Research Square

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

Machine learning radiomics models for preoperative prediction of postoperative early recurrence in gallbladder cancer

Qi Li
The First Affiliated Hospital of Xi'an Jiaotong University
Zhenqi Tang
The First Affiliated Hospital of Xi'an Jiaotong University
Min Yang
The First Affiliated Hospital of Xi'an Jiaotong University
Hengchao Liu
The First Affiliated Hospital of Xi'an Jiaotong University
Mai Xu
The First Affiliated Hospital of Xi'an Jiaotong University
Chen Chen
The First Affiliated Hospital of Xi'an Jiaotong University
Dong Zhang
The First Affiliated Hospital of Xi'an Jiaotong University
Jian Yang
The First Affiliated Hospital of Xi'an Jiaotong University
Zhimin Geng
gengzhimin@mail.xjtu.edu.cn

The First Affiliated Hospital of Xi'an Jiaotong University

Research Article

Keywords: Gallbladder cancer, Early recurrence, Contrast-enhanced computed tomography, Radiomics, Predictive model

Posted Date: June 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9593876/v1

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

Additional Declarations: No competing interests reported.

Abstract

Objective

Gallbladder cancer (GBC) is a high aggressive malignancy prone to postoperative early recurrence (ER). However, reliable tools for preoperative risk stratification remain limited. This study aimed to develop and validate a noninvasive predictive model integrating clinical variables and contrast-enhanced computed tomography-based radiomic features for ER prediction in GBC.

Methods

A retrospective cohort of 204 patients with pathologically confirmed GBC was included and randomly divided into a training cohort (n = 142) and a validation cohort (n = 62). Feature selection was performed using least absolute shrinkage and selection operator regression. Six machine learning algorithms were applied to construct clinical, radiomic, and combined models. Model performance was evaluated using the area under the receiver operating characteristic curve, and interpretability was assessed using Shapley additive explanations (SHAP).

Results

Among 204 patients, 67 (32.8%) experienced ER. Multivariate analysis identified CEA > 5.0 ng/mL, albumin-to-γ-glutamyl transpeptidase ratio, and imaging lymph node metastasis as independent predictors of ER (all $P < 0.05$). A total of 6 clinical features and 11 radiomic features (comprising first-order statistics, texture, and higher-order wavelet features) were selected. In the validation cohort, the clinical SVM model (AUC = 0.830) and radiomic XGB model (AUC = 0.786) achieved the best individual performances. XGB-based fusion models outperformed unimodal models, yielding AUCs of 0.865 for feature-level fusion and 0.879 for decision-level fusion.

Conclusion

The proposed multidimensional model demonstrates good performance in preoperatively identifying patients at high risk of early recurrence, which may assist individualized treatment planning and postoperative surveillance strategies.

Introduction

Gallbladder cancer (GBC), the most common and aggressive malignancy of the biliary tract, is associated with a dismal prognosis, evidenced by a global five-year overall survival (OS) rate of less than 10% 1–2. Owing to its vague clinical manifestations and rapid progression, most patients present with advanced-stage tumors, frequently complicated by hepatic invasion or regional lymph node metastasis.

Although curative-intent resection remains the only potentially effective treatment, postoperative early recurrence (ER) poses a significant challenge, occurring in over half of patients and substantially compromising long-term survival 3–4.

Conventional clinicopathological parameters alone are inadequate for accurately assessing the risk of recurrence in GBC, highlighting the need for more advanced imaging-based risk stratification tools. Recent advances in artificial intelligence and radiomics have transformed medical imaging from qualitative interpretation to quantitative analysis, enabling the extraction of high-throughput features that capture tumor heterogeneity and underlying biological behavior. Radiomics has shown considerable promise across multiple cancer types for preoperative prediction of recurrence, evaluation of metastatic potential, and estimation of survival, demonstrating strong translational value 5–10.

Given these developments and the pressing clinical need, this study aimed to construct and validate a multidimensional predictive model that integrates preoperative clinical and radiomic features to enable precise and personalized stratification of postoperative ER risk in GBC.

Materials and Methods

Study population

This retrospective study included patients with pathologically confirmed GBC who underwent curative-intent resection at a high-volume medical center between January 2013 and July 2024. Patients were randomly allocated into a training cohort (n = 142) and a validation cohort (n = 62) using a 7:3 stratified random split (StratifiedShuffleSplit), ensuring balanced distributions of clinical and imaging characteristics across cohorts (**Supplementary Table 1**). The flowchart of patient selection is presented in Fig. 1.

Collection and Selection of Clinical Features

Data Collection

Preoperative demographic, laboratory, and imaging data were retrieved from electronic medical records. Demographic information included age and sex. Laboratory parameters comprised tumor markers (CEA, CA19-9, and CA-125), total bilirubin, and various inflammation- and nutrition-related indices including the neutrophil-to-lymphocyte ratio (NLR), monocyte-to-lymphocyte ratio (MLR), platelet-to-lymphocyte ratio (PLR), albumin-to-γ-glutamyl transpeptidase ratio (AGR), and the prognostic nutritional index (PNI). Imaging semantic features assessed included radiologic T-stage and N-stage, hepatic invasion, vascular invasion, perivascular soft tissue infiltration around the celiac artery, tumor location, morphology, presence of gallstones, enhancement phase, and enhancement pattern.

Feature Selection

Clinical feature selection was conducted within the training cohort. We applied least absolute shrinkage and selection operator (LASSO) regression with five-fold cross-validation to identify the optimal regularization parameter (λ). Candidate features were ranked according to the absolute values of their regression coefficients and were sequentially incorporated into the model. The feature subset that achieved the highest mean cross-validated area under the receiver operating characteristic curve (AUC) was retained as the optimal set of clinical features.

Extraction and Selection of Radiomic Features

Image Acquisition and Segmentation

All patients underwent multidetector CT (MDCT) within one month preoperatively, covering the liver and both kidneys. Contrast-enhanced CT included arterial (25 ~ 30 s) and portal venous (65 ~ 70 s) phases. Tumor volumes of interest (VOIs) were manually delineated by two radiologists and further refined by expanding 2 mm outward and contracting 1 mm inward using OpenCV in Python.

Preprocessing and Feature Extraction

Images and VOIs were resampled to 1 × 1 × 1 mm isotropic resolution using SimpleITK. Filters including square root, natural logarithm, local binary pattern (LBP), gradient, three-dimensional stationary wavelet transform (SWT), and Laplacian of Gaussian (LoG) were applied. Radiomic features were extracted with PyRadiomics (v3.0.1) and categorized as: first-order statistics, 3D shape features, texture features (GLCM, GLSZM, GLRLM, GLDM, NGTDM), and higher-order wavelet features.

Feature Dimensionality Reduction and Selection

In the training cohort, redundant and uninformative features were excluded if they exhibited excessive missing values, were constant, or had duplication rates exceeding 50%. Sequential feature selection was then performed using Pearson correlation (threshold > 0.75), univariate t-tests ($P < 0.05$), and LASSO regression with five-fold cross-validation. The feature set that yielded the highest mean cross-validated AUC was selected for final model development.

Model Development and Validation

To establish a robust predictive framework, a variety of supervised machine learning approaches—including Logistic Regression (LR), Support Vector Machine (SVM), Naïve Bayes (NB), Random Forest (RF), Light Gradient Boosting Machine (LGB), and eXtreme Gradient Boosting (XGB)—were implemented. Hyperparameter tuning was performed using randomized search with five-fold cross-validation. Separate

models were developed based on clinical and radiomic features, respectively. Feature-level fusion models were constructed by integrating selected clinical and radiomic variables, whereas decision-level fusion combined predicted probabilities from the base classifiers using logistic regression as a meta-learner. Model performance was assessed using AUC, accuracy, F1-score, sensitivity, and specificity.

Model Interpretability Analysis

Model interpretability was assessed using SHapley Additive exPlanations (SHAP). Global feature importance was derived from mean absolute SHAP values, and dependence plots were generated to visualize nonlinear relationships and feature interactions.

Statistical analysis

Statistical analyses were performed using SPSS software version 28.0 (IBM Corp., Armonk, NY, USA). Categorical variables were expressed as counts and percentages, continuous variables as mean ± SD or median (*IQR*). Comparisons used chi-square or Mann–Whitney U tests. Multivariable logistic regression identified independent predictors of ER. A two-sided $P < 0.05$ was considered statistically significant.

Results

Baseline Characteristics in the Whole Cohort

A total of 204 patients who underwent curative-intent resection for GBC were screened. Of these, 67 (32.8%) experienced ER and the median OS was 60.0 months. In the ER group, more than 90% of patients presented with imaging lymph node metastasis, and most tumors exhibited a wall-thickening morphology. In contrast, among patients without ER, 65% had imaging lymph node metastasis, with mass-forming lesions being more common. In both groups, the majority of patients showed no perivascular soft tissue around the celiac artery (89.6% *vs.* 93.4%, $P = 0.333$). Baseline clinicopathologic characteristics are summarized in **Supplementary Table 2**.

Analysis of Risk Factors for ER

Univariate analysis identified total bilirubin ($P < 0.001$), CEA > 5.0 ng/mL ($P < 0.001$), CA19-9 ($P < 0.001$), CA-125 ($P = 0.002$), PLR ($P = 0.003$), MLR ($P = 0.004$), PNI ($P = 0.038$), AGR ($P < 0.001$), imaging lymph node metastasis ($P < 0.001$), imaging liver invasion ($P < 0.001$), tumor location ($P = 0.001$), morphology ($P < 0.001$), and enhancement pattern ($P = 0.028$) as significantly associated with ER. Multivariable logistic regression analysis demonstrated that CEA > 5.0 ng/mL ($P = 0.013$, $OR = 2.833$, 95% CI: 1.247 ~ 6.433), AGR ($P = 0.038$, $OR = 0.656$, 95% CI: 0.441 ~ 0.977), and imaging lymph node metastasis ($P = 0.010$, $OR = 4.949$, 95% CI: 1.455 ~ 16.832) were independent predictors of ER (Table 1).

Table 1
Univariate and multivariate analysis of preoperative factors associated with early recurrence for gallbladder cancer after curative-intent resection in the training cohort

Variables	**Univariate analysis**		**Multivariate analysis**	
	***OR* (95%*CI*)**	***P***	***OR* (95%*CI*)**	***P***
Sex				
Male *vs.* Female	0.820(0.443 ~ 1.518)	0.527		
Age (year)				
>65 *vs.*≤65	0.562(0.307 ~ 1.030)	0.062		
Total bilirubin (μmol/L)				
>34.2 *vs.*≤34.2	6.606(3.060 ~ 14.261)	< 0.001	2.470(0.803 ~ 7.599)	0.115
CEA (ng/mL)				
>5.0 *vs.*≤5.0	4.521(2.396 ~ 8.530)	< 0.001	2.886(1.272 ~ 6.547)	0.011
CA19-9 (U/mL)				
>39.0 vs ≤ 39.0	4.639(2.484 ~ 8.662)	< 0.001	2.099(0.893 ~ 4.932)	0.089
CA-125 (U/mL)				
>35.0 *vs.* ≤35.0	3.542(1.578 ~ 7.949)	0.002	1.453(0.512 ~ 4.117)	0.482
NLR	1.063(0.994 ~ 1.136)	0.074		
PLR	1.006(1.002 ~ 1.009)	0.002	1.002(0.997 ~ 1.006)	0.421
MLR	10.010(2.101 ~ 47.698)	0.004	1.177(0.138 ~ 10.025)	0.882
PNI	0.943(0.892 ~ 0.997)	0.038	1.007(0.931 ~ 1.088)	0.870
AGR	0.492(0.361 ~ 0.671)	< 0.001	0.660(0.443 ~ 0.982)	0.041

*NLR: neutrophil-to-lymphocyte ratio; PLR: platelet-to-lymphocyte ratio; MLR: monocyte-to-lymphocyte ratio PNI: prognostic nutritional index; AGR: albumin-to-γ-glutamyl transpeptidase ratio.

"–" indicates categories that are not comparable due to insufficient sample size.

Variables	Univariate analysis		Multivariate analysis	
	***OR* (95%*CI*)**	***P***	***OR* (95%*CI*)**	***P***
Imaging T stage				
T3 *vs.* ≤T1-T2	1.885(0.654 ~ 5.431)	0.240		
Imaging lymph node metastasis				
Yes *vs.* No	6.688(2.519 ~ 17.756)	< 0.001	4.862(1.433 ~ 16.497)	0.011
Imaging liver invasion				
Yes *vs.* No	4.192(2.086 ~ 8.425)	< 0.001	1.999(0.758 ~ 5.274)	0.162
Imaging vascular invasion				
Yes *vs.* No	-	-		
Perivascular soft tissue around the celiac artery				
Yes *vs.* No	1.659(0.590 ~ 4.667)	0.337		
*NLR: neutrophil-to-lymphocyte ratio; PLR: platelet-to-lymphocyte ratio; MLR: monocyte-to-lymphocyte ratio PNI: prognostic nutritional index; AGR: albumin-to-γ-glutamyl transpeptidase ratio.				
"–" indicates categories that are not comparable due to insufficient sample size.				

Unimodal Prediction Model Development

In the clinical model, six features were selected via LASSO regression for model development: imaging lymph node metastasis (|β| = 0.9677), imaging liver invasion (|β| = 0.9116), CEA (|β| = 0.7574), total bilirubin (|β| = 0.5222), enhancement pattern (|β| = 0.4461), and NLR (|β| = 0.4103). In the validation cohort, the SVM model achieved superior performance with an AUC of 0.837 (95% *CI*: 0.711 ~ 0.944; Fig. 2A **and B**).

The radiomic model integrated a set of 11 selected features, comprising first-order statistics, texture, and higher-order wavelet features (**Supplementary Table 3**). In the validation cohort, the XGB model demonstrated superior performance with an AUC of 0.786 (95% *CI*: 0.626 ~ 0.912; Fig. 2C **and D**). The predictive efficacy of both the clinical and radiomic models across the training and validation cohorts is summarized in Table 2.

Table 2
Comparison of the performance between clinical and radiomic models based on different machine learning algorithms in the training and validation cohorts

	Algorithms	Clinical models [AUC (95%CI)]	Radiomic models [AUC (95%CI)]
Training cohort	LGB	0.853(0.791 ~ 0.913)	0.888(0.832 ~ 0.939)
	LR	0.860(0.796 ~ 0.916)	0.849(0.778 ~ 0.915)
	NB	0.827(0.752 ~ 0.895)	0.833(0.755 ~ 0.900)
	RF	0.867(0.804 ~ 0.920)	0.907(0.846 ~ 0.959)
	SVM	0.847(0.783 ~ 0.905)	0.741(0.648 ~ 0.827)
	XGB	0.871(0.811 ~ 0.921)	0.842(0.773 ~ 0.909)
Validation cohort	LGB	0.836(0.692 ~ 0.949)	0.739(0.603 ~ 0.859)
	LR	0.788(0.649 ~ 0.904)	0.669(0.501 ~ 0.828)
	NB	0.786(0.656 ~ 0.898)	0.673(0.505 ~ 0.811)
	RF	0.836(0.698 ~ 0.955)	0.737(0.591 ~ 0.868)
	SVM	**0.837(0.711 ~ 0.944)**	0.742(0.596 ~ 0.871)
	XGB	0.834(0.701 ~ 0.944)	**0.786(0.626 ~ 0.912)**

*LGB, Light Gradient Boosting machine; LR, Logistic Regression; NB, Naive Bayes; RF, Random Forest; SVM, Support Vector Machine; XGB, Extreme Gradient Boosting.

Fusion Prediction Model Development

Given the superior performance of the XGB classifier in the radiomic model, it was selected for fusion modeling. In the validation cohort, the feature-level fusion model achieved an AUC of 0.865 (95% *CI*: 0.736 ~ 0.965), while the decision-level fusion model exhibited an AUC of 0.879 (95% *CI*: 0.778 ~ 0.958), both outperforming unimodal prediction models (Table 3).

Table 3
Comparison of the performance among four models based on eXtreme Gradient Boosting algorithm in the validation cohort

Model	AUC(95% CI)	Accuracy	F1-score	Sensitivity	Specificity
Clinical model	0.834(0.701 ~ 0.944)	0.790	0.683	0.700	0.833
Radiomic model	0.786(0.626 ~ 0.912)	0.710	0.625	0.720	0.690
Feature-level fusion model	0.865(0.736 ~ 0.965)	0.855	0.743	0.650	0.952
Decision-level fusion model	0.879(0.778 ~ 0.958)	0.822	0.667	0.550	0.952

Interpretability Analysis of XGB-Based Prediction Models

The XGB-based feature-level fusion model demonstrated the best overall performance within its category, achieving an AUC of 0.865 (Accuracy = 0.855; F1-score = 0.743) in the validation cohort. SHAP analysis revealed the relative contribution of each feature, identifying CEA, imaging lymph node metastasis, and NLR as the top three clinical predictors. The radiomic features, LBP Boundary Run Variance and Wavelet L1 HL Skewness, were also among the top five most influential predictors (Fig. 3A). SHAP dependence plots further illustrated the direction of feature effects: higher levels of CEA, NLR, and LBP Boundary Run Variance were all positively associated with an increased predicted probability of ER, whereas lower levels of Wavelet L1 HL and Boundary L3 HL Correlation exhibited a protective effect by reducing the predicted ER risk (Fig. 3B).

To further illustrate the individualized interpretability of the model, SHAP force plots were generated for two representative patients. For the non-ER patient, the prediction was pushed towards a negative outcome (low ER risk), primarily driven by CEA acting as the most significant negative predictor. However, the radiomic feature Wavelet L1 HL Skewness exerted the strongest positive influence, attempting to raise the predicted risk (Fig. 4A). Conversely, for the ER patient, the prediction was pushed towards the positive outcome (high ER risk). This high risk was strongly propelled by the radiomic feature LBP Boundary Run Variance, which served as the most influential positive predictor. Similar to the first patient, CEA was the most significant negative predictor, attempting to lower the predicted risk (Fig. 4B).

Discussion

Postoperative recurrence is the principal challenge following curative-intent resection for GBC, undermining long-term survival and occurring in up to half of patients 1, 3–4. ER, generally defined as recurrence within 6 to 12 months after surgery, reflects aggressive tumor biology and is strongly correlated with poor prognosis. Previous studies have shown that regardless of whether recurrence is

locoregional or distant, its timing substantially influences post-recurrence survival 3, 11. Patients with ER experience significantly shorter OS compared to those with late or no recurrence 3, 12. In the present study, 32.8% of patients experienced recurrence with a median OS of 60.0 months, which was markedly lower than that observed in patients without ER. These findings highlight the critical need for accurate preoperative identification of high-risk patients to guide personalized treatment strategies

To explore clinical variables associated with ER, we performed univariate and multivariate analyses and identified CEA > 5.0 ng/mL, AGR, and radiologically detected lymph node metastasis as independent risk factors for ER, consistent with previous studies 1, 3, 13–15. These factors reflect both tumor aggressiveness and host nutritional-inflammatory status, underscoring the multifactorial basis of recurrence. Previous efforts have explored preoperative clinical risk models, which demonstrated moderate predictive value and provided guidance for surgical decision-making in high-risk patients 12, 16. Leveraging a comprehensive set of clinical parameters, we applied LASSO regression to identify an optimal feature subset for predictive modeling. This process yielded six features including imaging lymph node metastasis, liver invasion, CEA, total bilirubin, enhancement pattern, and NLR, which were then incorporated into the machine learning algorithms. Among six algorithms, the SVM-based model achieved the best predictive performance (AUC = 0.837, 95% *CI*: 0.711 ~ 0.944), yet clinical variables alone may incompletely capture tumor heterogeneity.

Radiomics, an emerging noninvasive quantitative imaging analysis technique, facilitates the extraction of high-throughput features from routine medical images, thereby capturing the underlying microscopic heterogeneity of tumors. It has shown promising potential in predicting recurrence across multiple malignancies, such as pancreatic cancer, hepatocellular carcinoma, and colorectal cancer 5–7. In this study, the radiomic model derived from contrast-enhanced CT achieved robust performance (AUC = 0.786, 95% *CI*: 0.626 ~ 0.912), although its predictive power was slightly inferior to that of the clinical model. This difference may be attributed to the fact that radiomic features predominantly capture structural and textural characteristics, with limited reflection of systemic factors such as hematologic or nutritional status. Moreover, the high-dimensional nature of radiomic features carries an inherent risk of overfitting and susceptibility to noise, which may compromise model generalizability.

To effectively leverage the complementary strengths of the two data modalities, we established both feature-level and decision-level fusion models using the XGB algorithm. Both fusion approaches outperformed single-modality models in the validation cohort, achieving AUCs of 0.865 (95% *CI*: 0.736 ~ 0.965) and 0.879 (95% *CI*: 0.778 ~ 0.958), respectively. These findings indicated that combining clinical and radiomic features provides a more comprehensive and accurate assessment of ER risk in GBC, supporting their joint use in preoperative evaluation.

Despite the promising predictive performance of the fusion models, the inherent “black-box” nature of complex machine learning algorithms often limits their transparency and clinical applicability 17–18. Therefore, we applied SHAP analysis to quantify the contribution of each feature and visualize their

nonlinear effects on ER risk. This interpretability framework not only reinforces the model's clinical credibility but also facilitates physician understanding and patient-specific decision-making.

Several limitations should be acknowledged. First, the inherent weaknesses of this single-center, retrospective study design, coupled with a modest sample size, may introduce selection bias and critically limit external validity. Second, the lack of available genomic and molecular data constrained the exploration of the underlying biological mechanisms that drive ER. Third, radiomic features can be sensitive to imaging acquisition and reconstruction parameters; thus, multicenter external validation is necessary. Future research should prioritize prospective, multicenter designs that integrate clinical, radiomic, and molecular data to develop robust, generalizable predictive models and ultimately optimize postoperative management strategies in GBC.

Conclusion

This study developed and internally validated a multidimensional, noninvasive predictive model that integrates preoperative clinical and radiomic features to estimate individual risk of ER in GBC. This model demonstrates strong potential to support preoperative identification of high-risk patients, enabling optimized perioperative management and tailored therapeutic strategies. Prospective, multicenter external validation is warranted to confirm its generalizability and facilitate clinical implementation.

Abbreviations

GBC
gallbladder cancer
ER
early recurrence
SHAP
shapley additive explanations
OS
overall survival
NLR
neutrophil-to-lymphocyte ratio
MLR
monocyte-to-lymphocyte ratio
PLR
platelet-to-lymphocyte ratio
AGR
albumin-to-γ-glutamyl transpeptidase ratio
PNI
prognostic nutritional index
LASSO

least absolute shrinkage and selection operator
AUC
area under the receiver operating characteristic curve
VOIs
volumes of interest
SWT
stationary wavelet transform
LoG
Laplacian of Gaussian
LR
Logistic Regression
SVM
Support Vector Machine
NB
Naïve Bayes
RF
Random Forest
LGB
Light Gradient Boosting Machine
XGB
eXtreme Gradient Boosting

Declarations

Ethics approval and consent to participate

The study was approved by the ethics committee of the First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China (No: XJTU1AF2018LSK-095). Written informed consent was obtained from all individual participants and their families prior to their enrollment in the study. The study was conducted in accordance with the Declaration of Helsinki.

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests

Funding

This work was supported by the National Natural Science Foundation of China (No.62576271), Key Research and Development Program of Shaanxi Province (No. 2025SF-YBXM-386 and 2022-SF-606), and Institutional Foundation of The First Affiliated Hospital of Xi'an Jiaotong University (No.2024-QN-15).

Authors' contributions

Qi Li: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review and editing. Zhenqi Tang: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing. Min Yang: Data curation, Investigation, Validation, Visualization. Hengchao Liu: Methodology, Software, Validation, Visualization. Mai Xu: Methodology, Software, Validation, Visualization. Chen Chen: Methodology, Project administration, Supervision, Writing – review and editing. Dong Zhang: Methodology, Project administration, Supervision, Writing – review and editing. Jian Yang: Methodology, Project administration, Supervision, Writing – review and editing. Zhimin Geng: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Writing – review and editing.

Acknowledgements

We thank all clinicians and research staff for their assistance with data collection and image analysis. We are also grateful to the patients for their participation in this study.

References

1. Yang SQ, Feng H, Tian Y, et al. Unraveling early recurrence risk factors in gallbladder cancer: a systematic review and meta-analysis. Eur J Surg Oncol. 2024;50(7):108372.
2. Goetze TO. Gallbladder carcinoma: prognostic factors and therapeutic options. World J Gastroenterol. 2015;21(43):12211–7.
3. Margonis GA, Gani F, Buettner S, et al. Rates and patterns of recurrence after curative intent resection for gallbladder cancer: a multi-institution analysis from the US Extra-hepatic Biliary Malignancy Consortium. HPB (Oxford). 2016;18(11):872–8.
4. Kim WS, Choi DW, You DD, et al. Risk factors influencing recurrence, patterns of recurrence, and the efficacy of adjuvant therapy after radical resection for gallbladder carcinoma. J Gastrointest Surg. 2010;14(4):679–87.
5. Zhang Y, Lobo-Mueller EM, Karanicolas P, et al. Improving prognostic performance in resectable pancreatic ductal adenocarcinoma using radiomics and deep learning feature fusion in CT images. Sci Rep. 2021;11(1):1378.
6. Zheng BH, Liu LZ, Zhang ZZ, et al. Radiomics score: a potential prognostic imaging feature for postoperative survival of solitary hepatocellular carcinoma patients. BMC Cancer. 2018;18(1):1148.

7. Creasy JM, Cunanan KM, Chakraborty J, et al. Differences in liver parenchyma measurable with CT radiomics at initial colon resection in patients developing hepatic metastases from stage II/III colon cancer. Ann Surg Oncol. 2021;28(4):1982–9.
8. Gu J, Tong T, Xu D, et al. Deep learning radiomics of ultrasonography for predicting tumor and axillary lymph node status after neoadjuvant chemotherapy in breast cancer: a multicenter study. Cancer. 2023;129(3):356–66.
9. Huang Y, Liu Z, He L, et al. Radiomics signature: a biomarker for prediction of disease-free survival in early-stage non-small cell lung cancer. Radiology. 2016;281(3):947–57.
10. Shiradkar R, Ghose S, Jambor I, et al. Radiomic features from pretreatment biparametric MRI predict prostate cancer biochemical recurrence: preliminary findings. J Magn Reson Imaging. 2018;48(6):1626–36.
11. Jarnagin WR, Ruo L, Little SA, et al. Patterns of initial disease recurrence after resection of gallbladder carcinoma and hilar cholangiocarcinoma. Cancer. 2003;98(8):1689–700.
12. Shimizu Y, Ashida R, Sugiura T, et al. Early recurrence in resected gallbladder carcinoma: clinical impact and preoperative predictive score. Ann Surg Oncol. 2022;29(9):5447–57.
13. Kawashima J, Sahara K, Shen F, et al. Predicting recurrence after resection of stage I intrahepatic cholangiocarcinoma. J Gastrointest Surg. 2024;28(1):18–25.
14. Chen M, Li S, Topatana W, et al. Development and validation of a nomogram for predicting survival in gallbladder cancer patients with recurrence after surgery. Front Oncol. 2020;10:537789.
15. Park JS, Yoon DS, Kim KS, et al. Recurrence patterns and risk factors after curative resection for stage II gallbladder carcinoma. J Gastrointest Surg. 2007;11(5):631–7.
16. Sahara K, Tsilimigras DI, Kikuchi Y, et al. Defining and predicting early recurrence after resection for gallbladder cancer. Ann Surg Oncol. 2021;28(1):417–25.
17. Wang Y, Hu Z, Wang H. Clinical implications and interpretability of radiomics in brain tumors. Insights Imaging. 2025;16(1):77.
18. Najjar R. Redefining radiology: a review of artificial intelligence integration in medical imaging. Diagnostics (Basel). 2023;13(17):2760.

Figures

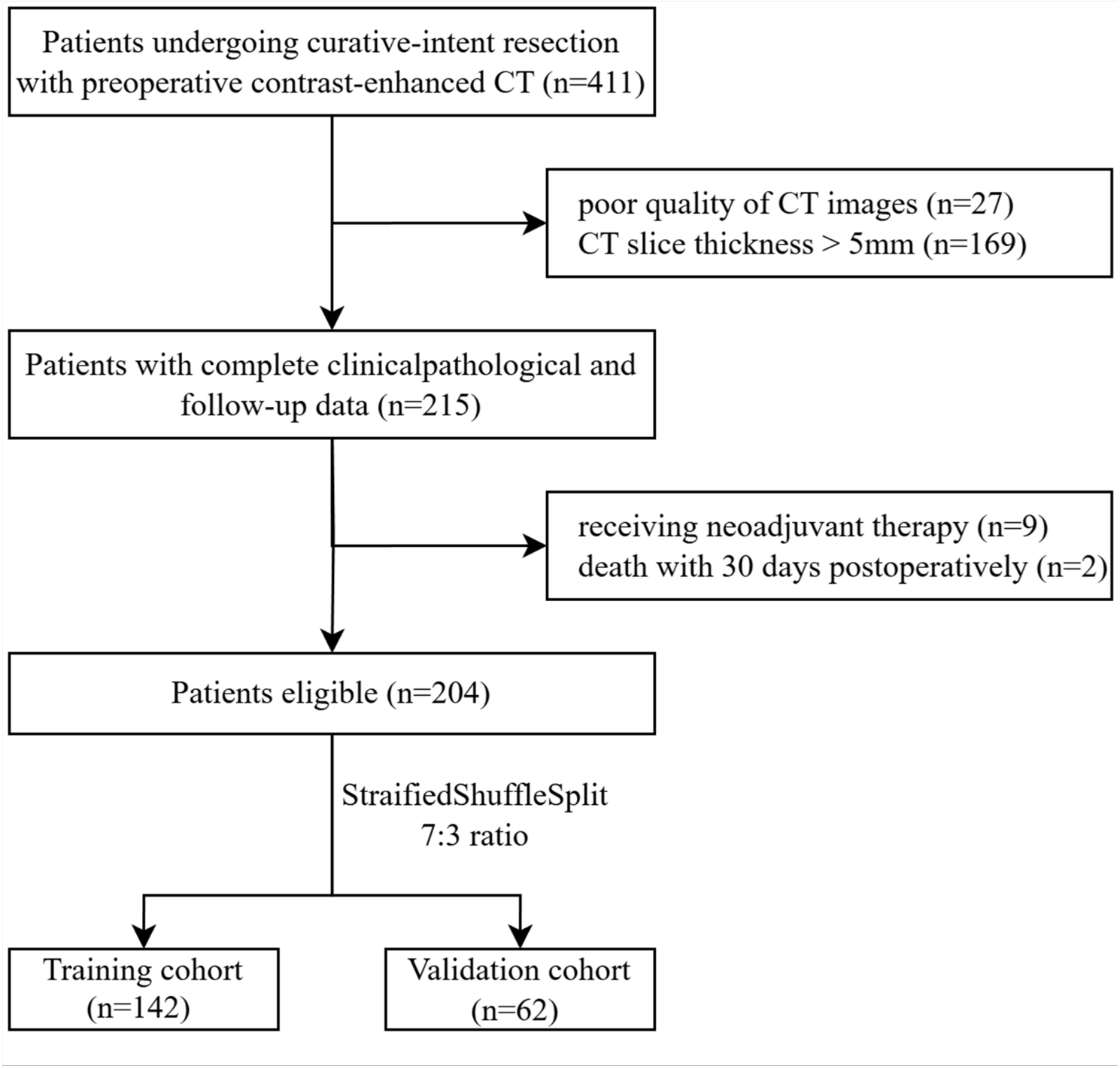


Figure 1

Study Flowchart

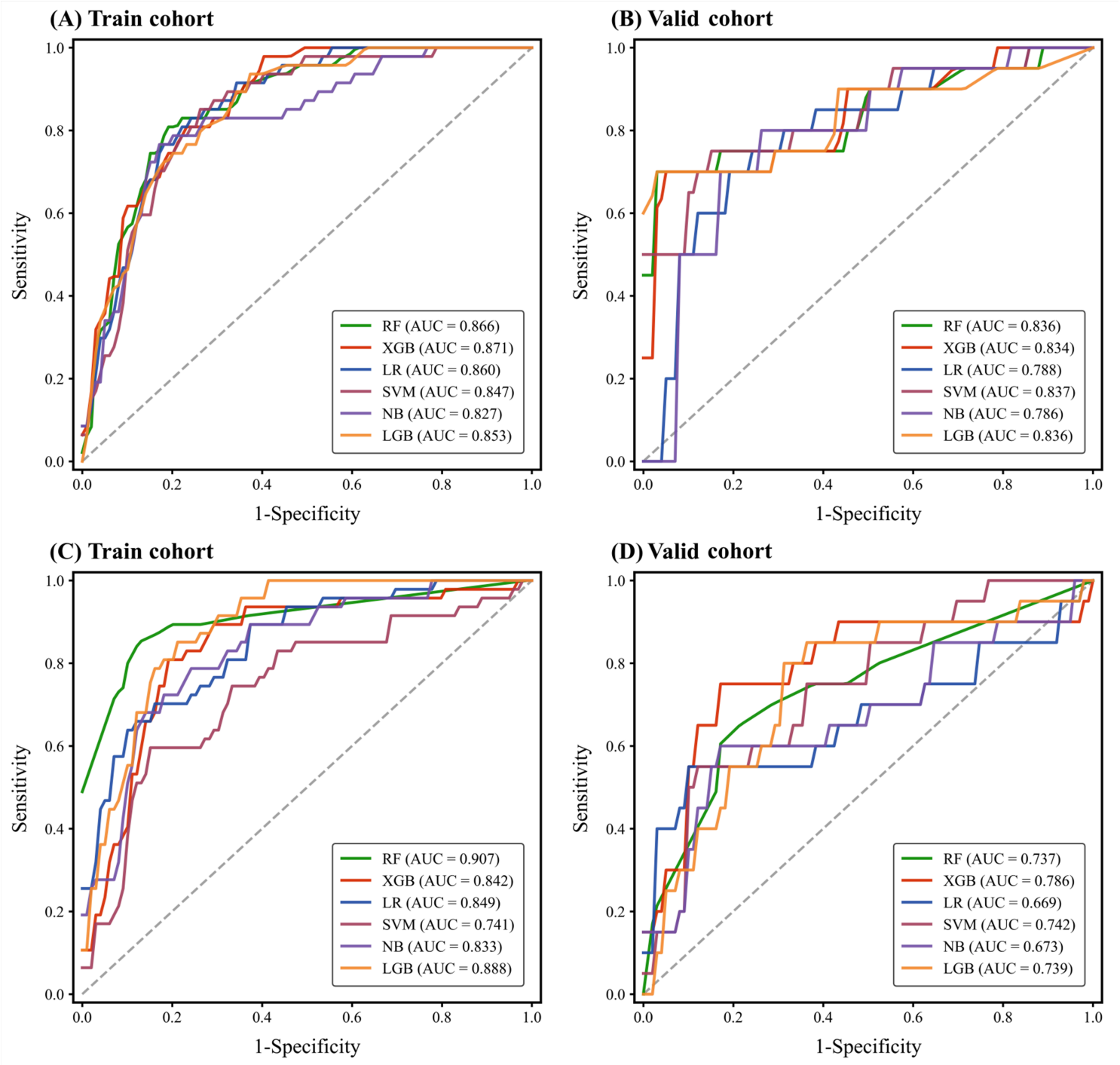


Figure 2

Receiver operating characteristic (ROC) curves of the prediction models constructed using six machine learning algorithms. (A) Clinical models in the training cohort. (B) Clinical models in the validation cohort. (C) Radiomic models in the training cohort. (D) Radiomic models in the validation cohort.

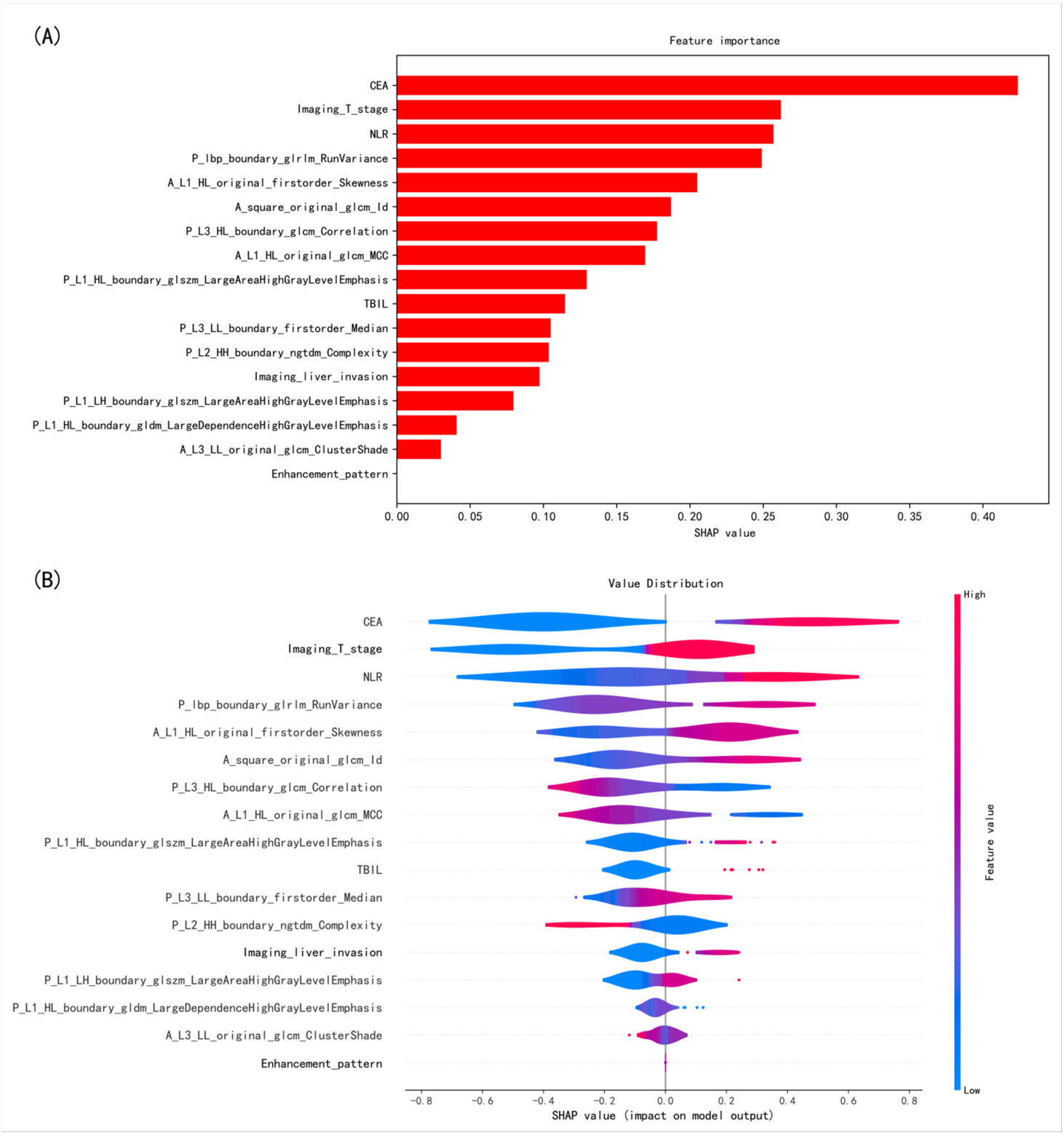


Figure 3

SHAP value analysis of modeling variables. (A) The SHAP bar figure shows the weight of features in constructing the models. (B) The SHAP dependence plot illustrates the individual effect of key features on the prediction probability.

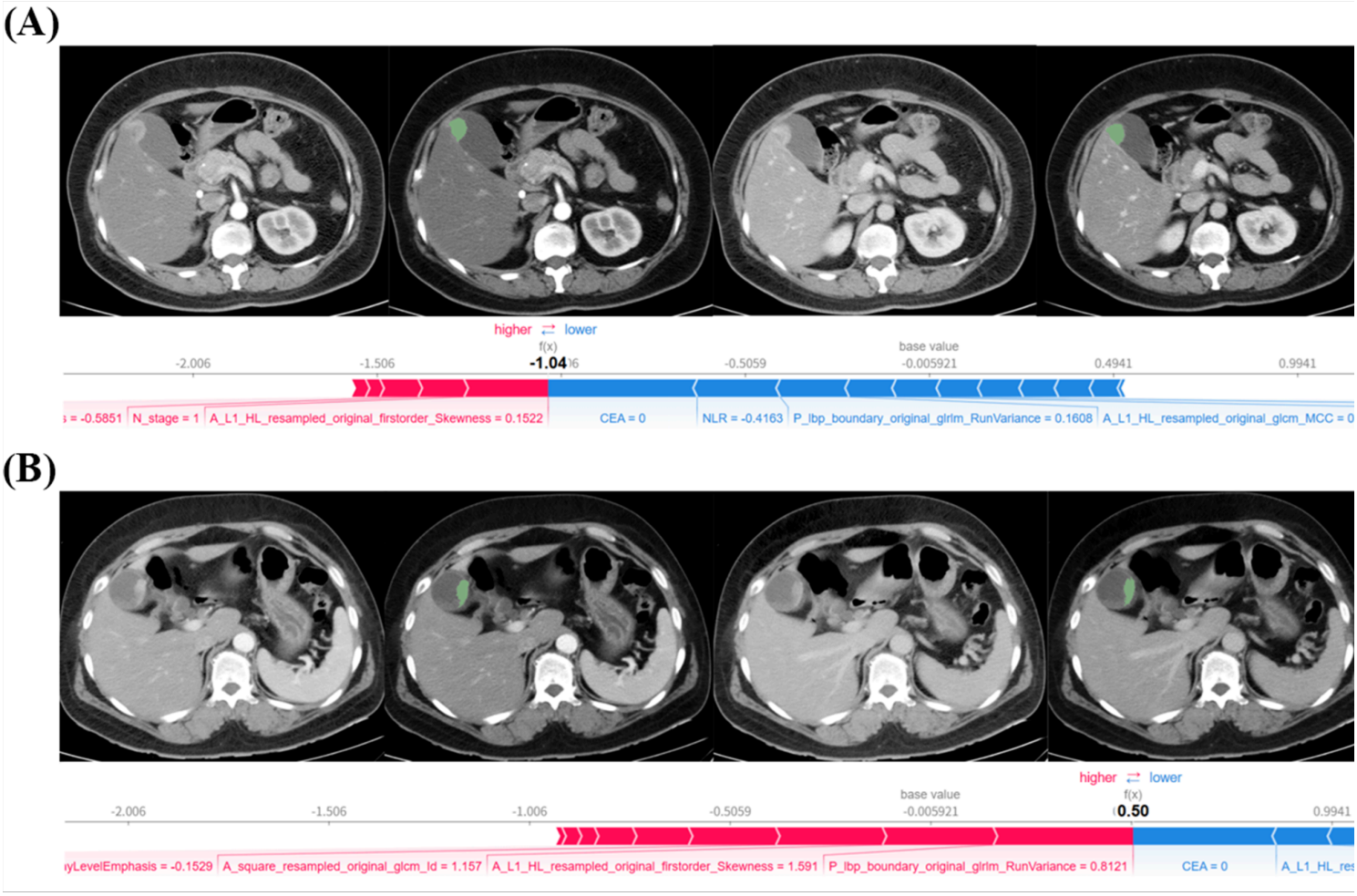


Figure 4

Individual prediction explanation using the SHAP force plot for a representative patient in the validation cohort

Supplementary Files

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

- SupplementaryMaterial.docx