Tumour Marker Prognostic Study

INTERNATIONAL JOURNAL OF SURGERY

OPEN

Noninvasive imaging evaluation of peritoneal recurrence and chemotherapy benefit in gastric cancer after gastrectomy: a multicenter study

Zepang Sun, MD, PhD[a], Wei Wang, MD, PhD[b,d], Weicai Huang, MD, PhD[a], Taojun Zhang, Msc[a], Chuanli Chen, Msc[c], Qingyu Yuan, Msc[c], Yan Chen, Msc[e], Kangneng Zhou, PhD[h], Zhen Han, MD, PhD[a], Hao Feng, Msc[a], Hao Chen, MD, PhD[a], Xiaokun Liang, PhD[f,g], Yanfeng Hu, MD, PhD[a], Jiang Yu, MD, PhD[a], Hao Liu, MD, PhD[a], Lequan Yu, PhD[i], Yikai Xu, MD, PhD[c,*], Guoxin Li[a,*], Yuming Jiang, MD, PhD[a,*]

Background: Peritoneal recurrence (PR) is the predominant pattern of relapse after curative-intent surgery in gastric cancer (GC) and indicates a dismal prognosis. Accurate prediction of PR is crucial for patient management and treatment. The authors aimed to develop a noninvasive imaging biomarker from computed tomography (CT) for PR evaluation, and investigate its associations with prognosis and chemotherapy benefit.

Methods: In this multicenter study including five independent cohorts of 2005 GC patients, the authors extracted 584 quantitative features from the intratumoral and peritumoral regions on contrast-enhanced CT images. The artificial intelligence algorithms were used to select significant PR-related features, and then integrated into a radiomic imaging signature. And improvements of diagnostic accuracy for PR by clinicians with the signature assistance were quantified. Using Shapley values, the authors determined the most relevant features and provided explanations to prediction. The authors further evaluated its predictive performance in prognosis and chemotherapy response.

Results: The developed radiomics signature had a consistently high accuracy in predicting PR in the training cohort (area under the curve: 0.732) and internal and Sun Yat-sen University Cancer Center validation cohorts (0.721 and 0.728). The radiomics signature was the most important feature in Shapley interpretation. The diagnostic accuracy of PR with the radiomics signature assistance was improved by 10.13–18.86% for clinicians ($P < 0.001$). Furthermore, it was also applicable in the survival prediction. In multivariable analysis, the radiomics signature remained an independent predictor for PR and prognosis ($P < 0.001$ for all). Importantly, patients with predicting high risk of PR from radiomics signature could gain survival benefit from adjuvant chemotherapy. By contrast, chemotherapy had no impact on survival for patients with a predicted low risk of PR.

Conclusion: The noninvasive and explainable model developed from preoperative CT images could accurately predict PR and chemotherapy benefit in patients with GC, which will allow the optimization of individual decision-making.

Keywords: artificial intelligence, chemotherapy benefit, gastric cancer, noninvasive evaluation, peritoneal recurrence

[a]Department of General Surgery and Guangdong Provincial Key Laboratory of Precision Medicine for Gastrointestinal Tumor, Nanfang Hospital, The First School of Clinical Medicine, Southern Medical University, [b]Department of Gastric Surgery, Sun Yat-sen University Cancer Center, [c]Department of Medical Imaging Center, Nanfang Hospital, Southern Medical University, Guangzhou Avenue North, Guangzhou, Guangdong, [d]State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, [e]Shenzhen Hospital of Integrated Traditional Chinese and Western Medicine, Shenzhen, [f]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, [g]Shenzhen Colleges of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen, Guangdong, [h]School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing and [i]The Department of Statistics and Actuarial Science, The University of Hong Kong, HKSAR, People's Republic of China

Z.S., W.W., W.H., and T.Z. are contributed equally as the first authors.

Y.X., G.L., and Y.J. are contributed equally as senior authors.

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

**Corresponding author. Address: Department of General Surgery and Guangdong Provincial Key Laboratory of Precision Medicine for Gastrointestinal Tumor, Nanfang Hospital, The First School of Clinical Medicine, Southern Medical University, Guangzhou, Guangdong 510515, People's Republic of China. Tel.:/fax.: +86 20 6278 7170. E-mail: jiangymbest@163.com (Y. Jiang) and E-mail: gzliguoxin@163.com (G. Li); Department of Medical Imaging Center, Nanfang Hospital, Southern Medical University, 1838 North Guangzhou Avenue, Guangzhou 510515, People's Republic of China. Tel.: +86 20 6278 7170. E-mail: Yikaivip@163.com (Y. Xu).*

International Journal of Surgery (2023) 109:2010–2024

Received 20 December 2022; Accepted 21 February 2023

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.lww.com/international-journal-of-surgery.

Published online 7 June 2023

http://dx.doi.org/10.1097/JS9.0000000000000328

Introduction

Peritoneal recurrence (PR) is one of the most common patterns of disease relapse and the primary cause of cancer-related death in gastric cancer (GC)[1–3]. The 5-year PR rate for patients with advanced GC after curative surgery is ~50%[4]. The median survival time from recurrence to death is 4–6 months[5,6]. PR indicates a poor prognosis and is often accompanied by complications, such as ascitic fluid, abdominal distension, and obstruction, which cause significant discomfort to GC patients[5–7]. The controlling strategies for PR in recent years have been more focused on preventing its recurrence by offering additional treatment alternatives[4,8–12]. Whether patients receive intensive treatment after surgery often depends on the clinical judgment of surgeons. However, such judgments often require extensive experience and are accompanied by randomness. Thus, there remains an unmet need to develop machine learning tools for prediction of PR risk to augment decision-making.

New strategies have been actively investigated in PR prevention, including hyperthermic intraperitoneal chemotherapy (HIPEC) and extensive intraoperative peritoneal lavage (EIPL)[11–17]. However, not all GC patients will develop PR after surgery, and thus a considerable number of patients are ineligible for these treatments. The findings of several investigations on HIPEC and EIPL remain controversial in survival benefit and suggest an increase in postoperative complications[11–17]. Lack of individualized risk prediction of PR to identify patients who may potentially benefit from these treatments could be the reason.

Although peritoneal metastasis can be diagnosed, prediction of future postoperative relapse is different and more challenging in clinic. Cytologic examination of peritoneal lavage was used to determine peritoneal metastasis, but it was reported to be invasive, time consuming, and expensive[18–20]. Other conventional imaging modalities, including computed tomography (CT), positron emission tomography-computed tomography (PET-CT), or ultrasonography, yielded unsatisfactory accuracy for PR prediction[21]. Thus, a novel biomarker to assess the risk of PR in patients with GC after surgery is urgently needed.

Previous data showed that radiomics, which extracted multidimensional quantitative data from standard medical imaging, could accurately predict cancer dissemination and prognosis[22–24]. Radiomics features from the center and contour of the tumor regions of interest (ROI) were governed by both cancer cell-intrinsic biological processes and tumor microenvironment, which were strongly correlated with tumor heterogeneity[25–28]. Therefore, we hypothesized that intratumoral and peritumoral radiomics features are excellent noninvasive biomarkers to predict PR and prognosis of GC.

This study aimed to develop a noninvasive and explainable artificial intelligence model to predict PR based on intratumoral and peritumoral radiomics features from preoperative CT images in GC patients after gastrectomy. The improvements of diagnostic accuracy for PR by clinicians with the assistance of noninvasive image model were further evaluated. We also investigated its performance in predicting prognosis and identifying which patients could benefit from adjuvant chemotherapy.

HIGHLIGHTS

- This is the first study using radiomics approach to develop a noninvasive biomarker from computed tomography images for predicting peritoneal recurrence (PR) of gastric cancer after gastrectomy.
- This noninvasive imaging biomarker improved the clinician's diagnosis of PR after gastrectomy by 10.13–18.86%.
- This noninvasive imaging biomarker can predict survival outcome and identify patients who can benefit from chemotherapy after gastrectomy.

Materials and methods

Study design and data retrieval

Five independent cohorts of 2005 patients diagnosed with stage I–III GC were enrolled in this study. All patients underwent contrast-enhanced abdominal CT examination within 2 weeks before surgery. The main inclusion and exclusion criteria are given in the Supplementary Methods (Supplemental Digital Content 1, http://links.lww.com/JS9/A669). The training cohort (Nanfang Hospital of Southern Medical University (SMUNFH) cohort 1, 433 patients) was retrospectively recruited from SMUNFH (Guangzhou, China) (2005–2010). The internal validation cohort (SMUNFH cohort 2, 471 patients) was retrospectively recruited from SMUNFH (2011–2014). And the prospective validation cohort (SMUNFH cohort 3, 136 patients) was prospectively recruited from SMUNFH (2017–2019). The training cohort comprised cohort 1 (SMUNFH cohort 1A) and cohort 2 (SMUNFH cohort 1B). Given that logistic regression performs best in a 1:1 case-to-noncase ratio, we designed training cohort 1 with 50 PR-positive patients and 50 PR-negative patients randomly selected from the 433 GC patients. The remaining 333 patients in training cohort 2 were assigned to penalized Cox regression. The Sun Yat-sen University Cancer Center (SYSUCC) validation cohort of 922 patients was collected from the SYSUCC (2008–2012). The TCIA cohort (43 patients) was obtained from The Cancer Immunome Atlas and The Cancer Genome Atlas databases (downloaded in 2022). This study had been reported in line with the REMARK criteria.

Baseline information was obtained from the medical record system. The threshold values of carcinoembryonic antigen (CEA) and carbohydrate antigen 19-9 (CA19-9) were 5.0 μg/ml and 37 U/ml, respectively. The pathologic stages were determined following the Eighth edition of the American Joint Committee on Cancer[29]. Adjuvant chemotherapy was mainly fluorouracil based. PR was diagnosed by CT, PET-CT, abdominal ultrasonography, or clinical signs, such as ascites or even reoperation during follow-up. Patients had a follow-up every 3 months for the first 2 years and every 6 months thereafter.

Radiomics feature extraction and reproducibility assessment

Portal venous-phase enhanced CT images were acquired using the multidetector-row CT system. Details about CT acquisition and image pre-treatment were presented in the Supplementary Methods (Supplemental Digital Content 1, http://links.lww.com/JS9/A669). Two-dimensional ROI of the tumors with the largest lesion diameter was manually delineated by two radiologists

(C.C. and Q.Y.) using the ITK-SNAP software. Any discrepancies were resolved by a senior radiologist (Y.X.). Both the center and contour of the tumor ROI were manually delineated, which was similar to the definition in our previous paper[30].

Radiomics features for ROIs were extracted and calculated using Matlab R2016a software using a package at GitHub (https://github.com/mvallieres/radiomics/). A total of 584 quantitative imaging features were computed for each tumor ROI, including the intratumoral and peritumoral areas. The reproducibility and robustness of imaging features were assessed according to intraclass and interclass correlation coefficients and shown in Supplementary Methods (Supplemental Digital Content 1, http://links.lww.com/JS9/A669). The radiomics features included shape features, first-order intensity features, and second and higher-order textural features. The detailed mathematical description of all extracted features was presented in Supplementary Methods (Supplemental Digital Content 1, http://links.lww.com/JS9/A669). This study followed the Image Biomarker Standardization Initiative (IBSI) guidelines, and the software used was IBSI-compliant[31].

PR-related feature selection and radiomics signature building

We designed a five-step protocol to select PR-related radiomics features based on the training cohort. Firstly, redundant features were removed using the Max-Relevance and Min-Redundancy (mRMR) algorithm. Secondly, Pearson correlation matrices were performed to eliminate the features with a low correlation with the PR in highly correlated paired-features. Thirdly, the features were further narrowed using the Least Absolute Shrinkage and Selector Operation (LASSO) logistic algorithm and the Support Vector Machine-Recursive Feature Elimination (SVM-RFE) algorithm with 10-fold cross-validation[32–35]. Fourthly, penalized Cox regression was used to further refine the features identified in the preceding steps. Finally, multiple features with nonzero coefficients at optimal lambda were integrated into a radiomics signature, also referred to as radiomics score (Rad-score). Patients were classified into high or low Rad-score group based on the threshold using X-tile in the training cohort, and then the validation cohorts were assigned the same threshold[36]. The detailed data processing is illustrated in Figure 1.

Model interpretability

To interpret how each radiomics feature and the integrated signature influenced the model prediction for PR, we used Shapley values[37]. The artificial intelligence SHapley Additive exPlanations (SHAP) provided a unified method for interpreting machine learning models. Based on the SHAP package in Python[38], we were able to get the importance of the radiomics features and other clinicopathologic characteristics with interpretations on how they participated in the prediction of PR.

Evaluation of the image model for predicting PR

We evaluated the performance of radiomics model in predicting PR using the receiver operating characteristic (ROC) curve, calibration curves, and decision curve analysis (DCA). The Hosmer–Lemeshow (H–L) test was used to evaluate the fit of the predicted and true values. And the prediction for 1, 3, and 5-year PR by radiomics signature was evaluated using time-dependent ROC analysis. The specific evaluation metrics were the area under the ROC curve (AUC), specificity, sensitivity, accuracy, positive predictive value (PPV), and negative predictive value (NPV). The multivariate analysis was also performed to assess this signature. Additionally, Fine and Gray's competing-risk regression analysis[39] was used to assess PR-free survival (PRFS), and deaths, local recurrence, or distant metastasis before PR were identified as competing events. We also assessed the ability of the radiomics signature to improve the accuracy of clinicians (Z.H., 5 years of experience, and H.L., 15 years of experience) for predicting PR. To evaluate the impact of the radiomics signature on clinician performance for individual prediction, two oncologists (Z.H. and H.L.) independently evaluated the PR status of each patient, and then re-evaluated recurrence status after considering the radiomics signature. And the statistical difference with or without assistance from radiomics signature for predicting PR was evaluated by independent sample *t*-test.

Evaluation of the image model for predicting prognosis and chemotherapy benefit

Performance of the radiomics signature for predicting disease-free survival (DFS) and overall survival (OS) was also assessed through time-dependent ROC analysis. Furthermore, we constructed the combined models, integrated radiomics signature and clinicopathological characteristics, to predict the PR and prognosis. Additionally, we investigated the ability of the radiomics signature in predicting the benefit of adjuvant chemotherapy in patients with stage II and III disease. Based on the aforementioned classification, we divided patients into low Rad-score or high Rad-score group. In two defined groups, a 1:1 nearest matching strategy of propensity score matching (PSM) was performed for patients with or without chemotherapy to balance confounding factors including age, sex, CEA, CA19-9, tumor grade, tumor location, T stage, N stage, tumor size, and histology type. We also evaluated the chemotherapy benefit in all patients of each group without PSM.

Statistical analysis

Continuous variables were expressed as mean ± standard deviation ($X \pm SD$), and compared among groups using *t*-test or Mann–Whitney test. Enumeration data were expressed as percentages and compared among groups by χ^2 or Fisher's exact test. Survival curves were generated according to the Kaplan–Meier method and compared using the log-rank test. Univariate and multivariate analyses were performed using the Cox proportional hazards model. Interaction between the radiomic signature and adjuvant chemotherapy was assessed by the Cox model. All statistical analyses were performed using R software (version 3.5.3), SPSS statistical software (version 26.0), and Python software (version 3.6). A two-sided $P < 0.05$ was considered statistically significant.

Results

Clinicopathological characteristics

The present study included 2005 patients with GC from five independent cohorts of three centers. The overall study design is shown in Figure 1. Patients ($n = 1826$) in the training cohort, internal validation

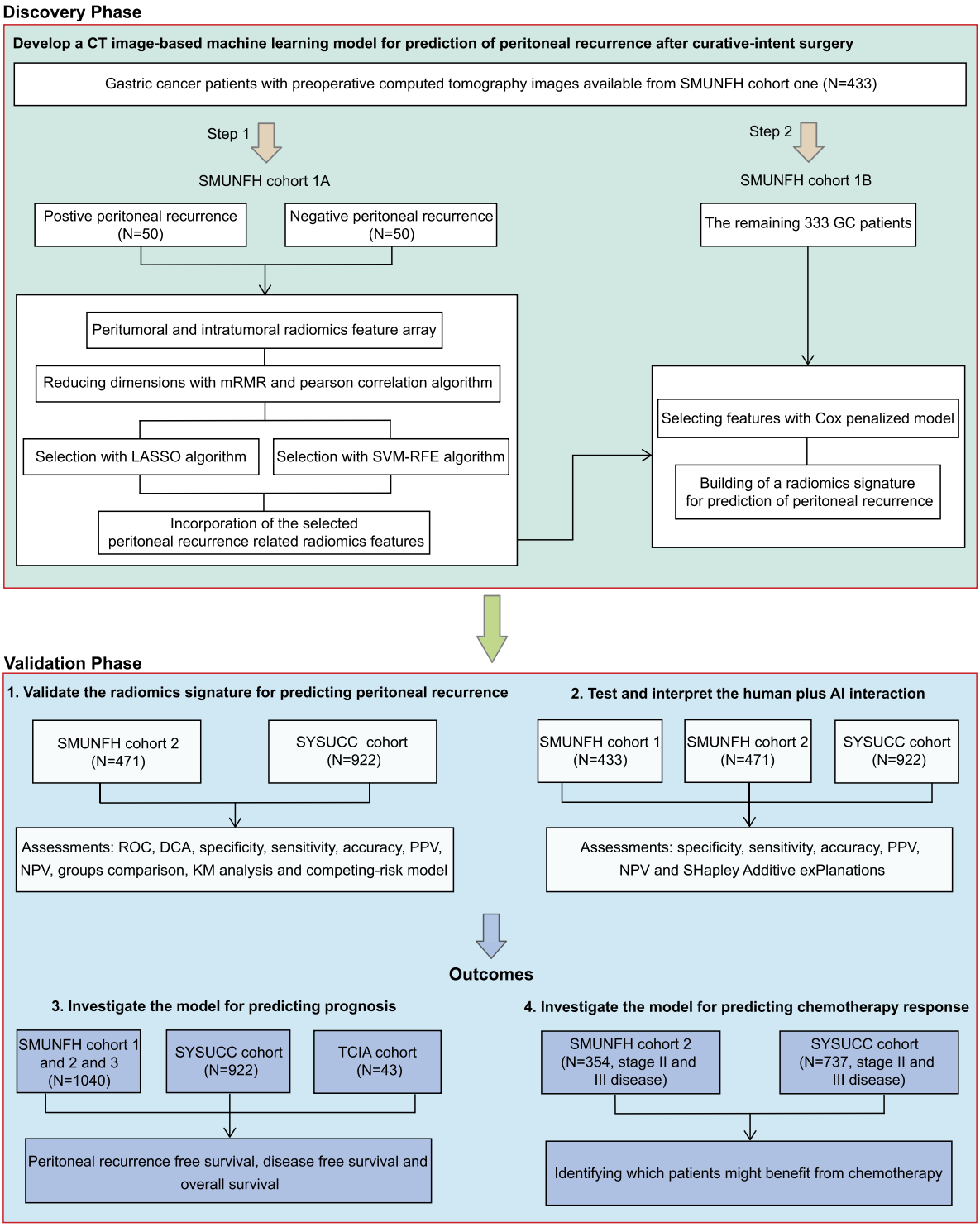


Figure 1. Study design. LASSO, Least Absolute Shrinkage and Selector Operation; mRMR, Max-Relevance and Min-Redundancy; SVM-RFE, Support Vector Machine-Recursive Feature Elimination.

cohort, and SYSUCC validation cohort were used to evaluate the predictive value of Rad-score for PR, survival outcome, and chemotherapy response. Moreover, patients in the prospective validation cohort ($n=136$) or TCIA cohort ($n=43$) were used to assess relationships between Rad-score and survival outcome.

Among the training cohort, internal validation cohort, and SYSUCC validation cohort, 1252 patients (68.6%) were male, and the median (interquartile range) age was 57.0 (49.0–64.0) years. Most patients [1404 (76.9%)] had stage II–III disease; 806 (44.1%) patients had received postoperative chemotherapy. The

median follow-up period (IQR) was 62 (47–73) months. There were 491 (26.9%) patients with PR and 379 (20.8%) patients with competing events till the last follow-up. The median duration (IQR) to PR was 21 (11–36) months, and the median time from PR to death was 5 months for all patients. Moreover, the median duration (IQR) to PR was 19 (11–37) months, 12 (5–28) months, and 24 (15–38) months in the training cohort, internal validation cohort, and SYYUCC validation cohort, separately.

Table 1

Clinicopathological characteristics of patients with gastric cancer in the training and validation cohorts.

	Training cohort ($N=433$)	Internal validation cohort ($N=471$)	SYSUCC validation cohort ($N=922$)
Variables	***N* (%)**	***N* (%)**	***N* (%)**
Age (years)			
< 60	262 (60.5)	281 (59.7)	520 (56.4)
≥ 60	171 (39.5)	190 (40.3)	402 (43.6)
Sex			
Female	137 (31.6)	149 (31.6)	288 (31.2)
Male	296 (68.4)	322 (68.4)	634 (68.8)
Tumor size			
< 4 cm	217 (50.1)	312 (66.2)	378 (41.0)
≥ 4 cm	216 (49.9)	159 (33.8)	544 (59.0)
Histology type			
Intestinal	205 (47.3)	211 (44.8)	320 (34.7)
Diffuse and mixed	228 (52.7)	260 (55.2)	602 (65.3)
Tumor grade			
Well	49 (11.3)	82 (17.4)	14 (1.5)
Moderate	116 (26.8)	103 (21.9)	150 (16.3)
Poor or undifferentiated	268 (61.9)	286 (60.7)	758 (82.2)
Tumor location			
Cardia	91 (21.0)	63 (13.4)	323 (35.0)
Body	88 (20.3)	86 (18.3)	179 (19.4)
Antrum	234 (54.0)	284 (60.3)	381 (41.3)
Whole	20 (4.6)	38 (8.1)	39 (4.2)
CEA			
Normal	388 (89.6)	420 (89.2)	746 (80.9)
Elevated	45 (10.4)	51 (10.8)	176 (19.1)
CA19-9			
Normal	380 (87.8)	390 (82.8)	753 (81.7)
Elevated	53 (12.2)	81 (17.2)	169 (18.3)
pT stage			
T1	103 (23.8)	126 (26.8)	129 (14.0)
T2	44 (10.2)	71 (15.1)	122 (13.2)
T3	52 (12.0)	14 (3.0)	203 (22.0)
T4a	193 (44.6)	154 (32.7)	413 (44.8)
T4b	41 (9.5)	106 (22.5)	55 (6.0)
pN stage			
N0	199 (46.0)	222 (47.1)	337 (36.6)
N1	78 (18.0)	108 (22.9)	164 (17.8)
N2	57 (13.2)	53 (11.3)	159 (17.2)
N3a	56 (12.9)	66 (14.0)	179 (19.4)
N3b	43 (9.9)	22 (4.7)	83 (9.0)
pM stage			
M (−)	433 (100)	471 (100)	922 (100)
M (+)	0 (0)	0 (0)	0 (0)
pTNM stage			
IA	92 (21.2)	102 (21.7)	98 (10.6)
IB	35 (8.1)	8 (1.7)	87 (9.4)
IIA	30 (6.9)	14 (3.0)	91 (9.9)
IIB	68 (15.7)	74 (15.7)	180 (19.5)
IIIA	98 (22.6)	181 (38.4)	195 (21.1)
IIIB	54 (12.5)	44 (9.3)	175 (19.0)
IIIC	56 (12.9)	48 (10.2)	96 (10.4)
Postoperative peritoneal recurrence			
Negative	330 (76.2)	329 (69.9)	676 (73.3)
Positive	103 (23.8)	142 (30.1)	246 (26.7)
Chemotherapy			
Yes	212 (49.0)	209 (44.4)	385 (41.8)
No	221 (51.0)	262 (55.6)	537 (58.2)

CA19-9, carbohydrate antigen 19-9; CEA, carcinoembryonic antigen; pTNM stage, based on the eighth edition NCCN guideline.

The median time from PR to death was 6 months, 3 months, and 5 months in the training cohort, internal validation cohort, and SYYUCC validation cohort, separately. The clinicopathological characteristics of patients were summarized in Table 1 and Table S1 (Supplemental Digital Content 1, http://links.lww.com/JS9/A669).

Table S2 (Supplemental Digital Content 1, http://links.lww.com/JS9/A669) lists the detailed clinicopathological features of the prospective validation cohort ($n = 136$). In this cohort, 94 patients (69.1%) were male, and the median age (interquartile range) was 55.0 (49.0–65.0) years; 49 patients (36.0%) had stage I GC; 37 patients (27.2%) had stage II GC; 50 patients (36.8%) had stage III GC. Only 11.7% (16/136) of patients suffered from PR at the last follow-up. Therefore, this cohort was merely used to evaluate the predictive value of the Rad-score for prognosis.

In the TCIA cohort, 36 patients (83.7%) were male, and the median age (interquartile range) was 67.0 (59.0–70.0) years. Most patients ($n = 42$; 97.7%) had stage II or III disease, except for one patient with stage I GC. And their clinicopathological characteristics are shown in Table S3 (Supplemental Digital Content 1, http://links.lww.com/JS9/A669).

Feature selection and radiomics signature establishment

Initially, a total of 584 radiomics features were extracted from the tumor ROI of each patient. After removing features with unsatisfactory intra- and interclass correlation coefficientss and redundant features, 157 independent features were retained for further analysis (Fig. S1, Supplemental Digital Content 1, http://links.lww.com/JS9/A669). Then, 36 PR-related features were identified using the LASSO logistic algorithm and SVM-RFE algorithm in training cohort 1. Finally, six robust radiomics features, including one intratumoral feature and five peritumoral features, were selected using penalized Cox regression in the training cohort 2 for construction of the PR-related Rad-score (Fig. 2). The selected features were GLCM_IMC2, Extent, Perimeter, NGTDM_Coarseness, GLRLM_LGRE_1.0, and GLSZM_LZHGE_2.5. For example, the Extent was defined as the ratio of pixels in the region to pixels in the total bounding box, returned as a scalar, and the Perimeter was defined as the distance around the boundary of the region. The definitions for GLCM_IMC2, NGTDM_Coarseness, GLRLM_LGRE_1.0, and GLSZM_LZHGE_2.5 are presented in the Supplementary Materials (Supplemental Digital Content 1, http://links.lww.com/JS9/A669).

Explanations via artificial intelligence SHAP

Risk estimates can be extracted from the prediction model by SHAP values to allow an explanation of risk on a global level. The most important features to predict PR were NGTDM_Coarseness, Perimeter, Extent, GLCM_IMC2, and GLRLM_LGRE_1.0 in SHAP interpretation. The seventh most important feature was GLSZM_LZHGE_2.5. T stage, N stage, age, sex, tumor size, location, tumor grade, histology type, CEA, and CA19-9 were not among the top five features (Fig. 3A). Furthermore, when six radiomics features were integrated into a signature, the importance had been further enhanced (Fig. 3B).

Assessment of the radiomics signature with PR

There was no significant difference in Rad-score [mean (95% CI)] between the training [−0.023 (−0.080 to 0.033)], internal validation [−0.011 (−0.055 to 0.033)], and SYSUCC validation [−0.049 (−0.087 to −0.012)] cohorts ($P = 0.44$). The optimal threshold generated by X-tile was −0.02 (Fig. S2, Supplemental Digital Content 1, http://links.lww.com/JS9/A669), and all 1826 patients were classified into high or low Rad-score group. Rad-score [mean (95% CI)] in PR-positive was significantly higher compared with PR-negative in the training [0.344 (0.238–0.451) vs. −0.138 (−0.200 to −0.076)], internal validation [0.233 (0.160–0.306) vs. −0.117 (−0.167 to −0.066)], and SYSUCC validation [0.292 (0.226–0.358) vs. −0.174 (−0.216 to −0.133)] cohorts (all $P < 0.001$) (Fig. 4A). The clinical characteristics' distribution of patients according to the PR status in the training and two validation cohorts was shown in Table S4 (Supplemental Digital Content 1, http://links.lww.com/JS9/A669). A significantly higher PR rate at 1, 3, and 5 years was found in patients with a high Rad-score in all three cohorts (all $P < 0.001$) (Fig. 4B). The risk of PR increased as the Rad-score increased with an odds ratio of 3.762–6.302 ($P < 0.001$) in the multivariable analysis [Table 2 and Table S5 (Supplemental Digital Content 1, http://links.lww.com/JS9/A669)].

Artificial intelligence interaction for diagnosis of PR

The Rad-score had a consistently high accuracy in predicting PR, with an AUC of 0.732 (0.677–0.787) in the training cohort, 0.721 (0.672–0.771) in the internal validation cohort, and 0.728 (0.692–0.764) in the SYSUCC validation cohort, which were superior to TNM stage (0.625) [Fig. 5A and Table S6 (Supplemental Digital Content 1, http://links.lww.com/JS9/A669)]. By analyzing the sub-queue of the training cohort, a comparable AUC was found in cohorts 1 (0.703, 0.601–0.806) and 2 (0.758, 0.688–0.827), which was also shown in Figure S3 (Supplemental Digital Content 1, http://links.lww.com/JS9/A669). The specificity, sensitivity, accuracy, PPV, and NPV of the Rad-score in the validation cohorts were 68.4–79.4%, 59.2–61.4%, 66.2–67.6%, 43.0–44.7%, and 79.5–83.4%, respectively, when the optimal Youden index was selected (Table S7, Supplemental Digital Content 1, http://links.lww.com/JS9/A669).

The AUC of the Rad-score for 1, 3, and 5-year PR evaluation was 0.770 (0.692–0.848), 0.773 (0.711–0.835), and 0.799 (0.744–0.854), respectively, in the training cohort. Similar AUC values for Rad-score were observed in the internal [0.786 (0.737–0.835) at 1-year, 0.779 (0.726–0.832) at 3-year, and 0.775 (0.716–0.834) at 5-year] and SYSUCC validation cohorts [0.784 (0.715–0.853) at 1-year, 0.723 (0.682–0.764) at 3-year, and 0.792 (0.751–0.833) at 5-year] (Table S8, Supplemental Digital Content 1, http://links.lww.com/JS9/A669). Rad-score was a better predictor for 1, 3, and 5-year PR than other clinicopathological characteristics (Figs S4–S6, Supplemental Digital Content 1, http://links.lww.com/JS9/A669).

The Rad-score had the best predictive performance for PR, and improvement in prediction was observed when combined with clinicopathological characteristics (Fig. 5A). DCA also showed that the Rad-score had a better net benefit than any clinical factor at probability thresholds ranging from 0.20 to 0.80 (Fig. 5B). Additionally, the combined models integrating Rad-score and clinical factors were developed for individualized prediction of

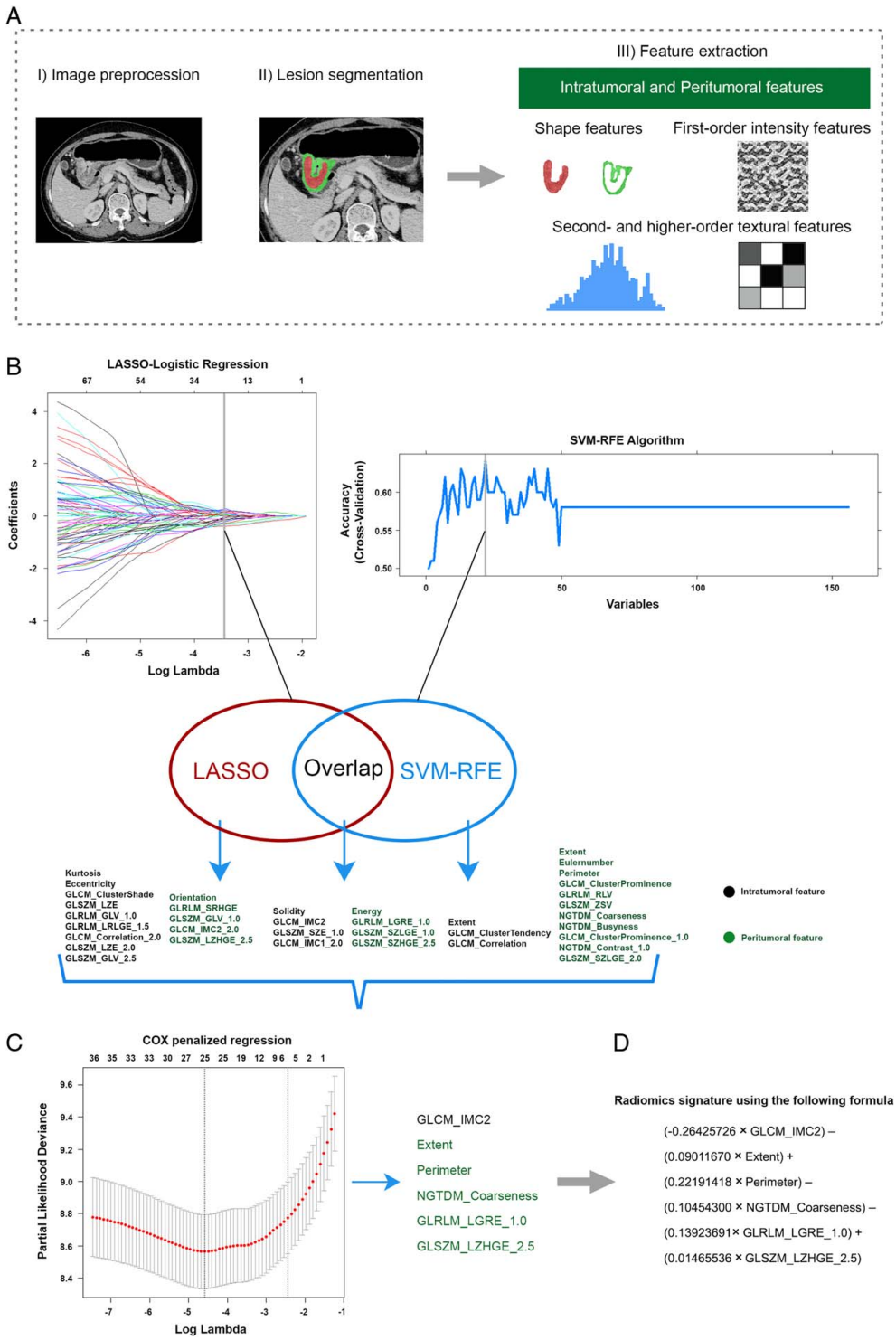


Figure 2. Intratumoral and peritumoral ROI segmentation and feature extraction (A). Cluster analysis of incorporation of peritoneal recurrence-related radiomics features that were selected from the LASSO or SVM-RFE algorithms in the training cohort 1 (B). Penalized Cox regression analysis to select robust radiomics features in the training cohort 2 (C). A radiomics signature is built based on the radiomics features (D). LASSO, Least Absolute Shrinkage and Selector Operation; SVM-RFE, Support Vector Machine-Recursive Feature Elimination.

PR, with good agreements between the predictive and actual probability in calibration curves (Figs S7 and S8, Supplemental Digital Content 1, http://links.lww.com/JS9/A669). The H–L test also showed a good fit between the predicted and true values [$P > 0.05$; Table S9 (Supplemental Digital Content 1, http://links.lww.com/JS9/A669)].

We also found that, with the assistance of radiomics signature, diagnostic performance of PR was significantly improved

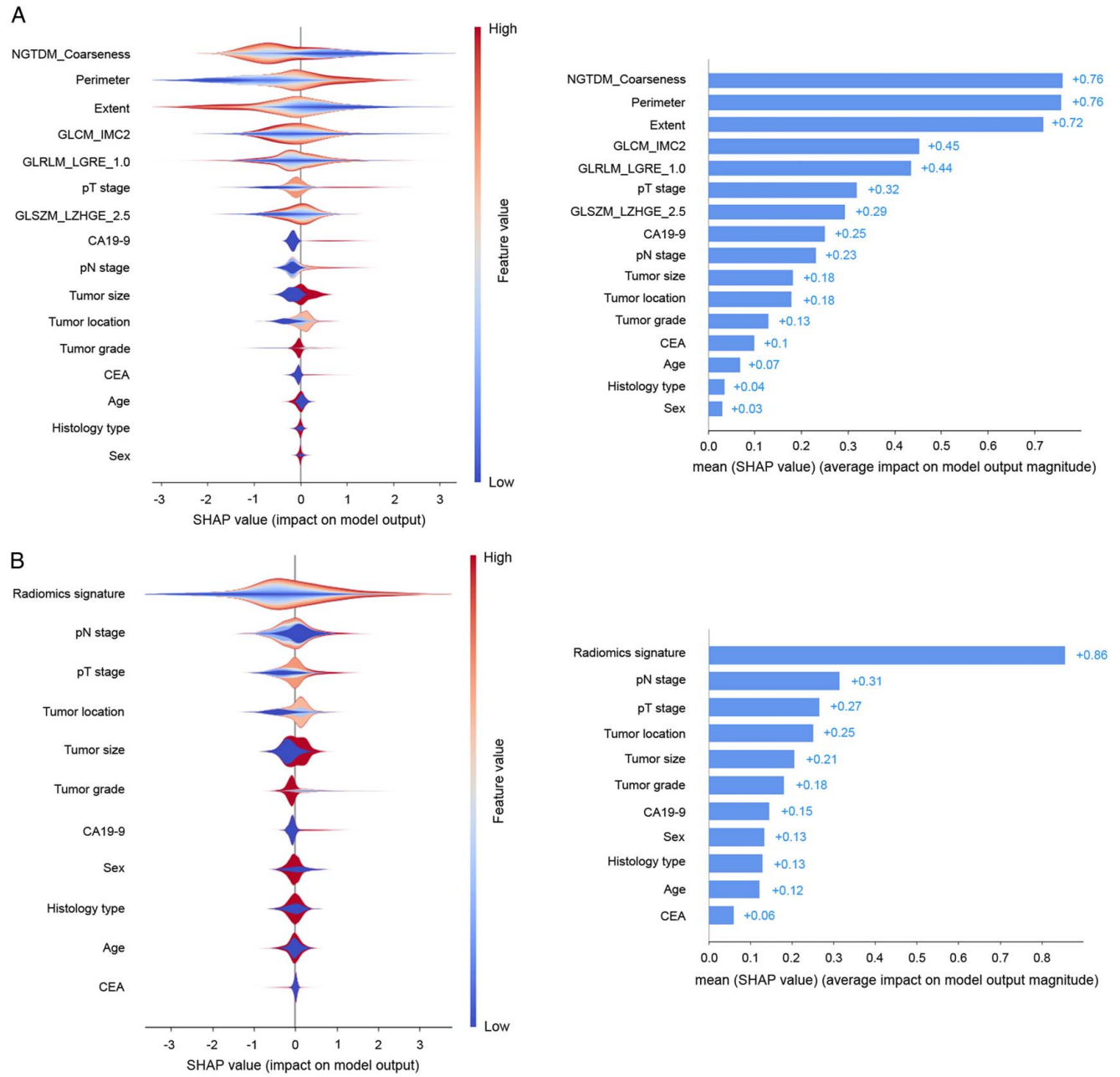


Figure 3. The features to predict the risk of peritoneal recurrence. SHAP for radiomics features and clinicopathologic characteristics (A). SHAP for radiomics signature and clinicopathologic characteristics (B). On the *x*-axis, the contribution of each feature is shown. A feature with a negative Shapley value will favorably impact the prediction (decrease the risk of peritoneal recurrence). The influence of the value of the feature itself is shown on the *y*-axis, for example, for radiomics signature, a high value (in red) is associated with a positive Shapley value that will increase the risk of peritoneal recurrence, while a low value (in blue) will decrease the Shapley value and the risk of peritoneal recurrence.

(Fig. 5C). For the junior clinician, the accuracy increased from 53.34 to 68.36% in the training cohort, 49.36 to 68.22% in the internal validation cohort, and from 50.76 to 67.46% in the SYSUCC validation cohort. Similarly, for the senior clinician, the accuracy increased from 60.74 to 71.35% in the training cohort, 60.12 to 70.25% in the internal validation cohort, and from 58.74 to 69.49% in the SYSUCC validation cohort. A significant statistical difference was observed in the diagnostic accuracy for predicting PR by clinicians with or without the assistance of noninvasive radiomics tools [$P < 0.001$; Table S10 (Supplemental Digital Content 1, http://links.lww.com/JS9/A669)]. These results indicated that with the help of this radiomics model, the diagnostic accuracy of junior clinicians could reach a level comparable to expert clinicians.

Prognostic value of the radiomics signature

The relationship between Rad-score and prognosis was then explored. Kaplan–Meier analysis revealed a significantly higher 1, 3, and 5-year PRFS, DFS, and OS rate in patients with a low

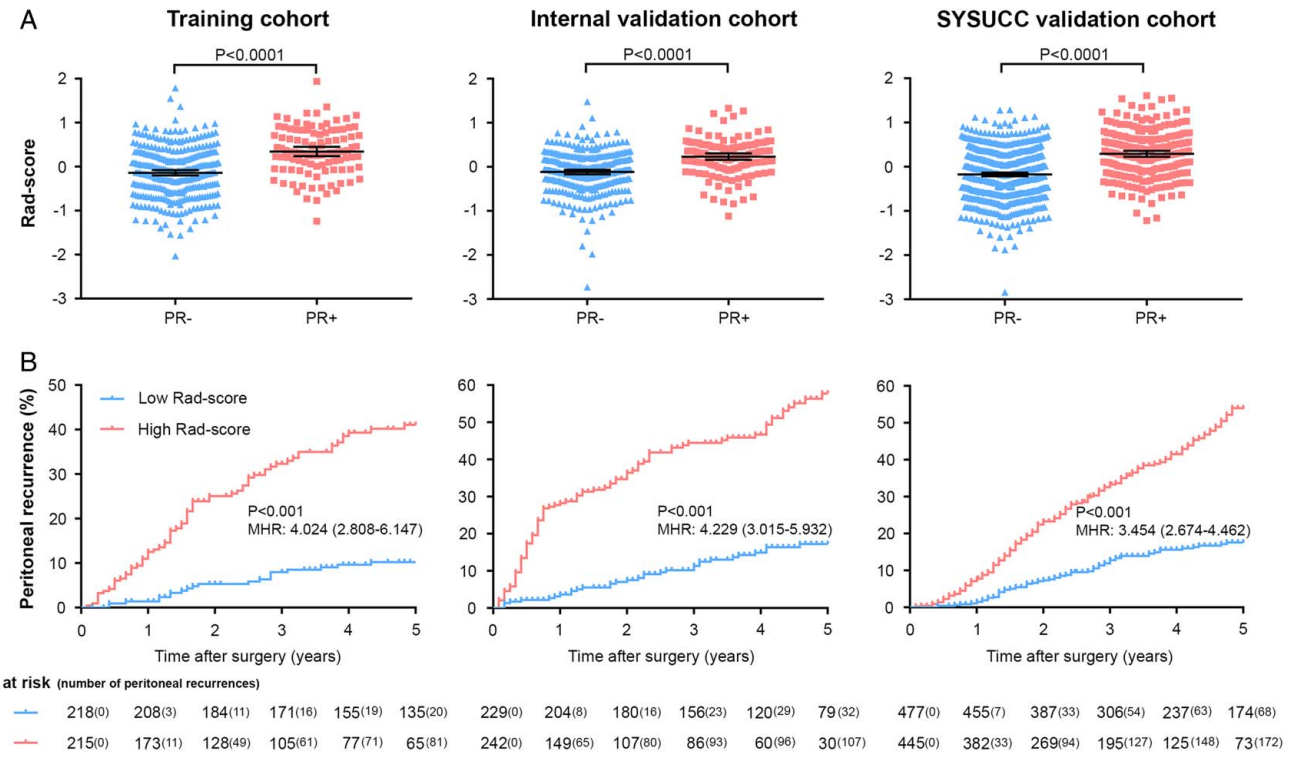


Figure 4. Distribution of radiomics scores regarding the status of peritoneal recurrence (PR− and PR+) in the training cohort, internal validation cohort, and SYSUCC validation cohort (A). Cumulative peritoneal recurrence rate stratified by the radiomics scores in the training cohort, internal validation cohort, and SYSUCC validation cohort using a Mantel–Haenszel's Hazard Ratio (MHR) test (B).

Table 2

Multivariable logistic analysis for postoperative peritoneal recurrence and multivariable Cox regression analysis for peritoneal recurrence-free survival in the training and validation cohorts.

	Multivariable logistic analysis		Multivariable cox analysis	
Variables	**OR (95% CI)**	***P***	**HR (95% CI)**	***P***
Training cohort				
Tumor size (< 4 cm vs. ≥ 4 cm)	0.884 (0.526–1.486)	0.642	1.110 (0.740–1.664)	0.614
CEA (< 5.0 μg/ml vs. ≥ 5.0 μg/ml)	–	–	1.897 (1.123–3.203)	**0.017**
CA19-9 (< 37 U/ml vs. ≥ 37 U/ml)	1.272 (0.649–2.492)	0.484	1.114 (0.662–1.876)	0.684
pTNM stage (IA vs. IB vs. IIA vs. IIB vs. IIIA vs. IIIB vs. IIIC)	1.257 (1.090–1.450)	**0.002**	1.255 (1.112–1.417)	< **0.001**
Rad-score	3.762 (2.340–6.047)	< **0.001**	3.390 (2.388–4.813)	< **0.001**
Internal validation cohort				
Histology type (intestinal vs. diffuse/mixed)	–	–	1.502 (1.063–2.122)	**0.021**
Tumor location (cardia vs. body vs. antrum vs. whole)	1.611 (1.217–2.132)	**0.001**	1.603 (1.287–1.997)	< **0.001**
CEA (< 5.0 μg/ml vs. ≥ 5.0 μg/ml)	–	–	0.992 (0.574–1.713)	0.976
CA19-9 (< 37 U/ml vs. ≥ 37 U/ml)	2.144 (1.231–3.733)	**0.007**	2.503 (1.612–3.886)	< **0.001**
pTNM stage (IA vs. IB vs. IIA vs. IIB vs. IIIA vs. IIIB vs. IIIC)	1.131 (0.997–1.283)	0.055	1.260 (1.139–1.393)	< **0.001**
Rad-score	6.302 (3.579–11.096)	< **0.001**	4.722 (3.304–6.747)	< **0.001**
SYSUCC validation cohort				
Sex (female vs. male)	1.299 (0.910–1.855)	0.150	1.328 (0.991–1.779)	0.057
Tumor location (cardia vs. body vs. antrum vs. whole)	–	–	1.047 (0.914–1.200)	0.510
Tumor size (< 4 cm vs. ≥ 4 cm)	0.783 (0.547–1.122)	0.183	0.952 (0.721–1.257)	0.727
CEA (< 5.0 μg/ml vs. ≥ 5.0 μg/ml)	–	–	0.972 (0.714–1.322)	0.856
CA19-9 (< 37 U/ml vs. ≥ 37 U/ml)	1.605 (1.089–2.364)	**0.017**	1.475 (1.096–1.985)	**0.010**
pTNM stage (IA vs. IB vs. IIA vs. IIB vs. IIIA vs. IIIB vs. IIIC)	1.267 (1.141–1.406)	< **0.001**	1.326 (1.214–1.447)	< **0.001**
Rad-score	4.836 (3.478–6.725)	< **0.001**	3.739 (2.972–4.703)	< **0.001**

Bold values denote significant variables.

CA19-9, carbohydrate antigen 19-9; CEA, carcinoembryonic antigen; CI, confidence interval; HR, hazard ratio; OR, odds ratio; Rad-score, based on continuous values.

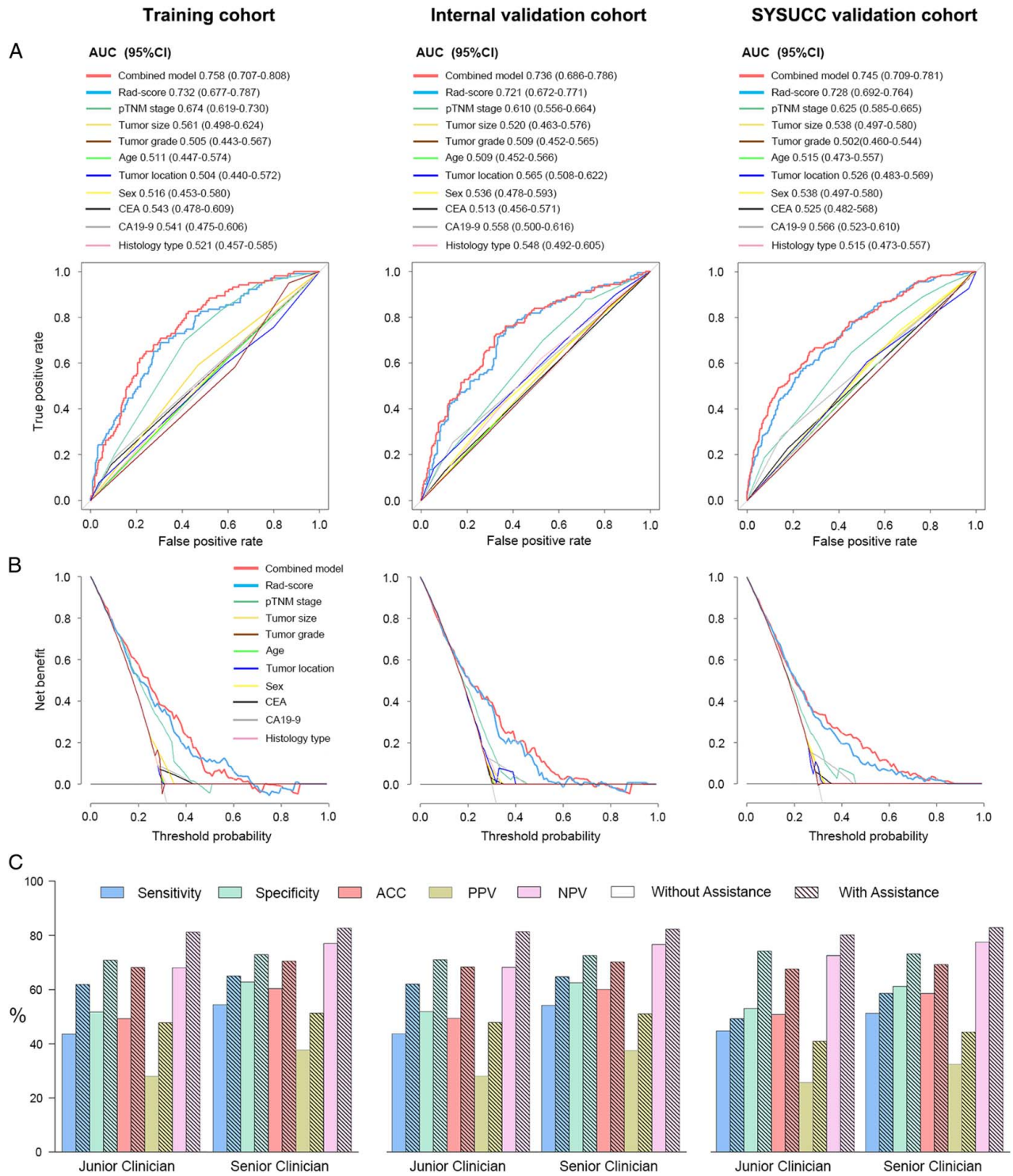


Figure 5. Receiver operating characteristic curves to predict peritoneal recurrence in the training cohort, internal validation cohort, and SYSUCC validation cohort (A). Decision curves to predict peritoneal recurrence in the training cohort, internal validation cohort, and SYSUCC validation cohort (B). Improvements in the performance of the clinician to predict peritoneal recurrence with the radiomics model's assistance. Striped bar indicates the result with the radiomics model's assistance (C).

Rad-score (Fig. 6), which was confirmed when patients were stratified by pathological stage (Fig. S9, Supplemental Digital Content 1, http://links.lww.com/JS9/A669). Furthermore, in multivariable analysis, the Rad-score was identified as the most important prognostic factor for PRFS (Table 2), DFS (Table S11, Supplemental Digital Content 1, http://links.lww.com/JS9/A669), and OS (Table S12, Supplemental Digital Content 1, http://links.lww.com/JS9/A669). Fine and Gray's competing-risk regression analysis also showed that the Rad-score remained a significant predictor for PFRS (Table S13, Supplemental Digital Content 1,

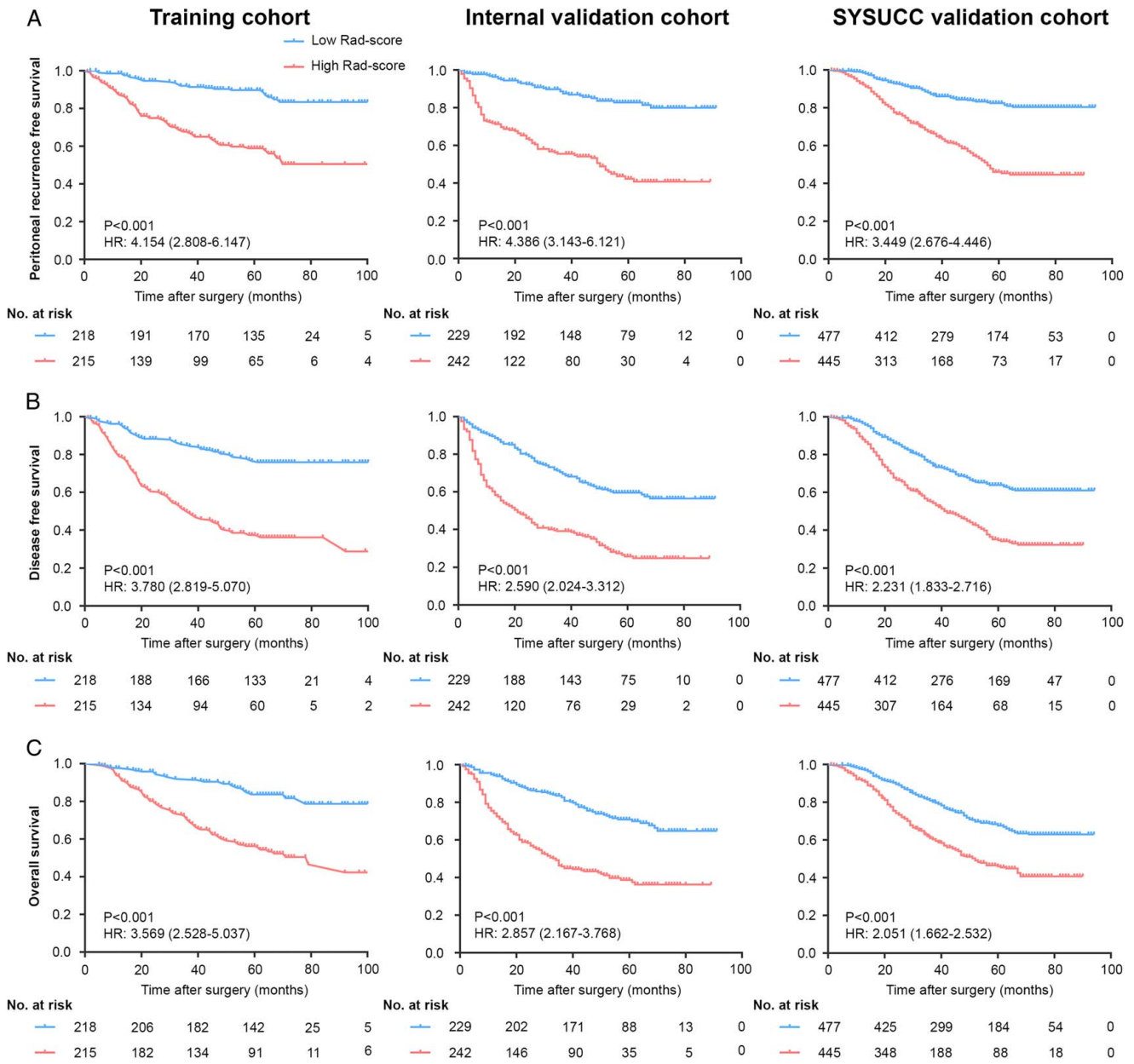
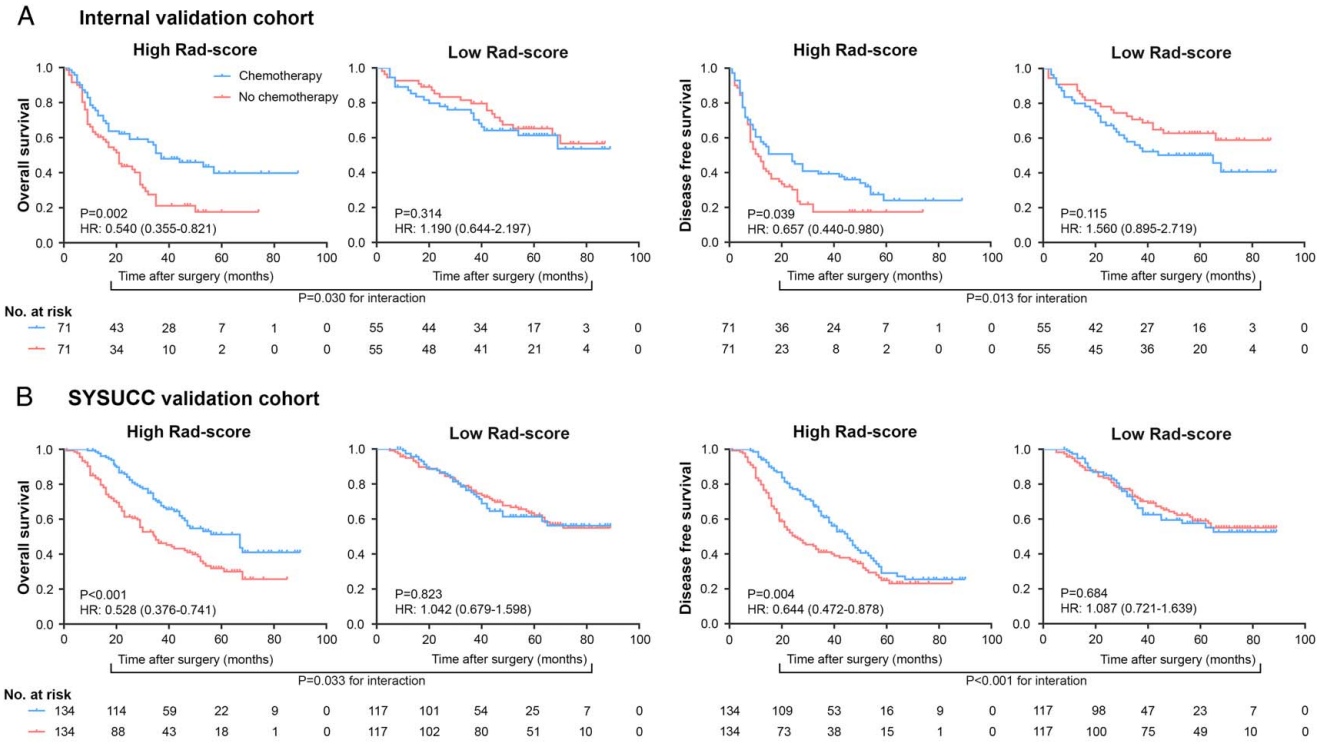


Figure 6. Kaplan–Meier plots for peritoneal recurrence-free survival (A), disease-free survival (B), and overall survival (C) in the training cohort, internal validation cohort, and SYSUCC validation cohort.

http://links.lww.com/JS9/A669). In addition, the Rad-score exhibited a consistently high accuracy for evaluating DFS and OS at 1, 3, and 5 years, with high AUCs in the training cohort (DFS: 0.773–0.823; OS: 0.801–0.827), internal validation cohort (DFS: 0.800–0.824; OS: 0.801–0.834), and SYSUCC validation cohort (DFS: 0.771–0.810; OS: 0.768–0.813) [Tables S14 and S15 (Supplemental Digital Content 1, http://links.lww.com/JS9/A669), Figs S10 and S11 (Supplemental Digital Content 1, http://links.lww.com/JS9/A669)]. Furthermore, as shown in Figure S12 (Supplemental Digital Content 1, http://links.lww.com/JS9/A669), the median DFS (18.4 vs. 9.7 months) and OS (34.3 vs. 21.0 months) were longer in GC patients with low Rad-score, compared to patients with high Rad-score, although the difference was not statistically significant. It may be attributed to the small sample size and short follow-up time. Besides, in the SMUNFH cohort 3 including prospective CT data, we found that patients with low Rad-score also had a better prognosis ($P < 0.006$) (Fig. S13, Supplemental Digital Content 1, http://links.lww.com/JS9/A669). In multivariable analysis, the Rad-score was also identified as an independent predictor for survival outcome (Table S16, Supplemental Digital Content 1, http://links.lww.com/JS9/A669).

Predictive value of the radiomics signature for chemotherapy response

We next investigated the association between Rad-score and adjuvant chemotherapy in stage II–III patients. Firstly, patients



C

Forest plot of effect of chemotherapy vs. no chemotherapy on OS and DFS among patients with stage II and stage III disease after PSM in internal validation cohort

Subgroups	Chemotherapy (Yes/No)	Outcomes	Median time (months) (CT/No CT)	HR (95% CI)	P value
High Rad-score	71/71	OS	37 (19-55) / 21 (15-26)	0.540 (0.355-0.821)	0.002
Low Rad-score	55/55	OS	62 (53-71) / 65 (57-73)[a]	1.190 (0.644-2.197)	0.314
High Rad-score	71/71	DFS	24 (13-35) / 11 (7-15)	0.657 (0.440-0.980)	0.039
Low Rad-score	55/55	DFS	52 (43-62) / 61 (53-70)[a]	1.560 (0.895-2.719)	0.115

D

Forest plot of effect of chemotherapy vs. no chemotherapy on OS and DFS among patients with stage II and stage III disease after PSM in SYSUCC validation cohort

Subgroups	Chemotherapy (Yes/No)	Outcomes	Median time (months) (CT/No CT)	HR (95% CI)	P value
High Rad-score	134/134	OS	67 (48-86) / 35 (25-44)	0.528 (0.376-0.741)	<0.001
Low Rad-score	117/117	OS	64 (59-71) / 65 (60-71)[a]	1.042 (0.679-1.589)	0.823
High Rad-score	134/134	DFS	45 (38-52) / 25 (16-33)	0.644 (0.472-0.878)	0.004
Low Rad-score	117/117	DFS	62 (56-68) / 64 (58-69)[a]	1.087 (0.721-1.639)	0.684

Figure 7. Chemotherapy benefits in stage II–III gastric cancer compared using disease-free survival (DFS) and overall survival (OS). Kaplan–Meier survival curves for patients with gastric cancer in different Rad-score subgroups, which were stratified by the receipt of chemotherapy. Internal validation cohort (A) and SYSUCC validation cohort (B). Forest plot of effect of chemotherapy versus no chemotherapy on OS and DFS among patients with stage II and stage III disease after propensity score matching (PSM) in the internal validation cohort (C) and SYSUCC validation cohort (D). [a]Indicates the mean time (95% CI). CT, chemotherapy.

were divided into high or low Rad-score groups, as mentioned above. After a 1:1 patient matching, clinical characteristics of patients with or without chemotherapy in each group were similar. We found that postoperative chemotherapy significantly increased OS and DFS in patients with high Rad-score in the internal validation cohort [OS:(HR: 0.540, $P = 0.002$); DFS:(HR: 0.657, $P = 0.039$)] and SYSUCC validation cohort [OS:(HR: 0.528, $P < 0.001$); DFS:(HR: 0.644, $P = 0.004$)] (Fig. 7). By contrast, chemotherapy had no impact on survival for patients with low Rad-score. A formal statistical interaction test was then performed between the Rad-score and chemotherapy, which confirmed a significant interaction regarding the impact on DFS and OS. We also conducted the above analyses using all stage II–III patients

without matching and obtained similar results (Fig. S14, Supplemental Digital Content 1, http://links.lww.com/JS9/A669). However, the predictive ability of radiomics signature for chemotherapy response had not been observed in patients with stage I disease (Fig. S15, Supplemental Digital Content 1, http://links.lww.com/JS9/A669).

Discussion

In this retrospective study, we developed and validated an explainable CT image-based artificial intelligence model that enabled accurate prediction of PR and prognosis after radical surgery in GC. This image model could provide assistance to clinicians in the judgment of PR. We further observed that it could identify patients who could benefit from adjuvant chemotherapy in stage II and III GC. Given the heterogeneous outcomes of patients with similar clinicopathologic characteristics, the developed image model could guide decision-making for individualized treatment and follow-up management of patients with GC after surgery.

PR is the most common pattern of relapse and indicates a dismal prognosis in patients with GC after radical gastrectomy, which accounts for 32–54% of all GC recurrence and has a median survival time of 4–6 months[40–42]. A long-running controversy is present on whether local treatments can prevent PR in patients with GC after curative-intent surgery, including intraperitoneal chemotherapy (IPC) and EIPL[8–17,43]. Masuda *et al*.[15] revealed that EIPL could prevent peritoneal metastasis and improve survival in patients with GC after potentially curative surgery. However, data from Yang *et al*.[16] suggested that EIPL had no survival benefit for patients with GC undergoing curative gastrectomy. IPC has been reported to eliminate potential free cancer cells from the peritoneal cavity[20]. Several studies had found that IPC could reduce peritoneal metastases and improve the survival outcome of patients with GC[7–14]. Whereas, not all GC patients develop PR after surgery and a considerable number of patients are ineligible for IPC therapy[11–13]. Additionally, IPC is an invasive procedure that is costly and has the potential to exacerbate postoperative complications[9–11]. Thus, the key question is how to identify patients at a high risk of PR who could potentially benefit from intensified treatments.

To our knowledge, this is the first study using radiomics analysis of CT images to predict PR in patients with GC after gastrectomy. Additionally, unlike other radiomics researches using a single algorithm[23,24,44], our study combined SVM-RFE, LASSO, and penalized Cox regression algorithm. This strategy could maximize the likelihood of reserving important features[45] to generate a more robust radiomics signature. And we provided the radiomics features and integrated signature with explanations on how they participated in the prediction with a unified method of SHAP. In this multicenter study, we enrolled a cohort of nearly 2000 patients and developed an artificial intelligence model to predict PR based on widely available preoperative CT images. Our image model exhibited a high accuracy in predicting PR for patients with GC after gastrectomy. Furthermore, the diagnostic accuracy of PR with the model assistance was improved by 10.13–18.86% for clinicians with 5–15 years of clinical experience. Therefore, this model could reassure clinicians to select the appropriate patients for intensified treatments.

Beyond predicting PR, our radiomics signature demonstrated a good predictive ability for DFS and OS in GC after curative-intent surgery. This signature predicted survival outcomes independently of clinicopathological characteristics. And the signature could stratify patients with the same clinical characteristics into subgroups with different survival risks. Additionally, the discriminating performance was improved when the signature was integrated into the combined model along with the clinicopathological factors. Further, our image model can identify patients who can benefit from postoperative adjuvant chemotherapy in stage II and III GC. Patients with a signature indicating high risk of PR could gain significant survival benefit from adjuvant chemotherapy, while patients with a signature indicating low risk of PR gained no benefit. Using our image model, low-risk patients can be protected from the side effects of adjuvant treatments, and conversely, high-risk patients would receive aggressive regimens and frequent surveillance to prevent cancer recurrences.

However, the present study still has several limitations. First, the primary point is its retrospective nature. Although the temporal sampling strategy and a cohort with prospective CT data are adopted to mimic a prospective situation, a prospective trial is a future requirement. Second, the model was developed and validated using a multi-cohort, most of whom were Chinese, necessitating validation with other diverse populations and ethnic groups. Third, CT images were obtained from several scanners with varied acquisition protocols, while this could conversely improve the possibility of reproducibility in multi-institution.

In conclusion, we developed a noninvasive artificial intelligence model based on intratumoral and peritumoral features from preoperative CT images, which exhibited a high accuracy for predicting PR of GC after curative-intent surgery. Furthermore, it could predict prognosis and identify patients who can benefit from postoperative adjuvant chemotherapy.

Ethical approval

Five independent cohorts of 2005 patients diagnosed with stage I–III GC were enrolled in this study, including the training cohort (433 patients), the internal validation cohort (471 patients), the SYSUCC validation cohort (922 patients), a prospective validation cohort (136 patients), and the TCIA cohort (43 patients). This study was retrospectively reviewed and approved by the Ethics Committee of Nanfang Hospital of Southern Medical University and Sun Yat-sen University Cancer Center. The approval number was NFEC-2017–171. All procedures involving human participants were in accordance with the Declaration of Helsinki.

Sources of funding

This work was supported by grants from the National Natural Science Foundation of China (82102156), National Natural Science Foundation of China (81872013), Guangdong Provincial Key Laboratory of Precision Medicine for Gastrointestinal Cancer (2020B121201004), the Guangdong Provincial Major Talents Project (No. 2019JC05Y361), Natural Science Foundation of Guangdong Province (2019A1515011445) and the Key-Area Research and Development Program of Guangdong Province (grant number: 2021B0101420005).

Author contribution

Z.S., Y.J., and G.L.: conceptualization; Z.H. and Z.S.: data curation; H.C. and J.Y.: formal analysis; H.F. and H.L.: investigation; Y.X., C.C., and Q.Y.: methodology; L.Y., K.Z., and X.L.: software; Y.J. and G.L.: project administration; Y.H.: supervision; Z.S., W.W., W.H., and T.Z.: writing – original draft; Z.S., Y.C., W.H., and Y.J.: writing – review and editing; Y.J. and G.L.: funding acquisition.

Conflicts of interest disclosure

The authors declare that there are no conflicts of interest.

Research registration unique identifying number (UIN)

1. 1 Name of the registry: Medical Ethics Committee of Nanfang Hospital.
2. 2 Unique identifying number or registration ID: NFEC-2017–171 or GIPMCS-1701.
3. 3 Hyperlink to your specific registration: Medical Ethics Committee of Nanfang Hospital [Radiomics for Prediction of Lymph Node Metastasis in Gastric Cancer (RPLNM) (GIPMCS-1701) – Full Text View – ClinicalTrials.gov].

Guarantor

All the authors took responsibility of the final manuscript and approved it for publication.

Data availability statement

All data generated for this study are included in the article or Supplementary Material (Supplemental Digital Content 1, http://links.lww.com/JS9/A669) and from the corresponding author upon reasonable request.

Acknowledgements

None.

References

[1] Yoo CH, Noh SH, Shin DW, *et al*. Recurrence following curative resection for gastric carcinoma. Br J Surg 2000;87:236–42.
[2] Sasako M, Sano T, Yamamoto S, *et al*. D2 lymphadenectomy alone or with para-aortic nodal dissection for gastric cancer. N Engl J Med 2008;359:453–62.
[3] Isobe Y, Nashimoto A, Akazawa K, *et al*. Gastric cancer treatment in Japan: 2008 annual report of the JGCA nationwide registry. Gastric Cancer 2011;14:301–16.
[4] Lei Z, Wang J, Li Z, *et al*. Hyperthermic intraperitoneal chemotherapy for gastric cancer with peritoneal metastasis: a multicenter propensity score-matched cohort study. Chin J Cancer Res 2020;32:794–803.
[5] D'Angelica M, Gonen M, Brennan MF, *et al*. Patterns of initial recurrence in completely resected gastric adenocarcinoma. Ann Surg 2004;240:808–16.
[6] Kikuchi H, Kamiya K, Hiramatsu Y, *et al*. Laparoscopic narrow-band imaging for the diagnosis of peritoneal metastasis in gastric cancer. Ann Surg Oncol 2014;21:3954–62.
[7] Chen Y, Zhou Q, Wang H, *et al*. Predicting peritoneal dissemination of gastric cancer in the era of precision medicine: molecular characterization and biomarkers. Cancers (Basel) 2020;12:2236.
[8] Bonnot PE, Piessen G, Kepenekian V, *et al*. Cytoreductive surgery with or without hyperthermic intraperitoneal chemotherapy for gastric cancer with peritoneal metastases (CYTO-CHIP study): a propensity score analysis. J Clin Oncol 2019;37:2028–40.
[9] Kim DW, Jee YS, Kim CH, *et al*. Multicenter retrospective analysis of intraperitoneal paclitaxel and systemic chemotherapy for advanced gastric cancer with peritoneal metastasis. J Gastric Cancer 2020;20:50–9.
[10] Shi B, Lin H, Zhang M, *et al*. Gene regulation and targeted therapy in gastric cancer peritoneal metastasis: radiological findings from dual energy CT and PET/CT. J Vis Exp 2018;131:e56526.
[11] Huang O, Lu X, Xu X, *et al*. Fibrin-sealant-delivered cisplatin chemotherapy versus cisplatin hyperthermic intraperitoneal perfusion chemotherapy for locally advanced gastric cancer without peritoneal metastases: a randomized phase-II clinical trial with a 40-month follow-up. Cell Biochem Biophys 2015;71:1171–80.
[12] Zhu L, Xu Z, Wu Y, *et al*. Prophylactic chemotherapeutic hyperthermic intraperitoneal perfusion reduces peritoneal metastasis in gastric cancer: a retrospective clinical study. BMC Cancer 2020;20:827.
[13] Beeharry MK, Zhu ZL, Liu WT, *et al*. Prophylactic HIPEC with radical D2 gastrectomy improves survival and peritoneal recurrence rates for locally advanced gastric cancer: personal experience from a randomized case control study. BMC Cancer 2019;19:932.
[14] Eveno C, Pocard M. Randomized controlled trials evaluating cytoreductive surgery (CRS) and hyperthermic intraperitoneal chemotherapy (HIPEC) in prevention and therapy of peritoneal metastasis: a systematic review. Pleura Peritoneum 2016;1:169–82.
[15] Masuda T, Kuramoto M, Shimada S, *et al*. The effect of extensive intraoperative peritoneal lavage therapy (EIPL) on stage III B + C and cytology-positive gastric cancer patients. Int J Clin Oncol 2016;21:289–94.
[16] Yang HK, Ji J, Han SU, *et al*. Extensive peritoneal lavage with saline after curative gastrectomy for gastric cancer (EXPEL): a multicentre randomised controlled trial. Lancet Gastroenterol Hepatol 2021;6:120–7.
[17] Chia DKA, So JBY. Recent advances in intra-peritoneal chemotherapy for gastric cancer. J Gastric Cancer 2020;20:115–26.
[18] Hasbahceci M, Akcakaya A, Guler B, *et al*. Use of peritoneal washing cytology for the detection of free peritoneal cancer cells before and after surgical treatment of gastric adenocarcinoma. J Cancer Res Ther 2018;14:1225–9.
[19] Allen CJ, Newhook TE, Vreeland TJ, *et al*. Yield of peritoneal cytology in staging patients with gastric and gastroesophageal cancer. J Surg Oncol 2019;120:1350–7.
[20] Kolomanska MM, Gluszek S. Free cancer cells in gastric cancer - methods of detection, clinical and prognostic importance (meta-analysis). Contemp Oncol (Pozn) 2020;24:67–74.
[21] Wang Z, Chen JQ. Imaging in assessing hepatic and peritoneal metastases of gastric cancer: a systematic review. BMC Gastroenterol 2011;11:19.
[22] Jiang Y, Liang X, Wang W, *et al*. Noninvasive prediction of occult peritoneal metastasis in gastric cancer using deep learning. JAMA Netw Open 2021;4:e2032269.
[23] Sun Z, Jiang Y, Chen C, *et al*. Radiomics signature based on computed tomography images for the preoperative prediction of lymph node metastasis at individual stations in gastric cancer: a multicenter study. Radiother Oncol 2021;165:179–90.
[24] Jiang Y, Chen C, Xie J, *et al*. Radiomics signature of computed tomography imaging for prediction of survival and chemotherapeutic benefits in gastric cancer. EBioMedicine 2018;36:171–82.
[25] Mikula-Pietrasik J, Uruski P, Tykarski A, *et al*. The peritoneal "soil" for a cancerous "seed": a comprehensive review of the pathogenesis of intraperitoneal cancer metastases. Cell Mol Life Sci 2018;75:509–25.
[26] Jiang YM, Liang XK, Han Z, *et al*. Radiographical assessment of tumour stroma and treatment outcomes using deep learning: a retrospective, multicohort study. Lancet Digit Health 2021;3:E371–82.
[27] Chen Y, Sun ZP, Chen WL, *et al*. The immune subtypes and landscape of gastric cancer and to predict based on the whole-slide images using deep learning. Front Immunol 2021;12:685992.
[28] Lambin P, Leijenaar RTH, Deist TM, *et al*. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749–62.
[29] Amin MB, Greene FL, Edge SB, *et al*. The Eighth Edition AJCC Cancer Staging Manual: continuing to build a bridge from a population-based to

a more "personalized" approach to cancer staging. CA Cancer J Clin 2017;67:93–9.
[30] Jiang Y, Wang H, Wu J, *et al*. Noninvasive imaging evaluation of tumor immune microenvironment to predict outcomes in gastric cancer. Ann Oncol 2020;31:760–8.
[31] Zwanenburg A, Vallieres M, Abdalah MA, *et al*. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. Radiology 2020;295:328–8.
[32] Wang HS, Li GD, Tsai CL. Regression coefficient and autoregressive order shrinkage and selection via the lasso. J R Stat Soc B 2007;69:63–78.
[33] Choi H, Yeo D, Kwon S, *et al*. Gene selection and prediction for cancer classification using support vector machines with a reject option. Comput Stat Data Anal 2011;55:1897–908.
[34] Huang ML, Hung YH, Lee WM, *et al*. SVM-RFE based feature selection and taguchi parameters optimization for multiclass SVM classifier. ScientificWorldJournal 2014;2014:795624.
[35] Duan KB, Rajapakse JC, Wang HY, *et al*. Multiple SVM-RFE for gene selection in cancer classification with expression data. IEEE T Nanobioscience 2005;4:228–34.
[36] Camp RL, Dolled-Filhart M, Rimm DL. X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. Clin Cancer Res 2004;10:7252–9.
[37] Lundberg S, Lee SI. A Unified Approach to Interpreting Model Predictions, 2017.
[38] Lundberg S. slundberg/shap, 2020. Accessed 23 January 2020. https://github. com/ slundberg/ shap
[39] Scrucca L, Santucci A, Aversa F. Competing risk analysis using R: an easy guide for clinicians. Bone Marrow Transplant 2007;40:381–7.
[40] Chen DX, Liu ZYZ, Liu WJ, *et al*. Predicting postoperative peritoneal metastasis in gastric cancer with serosal invasion using a collagen nomogram. Nat Commun 2021;12:179.
[41] Han TS, Kong SH, Lee HJ, *et al*. Dissemination of free cancer cells from the gastric lumen and from perigastric lymphovascular pedicles during radical gastric cancer surgery. Ann Surg Oncol 2011;18:2818–25.
[42] Lee IS, Lee H, Hur H, *et al*. Transcriptomic profiling identifies a risk stratification signature for predicting peritoneal recurrence and micrometastasis in gastric cancer. Clin Cancer Res 2021;27:2292–300.
[43] Guo J, Xu AM, Sun XW, *et al*. Combined surgery and extensive intraoperative peritoneal lavage vs surgery alone for treatment of locally advanced gastric cancer: The SEIPLUS Randomized Clinical Trial. JAMA Surg 2019;154:610–6.
[44] Jiang Y, Liu W, Li T, *et al*. Prognostic and predictive value of p21-activated kinase 6 associated support vector machine classifier in gastric cancer treated by 5-fluorouracil/oxaliplatin chemotherapy. EBioMedicine 2017;22:78–88.
[45] Vidyasagar M. Identifying predictive features in drug response using machine learning: opportunities and challenges. Annu Rev Pharmacol 2015;55:15–34.