Original Article

The predictive potential of contrast-enhanced computed tomography based radiomics in the preoperative staging of cT4 gastric cancer

Bo Liu[1]^, Dengyun Zhang[1], He Wang[1], Hexiang Wang[2], Pengfei Zhang[3], Dawei Zhang[1], Qun Zhang[4], Jian Zhang[1]^

[1]Department of Gastrointestinal Surgery, the Affiliated Hospital of Qingdao University, Qingdao, China; [2]Department of Radiology, the Affiliated Hospital of Qingdao University, Qingdao, China; [3]Department of Radiology, the Weihai Wendeng District People's Hospital, Weihai, China; [4]Multidisciplinary Centre, the Institute of High Energy Physics of the Chinese Academy of Sciences, Beijing, China

Contributions: (I) Conception and design: B Liu, J Zhang; (II) Administrative support: J Zhang; (III) Provision of study materials or patients: B Liu, He Wang, J Zhang; (IV) Collection and assembly of data: B Liu, D Zhang, He Wang, D Zhang, P Zhang; (V) Data analysis and interpretation: B Liu, De Zhang, Hexiang Wang, Q Zhang; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Jian Zhang. Department of Gastrointestinal Surgery, the Affiliated Hospital of Qingdao University, No. 91 Shanghai Road, Pingdu District, Qingdao 266700, Shandong, China. Email: qdfypd@126.com; Zhangjian@qduhospital.cn.

Background: The accuracy of preoperative staging is crucial for cT4 stage gastric cancer patients. The aim of this study was to develop the radiomics model and evaluate its predictive potential for differentiating preoperative cT4 stage gastric cancer patients into pT4b and no-pT4b patients.

Methods: A multicenter retrospective analysis of 704 gastric cancer patients with preoperative contrast-enhanced computed tomography (CE-CT) staging cT4 between January 2008 and December 2021. These patients were divided into the training cohort (478 patients, the Affiliated Hospital of Qingdao University) and validation cohort (226 patients, the Weihai Wendeng District People's Hospital). According to the pathological stage of the tumors, the patients were divided into pT4b or no-pT4b stage. In the training cohort, the clinical and radiomics features were analyzed to construct the clinical model, tri-phase radiomics signatures and nomogram. Two kinds of methods were employed to achieve dimensionality reduction: (I) the least absolute shrinkage and selection operator (LASSO); and (II) the minimum redundancy maximum relevance (mRMR) algorithms. We utilized Logistic regression, support vector machine (SVM), Decision tree and Adaptive boosted tree (AdaBoost) algorithms as the machine learning classifiers. The nomogram was constructed on the clinical characteristics and the Rad-score. The performance of the models was evaluated by receiver operating characteristic (ROC) area under the curve (AUC), Decision Curve Analysis (DCA) curve and calibration curve.

Results: The 345 pT4b and 359 no-pT4b stage patients were included in this study. In the validation cohort, the AUC of the clinical model was 0.793 (95% CI: 0.732–0.855). The tri-phase radiomics features combined with the SVM algorithm was the best radiomics signature with an AUC of 0.862 (95% CI: 0.812–0.912). The nomogram was the best predictive model of all with an AUC of 0.893 (95% CI: 0.834–0.927). In the training and validation cohorts, the calibration curves and DCA curves of the nomogram showed satisfactory result.

Conclusions: CE-CT-based radiomics nomogram offers good accuracy and stability in differentiating preoperative cT4 stage gastric cancer patients into pT4b and non-pT4b stages, which has a great clinical relevance for selecting the course of treatment for cT4 stage gastric cancer patients.

Keywords: Gastric cancer; contrast-enhanced computed tomography (CE-CT); radiomics; T staging

^ ORCID: Bo Liu, 0000-0003-2998-873X; Jian Zhang, 0000-0001-6278-2384.

Submitted Mar 28, 2022. Accepted for publication Jul 29, 2022.
doi: 10.21037/qims-22-286
View this article at: https://dx.doi.org/10.21037/qims-22-286

Introduction

The eighth edition of tumor node metastasis (TNM) classification is currently the most authoritative and generalized staging system for gastric cancer (1,2). According to this staging system, T4b stage gastric cancer is defined as the tumor infiltrating the serosa and invading adjacent structures and organs; accounting for 6–27% of all gastric cancers. Such patients are known to have an unsatisfactory response to therapeutics and a poor prognosis (1,3). Pathological T4b (pT4b) stage gastric cancer patients often require extended radical gastrectomy or multi-organ resection to achieve R0 resection (1,4). Despite undergoing radical resection, the 5-year overall survival rate for these patients is only 10–32% (4). The eighth edition of American Joint Committee on Cancer (AJCC) and Chinese Society of Clinical Oncology (CSCO) gastric cancer guidelines (1,2) recommend the choice of surgery or adjuvant therapy based on tumor invasion and difficulty of surgery, among which neoadjuvant therapy can effectively achieve tumor downstaging and improve the chance of radical surgery. In clinical practice, some patients with preoperative cT4 stage received surgery due to the tumor didn't invade important blood vessels, distant metastasis and other unresectable factors (2,5). Because of the overestimation of tumor staging or the inability to precisely identify the intraoperative pathological properties of adhesions present at multiple locations to perform extended resection directly, only 42–66% of cT4b patients have the same pathological stage (6,7). This suggests that better diagnostic methods are needed for the accurate identification of T4b stage. Therefore, in order to enable an accurate and personalized diagnosis and treatment, the precise staging of cT4 gastric cancer is very critical (5,8).

Contrast-enhanced computed tomography (CE-CT) imaging is widely used in the diagnosis and treatment of gastric cancer and evaluates the depth of tumor invasion (9,10). The eighth edition of the TNM staging system for gastric cancer also recommends CE-CT as one of the first-line imaging examination (1). At present, the application of imaging examinations relies on the subjective visual assessment by radiologists, which is mostly limited to the morphological characteristics of the tumors. Emerging radiomics has improved the predictive accuracy of oncological diagnosis and treatment by converting images into massive data and using non-invasive analysis and data mining (11-13). The radiomics research in the T staging of gastric cancer mainly focuses on the application of CT and other imaging technologies to extract conventional or deep learning radiomics features and combine this with different algorithms to model and distinguish between T1, T2, T3 and T4a stages (11,12,14-17). Currently, there are no studies reporting on radiomics methods for the identification of pT4b stage gastric cancer. The purpose of this multicenter study was to evaluate the predictive potential of radiomics models based on CE-CT images to identify pT4b stage patients among cT4 stage gastric cancer patients. We present the following article in accordance with the TRIPOD reporting checklist (available at https://qims.amegroups.com/article/view/10.21037/qims-22-286/rc).

Methods

Patients

This study retrospectively analyzed the clinical and imaging data from 704 gastric cancer patients of dual center that underwent gastrectomy or submitted specimens to obtain the precise pathological stage and were staged as cT4 by preoperative CE-CT imaging, between January 2008 to December 2022. These patients were divided into the training cohort (478 patients, the Affiliated Hospital of Qingdao University), and validation cohort (226 patients, the Weihai Wendeng District People's Hospital). Following were the patient inclusion criteria: (I) diagnosed with gastric cancer by preoperative endoscopic biopsy; (II) the CE-CT examination was completed within 15 days before the surgery, and the staging was cT4NxM0; (III) the clinical, pathological and imaging data of the patient were complete; (IV) there was no concurrent tumor in the gastric cancer lesion. Following were the patient exclusion criteria: (I) the quality of CE-CT images was poor; (II) there was an obvious distant metastasis or peritoneal spread; (III) anti-tumor therapy was performed before surgery, history of gastric surgery; (IV) the tumor area could not be identified and segmented by CE-CT imaging. The flowchart

displaying the details regarding patient enrollment in this study is shown in *Figure 1A*. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Institution's Ethics Committee of the Affiliated Hospital of Qingdao University and the Weihai Wendeng District People's Hospital, and individual consent for this retrospective analysis was waived.

Clinical characteristics and pathological diagnosis

We collected the clinical and pathological data, including gender, age, Borrmann type, histological differentiated type, tumor location, carbohydrate Antigen 724 (CA724) status, carbohydrate Antigen 199 (CA199) status, carcinoembryonic antigen (CEA) status, pT stage, cT stage, and clinical N stage (cN stage).

The diagnosis of pT staging and cT staging was performed according to the AJCC TNM staging criteria (1). The pathological sections of the specimens were analyzed, and for the pathological specimens marked as biopsy of the invasion site, the operation records were verified. Patients were divided into pT4b stage and no-pT4b stage using pathological results as the gold standard. CE-CT images were analyzed in a double-blind manner by two senior radiologists A and B, with 10 and 17 years of experience, respectively. The patients were then divided into cT4a stage and cT4b stage, and consensus was reached through negotiation when there was any disagreement. The analysis was performed in a double-blind fashion, wherein the radiologists only knew of the suspicious gastric cancer patients.

Image acquisition and lesion segmentation

The 64-slice CT machine (IQon Spectral, Philips or Discovery CT750, GE Medical Systems or other CT scanners in the dual center) was utilized for scanning. All the patients were placed in the supine position. The scanning range was from the top of the diaphragm to the plane of the anterior superior iliac spine. The following conditions were used for CT scanning: (I) tube current: 200 to 260 mA; (II) tube voltage: 120 kV; and (III) matrix: 512×512. 1.5 mL/kg of iopromide was injected into the patient's antecubital vein at a flow rate of 3.0 mL/s, and arterial phase (A-phase), venous phase (V-phase) and delayed phase (D-phase) images were obtained at the delays of 30, 70, and 110 s.

The ITK-SNAP (http://www.itksnap.org) software was used to segment the whole tumor regions in all the slices to generate three-dimensional volume of interest (3D VOI). Tri-phase (A-phase, V-phase and D-phase) CE-CT images were used to segment the whole tumors. Large vessels, air cavities, and adjacent fluid were excluded. Two senior radiologists C and D, with 8 and 13 years of experience, respectively, performed VOI segmentation and radiomics features extraction for each of the phases. The radiologists only knew that the patient was diagnosed with gastric cancer. The radiologist C delineated the lesions to generate tri-phase of VOI and extracted radiomics features. One week later, 30 cases were randomly selected, and radiologist C performed the delineation and features extraction for these 30 cases again. Radiologist D also performed delineation and features extraction for these 30 cases. Intra-observer correlation coefficient (ICC) was calculated using features extracted from the two VOIs delineated by the radiologist C, inter-ICC was calculated using features extracted from the VOIs delineated by the radiologists C and D. Radiomics features with inter- and intra-observer ICCs >0.75 were used for further analysis. Lesion segmentation and the radiomics study workflow are shown in *Figure 1B*.

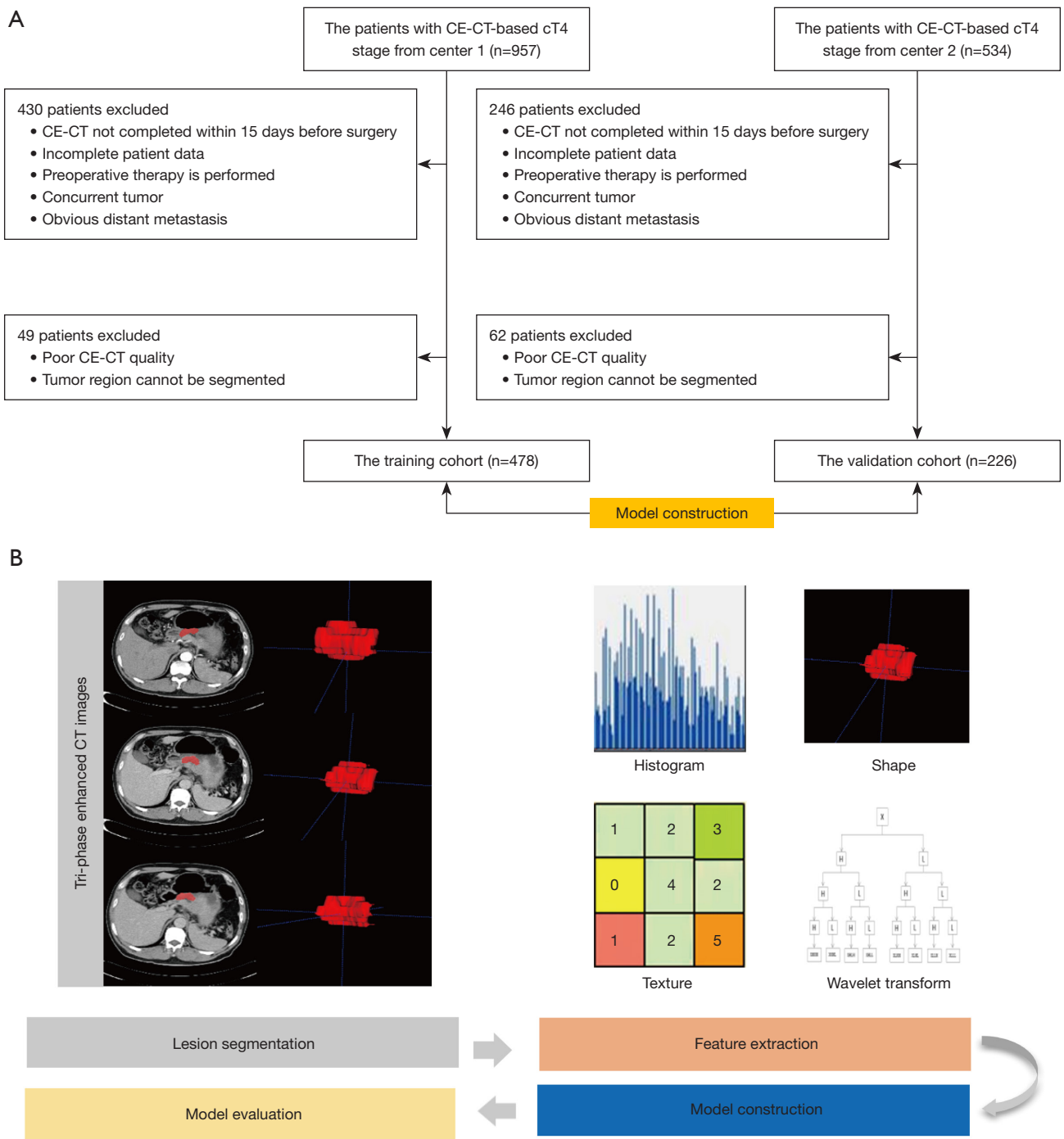
Data preprocessing and feature extraction

Image preprocessing is required before feature extraction. Firstly, the VOI was normalized by $\mu \pm 3\sigma$ to highlight the differences between the categories (18). Secondly, in order to save computation time and improve the signal-to-noise ratio for texture analysis, gray-level quantization of the images was performed (19). The VOI was resampled isotopically to in-plane resolution (0.5 mm * 0.5 mm) using cubic interpolation to ensure that the acquired 3D feature scale and orientation remained the same (20). A commercial software (AnalysisKit v.: 1.0.3; GE Healthcare, China) was utilized to extract the radiomics features. Finally, 4 types and 396 radiomics features were extracted (21-23): including, 335 GLRLM features, 10 morphological features, 42 histogram, and 10 Haralick features. The Combat compensation method was used to reduce the influence of different CT scanners and protocols, and the uniqueness of texture features was preserved (24). Next, the radiomics features pre-processing was performed, and the missing values were replaced by the median and normalized by z-score. By extracting the VOIs of the three-phase CE-CT, the radiomics feature sets of A-phase, V-phase and D-phase were generated respectively, and the above three sets of radiomics features were combined to form a total tri-phase radiomics features set, consisting of a total of 4 sets of

radiomics features.

Model construction

R 4.0.0 (https://www.Rproject.org) was used to select radiomics features and utilized machine learning classifiers to build models. The top 20 features with the best performance were selected by the minimum redundancy maximum relevance (mRMR) algorithm for the above 4 sets of radiomics features, and then the least absolute shrinkage and selection operator (LASSO) algorithm was used to reduce the dimension again. We obtained 4 radiomics



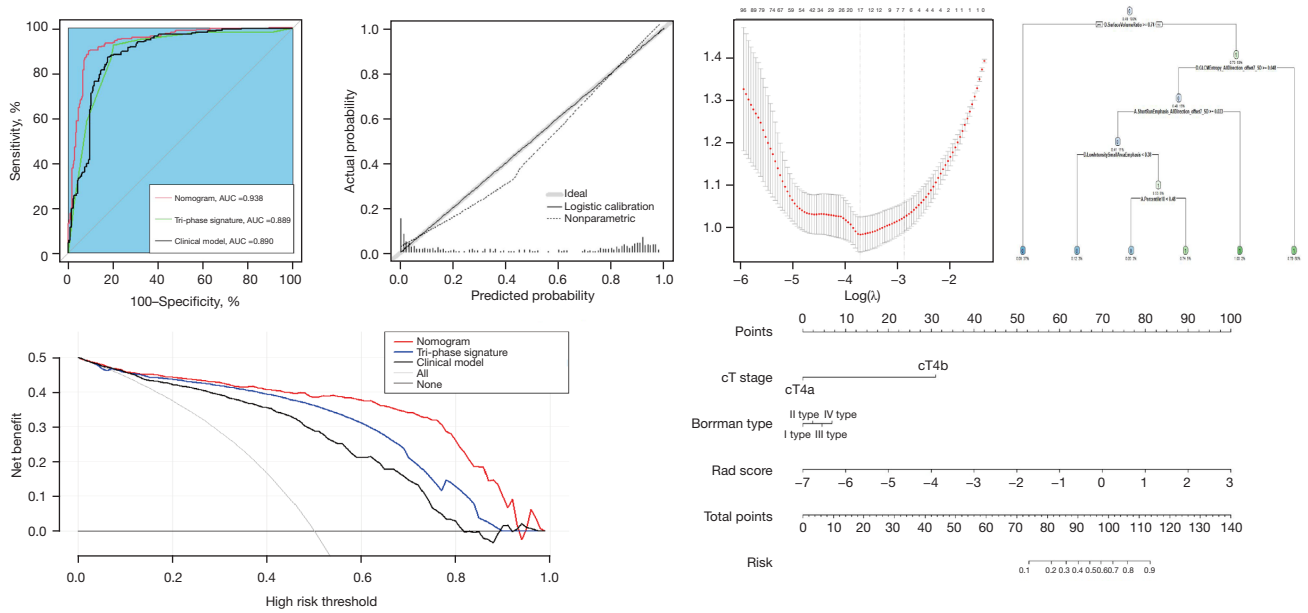


Figure 1 Data sources and research process. (A) Patient enrollment in this study, these patients were divided into the training cohort (center 1, the Affiliated Hospital of Qingdao University) and validation cohort (center 2, the Weihai Wendeng District People's Hospital). (B) The framework for this research includes lesion segmentation, feature extraction, model construction and model evaluation. CE-CT, contrast-enhanced computed tomography; AUC, area under the curve.

feature subsets of A-phase, V-phase, D-phase and tri-phase, and drew the heat map of the subset features, and obtained the radiomics score (Rad-score) for each of the subsets, according to the linear combination of the corresponding LASSO weight coefficients. The distribution of Rad-score in pT4b and no-pT4b patients is represented by a box plot. In this study, the following four machine learning classifier algorithms were used to construct the radiomics signatures: Logistic Regression algorithm; Support Vector Machine (SVM) algorithm; AdaBoost algorithm; and Decision Tree algorithm. Each of the classifier combined various radiomics feature subsets to form various radiomics signatures. Each radiomics signature generated evaluation indicators such as area under the ROC curve (AUC) in the training and validation cohorts. In the validation cohort, the machine learning classifier with the highest predictive performance was considered as the most suitable classifier algorithm.

We used univariate and multivariate logistic regression to analyze the clinical characteristics associated with pT4b, and generate a clinical model. The nomogram was constructed from the clinical characteristics and the Rad-score of the best-performing radiomics features subset by the multivariate logistic regression analysis.

Model evaluation

The AUC, precision, specificity, sensitivity, negative predictive value (NPV) and positive predictive value (PPV) of each model in the training and validation cohort were used to evaluate the performance of the models. Decision Curve Analysis (DCA) was used to evaluate the clinical net benefit of the model. Calibration curves and Hosmer-Lemeshow tests were used to evaluate the agreement between the actual and the predicted results. DeLong's test was used to compare the differences in AUC between the models. P value <0.05 was considered as statistically significant.

Statistical analysis

R software (https://www.Rproject.org, version: 4.0.0) was used for all the statistical analyses. Fisher's test and χ^2 test were applied to compare the categorical variables between different groups.

Results

Clinical characteristics

A total of 704 patients were included in this dual center study. According to preoperative CE-CT evaluation, 314 patients were cT4a stage and 390 patients were cT4b stage. According to the pathological results, 345 patients were classified as pT4b stage and 359 patients were classified as no-pT4b stage. The clinic-pathological characteristics of the patients are shown in *Table 1*. There were no statistical differences in the clinical characteristics between the training and validation cohorts (P values >0.05). According to univariate and multivariate logistic regression analysis, cT stage and Borrmann type were predictors of pT4b stage patients with statistical difference (P<0.05), as shown in *Table 2*.

Radiomics feature selection

Table 3 shows the data for the training cohort, including the radiomics feature subset of A-phase, V-phase, D-phase and tri-phase and their associated LASSO coefficients. The correlation between radiomics features and pathological staging is presented in the heatmaps, as shown in *Figure 2*. It can be seen that there were many sophisticated correlations between radiomics features and the pathological staging.

The predictive performance of the models

The machine learning classifier algorithms were crossed with each subset of radiomics features to construct radiomics signatures, as shown in *Table 4*. In the validation cohort, SVM algorithm combined with tri-phase radiomics features achieved the highest AUC. Therefore, the SVM algorithm was selected as the most suitable machine learning classifier. The tri-phase radiomics feature subset was considered as the best predictive feature. The predictive performance of the SVM algorithm, tri-phase radiomics features and the cross-constructed models is shown in *Table 5*. The SVM algorithm combined with tri-phase radiomics features constituted the best radiomics signature (referred to as the tri-phase signature), with AUCs of 0.889 (95% CI: 0.858–0.920) and 0.862 (95% CI: 0.812–0.912) in the training and validation cohorts, respectively. *Figure 3* shows the differences in the Rad-score distribution of patients with pT4b and no-pT4b stages. Patients with pT4b stage had a higher Rad-score than those with no-pT4b stage, and the difference was statistically significant (P values <0.05).

Based on multivariate logistic regression analysis, the Borrmann type and cT stage were used to construct a clinical model, and then the Borrmann type, cT stage and tri-phase Rad-score were used to construct a nomogram, as shown in *Figure 4A*. The predictive performance of the clinical model and nomogram is shown in *Table 6*. The AUC of the clinical model was 0.890 (95% CI: 0.860–0.921) and 0.793 (95% CI: 0.732–0.855) in the training and validation cohorts, respectively. The AUC of the nomogram were 0.938 (95% CI: 0.916–0.961) and 0.893 (95% CI: 0.834–0.927) in the training and validation cohorts, respectively.

The AUC values of the nomogram, clinical model, and tri-phase signature were compared by Delong test in pairs, and the results are shown in *Table 7*. The nomogram achieved the highest AUC in both the training and validation cohorts, which was statistically different from the other models (P value <0.05). As for the clinical model and the tri-phase signature, there was no statistical difference in AUC between the two models in the training cohort (P value =0.164). In the validation cohort, the AUC of the tri-phase signature was higher than the clinical model, which was statistically different (P value =0.038). In summary, the performance of the predictive models were as follows: nomogram > tri-phase signature > clinical model. The ROC curves of the three models are shown in *Figure 4B,4C*.

As shown in *Figure 5A,5B*, the calibration curve of the nomogram showed a good calibration (Hosmer-Lemeshow test, P value =0.291 and 0.165 in the training and validation cohorts, respectively). The DCA curves for evaluating the clinical application value of the models are shown in *Figure 5C,5D*. All the three models achieved satisfactory clinical utility. In the validation cohort, when the risk thresholds were between 0–0.70, 0–0.90 and 0–0.90, using the clinical model, tri-phase signature, and nomogram to predict pT4b stage patients rendered greater net positive clinical benefit than treating all or no patients. Following was the net positive clinical benefit when the different models were compared over a wide range of risk thresholds: nomogram > tri-phase signature > clinical model, indicating that the nomogram had the best clinical utility.

Discussion

With continuous advances in the personalized diagnosis and treatment of gastric cancer, the need for accurate image-based tumor staging has significantly increased over the recent years (2). In this multicenter research, a

Table 1 Clinical characteristics of the training and validation cohorts

Clinical characteristics	Training cohort (n=478)	Validation cohort (n=226)	χ^2 value	P value
Age, n (%)				
<60 years	228 (47.70)	89 (39.38)	4.289	0.058
≥60 years	250 (52.30)	137 (60.62)		
Sex, n (%)				
Female	146 (30.50)	78 (34.51)	1.114	0.291
Male	332 (69.50)	148 (65.49)		
Tumor location, n (%)				
Fundus	80 (16.70)	32 (14.16)	3.040	0.645
Body	177 (37.00)	99 (43.81)		
Antrum	221 (46.20)	95 (42.04)		
Borrmann type, n (%)				
Type I	42 (8.80)	30 (13.27)	6.220	0.101
Type II	174 (36.40)	87 (38.50)		
Type III	171 (35.80)	63 (27.88)		
Type IV	91 (19.00)	46 (20.35)		
Histological differentiated type, n (%)				
Well differentiated	284 (59.40)	125 (55.30)	1.062	0.303
Undifferentiated	194 (40.60)	101 (44.70)		
CEA, n (%)				
≤5.00 ng/mL	393 (82.20)	180 (79.65)	0.670	0.413
>5.00 ng/mL	85 (17.80)	46 (20.35)		
CA199, n (%)				
≤37.00 U/mL	440 (92.10)	194 (85.90)	0.556	0.995
>37.00 U/mL	38 (7.90)	32 (14.10)		
CA724, n (%)				
≤5.00 U/mL	358 (74.90)	169 (74.78)	0.522	0.973
>5.00 U/mL	120 (25.10)	57 (25.22)		
cT stage, n (%)				
cT4a	213 (44.60)	101 (44.69)	0.022	0.974
cT4b	265 (55.40)	125 (55.31)		
cN stage, n (%)				
cN0	108 (22.60)	58 (25.66)	0.833	0.842
cN1	175 (36.60)	80 (35.40)		
cN2	148 (31.00)	66 (29.20)		
cN3	47 (9.80)	22 (9.73)		
pT stage, n (%)				
No-pT4b	244 (51.00)	115 (50.88)	0.002	0.968
pT4b	234 (49.00)	111 (49.12)		

CEA, carcinoembryonic antigen; CA199, carbohydrate antigen 199; CA724, carbohydrate antigen 724; cT stage, clinical T stage; cN stage, clinical T stage; pT stage, pathological T stage.

Table 2 Univariate and multivariate logistic regression analysis of clinical characteristics in the training cohort

Clinical characteristics	Univariate analysis		Multivariate analysis	
	OR (95% CI)	P value	OR (95% CI)	P value
Age				
<60 years*				
≥60 years	0.78 (0.545–1.118)	0.176		
Sex				
Female*				
Male	0.905 (0.613–1.336)	0.616		
Tumor location				
Fundus*		0.007		0.579
Body	0.563 (0.330–0.959)	0.035	0.938 (0.398–2.210)	0.883
Antrum	1.041 (0.623–1.739)	0.878	1.442 (0.630–3.299)	0.386
Borrmann type				
Type I*		0.001		0.016
Type II	0.815 (0.411–1.614)	0.557	0.298 (0.097–0.915)	0.034
Type III	1.628 (0.824–3.217)	0.161	0.646 (0.204–2.041)	0.457
Type IV	2.133(1.015–4.484)	0.046	1.116 (0.339–3.670)	0.857
Histological differentiated type				
Well differentiated*				
Undifferentiated	0.732 (0.507–1.056)	0.095		
CEA, ng/mL				
≤5.00*				
>5.00	0.861 (0.538–1.378)	0.532		
CA199, U/mL				
≤37.00*				
>37.00	0.515 (0.257–1.033)	0.062		
CA724, U/mL				
≤5.00*				
>5.00	0.925 (0.612–1.400)	0.713		
cT stage				
cT4a*				
cT4b	47.755 (26.903–84.769)	0.000	60.226 (31.961–113.486)	0.000
cN stage				
cN0*		0.294		
cN1	1.554 (0.959–2.519)	0.073		
cN2	1.511 (0.918–2.489)	0.105		
cN3	1.304 (0.657–2.591)	0.448		

*, as a reference item. OR, odds ratio; 95% CI, 95% confidence interval; CEA, carcinoembryonic antigen; CA199, carbohydrate antigen 199; CA724, carbohydrate antigen 724; cT stage, clinical T stage; cN stage, clinical N stage.

Table 3 Radiomics features and related LASSO weight coefficients associated with pT4b stage for LASSO dimensionality reduction

Radiomics features	Weight coefficients
A-phase radiomics features	
(Intercept)	6.942
A.Percentile10	0.195
A.RelativeDeviation	0.006
A.skewness	0.076
A.GLCMEntropy_AllDirection_offset1_SD	−0.066
A.GLCMEntropy_angle90_offset4	−4.224
A.HaralickCorrelation_angle0_offset1	−0.041
A.sumVariance	−0.079
A.GreyLevelNonuniformity_angle90_offset1	−0.071
A.LongRunEmphasis_angle0_offset1	0.168
A.LongRunEmphasis_angle135_offset1	−0.016
A.LongRunEmphasis_angle135_offset4	−0.064
A.LongRunEmphasis_angle90_offset4	−0.078
A.ShortRunEmphasis_AllDirection_offset7_SD	−0.116
A.Compactness2	−0.647
A.SurfaceVolumeRatio	−2.997
A.LowIntensityEmphasis	0.045
V-phase radiomics features	
(Intercept)	−7.848
V.histogramEntropy	−1.136
V.GLCMEntropy_AllDirection_offset4	−2.970
V.HaralickCorrelation_angle0_offset1	−0.112
V.ShortRunEmphasis_AllDirection_offset4_SD	−0.022
V.ShortRunEmphasis_angle0_offset1	12.485
V.SurfaceVolumeRatio	−2.350
V.LargeAreaEmphasis	1.149
V.LowIntensityLargeAreaEmphasis	0.099
Tri-phase radiomics features	
(Intercept)	5.150
A.Percentile10	0.147
A.GLCMEntropy_AllDirection_offset1_SD	−0.017
A.GLCMEntropy_angle90_offset4	−0.866
A.LongRunEmphasis_angle0_offset1	0.038
A.ShortRunEmphasis_AllDirection_offset7_SD	−0.064
A.Compactness2	−0.063
V.histogramEntropy	−0.794
V.GLCMEntropy_AllDirection_offset4	−1.527
V.GLCMEntropy_angle45_offset4	−0.532
V.HaralickCorrelation_angle0_offset1	−0.072
V.Inertia_angle45_offset7	−0.006
V.SurfaceVolumeRatio	−0.596
V.LargeAreaEmphasis	0.872
D.GLCMEntropy_AllDirection_offset7_SD	−0.166
D.GLCMEntropy_angle45_offset1	−0.712
D.SurfaceVolumeRatio	−1.780
D.LowIntensitySmallAreaEmphasis	0.208
D-phase radiomics features	
(Intercept)	4.942
D.RMS	−0.097
D.GLCMEntropy_AllDirection_offset7_SD	−0.149
D.GLCMEntropy_angle0_offset4	−1.438
D.GLCMEntropy_angle45_offset1	−1.801
D.HaralickCorrelation_angle0_offset1	−0.023
D.ShortRunEmphasis_AllDirection_offset4_SD	−0.019
D.ShortRunEmphasis_AllDirection_offset7_SD	−0.031
D.SurfaceVolumeRatio	−2.481
D.LowIntensitySmallAreaEmphasis	0.202

A-phase, arterial phase; V-phase, venous phase; D-phase, delayed phase; Tri-phase, arterial phase and venous phase and delayed phase combinative; LASSO, the least absolute shrinkage and selection operator.

nomogram combining the clinical features and Rad-score was constructed to identify pT4b stage patients among preoperative cT4b stage gastric cancer patients. The model was expected to have a high performance and clinical application potential.

Radiomics is a post-processing technology based on big data (25). It can objectively evaluate the heterogeneity in tumors by mining the massive information behind the images that cannot be detected by the naked eye. Currently,

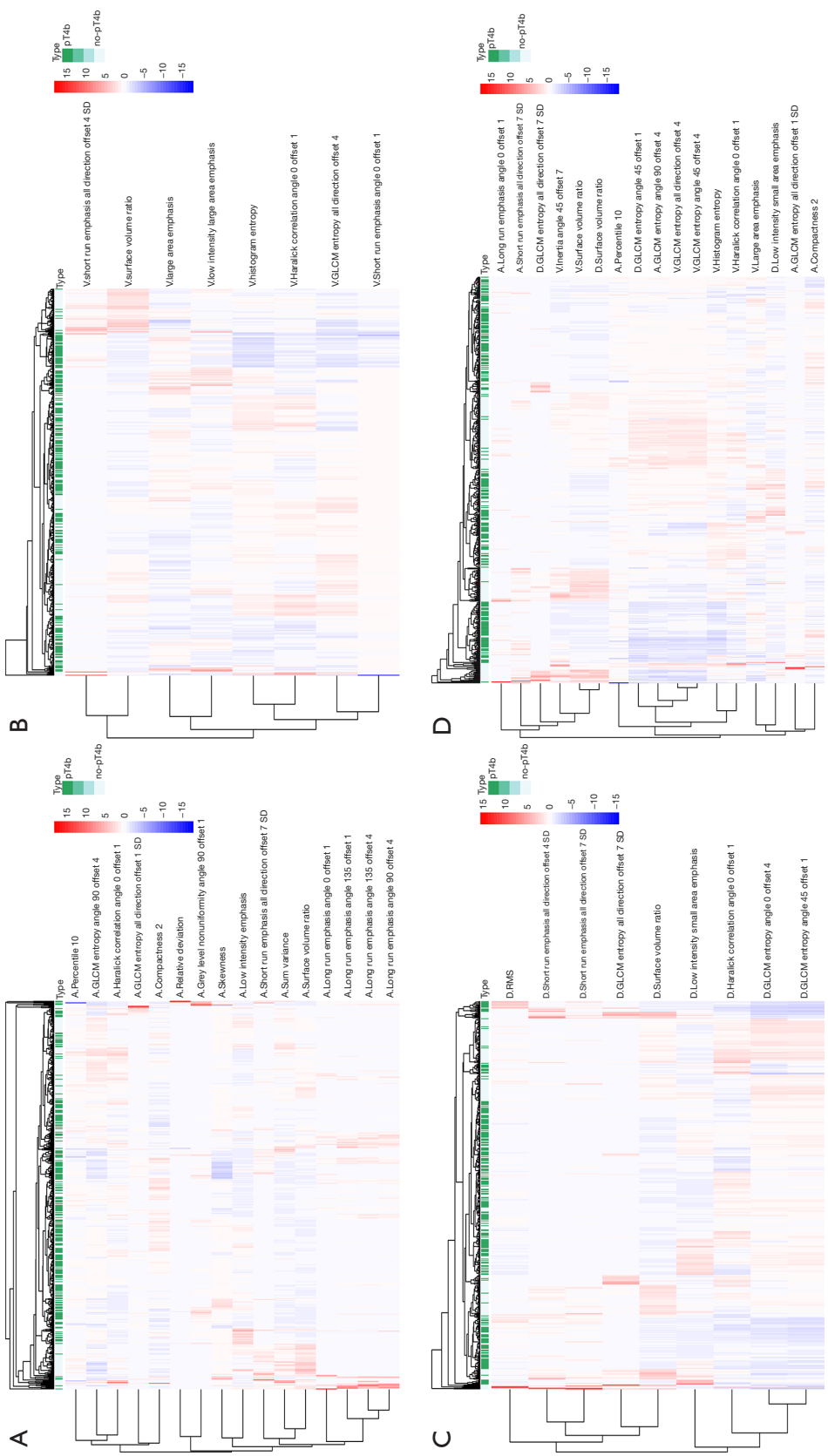


Figure 2 Radiomics feature analysis. (A) Heatmaps of A-phase radiomics features. (B) Heatmaps of V-phase radiomics features. (C) Heatmaps of D-phase radiomics features. (D) Heatmaps of tri-phase radiomics features. Heatmap depicting hierarchical clustering of Spearman's correlation coefficients between radiomics data. Red means positive correlation, blue means negative correlation. A-phase, arterial phase; V-phase, venous phase; D-phase, delayed phase.

Table 4 The AUC of radiomics signatures constructed by crossing machine learning classifier algorithms with radiomics features

Radiomics signatures' AUC	Logistic regression	Decision tree	SVM	AdaBoost
A-phase radiomics features				
Training cohort	0.867	0.851	0.947	0.898
Validation cohort	0.766	0.772	0.758	0.740
V-phase radiomics features				
Training cohort	0.755	0.806	0.866	0.811
Validation cohort	0.700	0.708	0.754	0.754
D-phase radiomics features				
Training cohort	0.862	0.851	0.898	0.882
Validation cohort	0.756	0.725	0.767	0.729
Tri-phase radiomics features				
Training cohort	0.859	0.849	0.889	0.883
Validation cohort	0.771	0.733	0.862	0.733

AUC, area under the curve; SVM, Support Vector Machine.

Table 5 Predictive performance of the radiomics signatures constructed by crossing the SVM classifier algorithm and the tri-phase radiomics features

Model's performance	AUC	95% CI	Accuracy	Sensitivity	Specificity	PPV	NPV
Logistic regression (tri-phase features)							
Training cohort	0.859	0.825–0.893	0.797	0.859	0.738	0.758	0.845
Validation cohort	0.771	0.706–0.836	0.720	0.634	0.802	0.753	0.697
Decision tree (tri-phase features)							
Training cohort	0.849	0.815–0.884	0.814	0.791	0.836	0.822	0.806
Validation cohort	0.733	0.666–0.800	0.725	0.703	0.745	0.724	0.725
SVM (tri-phase features)							
Training cohort	0.889	0.858–0.920	0.860	0.923	0.799	0.815	0.915
Validation cohort	0.862	0.812–0.912	0.816	0.871	0.764	0.779	0.862
AdaBoost (tri-phase features)							
Training cohort	0.883	0.853–0.913	0.805	0.821	0.791	0.790	0.821
Validation cohort	0.733	0.668–0.798	0.681	0.663	0.698	0.677	0.685
A-phase features (SVM)							
Training cohort	0.947	0.925–0.970	0.897	0.936	0.861	0.866	0.933
Validation cohort	0.758	0.692–0.825	0.705	0.723	0.689	0.689	0.723
V-phase features (SVM)							
Training cohort	0.866	0.797–0.870	0.764	0.910	0.623	0.698	0.879
Validation cohort	0.754	0.663–0.786	0.650	0.802	0.487	0.592	0.689
D-phase features (SVM)							
Training cohort	0.898	0.869–0.927	0.831	0.949	0.717	0.763	0.936
Validation cohort	0.767	0.702–0.833	0.691	0.772	0.613	0.655	0.739
Tri-phase features (SVM)							
Training cohort	0.889	0.858–0.920	0.860	0.923	0.799	0.815	0.915
Validation cohort	0.862	0.812–0.912	0.816	0.871	0.764	0.779	0.862

SVM, Support Vector Machine; AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value.

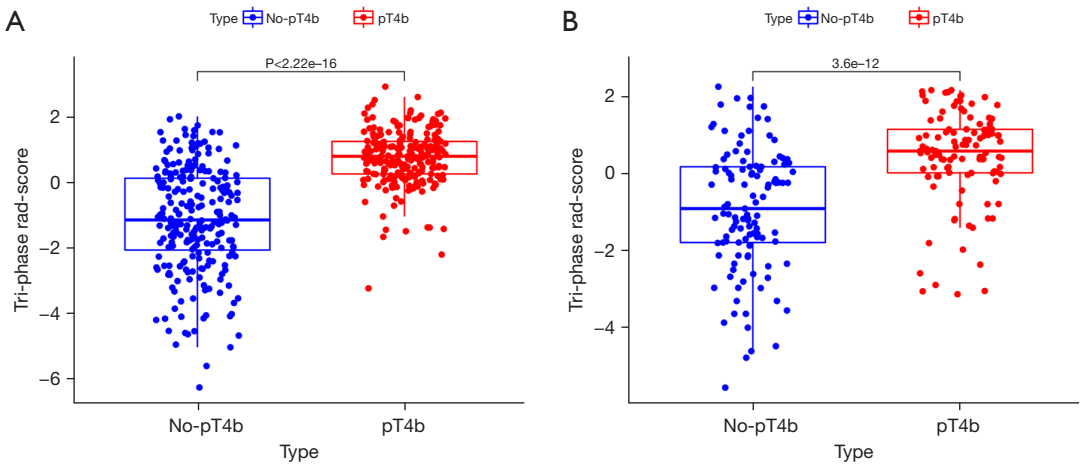


Figure 3 Rad-score construction. (A) Boxplots showing the distribution of tri-phase Rad-score over no-pT4b and pT4b in the training cohort. (B) Boxplots showing the distribution of tri-phase Rad-score over no-pT4b and pT4b in the validation cohort. No-pT4b and pT4b stage are well differentiated by tri-phase Rad-score with statistically significant (P<0.05). Rad-score, radiomics score.

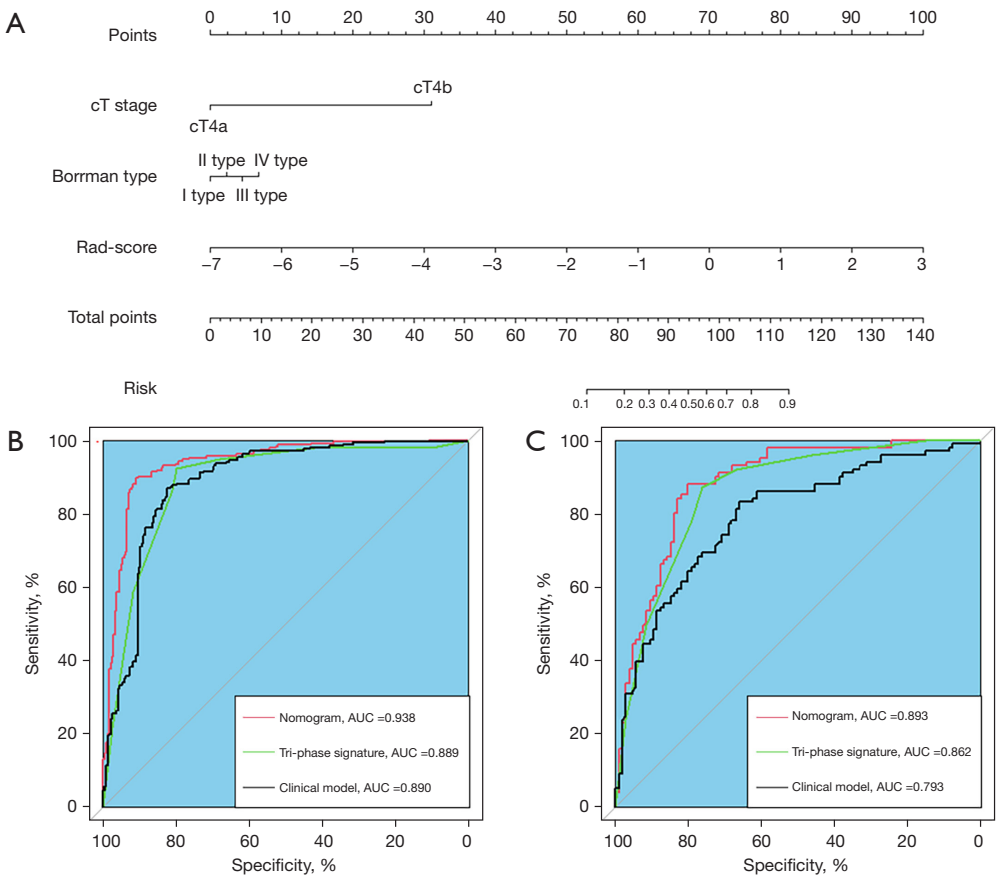


Figure 4 Model construction. (A) Nomogram constructed from the tri-phase Rad-score and clinical characteristics (Borrmann type and cT stage). (B) ROC curves of nomogram, clinical model, and tri-phase model for predicting pT4b stage in the training cohort. (C) ROC curves of nomogram, clinical model, and tri-phase model for predicting pT4b stage in the validation cohort. AUC, area under the curve; ROC, receiver operating characteristic.

Table 6 Predictive performance of the clinical model and nomogram

Model's performance	AUC	95% CI	Accuracy	Sensitivity	Specificity	PPV	NPV
Clinical model							
Training cohort	0.890	0.860–0.921	0.818	0.902	0.738	0.767	0.887
Validation cohort	0.793	0.732–0.855	0.739	0.822	0.660	0.697	0.795
Nomogram							
Training cohort	0.938	0.916–0.961	0.891	0.902	0.881	0.879	0.903
Validation cohort	0.893	0.834–0.927	0.812	0.782	0.840	0.823	0.802

AUC, area under the curve; PPV, positive predictive value; NPV, negative predictive value.

Table 7 Delong test for AUC of the clinical model, tri-phase signature and nomogram

Model comparison	AUC	P value
Training cohort		
Nomogram	0.938	0.000
Clinical model	0.890	
Nomogram	0.938	0.000
Tri-phase signature	0.889	
Clinical model	0.890	0.164
Tri-phase signature	0.889	
Validation cohort		
Nomogram	0.893	0.009
Clinical model	0.793	
Nomogram	0.893	0.000
Tri-phase signature	0.862	
Clinical model	0.793	0.038
Tri-phase signature	0.862	

AUC, area under the curve.

radiomics has been reported to play an important role in the diagnosis and treatment of cancer, as well as in predicting the prognosis of cancer patients (11,13). The current radiomics studies on gastric cancer T staging have been focused on the application of different imaging techniques, such as: CT (15), dual-energy CT (16) and CE-CT (26), etc. It involves the extraction of different types of image features such as deep Learning Features (27); texture features (16); and spleen radiomic features (15), using different computer algorithms for modeling, such as: SVM algorithm (14); Random Forest algorithm (26); and Convolutional Neural Network algorithm (27), etc. These studies have been able to predict the patients with T1, T2, T3, and T4a stage, however, currently there are no studies reporting the use of radiomics methods to identify pT4b stage of gastric cancer.

Gastric cancer growth and invasion are mainly dependent on the extent of tumor angiogenesis (10), CE-CT imaging enables the comprehensively display of tumor blood vessels and perfusion, reflecting tumor heterogeneity (28). The arterial phase reflects tumor blood supply and functional capillary density, while the venous phase reflects dysfunctional neovascularization and differences in the contrast agent distribution (28). In this study, four subsets of radiomics features were selected based on tri-phase CE-CT, and the tri-phase radiomics feature subset with the most features demonstrated better predictive performance than the rest. In addition, in this study, we segmented the whole tumors to generate 3D VOI. Previous studies have shown that 3D VOI could extract better radiomics features than 2D VOI and avoid data bias due to tumor heterogeneity (29,30). In terms of the algorithms, this study used mRMR combined with the Lasso algorithm to reduce dimensionality. Among them, the mRMR algorithm played a major role in solving the problem due to the redundancy between features: the best performance m features obtained in terms of the correlations between features and target events may not necessarily yield the best performing model. Therefore, the mRMR algorithm was utilized to ensure the maximum correlation between features, as well as minimum redundancy between them (31). Lasso algorithm reduced the number of features based on model complexity to reduce overfitting (32). Due to the longtime-span of the current study, there were some variations in the protocols and parameters between the different CT scanners. In order to correct for the differences in the

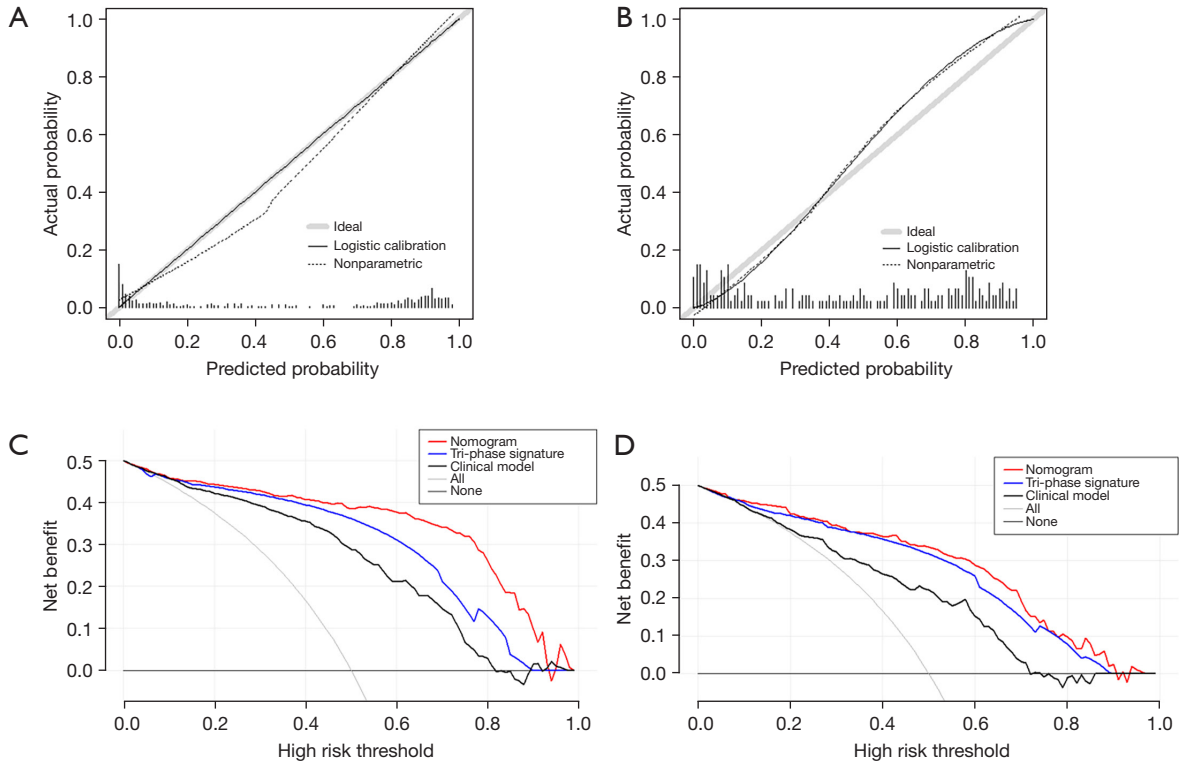


Figure 5 Model evaluation. Calibration curves of the nomogram in the training cohort (A) and validation cohort (B). The calibration curves described the calibration in agreement between predicted and observed outcome. The 45-degree reference line means a perfect calibration with the outcome by ideal model. The solid line is the performance of the nomogram. The dotted line is the bootstrap-corrected performance of the nomogram, with a scatter estimate for future accuracy. (C) Decision curve analysis of the nomogram, clinical model and tri-phase model in the training cohort. (D) Decision curve analysis of the nomogram, clinical model and tri-phase model in the validation cohort. The nomogram had the best net benefit when the threshold probability is in a large range, compared with the simple strategies of no patients (horizontal black line) or all patients (grey line) and other methods.

image quality, the radiomics features were pre-processed using the combat compensation method, which could realign radiomic features derived from different CT imaging protocols and machines without changing the meaning of the features, so that we could generate stable and reproducible data in our study (33). SVM is a stable supervised machine learning classifier that combines dimensionality enhancement with linearization. Because adding dimensions can lead to a "dimension disaster", the SVM algorithm was used to solve the problem associated with increasing the dimension, which might increase the computational complexity through the kernel function expansion theorem (34).

Logistic regression analysis of the clinical characteristics showed that gastric cancer patients with cT4b stage and Borrmann type IV were more likely to be pT4b stage patients. Borrmann type is the classic gross type of gastric cancer. Among them, type IV gastric cancer is most common to invade the serosa and the surrounding organs. Previously, when Borrmann type IV patients were treated, more than 70% of the patients belonged to the pT4a stage, and nearly 20% of the patients belonged to the pT4b stage (35). With the recent advances in CT imaging technology, the ability of CE-CT imaging to distinguish between the 5-layer structure of the gastric wall has significantly improved (36). If gastric cancer has invaded into the surrounding or distant organs, specific imaging patterns can be seen, including, low-density mass shadow in the hepatoduodenal ligament, disappearance of fat space, finger-like insertion or direct infiltration, etc. (2,37). Previous study (38) has shown that the diagnostic accuracy of T4b stage gastric cancer by CT is 75%, which is similar to the accuracy of the clinical model (73.91%) in our study.

As for the comparison of model's performance, the tri-

phase signature outperformed the clinical model in the validation cohort. Interestingly, there was no statistical difference between the two models in the training cohort. We speculate on possible reason explaining the similar performance between the models: most of the tumors in this study had already penetrated into the serosa and was often accompanied by edema and inflammatory changes around the tumor (38), which might cause inaccurate delineation of the target area and add bias to the extracted radiomic features, further affect the performance of radiomics model. The nomogram combining Rad-score and clinical characteristics obtained the best predictive ability. It showed that the combination of clinical and radiomics methods could enhance the predictive ability. With the advances in automatic delineation and artificial intelligence technologies, it is now possible to obtain more radiomics features by automatically identifying and expanding the delineation area, in order to improve the predictive ability. For example, Li *et al.* (39) extracted more than 7,000 radiomics features for modeling by semi-automatically delineating lesions and the whole stomach and integrating large-scale imaging factors. This achieved good predictive performance in gastric cancer prognosis-related studies. Therefore, more in-depth research is needed in the future to explore the application potential of radiomics in predicting T4b stage in gastric cancer.

The nomogram in this study provides a personalized reference potential for preoperative prediction of the pathological status of the tumor. The AJCC gastric cancer guidelines (1) recommend that cT4b stage patients should discuss the treatment options with the multi-disciplinary treatment (MDT), which achieved a personalized diagnosis and treatment plan, fill in the shortcomings of the disciplines. Therefore, when faced with decision-making difficulties such as deciding whether the tumor can be completely resected, the results from this study may provide some useful reference for decision-making.

However, this study has several limitations. Firstly, this is a retrospective study, causing bias in case selection, and the model generalization still needs to be tested. Secondly, this study only uses machine learning algorithms to build models and lacks advanced dimensionality reduction techniques. Thirdly, the segmentation of tumors with ill-defined boundaries has been controversial, although good ICCs were achieved with the radiomics features employed in this study. Also, differences in the measurement of segmented areas were not compared. Combining these two methods could further improve the reproducibility of the delineated regions (12).

Conclusions

In this multicenter study, the CE-CT-based models were constructed to non-invasively identify pT4b stage patients among preoperative cT4 stage gastric cancer patients. Preoperative CE-CT-based cT stage and Borrmann type were independent predictors of gastric cancer pT4b stage patients. The nomogram combined with tri-phase radiomics Rad-score and clinical characteristics achieved the best predictive performance, which could provide personalized clinical decision support for cT4 gastric cancer patients.

Acknowledgments

Funding: This work was supported by National Natural Science Foundation of China (No. 81770631).

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://qims.amegroups.com/article/view/10.21037/qims-22-286/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://qims.amegroups.com/article/view/10.21037/qims-22-286/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). This study was approved by the Institution's Ethics Committee of the Affiliated Hospital of Qingdao University and the Weihai Wendeng District People's Hospital, and individual consent for this retrospective analysis was waived.

References

1. Coupland S, Barnhill R, Conway M, Damato BE, Esmaeli B, Albert DM, Auwhädrich C, Chevezbarrios P, Grossniklaus HE, Heegaard S. The AJCC TNM Cancer Staging Manual. Eighth edition. Conjunctival Melanoma 2017.
2. Wang FH, Zhang XT, Li YF, Tang L, Qu XJ, Ying JE, et al. The Chinese Society of Clinical Oncology (CSCO): Clinical guidelines for the diagnosis and treatment of gastric cancer, 2021. Cancer Commun (Lond) 2021;41:747-95.
3. Kunisaki C, Akiyama H, Nomura M, Matsuda G, Otsuka Y, Ono HA, Nagahori Y, Takahashi M, Kito F, Shimada H. Surgical outcomes in patients with T4 gastric carcinoma. J Am Coll Surg 2006;202:223-30.
4. Carboni F, Lepiane P, Santoro R, Lorusso R, Mancini P, Sperduti I, Carlini M, Santoro E. Extended multiorgan resection for T4 gastric carcinoma: 25-year experience. J Surg Oncol 2005;90:95-100.
5. Kim HD, Lee JS, Yook JH, Ryu MH, Park YK, Kim JY, Kim YW, Oh SC, Kim JG, Cheong JH, Jeong O, Noh SH, Kang YK. Radiological criteria for selecting candidates for neoadjuvant chemotherapy for gastric cancer: an exploratory analysis from the PRODIGY study. Gastric Cancer 2022;25:170-9.
6. Mita K, Ito H, Katsube T, Tsuboi A, Yamazaki N, Asakawa H, Hayashi T, Fujino K. Prognostic Factors Affecting Survival After Multivisceral Resection in Patients with Clinical T4b Gastric Cancer. J Gastrointest Surg 2017;21:1993-9.
7. Li MZ, Deng L, Wang JJ, Xiao LB, Wu WH, Yang SB, Li WF. Surgical outcomes and prognostic factors of T4 gastric cancer patients without distant metastasis. PLoS One 2014;9:e107061.
8. Hallinan JT, Venkatesh SK. Gastric carcinoma: imaging diagnosis, staging and assessment of treatment response. Cancer Imaging 2013;13:212-27.
9. Kwee RM, Kwee TC. Imaging in local staging of gastric cancer: a systematic review. J Clin Oncol 2007;25:2107-16.
10. Watanabe H, Hayano K, Ohira G, Imanishi S, Hanaoka T, Hirata A, Kano M, Matsubara H. Quantification of Structural Heterogeneity Using Fractal Analysis of Contrast-Enhanced CT Image to Predict Survival in Gastric Cancer Patients. Dig Dis Sci 2021;66:2069-74.
11. Yip SS, Aerts HJ. Applications and limitations of radiomics. Phys Med Biol 2016;61:R150-66.
12. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016;278:563-77.
13. O'Connor JP, Aboagye EO, Adams JE, Aerts HJ, Barrington SF, Beer AJ, et al. Imaging biomarker roadmap for cancer studies. Nat Rev Clin Oncol 2017;14:169-86.
14. Sun RJ, Fang MJ, Tang L, Li XT, Lu QY, Dong D, Tian J, Sun YS. CT-based deep learning radiomics analysis for evaluation of serosa invasion in advanced gastric cancer. Eur J Radiol 2020;132:109277.
15. Pan B, Zhang W, Chen W, Zheng J, Yang X, Sun J, Sun X, Chen X, Shen X. Establishment of the Radiologic Tumor Invasion Index Based on Radiomics Splenic Features and Clinical Factors to Predict Serous Invasion of Gastric Cancer. Front Oncol 2021;11:682456.
16. Wang L, Zhang Y, Chen Y, Tan J, Wang L, Zhang J, Yang C, Ma Q, Ge Y, Xu Z, Pan Z, Du L, Yan F, Yao W, Zhang H. The Performance of a Dual-Energy CT Derived Radiomics Model in Differentiating Serosal Invasion for Advanced Gastric Cancer Patients After Neoadjuvant Chemotherapy: Iodine Map Combined With 120-kV Equivalent Mixed Images. Front Oncol 2021;10:562945.
17. Chang X, Guo X, Li X, Han X, Li X, Liu X, Ren J. Potential Value of Radiomics in the Identification of Stage T3 and T4a Esophagogastric Junction Adenocarcinoma Based on Contrast-Enhanced CT Images. Front Oncol 2021;11:627947.
18. Shafiq-Ul-Hassan M, Latifi K, Zhang G, Ullah G, Gillies R, Moros E. Voxel size and gray level normalization of CT radiomic features in lung cancer. Sci Rep 2018;8:10545.
19. Gao X, Ma T, Bai S, Liu Y, Zhang Y, Wu Y, Li H, Ye Z. A CT-based radiomics signature for evaluating tumor infiltrating Treg cells and outcome prediction of gastric cancer. Ann Transl Med 2020;8:469.
20. Depeursinge A, Foncubierta-Rodriguez A, Van De Ville D, Müller H. Three-dimensional solid texture analysis in biomedical imaging: review and opportunities. Med Image Anal 2014;18:176-96.
21. Nyflot MJ, Yang F, Byrd D, Bowen SR, Sandison GA, Kinahan PE. Quantitative radiomics: impact of stochastic effects on textural feature analysis implies the need for standards. J Med Imaging (Bellingham) 2015;2:041002.
22. Yamashiro T, Matsuoka S, Estépar RS, Bartholmai BJ, Diaz A, Ross JC, Murayama S, Silverman EK, Hatabu H, Washko GR. Kurtosis and skewness of density histograms on inspiratory and expiratory CT scans in smokers. COPD 2011;8:13-20.

23. Löfstedt T, Brynolfsson P, Asklund T, Nyholm T, Garpebring A. Gray-level invariant Haralick texture features. PLoS One 2019;14:e0212110.
24. Orlhac F, Frouin F, Nioche C, Ayache N, Buvat I. Validation of A Method to Compensate Multicenter Effects Affecting CT Radiomics. Radiology 2019;291:53-9.
25. Attanasio S, Forte SM, Restante G, Gabelloni M, Guglielmi G, Neri E. Artificial intelligence, radiomics and other horizons in body composition assessment. Quant Imaging Med Surg 2020;10:1650-60.
26. Wang Y, Liu W, Yu Y, Liu JJ, Jiang L, Xue HD, Lei J, Jin Z, Yu JC. Prediction of the Depth of Tumor Invasion in Gastric Cancer: Potential Role of CT Radiomics. Acad Radiol 2020;27:1077-84.
27. Zheng L, Zhang X, Hu J, Gao Y, Zhang X, Zhang M, Li S, Zhou X, Niu T, Lu Y, Wang D. Establishment and Applicability of a Diagnostic System for Advanced Gastric Cancer T Staging Based on a Faster Region-Based Convolutional Neural Network. Front Oncol 2020;10:1238.
28. Chen XH, Ren K, Liang P, Chai YR, Chen KS, Gao JB. Spectral computed tomography in advanced gastric cancer: Can iodine concentration non-invasively assess angiogenesis? World J Gastroenterol 2017;23:1666-75.
29. Zhang L, Kang L, Li G, Zhang X, Ren J, Shi Z, Li J, Yu S. Computed tomography-based radiomics model for discriminating the risk stratification of gastrointestinal stromal tumors. Radiol Med 2020;125:465-73.
30. Liu B, Liu H, Zhang L, Song Y, Yang S, Zheng Z, Zhao J, Hou F, Zhang J. Value of contrast-enhanced CT based radiomic machine learning algorithm in differentiating gastrointestinal stromal tumors with KIT exon 11 mutation: a two-center study. Diagn Interv Radiol 2022;28:29-38.
31. Eroğlu Y, Yildirim M, Çinar A. Convolutional Neural Networks based classification of breast ultrasonography images by hybrid method with respect to benign, malignant, and normal using mRMR. Comput Biol Med 2021;133:104407.
32. McEligot AJ, Poynor V, Sharma R, Panangadan A. Logistic LASSO Regression for Dietary Intakes and Breast Cancer. Nutrients 2020;12:2652.
33. Orlhac F, Lecler A, Savatovski J, Goya-Outi J, Nioche C, Charbonneau F, Ayache N, Frouin F, Duron L, Buvat I. How can we combat multicenter variability in MR radiomics? Validation of a correction procedure. Eur Radiol 2021;31:2272-80.
34. Uddin S, Khan A, Hossain ME, Moni MA. Comparing different supervised machine learning algorithms for disease prediction. BMC Med Inform Decis Mak 2019;19:281.
35. Yamashita K, Hosoda K, Katada N, Moriya H, Mieno H, Higuchi K, Sasaki T, Katada C, Sakuramoto S, Tanabe S, Koizumi W, Kikuchi S, Watanabe M. Survival outcome of Borrmann type IV gastric cancer potentially improved by multimodality treatment. Anticancer Res 2015;35:897-906.
36. Lee IJ, Lee JM, Kim SH, Chang S, Han JK, Choi BI, Lee HJ, Yang HK, Lee KU. Helical CT evaluation of the preoperative staging of gastric cancer in the remnant stomach. AJR Am J Roentgenol 2009;192:902-8.
37. Chen CY, Hsu JS, Wu DC, Kang WY, Hsieh JS, Jaw TS, Wu MT, Liu GC. Gastric cancer: preoperative local staging with 3D multi-detector row CT--correlation with surgical and histopathologic results. Radiology 2007;242:472-82.
38. Makino T, Fujiwara Y, Takiguchi S, Tsuboyama T, Kim T, Nushijima Y, Yamasaki M, Miyata H, Nakajima K, Mori M, Doki Y. Preoperative T staging of gastric cancer by multi-detector row computed tomography. Surgery 2011;149:672-9.
39. Li Q, Qi L, Feng QX, Liu C, Sun SW, Zhang J, Yang G, Ge YQ, Zhang YD, Liu XS. Machine Learning-Based Computational Models Derived From Large-Scale Radiographic-Radiomic Images Can Help Predict Adverse Histopathological Status of Gastric Cancer. Clin Transl Gastroenterol 2019;10:e00079.

Cite this article as: Liu B, Zhang D, Wang H, Wang H, Zhang P, Zhang D, Zhang Q, Zhang J. The predictive potential of contrast-enhanced computed tomography based radiomics in the preoperative staging of cT4 gastric cancer. Quant Imaging Med Surg 2022;12(11):5222-5238. doi:10.21037/qims-22-286

