the training cohort (Fig. 6c), the radiomics-clinical model demonstrated a higher net benefit than both the clinical and radiomics models, indicating superior predictive utility. This advantage was consistently preserved in the external validation cohort (Fig. 6d), highlighting the model's generalizability and robustness. Calibration curves were used to assess the agreement between predicted and actual outcomes. In the training cohort (Fig. 6e), the radiomics-clinical model showed near-perfect alignment between predicted and observed probabilities, outperforming both the clinical and radiomics models. Similarly, in the validation cohort (Fig. 6f), the model maintained strong calibration, with minimal deviation from the ideal 45-degree reference line, confirming stable predictive performance across different populations.

Model explanation and application

To enhance the interpretability of the RF-based radiomics-clinical model, we utilized SHAP to evaluate feature importance and individual prediction contributions. SHAP values were calculated for all predictive features, revealing that GBWT had the highest predictive

Table 3 Different machine learning models' performance

Machine learning models		Radiomics models			Clinical models			Radiomics-Clinical models		
		AUC	Accuracy	Brier score	AUC	Accuracy	Brier score	AUC	Accuracy	Brier score
RF	Training	0.906	0.834	0.130	0.932	0.861	0.109	0.938	0.856	0.106
	Validation	0.840	0.764	0.162	0.867	0.802	0.148	0.874	0.797	0.146
XGboost	Training	0.869	0.799	0.147	0.926	0.859	0.106	0.932	0.864	0.100
	Validation	0.823	0.753	0.169	0.867	0.794	0.146	0.861	0.791	0.151
DT	Training	0.770	0.767	0.220	0.846	0.779	0.159	0.846	0.775	0.162
	Validation	0.737	0.692	0.224	0.802	0.736	0.182	0.771	0.713	0.194
NB	Training	0.730	0.659	0.242	0.760	0.648	0.290	0.778	0.659	0.286
	Validation	0.759	0.693	0.227	0.758	0.654	0.285	0.767	0.672	0.283
LR	Training	0.717	0.655	0.207	0.749	0.681	0.196	0.771	0.695	0.188
	Validation	0.758	0.687	0.196	0.741	0.669	0.197	0.761	0.683	0.192
SVM	Training	0.717	0.646	0.208	0.750	0.677	0.197	0.770	0.687	0.189
	Validation	0.757	0.678	0.197	0.745	0.681	0.198	0.764	0.682	0.192

value (SHAP value = 0.116), followed by the Rad-Score (SHAP value = 0.099) (Fig. 7a). Violin plots (Fig. 7b) visualized the correlation between SHAP values and feature magnitude, with higher feature values (yellow dots) associated with an increased risk of DLC, whereas lower values (purple dots) correlated with lower risk. Interestingly, while most biomarkers showed positive associations with DLC risk, two biochemical markers—LDL-C and apoA-1—exhibited an inverse correlation with DLC development, suggesting a potential protective effect.

To further illustrate individualized risk assessment, Fig. 7c and d present SHAP-based predictions for two representative patients: Fig. 7c depicts a patient who developed DLC, with SHAP values identifying high-risk features contributing to the prediction, while Fig. 7d illustrates a patient who underwent NDLC, with SHAP analysis highlighting protective factors that reduced the difficulty risk. In these plots, yellow features indicate increased DLC risk, whereas purple features indicate reduced risk. The length of each SHAP arrow represents the magnitude of a feature's contribution to the final prediction, where a longer arrow indicates a stronger effect. These findings confirm that SHAP analysis provides a transparent, interpretable framework for understanding model predictions at both the population and individual levels, enhancing its potential for clinical application and decision-making.

Finally, we developed a risk calculator on an online platform (http://123.56.229.150:1003), providing access to web tools tailored to the prediction of DLC. By inputting radiomics and clinical features data directly into designated text fields on the webpage, users can obtain the desired prediction outcomes conveniently.

Discussion

LC is the standard surgical approach for cholelithiasis with cholecystitis and a routine procedure in hepatobiliary surgery [1]. However, even experienced surgeons may encounter intraoperative complications, including bile duct injury, and occasionally require conversion to open surgery [5, 8]. Therefore, preoperative prediction of potential difficult cases holds significant clinical importance. This study developed and validated a radiomics-clinical predictive model for DLC based on preoperative CT images and clinical characteristics using various machine learning models. CT radiomics features and clinically significant data were obtained from 2,055 patients at our medical facility, including 646 DLC patients and 1,409 NDLC patients. The application of deep learning-based super-resolution reconstruction enhanced the quality of low-resolution images by generating high-resolution versions. Among the six machine learning models evaluated, the radiomics-clinical RF model demonstrated optimal performance, achieving

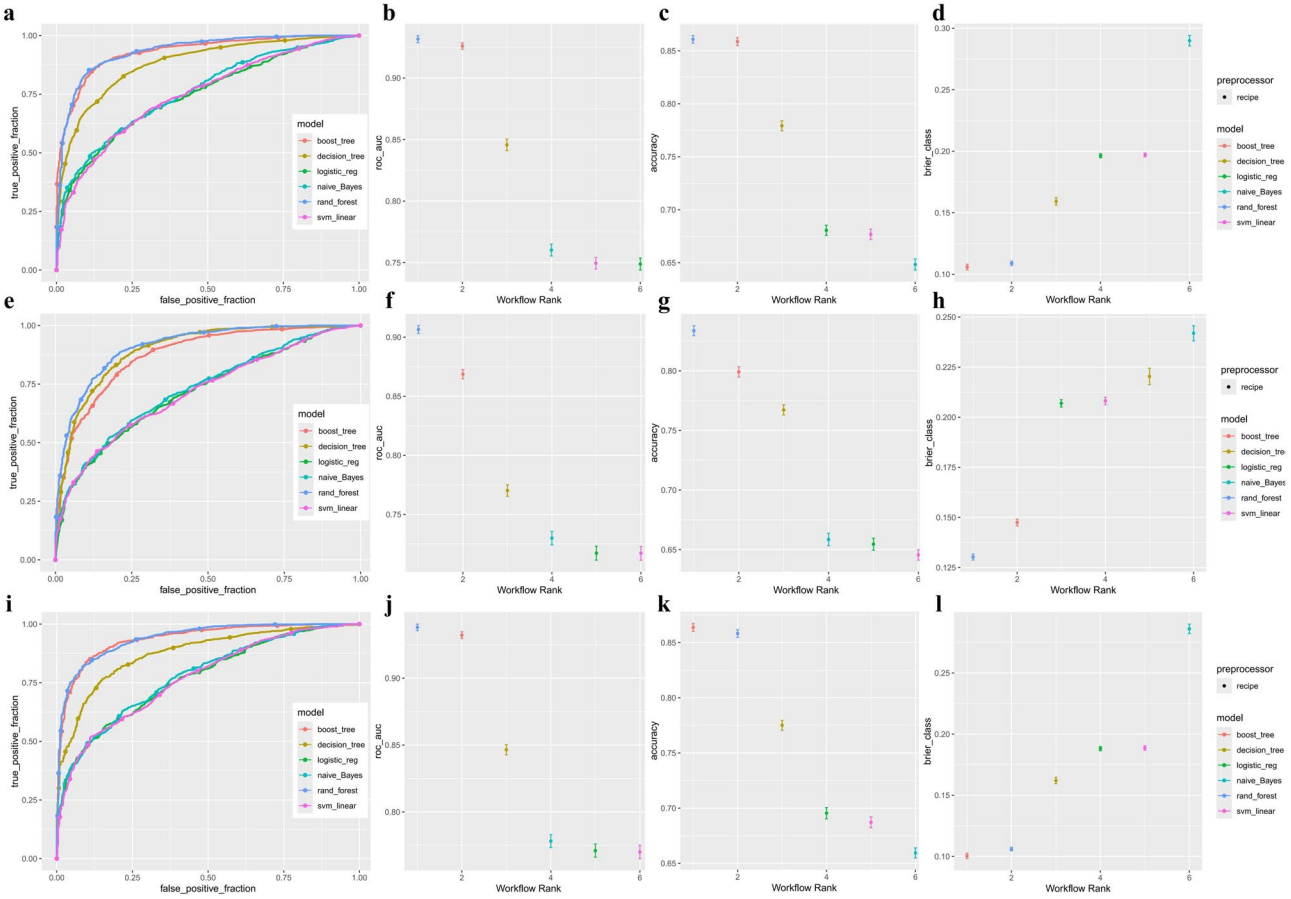

Fig. 5 The performance of the six models in the training cohort. **(a) - (d)** show the AUC, accuracy and brier score of clinical models; **(e)- (h)** show the AUC, accuracy and brier score of radiomics models; **(i) - (l)** show the AUC, accuracy and brier score of radiomics-clinical models

a significantly higher AUC of 0.938 and an accuracy of 85.60% in the training cohort, and an AUC of 0.874 with an accuracy of 79.7% in the validation cohort. DCA and calibration curves further demonstrated the strong clinical utility and reliability of our model. We employed SHAP analysis to generate intuitive visual representations elucidating the relative contributions of input variables. Ultimately, Rad-Score and nine clinical characteristics—GBWT, Fbg, ALP, LDL-C, CRP, PLR, type of acute or chronic disease, apoA1, and SDGB—were identified as pivotal risk factors for distinguishing DLC cases. Based on these findings, we recommend that clinicians use this predictive model preoperatively to improve surgical planning and preparation. This study confirms that AI-driven models offer reliable, cost-effective, and clinically actionable insights to support medical decision-making.

Our study pioneers the application of radiomics in DLC prediction, establishing a novel paradigm for forecasting surgical outcomes in cholelithiasis with cholecystitis. The inclusion of radiomic features encapsulated in the Rad-Score adds substantial value beyond semiquantitative morphological parameters by capturing subtle but critical patterns in preoperative imaging that are often overlooked by human observers. Specifically, radiomic texture analysis provides a means to quantify inflammatory and fibrotic changes, yielding a more comprehensive understanding of disease progression. The best-performing radiomics model in our study achieved an AUC of 0.906 (accuracy: 83.4%) in the training cohort and an AUC of 0.840 (accuracy: 76.4%) in the validation cohort, demonstrating robust predictive performance. Notably, in the SHAP analysis, Rad-Score emerged as the second most influential feature among the ten variables in the radiomics-clinical model, underscoring its strong predictive capacity. Our feature selection process also preserved two semiquantitative morphological parameters—GBWT (SHAP value = 0.116) and SDGB (SHAP value = 0.019)—which were previously associated with DLC predictive efficacy [4], thereby reaffirming their biological relevance through rigorous cross-validation. While other morphological parameters, such as LDGB and CBDD, demonstrated modest but statistically significant associations, they were ultimately excluded from the final model.

The multistage feature selection protocol employed in this study represents a significant methodological

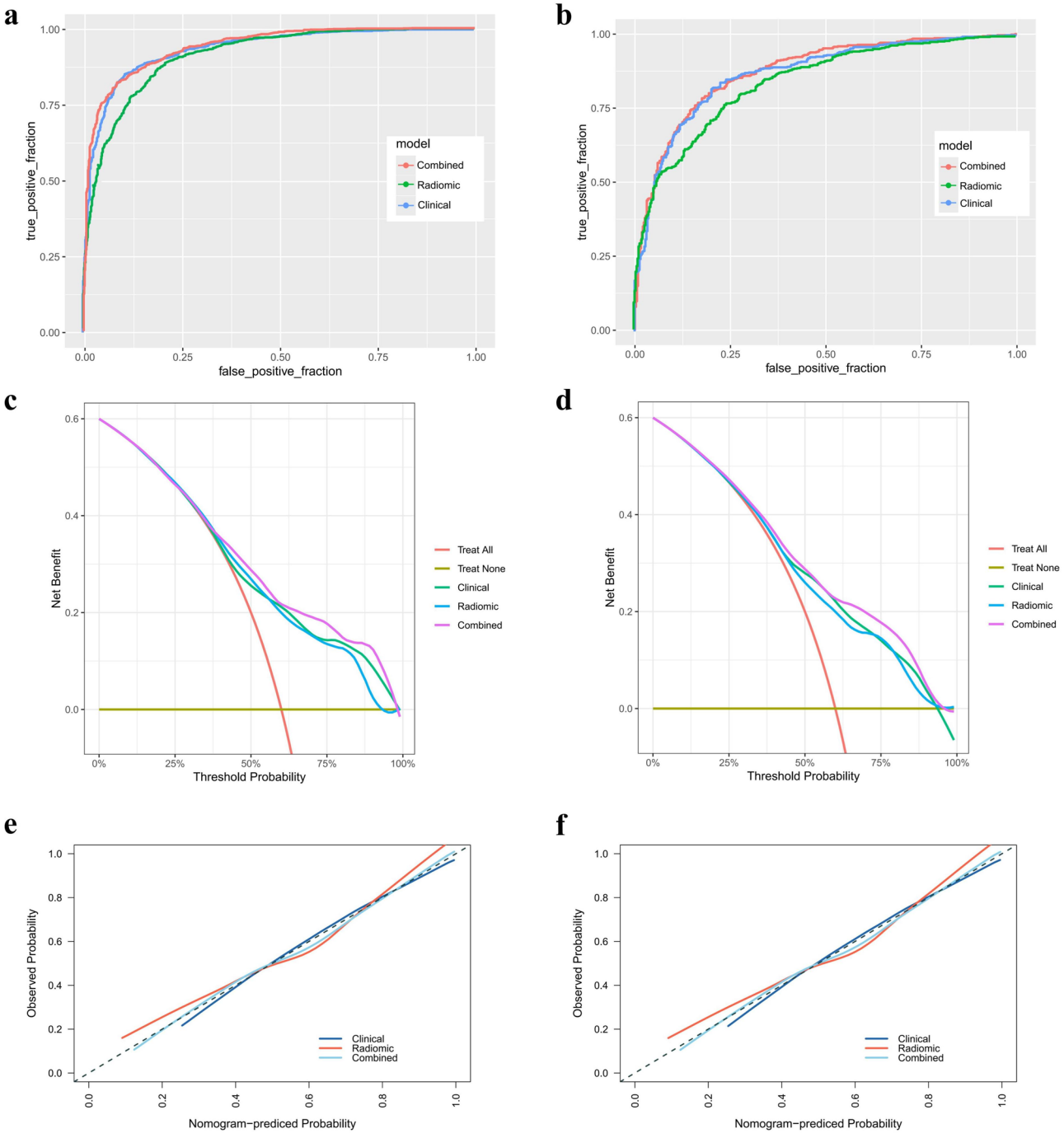

Fig. 6 The performance of the RF models in the training cohort **(a)** and validation cohort**(b)**. The decision curve analysis (DCA) of radiomics-clinical (combined) RF models in the training cohort **(c)** and validation cohort **(d)**. The calibration curves of radiomics-clinical (combined) RF models in the training cohort **(e)** and validation cohort **(f)**

advancement over conventional approaches. First, hypothesis-driven statistical filtering retained only variables demonstrating highly significant intergroup differences ($p < 0.001$) between DLC and NDLC cohorts, thereby substantially reducing the risk of Type I error. Next, we systematically eliminated collinear variables by excluding feature pairs with Spearman's correlation coefficients > 0.80, preemptively addressing multicollinearity-induced coefficient distortion before machine learning processing. Finally, a dual-algorithm consensus approach combining Boruta's all-relevant feature selection method with LASSO regularization was implemented. This orthogonal methodology effectively mitigated overfitting risks while identifying nine

Table 4 Performance of different predictive models

First author	Year of publication	Number of predictors	AUC
Kama	2001	6	0.83
Lipman	2007	6	0.83
Randhawa	2009	9	0.82
Gupta	2013	9	0.86
Kim	2014	3	0.83
Bourgouin	2016	5	0.80
Beksac	2016	4	0.77
Sutcliffe	2016	6	0.80
Hu	2016	5	0.87
Raza	2019	8	0.86
Nassar	2019	8	0.78
Wennmacker	2019	3	0.86
Tongyoo	2021	9	0.82
Ramírez-Giraldo	2022	8	0.88
Our study	**2025**	**10**	**0.94**

high-fidelity predictors through the intersection of both algorithms' results. The sequential execution of statistical filtering, correlation-based elimination, and machine learning driven feature selection proved particularly effective in eliminating 54 redundant variables while preserving the most informative predictors.

Beyond radiomic features, six serological markers and the type of acute or chronic disease were included in the final model. Consistent with previous literature, Fbg and ALP emerged as the top two serological predictors, second only to GBWT and the Rad-Score in predictive importance [24]. Elevated ALP levels likely reflect the presence of Mirizzi syndrome, in which impacted gallstones cause bile duct compression. Although Fbg is primarily recognized as a coagulation marker, it is also influenced by leukocyte-associated inflammatory responses, making it a key biomarker in the context of DLC [31, 32]. Other inflammatory markers, such as PLR and CRP, also demonstrated strong associations with DLC occurrence, aligning with prior studies linking systemic inflammation to surgical complexity. Moreover, our findings corroborate the well-documented association between DLC and acute cholecystitis, as 33.28% of DLC cases in our cohort originated from acute presentations compared to only 5.61% in the NDLC group ($p < 0.001$), further reinforcing this pathophysiological relationship with contemporary clinical data.

Interestingly, our analysis identified two previously unreported biomarkers, LDL-C and apoA1, which exhibited significant negative associations with the occurrence of DLC. A systematic re-evaluation of lipid metabolism parameters—including cholinesterase, triglycerides, total cholesterol, HDL-C, LDL-C, apoA1, and apoB—revealed a consistent pattern: DLC patients had significantly lower mean levels of all lipid biomarkers compared to NDLC controls (all $p < 0.001$). This paradoxical lipid profile closely resembles the well-documented "Lipid Paradox" observed in cardiovascular diseases and has also been reported in other conditions such as rheumatoid arthritis and stroke [33–35]. Moreover, a "U-shaped" association between LDL-C levels and mortality risk, mediated by CRP levels, has been identified in acute myocardial infarction patients, challenging the traditional notion that "lower is always better" for LDL-C management. This phenomenon may be explained by inflammation-induced cholesterol catabolism, leading to increased LDL receptor expression and enhanced LDL oxidation [36]. Oxidized LDL, in turn, exacerbates inflammation, ultimately contributing to disease progression and poorer outcomes [37]. However, this mechanism has not been extensively studied in hepatobiliary diseases, which suggests potential compensatory metabolic adaptations during the progression of cholecystitis. Another possible explanation for the observed reduction in lipid levels is dietary restriction, as patients with cholecystitis often limit their fat intake, further contributing to decreased blood lipid levels.

Despite its strengths, this study has several limitations. (I) Although our model demonstrates satisfactory performance in a large cohort of 2,055 patients, it requires validation through multicenter and prospective studies to ensure broader generalizability. Especially for the definition of DLC which has not be externally validated and could introduce subjective bias. (II) This study primarily focused on retrospective analysis of objective clinical indicators, such as age, BMI, and hematological tests, but did not include an in-depth assessment of descriptive indicators, particularly detailed intraoperative findings and symptom onset characteristics. (III) The radiomics methodologies used in this study can be further refined. The integration of deep learning and automated segmentation techniques may further improve predictive accuracy and offer a more convenient prediction method for clinical application. Future efforts should focus on external multicenter validation and prospective data collection together with the development of hybrid deep learning-radiomics-clinical models and automated segmentation tools which could leverage patient imaging data to provide even more reliable clinical decision support.

Conclusions

This study highlights the significant value of CT-based radiomics in preoperatively predicting the difficulty of LC. By integrating machine learning algorithms with preoperative clinical variables, we successfully developed the first radiomics-clinical random forest model for accurately predicting DLC. This model provides a practical tool for personalized and precise surgical planning, enabling clinicians to anticipate challenging cases and optimize preoperative preparation.

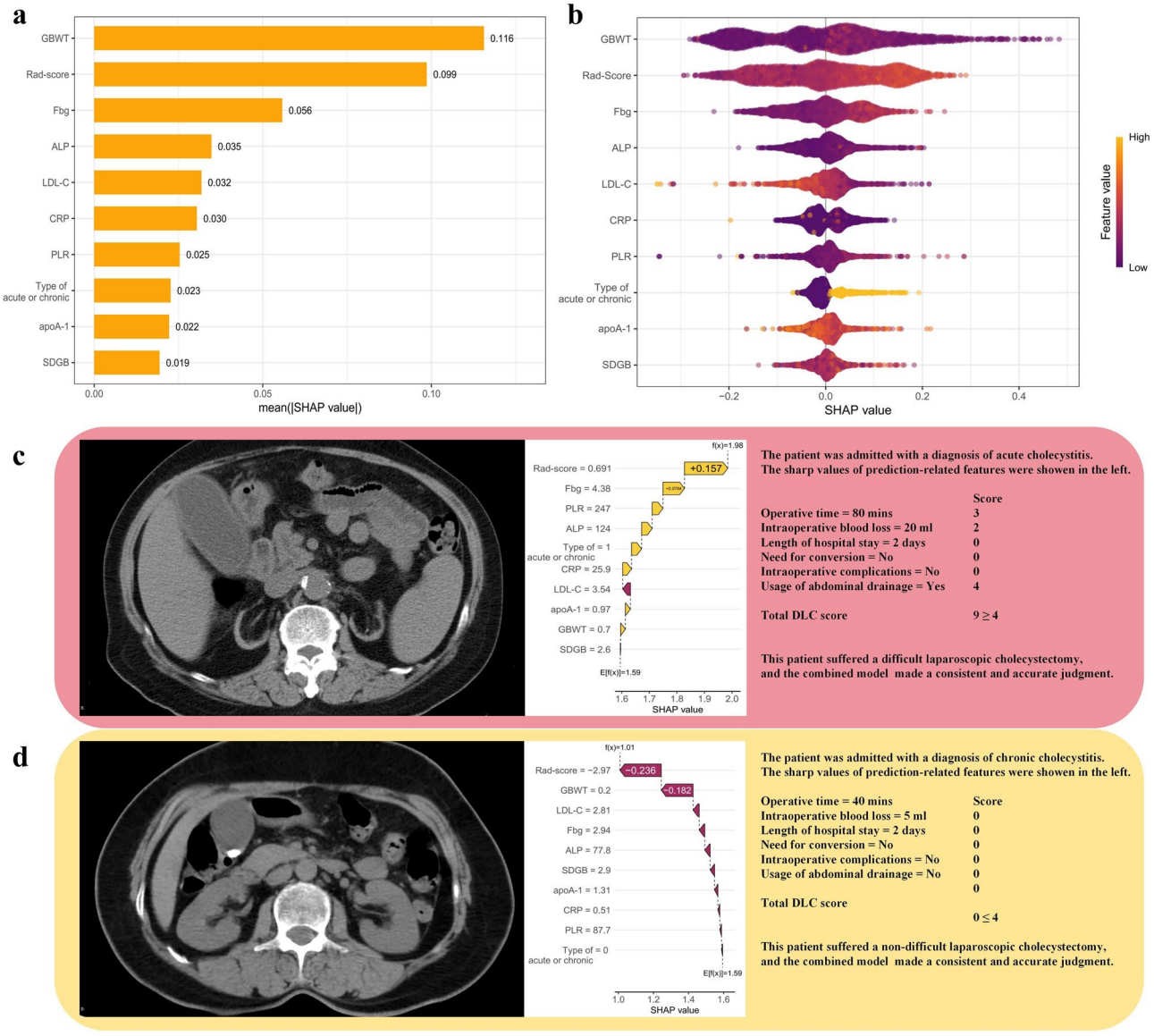

Fig. 7 The SHAP value of features maintained in the final combined RF model**(a)**, **(b)**. The original CT image, clinical data, and corresponding SHAP values for a representative DLC patient **(c)** and a representative NDLC patient **(d)**

Abbreviations

LC	Laparoscopic cholecystectomy
ERCP	Endoscopic retrograde cholangiopancreatography
LDGB	Long diameter of the gallbladder
SDGB	Short diameter of the gallbladder
GBWT	Gallbladder wall thickness
CBDD	Common bile duct diameter
DLC	Difficult laparoscopic cholecystectomy
NDLC	Non-difficult laparoscopic cholecystectomy
ROI	Region-of-interest
mRMR	Minimum redundancy maximum relevance
LASSO	The least absolute shrinkage and selection operator
SMOTE	The Synthetic Minority Oversampling Technique
RF	Random forest
AUC	Area under the receiver operating characteristic curve
DCA	Decision curve analysis
SHAP	SHapley Additive exPlanations
CRP	C-Reactive Protein
WBC	White Blood Cell Count
PLR	Platelet-to-lymphocyte ratio
NLR	Neutrophil-to-Lymphocyte Ratio
LDL-C	Lower Low-Density Lipoprotein Cholesterol
apoA-1	Apolipoprotein A-1
ALP	Alkaline Phosphatase
Fbg	Fibrinogen
DBIL	Direct bilirubin
ALT	Alanine aminotransferase
PT	Prothrombin time
TBIL	Total bilirubin
AST	Aspartate aminotransferase
HDL-C	High-density lipoprotein cholesterol
apoB	Apolipoprotein B
PT-INR	PT-international normalized ratio
PTA	Prothrombin activity

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13017-025-00635-1.

Supplementary Material 1

Acknowledgements
Not applicable.

Author contributions
RTS, CLL, CQ, JYC, YMJ and AYH designed the study. RTS, KL, KL, WYB, BT, WYH led the study data collection. RTS, CLL, CZZ conducted statistical analysis. XNY, JFC segmented the ROIs. RTS and CQ provided oversight on study data management. RTS, CLL, CQ, KL, KL, JYC lysed and interpreted all study data. RTS, CLL, CQ YMJ and AYH drafted the manuscript.All the authors reviewed the manuscript.

Funding
This study was supported by the project supported by the Taishan Scholar Foundation of Shandong Province, the Shandong Provincial Natural Science Foundation (ZR2024QH265), the Shandong Provincial Medical and Health Science and Technology Development Plan (202425020784), the Qingdao Natural Science Foundation(24-4-4-zrjj-103-jch).

Data availability
The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
This study involving humans was approved by the ethical review board of Qingdao University (No. QYFY WZLL 29872), and was conducted in accordance with the Declaration of Helsinki. The requirement for informed consent was waived by the institutional review board because of the anonymous nature of the data.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Hepatobiliary and Pancreatic Surgery, The Affiliated Hospital of Qingdao University, Qingdao 266003, Shandong, China
[2]Department of Radiology, The Affiliated Hospital of Qingdao University, Qingdao 266003, Shandong, China

Received: 29 April 2025 / Accepted: 6 July 2025
Published online: 14 July 2025

References

1. Gallaher JR, Charles A. Acute cholecystitis: A review. JAMA. 2022;327(10):965–75.
2. Okamoto K, Suzuki K, Takada T, Strasberg SM, Asbun HJ, Endo I, et al. Tokyo guidelines 2018: flowchart for the management of acute cholecystitis. J Hepatobiliary Pancreat Sci. 2018;25(1):55–72.
3. Coccolini F, Catena F, Pisano M, Gheza F, Fagiuoli S, Di Saverio S, et al. Open versus laparoscopic cholecystectomy in acute cholecystitis. Systematic review and meta-analysis. Int J Surg. 2015;18:196–204.
4. Ramírez-Giraldo C, Isaza-Restrepo A, Conde Monroy D, Castillo-Barbosa AC, Rubio-Avilez JJ, Van-Londoño I. What is the best score for predicting difficult laparoscopic cholecystectomy? A diagnostic trial study. Int J Surg. 2023;109(7):1871–9.
5. Wakabayashi G, Iwashita Y, Hibi T, Takada T, Strasberg SM, Asbun HJ, et al. Tokyo guidelines 2018: surgical management of acute cholecystitis: safe steps in laparoscopic cholecystectomy for acute cholecystitis (with videos). J Hepatobiliary Pancreat Sci. 2018;25(1):73–86.
6. Paul Wright G, Stilwell K, Johnson J, Hefty MT, Chung MH. Predicting length of stay and conversion to open cholecystectomy for acute cholecystitis using the 2013 Tokyo guidelines in a US population. J Hepatobiliary Pancreat Sci. 2015;22(11):795–801.
7. Tongyoo A, Chotiyasilp P, Sriussadaporn E, Limpavitayaporn P, Mingmalairak C. The pre-operative predictive model for difficult elective laparoscopic cholecystectomy: A modification. Asian J Surg. 2021;44(4):656–61.
8. Egawa N, Miyoshi A, Manabe T, Sadashima E, Koga H, Sato H, et al. Clinical evaluation of a surgical difficulty score for laparoscopic cholecystectomy for acute cholecystitis proposed in the Tokyo guidelines 2018. J Hepatobiliary Pancreat Sci. 2023;30(5):625–32.
9. Vannucci M, Laracca GG, Mercantini P, Perretta S, Padoy N, Dallemagne B, et al. Statistical models to preoperatively predict operative difficulty in laparoscopic cholecystectomy: A systematic review. Surgery. 2022;171(5):1158–67.
10. Hu ASY, Menon R, Gunnarsson R, de Costa A. Risk factors for conversion of laparoscopic cholecystectomy to open surgery - A systematic literature review of 30 studies. Am J Surg. 2017;214(5):920–30.
11. Manuel-Vázquez A, Latorre-Fragua R, Alcázar C, Requena PM, de la Plaza R, Blanco Fernández G, et al. Reaching a consensus on the definition of difficult cholecystectomy among Spanish experts. A Delphi project. A qualitative study. Int J Surg. 2022;102:106649.
12. Adachi T, Eguchi S, Muto Y. Pathophysiology and pathology of acute cholecystitis: A secondary publication of the Japanese version from 1992. J Hepatobiliary Pancreat Sci. 2022;29(2):212–6.
13. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the Bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol. 2017;14(12):749–62.
14. Guiot J, Vaidyanathan A, Deprez L, Zerka F, Danthine D, Frix AN, et al. A review in radiomics: making personalized medicine a reality via routine imaging. Med Res Rev. 2022;42(1):426–40.
15. Seshadri A, Peitzman AB. The difficult cholecystectomy: what you need to know. J Trauma Acute Care Surg. 2024;97(3):325–36.
16. Yi X, Walia E, Babyn P. Generative adversarial network in medical imaging: A review. Med Image Anal. 2019;58:101552.
17. Qiu D, Cheng Y, Wang X. Medical image super-resolution reconstruction algorithms based on deep learning: A survey. Comput Methods Programs Biomed. 2023;238:107590.
18. Xing X, Li L, Sun M, Yang J, Zhu X, Peng F, et al. Deep-learning-based 3D super-resolution CT radiomics model: predict the possibility of the micropapillary/solid component of lung adenocarcinoma. Heliyon. 2024;10(13):e34163.
19. Yan F, Chen X, Quan X, Wang L, Wei X, Zhu J. Association between the stress hyperglycemia ratio and 28-day all-cause mortality in critically ill patients with sepsis: a retrospective cohort study and predictive model establishment based on machine learning. Cardiovasc Diabetol. 2024;23(1):163.
20. Ahanger AB, Aalam SW, Masoodi TA, Shah A, Khan MA, Bhat AA, et al. Radiogenomics and machine learning predict oncogenic signaling pathways in glioblastoma. J Transl Med. 2025;23(1):121.
21. Bernard D, Doumard E, Ader I, Kemoun P, Pagès JC, Galinier A, et al. Explainable machine learning framework to predict personalized physiological aging. Aging Cell. 2023;22(8):e13872.
22. Zuo D, Yang L, Jin Y, Qi H, Liu Y, Ren L. Machine learning-based models for the prediction of breast cancer recurrence risk. BMC Med Inf Decis Mak. 2023;23(1):276.
23. Ma Y, Luo M, Guan G, Liu X, Cui X, Luo F. An explainable predictive machine learning model of gangrenous cholecystitis based on clinical data: a retrospective single center study. World J Emerg Surg. 2025;20(1):1.
24. Bourgouin S, Mancini J, Monchal T, Calvary R, Bordes J, Balandraud P. How to predict difficult laparoscopic cholecystectomy? Proposal for a simple preoperative scoring system. Am J Surg. 2016;212(5):873–81.
25. Nassar AHM, Ashkar KA, Mohamed AY, Hafiz AA. Is laparoscopic cholecystectomy possible without video technology? Minim Invasive Therapy. 1995;4(2):63–5.
26. Törnqvist B, Waage A, Zheng Z, Ye W, Nilsson M. Severity of acute cholecystitis and risk of iatrogenic bile duct injury during cholecystectomy, a Population-Based Case-Control study. World J Surg. 2016;40(5):1060–7.
27. Bugiantella W, Vedovati MC, Becattini C, Canger RC, Avenia N, Rondelli F. To drain or not to drain elective uncomplicated laparoscopic cholecystectomy? A systematic review and meta-analysis. J Hepatobiliary Pancreat Sci. 2014;21(11):787–94.

28. Micić D, Stanković S, Lalić N, Đukić V, Polovina S. Prognostic value of preoperative Neutrophil-to-lymphocyte ratio for prediction of severe cholecystitis. J Med Biochem. 2018;37(2):121–7.
29. Molasy B, Frydrych M, Kubala-Kukuś A, Głuszek S. The use of Neutrophil-to-Lymphocyte ratio, Monocyte-to-Lymphocyte ratio and Platelets-to-Lymphocyte ratio in the assessment of the risk of conversion and complications after cholecystectomy performed due to symptomatic cholelithiasis. Ther Clin Risk Manag. 2024;20:363–71.
30. Shimoda M, Kuboyama Y, Suzuki S. Laparoscopic bailout surgery effective procedure for patients with difficult laparoscopic cholecystectomy. Updates Surg. 2022;74(5):1611–6.
31. Wolberg AS. Fibrinogen and fibrin: synthesis, structure, and function in health and disease. J Thromb Haemost. 2023;21(11):3005–15.
32. Han CY, Pichon TJ, Wang X, Ringgold KM, St John AE, Stern SA, et al. Leukocyte activation primes fibrinogen for proteolysis by mitochondrial oxidative stress. Redox Biol. 2022;51:102263.
33. Giles JT, Wasko MCM, Chung CP, Szklo M, Blumenthal RS, Kao A, et al. Exploring the lipid paradox theory in rheumatoid arthritis: associations of low Circulating low-Density lipoprotein concentration with subclinical coronary atherosclerosis. Arthritis Rheumatol. 2019;71(9):1426–36.
34. Zeng G, Zhang C, Song Y, Zhang Z, Xu J, Liu Z, et al. The potential impact of inflammation on the lipid paradox in patients with acute myocardial infarction: a multicenter study. BMC Med. 2024;22(1):599.
35. Chen ZM, Gu HQ, Mo JL, Yang KX, Jiang YY, Yang X, et al. U-shaped association between low-density lipoprotein cholesterol levels and risk of all-cause mortality mediated by post-stroke infection in acute ischemic stroke. Sci Bull (Beijing). 2023;68(12):1327–35.
36. Robertson J, Peters MJ, McInnes IB, Sattar N. Changes in lipid levels with inflammation and therapy in RA: a maturing paradigm. Nat Rev Rheumatol. 2013;9(9):513–23.
37. Obermayer G, Afonyushkin T, Binder CJ. Oxidized low-density lipoprotein in inflammation-driven thrombosis. J Thromb Haemost. 2018;16(3):418–28.

Publisher's note