Radiotherapy and Oncology 145 (2020) 13–20

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

Development and validation of a CT-based radiomic nomogram for preoperative prediction of early recurrence in advanced gastric cancer

Wenjuan Zhang [a,b,c,d,1], Mengjie Fang [d,e,1], Di Dong [d,e,1], Xiaoxiao Wang [f], Xiaoai Ke [a,b,c], Liwen Zhang [d,e], Chaoen Hu [d,e], Lingyun Guo [b,g], Xiaoying Guan [b,h], Junlin Zhou [a,b,c,*], Xiuhong Shan [f,*], Jie Tian [d,e,i,*]

[a] Department of Radiology, Lanzhou University Second Hospital; [b] Second Clinical School, Lanzhou University; [c] Key Laboratory of Medical Imaging of Gansu Province; [d] CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences; [e] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing; [f] Department of Radiology, Affiliated People's Hospital of Jiangsu University, Zhenjiang; [g] Department of General Surgery; [h] Department of Pathology, Lanzhou University Second Hospital; and [i] Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, Beihang University, Beijing, PR China

ARTICLE INFO

Article history:
Received 15 March 2019
Received in revised form 21 November 2019
Accepted 23 November 2019

Keywords:
Gastric cancer
Computed tomography
Radiomics
Deep learning
Prognosis

ABSTRACT

Background: In the clinical management of advanced gastric cancer (AGC), preoperative identification of early recurrence after curative resection is essential. Thus, we aimed to create a CT-based radiomic model to predict early recurrence in AGC patients preoperatively.

Materials and methods: We enrolled 669 consecutive patients (302 in the training set, 219 in the internal test set and 148 in the external test set) with clinicopathologically confirmed AGC from two centers. Radiomic features were extracted from preoperative diagnostic CT images. Machine learning methods were applied to shrink feature size and build a predictive radiomic signature. We incorporated the radiomic signature and clinical risk factors into a nomogram using multivariable logistic regression analysis. The area under the curve (AUC) of operating characteristics (ROC), accuracy, and calibration curves were assessed to evaluate the nomogram's performance in discriminating early recurrence.

Results: A radiomic signature, including three hand crafted features and six deep learning features, was significantly associated with early recurrence (p-value <0.0001 for all sets). In addition, clinical N stage, carbohydrate antigen 199 levels, carcinoembryonic antigen levels, and Borrmann type were considered useful predictors for early recurrence. The nomogram, combining all these predictors, showed powerful prognostic ability in the training set and two test sets with AUCs of 0.831 (95% CI, 0.786–0.876), 0.826 (0.772–0.880) and 0.806 (0.732–0.881), respectively. The predicted risk yielded good agreement with the observed recurrence probability.

Conclusions: By incorporating a radiomic signature and clinical risk factors, we created a radiomic nomogram to predict early recurrence in patients with AGC, preoperatively, which may serve as a potential tool to guide personalized treatment.

 Radiotherapy and Oncology 145 (2020) 13–20

Gastric cancer (GC) is the fourth most common type of malignant disease, and it ranks as the second leading cause of cancer-related death worldwide, with 77% of cases occurring in developing countries [1]. Although the diagnostic procedures and multidisciplinary treatment strategies for patients with resectable locally advanced GC (AGC) have been greatly improved, the survival remains unsatisfactory owing to a high incidence of recurrence [2,3]. More than 50% of recurrences and tumor-related deaths occurred within 1 year after surgery [4,5] with the median recurrence-free survival of approximately 10.8 months [6]. For this reason, the preoperative estimation of the risk of early recurrence is relevant in clinical practice. Although surgical resection is regarded as the primary curative treatment for AGC, more aggressive treatment regimens, including preoperative and postoperative chemotherapy and chemoradiation therapy, are effective in prolonging relapse-free survival and overall survival [7–9]. However, not all patients benefit from these aggressive treatments [10,11], which may be due to failure in identifying groups at high risk for early recurrence. Additional adjuvant therapies and more frequent follow-up would also provide an enduring benefit for patients at high risk for recurrence [9,12]. Therefore, to improve prognosis, there is a critical need to develop better biomarkers to predict

* Corresponding authors at: Department of Radiology, Lanzhou University Second Hospital, Lanzhou 730030, PR China (J. Zhou). Department of Radiology, Affiliated People's Hospital of Jiangsu University, Zhenjiang 212002, PR China (X. Shan). CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, PR China (J. Tian).

E-mail addresses: ery_zhoujl@lzu.edu.cn (J. Zhou), xhongshan@hotmail.com (X. Shan), tian@ieee.org (J. Tian).

[1] Wenjuan Zhang, Mengjie Fang and Di Dong contributed equally to this work.

https://doi.org/10.1016/j.radonc.2019.11.023
0167-8140/

the risk of early recurrence in order to stratify and identify those who may benefit from aggressive treatment.

Several risk factors, such as pathological characteristics, biologic/genetic biomarkers, and tumor-node-metastasis (TNM) staging [13,14] have been demonstrated to be associated with the recurrence of AGC after curative resection. However, there has been shown to be large difference in the clinical outcomes of the patients with same disease stage who undergo similar treatment strategies [14,15]. These findings revealed the limitation of the TNM staging system in providing adequate prognostic information. Moreover, preoperative pathological characteristics and biologic/-genetic biomarkers can only be collected through biopsy, which has not been recommended in routine clinical care due to the potential risk of tumor cell dissemination and seeding and inconsistencies with the final pathology [16].

Non-invasive computed tomography (CT) has been widely used for differential diagnosis, preoperative assessment, and therapeutic evaluation in patients with GC [17,18]. In contrast to conventional imaging features, "radiomics" has been introduced because it is considered to have the potential ability to reveal disease characteristics invisible to the naked eye [19]. Radiomics is a process that quantifies medical images as high-dimensional, mineable data through advanced feature extraction procedures (e.g. hand-crafted texture features and deep learning features); subsequent data analysis is then performed to support the clinical decision [20,21]. Radiomics has been widely used to for tumor detection, tumor subtype classification, prognosis prediction, and therapeutic response evaluation [16,19–21]. Recently, several studies have explored the value of radiomics in the prognosis of GC. Giganti et al. [22] investigated the association between preoperative CT textures and overall survival in 56 GC patients and found that CT textures were significantly associated with overall survival. Li et al. [23] and Jiang et al. [24] investigated the prognostic significance of radiomic signature in GC patients who underwent radical resection. Their findings indicated that the radiomic signature could predict overall survival. Jiang's study [24] also demonstrated radiomic could predict which patients with stage II and III GC benefited more from chemotherapy. However, to the best of our knowledge, there are currently no reports of radiomics and deep learning study on predicting early recurrence in AGC.

The objective of this retrospective study was to create a prognostic model, which incorporates radiomic features and clinical risk factors that can preoperatively predict early recurrence in patients with AGC who undergo curative surgery; such a tool would help identify high-risk patients who require aggressive treatment and close follow-up.

Materials and methods

Patients

Ethical approval was obtained for this retrospective analysis (from the Ethics Committees at Lanzhou University Second Hospital [approval date: 11 March, 2017] and Jiangsu University [approval date: 16 September, 2018]), and the informed consent requirement was waived. The protocol conducted in accordance with the ethical principles of the Declaration of Helsinki and also the Good Clinical Practice guidelines, as defined by the International Conference on Harmonisation. This study included 521 consecutive patients who were pathologically diagnosed with AGC and who underwent radical gastrectomy with D2 lymphadenectomy between January 2013 and March 2017 at center 1 (Lanzhou University Second Hospital). Detailed descriptions of the inclusion and conclusion criteria, and the recruitment process are shown in Fig. S1. These patients were divided into a training set and an internal test set. The training set contained 302 patients treated from January 2013 to December 2015, with 120 (39.74%) patients having early recurrence. The internal test set included 219 patients treated between January 2016 and March 2017, which included 84 (38.36%) patients with early recurrence. Furthermore, we also collected an external test set from another center (center 2, Affiliated People's Hospital of Jiangsu University), including 148 patients (52 [35.14%] with early recurrence) who were treated between February 2014 and November 2016. Noted that we also used the randomized assignment of the training and testing sets (ratio 3:2) to assess the reproducibility of our method. Therefore, the model in this paper was trained and evaluated further 5 times on different patient assignments. The details on clinical characteristics of patients were described in Supplementary Material.

Follow-up

The time to recurrence (<1 year) was chosen as the endpoint for early recurrence. All patients underwent follow-up at least 1 year following radical gastrectomy. Patients were monitored every 3–6 months during the first 2 years and every 6–12 months for the subsequent 3–5 years. The median follow-up times were 20 months for center 1 and 26 months for center 2, respectively.

Recurrence of gastric cancer was categorized as: locoregional, hematogenous, peritoneal, and distant lymphatic. At every follow-up visit, we obtained a medical history and performed a physical examination, which included routine blood tests to check liver function and several tumor markers; chest X-ray, abdominal ultrasound, CT scan, other imaging work-ups, and upper endoscopy were also conducted at regular intervals. Recurrence was confirmed by clinical imaging, gastroscopic biopsy, cytological examination of ascites, or intraoperative findings of reoperation.

Image acquisition and segmentation

We retrieved CT images from the picture archiving and communication system (PACS) (Carestream, Langley, British Columbia, Canada). Analysis was done on portal venous-phase CT images, considering its well differentiation between tumor and the adjacent normal tissue. Manual tumor segmentation was performed on ITK-SNAP software (version 3.6.0; www.itksnap.org) on venous-phase CT images. CT scanner, acquisition parameters, image segmentation and construction of two additional sets (i.e. the re-segmentation set and the simulated slice-thickness set) are described in Supplementary Material.

Radiomic features extraction

Two-step image standardization was implemented before feature extraction: bicubic resampling was used to standardize the image scale in the slice, resulting in a pixel size of 0.5 mm × 0.5 mm; the pixel intensity translation was used to minimize the discrepancy of intensity distributions among patients. As shown in Fig. S2, both hand-crafted features and deep learning features were extracted from the ROIs to quantify the tumor phenotype. For each patient, 115 hand-crafted and 112 deep learning radiomic features were extracted from the two-dimensional largest tumor region. The feature extraction algorithms were standardized by referring to the Image Biomarker Standardisation Initiative (IBSI) [25]. The hand-crafted features contained four different groups of features: shape, histogram, gray-level co-occurrence matrix (GLCM), and gray-level run-length matrix (GLRLM). Additionally, a deep convolutional neural networks (DCNN) containing 8 weighted layers was constructed and trained to extract deep learning features (Figs. S2b and S3) based on the training set. The detailed introduction of the features is described in Supplementary Material. All the extracted features of each patient were normalized by z-score method based on the parameters calculated in the training set.

Table 1
Clinical characteristics of patients in the training and test sets.

		Training set				Internal test set				External test set			
		Overall	Non-recurrence	Recurrence	*p*-value	Overall	Non-recurrence	Recurrence	*p*-value	Overall	Non-recurrence	Recurrence	*p*-value
Age		54.90 ± 9.32	54.86 ± 10.00	54.97 ± 8.22	0.9174	55.19 ± 10.27	54.78 ± 10.00	55.86 ± 10.71	0.4580	62.08 ± 9.11	62.35 ± 9.30	61.58 ± 8.82	0.6166
Sex	Female	74	50	24	0.1396	57	40	17	0.1235	40	29	11	0.2364
	Male	228	132	96		162	95	67		108	67	41	
Histological grade	Well differentiated	21	16	5	0.1768	16	14	2	0.0010	1	0	1	0.1675
	Moderately differentiated	135	84	51		100	70	30		53	38	15	
	Poorly differentiated	146	82	64		103	51	52		94	58	36	
Borrmann	I/II	79	61	18	0.0003	67	59	8	<0.0001	52	40	12	0.0237
	III/IV	223	121	102		152	76	76		96	56	40	
Lauren	Intestinal	124	77	47	0.8592	91	64	27	0.0049	43	33	10	0.1484
	Diffuse	109	64	45		73	34	39		64	39	25	
	Mixed	69	41	28		55	37	18		41	24	17	
Tumor location	Antrum	143	83	60	0.7223	106	73	33	0.1025	63	38	25	0.6041
	Body	85	52	33		59	32	27		37	25	12	
	Cardia	74	47	27		54	30	24		48	33	15	
CEA	Normal	234	151	83	0.0050	156	104	52	0.0162	102	72	30	0.0299
	Abnormal	68	31	37		63	31	32		46	24	22	
CA199	Normal	237	157	80	0.0001	173	112	61	0.0876	114	78	36	0.0970
	Abnormal	65	25	40		46	23	23		34	18	16	
CA724	Normal	214	135	79	0.1185	161	104	57	0.1344				
	Abnormal	88	47	41		58	31	27					
CA125	Normal	280	171	109	0.3069	197	128	69	0.0024	136	91	45	0.1497
	Abnormal	22	11	11		22	7	15		12	5	7	
cN stage	N0/N1	146	114	32	<0.0001	112	93	19	<0.0001	91	69	22	0.0004
	N1/N2	156	68	88		107	42	65		57	27	30	
cT stage	T2	57	49	8	<0.0001	60	55	5	<0.0001	19	11	8	0.0171
	T3	159	101	58		111	57	54		80	60	20	
	T4a/T4b	86	32	54		48	23	25		49	25	24	
cTNM stage	I	38	34	4	<0.0001	31	30	1	<0.0001	13	8	5	0.0375
	II	19	15	4		29	25	4		6	3	3	
	III	41	32	9		27	19	8		25	22	3	
	IVa/IVb	204	101	103		132	61	71		104	63	41	

CEA, carcinoembryonic antigen; CA199, carbohydrate antigen 199; CA724, carbohydrate antigen 724; CA125, carbohydrate antigen 125; cN stage, clinical N stage; cT stage, clinical N stage; cTNM stage, clinical TNM stage.

Feature selection and radiomic signature building

Based on the training set, we performed feature selection and built a radiomic signature as an independent predictor for early recurrence.

Firstly, based on the re-segmentation set and the simulated slice-thickness set respectively, we calculated the intra-/inter-class correlation coefficients (ICCs) and the coefficient of variation (CV) to evaluate the reproducibility and robustness of feature extraction. Only the features which yielded ICCs of greater than 0.8 and CVs of less than 15% were entered in the process of signature building [26]. Secondly, we used the consensus clustering [27] with Pearson correlation as the distance metric and performed 10,000 resampling iterations to separate the features into several groups [28]. To eliminate redundant features, only the features with the highest average consensus index in each cluster, named the most representative medoid features, were retained. We chose the number of clusters which made the medoid features yielding significant correlation with all the intra-cluster features (Pearson correlation coefficients >0.8). Finally, the remaining features were used to construct the radiomic signature based on the logistic regression model. Backward step-wise selection was conducted with Akaike's information criterion (AIC). The variance inflation factor (VIF) was used to check the degree of multi-collinearity of each variable in the regression model. The variables with VIF > 5 were excluded [29].

Statistical analysis

We used univariate analysis to assess the relationship between patients' characteristics and early recurrence. Differences in patient characteristics between the different groups were assessed using the Mann–Whitney *U* test or independent t-test for continuous variables, and the chi-square test or Fisher's exact test for categorical variables. Noted that the categorical variables were grouped to eliminate sparse categories and to meet the clinical reasoning. These decisions were made before modelling.

Using the clinical characteristics and radiomic signature as input, multivariable logistic regression analysis with variable selection was implemented to identify the powerful combination of these predictors. Then, we built a quantitative radiomic nomogram to predict the individual probability of post-operation progression. For comparison, this multivariable analysis procedure was also performed to build a clinical model with only the clinical characteristics.

The calibration curves, with the Hosmer–Lemeshow test, were used to assess the agreement between the predicted risks and the actual results. We evaluated the radiomic nomogram using the area under the curve (AUC) of operating characteristics (ROC) with 95% confidence interval (95% CI) and used the DeLong test to compare different ROCs. We selected the point which yielded the highest Youden's index (i.e. specificity + sensitivity − 1) on the ROC curve of the training set as the optimal cut-off value, and used it to classify the patients into the high-risk group or the low-risk group. Accuracy was calculated to assess the prediction performance. Kappa test was used to determine intra-/inter-reader agreement for the models.

The software used for modeling and statistical analysis is reported in Supplementary Material.

Results

Clinical characteristics of the three sets are listed in Table 1. Borrmann type, CEA, CA199, clinical T stage (cT), clinical N stage (cN), and overall stage, differed significantly between the non-early recurrence group and the early recurrence group in the training set (p-value <0.05).

Most of the radiomic features (130/227) were demonstrated to have inter-/intra-reader agreement and robustness for the varying slice-thickness. A heatmap of these features and unsupervised cluster partitioning are shown in Fig. S4. A significant association between these features and early recurrence was observed. We obtained 22 distinct feature clusters using the consensus clustering (Fig. S5). Based on the medoid features, we built the radiomic signature using the logistic regression model. After the backward elimination and the multi-collinearity analysis, the final radiomic features in our signature contained three hand crafted features and six deep features (Table S1). The radiomic signature showed a good performance for discriminating early recurrence with AUCs of 0.785 (95% CI, 0.733–0.837) in the training set, 0.764 (0.701–0.827) in the internal test set and 0.769 (0.686–0.852) in the external test set.

Beginning with the radiomic signature and the clinical characteristics, multivariable analysis demonstrated that the radiomic signature, Borrmann type, CEA, CA199, and cN stage remained important predictors after adjustment for cofactors (Table S2). A radiomic nomogram was then built by using the regression coefficients for prediction of early recurrence (Fig. 1). Meanwhile, the clinical model incorporated sex, Borrmann type, CEA, CA199, cN stage, and cT stage (Table S3).

The nomogram had powerful prognostic ability in all the three sets with AUCs of 0.831 (95% CI, 0.786–0.876), 0.826 (0.772–0.880) and 0.806 (0.732–0.881) (Fig. 2). The DeLong test indicated that no statistical difference existed between the ROCs of the nomogram in the two test sets and the training set (p-value = 0.8840 and 0.5736). Furthermore, we randomly split our dataset into paired training (60%) and test sets (40%) to generate 5 training/test set pairs. We trained our predictive model and evaluated it repeatedly, generating a total of 5 validated AUCs ranging from 0.792 to 0.841. By using the cut-off value of 0.3488, the nomogram yielded accuracies of 0.748 (95% CI, 0.695–0.796; sensitivity, 0.800; specificity, 0.714) in training set, 0.726 (0.662–0.784; 0.750; 0.711) in internal test set and 0.723 (0.644–0.793; 0.827; 0.667) in external test set. The radiomic nomogram had better prognostic performance than the radiomic signature and the clinical model, which are listed in Table 2 in detail. The results of Kappa tests demonstrated there were substantial intra-/inter-reader agreement for the radiomic signature (Kappa values, 0.780 and 0.615). Moreover, the radiomic nomogram yielded even more excellent agreement as Kappa values equaled to 0.932 in both conditions.

The stratified analysis is shown in Supplementary Material. The results indicated that the performance of radiomic nomogram was not affected by patient age, sex, tumor location, histological grade, operative type, surgical approach, postoperative chemotherapy, or

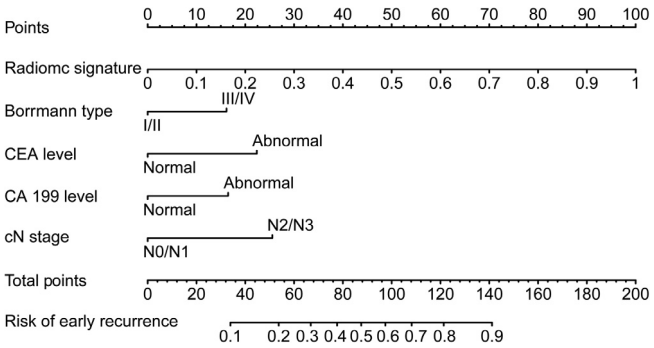

Fig. 1. Radiomic nomogram based on radiomic signature and clinical factors.

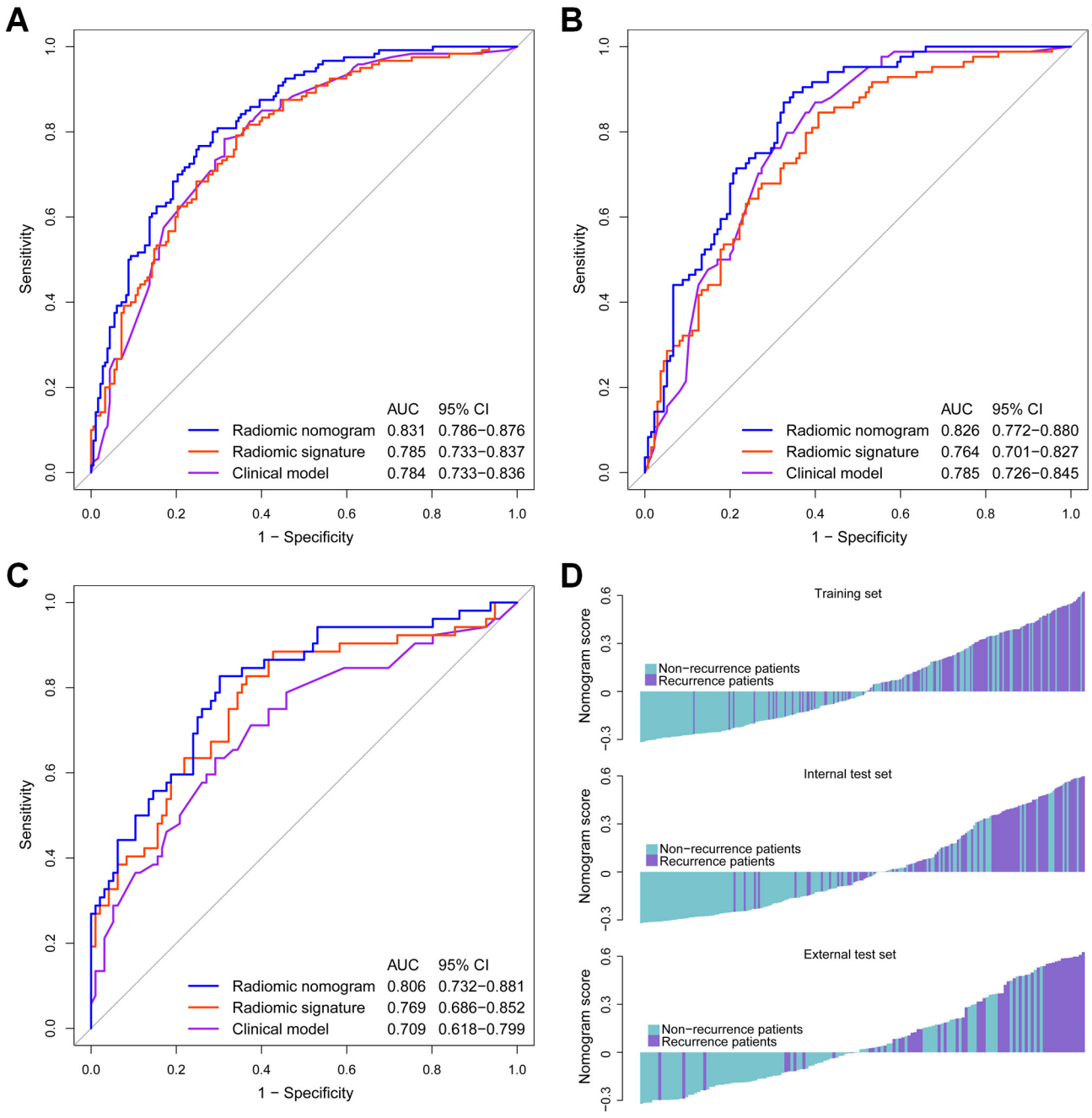

Fig. 2. ROC curves in the training (A), internal test (B) and external test set (C), and patients' radiomic nomogram scores (D). The scores have been subtracted by the cut-off value.

Table 2
Prognostic performances of models on the training and test sets.

	Training set				Internal test set				External test set			
	AUC (95% CI)	Accuracy (95% CI)	Sensitivity	Specificity	AUC (95% CI)	Accuracy (95% CI)	Sensitivity	Specificity	AUC (95% CI)	Accuracy (95% CI)	Sensitivity	Specificity
Radiomic nomogram	0.831 (0.786–0.876)	0.748 (0.695–0.796)	0.800	0.714	0.826 (0.772–0.880)	0.726 (0.662–0.784)	0.750	0.711	0.806 (0.732–0.881)	0.723 (0.644–0.793)	0.827	0.667
Radiomic signature	0.785 (0.733–0.837)	0.709 (0.654–0.759)	0.808	0.643	0.764 (0.701–0.827)	0.708 (0.643–0.767)	0.679	0.726	0.769 (0.686–0.852)	0.676 (0.594–0.750)	0.885	0.563
Clinical model	0.784 (0.733–0.836)	0.725 (0.671–0.775)	0.783	0.687	0.785 (0.726–0.845)	0.722 (0.657–0.780)	0.762	0.696	0.709 (0.618–0.799)	0.655 (0.573–0.732)	0.673	0.646

AUC, area under the curve; CI, confidence interval.

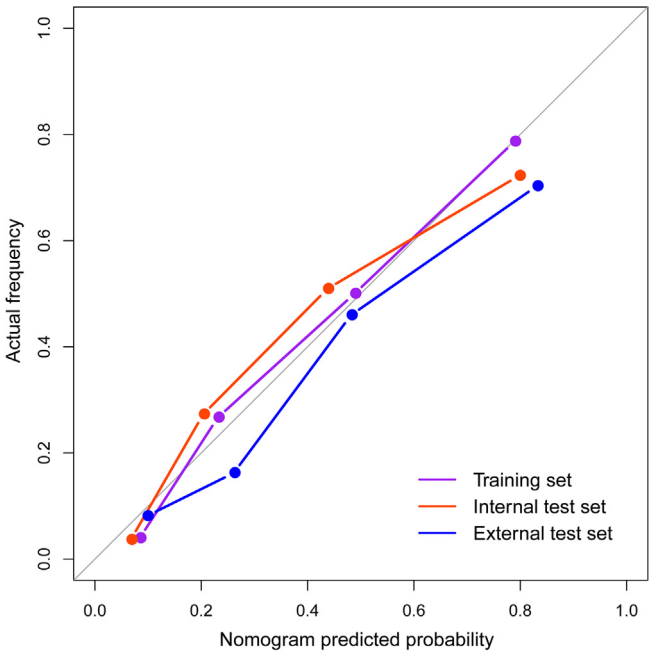

Fig. 3. Calibration curves of the radiomic nomogram in each set.

a recurrent pattern (DeLong test p-value > 0.05) (Figs. S6 and S7), which implied its great generalization ability.

The calibration curve of the radiomic nomogram showed good agreement between the predictive risk and the observed recurrence probability in the three sets (Fig. 3). The Hosmer–Lemeshow test was not significant (p-value = 0.4667, 0.1372 and 0.0937), suggesting there is no significant departure.

The performance of each risk score is shown in Figs. 4 and S8. In these figures, we separated patients into two groups—high-risk and low-risk—according to the corresponding nomogram scores and the cut-off value. The Kaplan-Meier curves of disease-free survival (DFS) for the two groups were statistically divergent in all the three sets.

Discussion

Preoperative prediction of early recurrence in AGC is important for clinical practice. Thus, we developed and validated a novel, CT-based radiomic nomogram to preoperatively predict early recurrence in patients with AGC following curative resection. The nomogram, incorporating the radiomic signature, Borrmann type, CEA, CA199, and cN stage, successfully identify patients at high risk of early recurrence. Furthermore, the nomogram provided better predictive accuracy than the clinical factor-based model and radiomic signature alone, demonstrating the incremental value of the radiomic nomogram to the current diagnostic management of AGC. Moreover, our nomogram is easy to use, and it could serve as a preoperative tool for individualized prediction of prognosis and for guiding treatment in patients with AGC.

The radiomic signature consisted of three hand crafted features and six deep learning feature. The hand crafted textural features measure the second-order or the high-order intensity patterns in the image, and therefore are able to reflect the heterogeneity of the tumor imaging phenotype. Deep learning models can learn high-level image features hierarchically, which are built upon low-level features, layer-by-layer (Fig. S2b). Six deep learning features were selected and incorporated into our radiomic signature. They were calculated based on the corresponding deep learning feature maps, which were generated from the trained convolutional architecture (Fig. S3). As shown in Fig. S9, the feature maps might highlight the region at high risk for recurrence. Considering the operability and practicability, we chose the two-dimensional features (rather than the three-dimensional features) to quantify the tumor phenotype characteristics. Experimental results demonstrated that the features, as well as the developed predictive models, had satisfactory reproducibility and prognostic ability. Furthermore, based on the CT images of 30 randomly selected patients, we compared the radiomic features extracted from the single slice and the three-dimensional multiple slices, and found that the values of most features (170/227) had high correlations (Pearson correlation coefficients >0.8) between the two approaches. This implied that the two-dimensional features might be able to in a certain extent reflect the whole tumor. In radiomics pipeline, the process of image interpolation is needed before feature extraction to increase the stableness of features and models on different cohorts. A lot of researches have been done on the advanced interpolation algorithms and the performance analysis of different interpolation strategies [25,30–32]. Mackin et al. implemented Butterworth filtering following interpolation to reduce the variability in features [32]. Whybra et al. investigated the sensitivity of radiomic features to different interpolation algorithms and scales, and developed a correction technique to increase features' robustness [31]. In this study, to avoid introducing a large number of artificial information, which might be produced in three-dimensional isotropic interpolation when the spacing between slices is too large, we implemented the in-slice interpolation to coordinate the two-dimensional feature extraction. Meanwhile, we utilized a simulated slice-thickness set to remove the features sensitive to the varying slice-thickness. The developed nomogram showed stable prognostic performance in the validation experiment. It should be noted that, since many

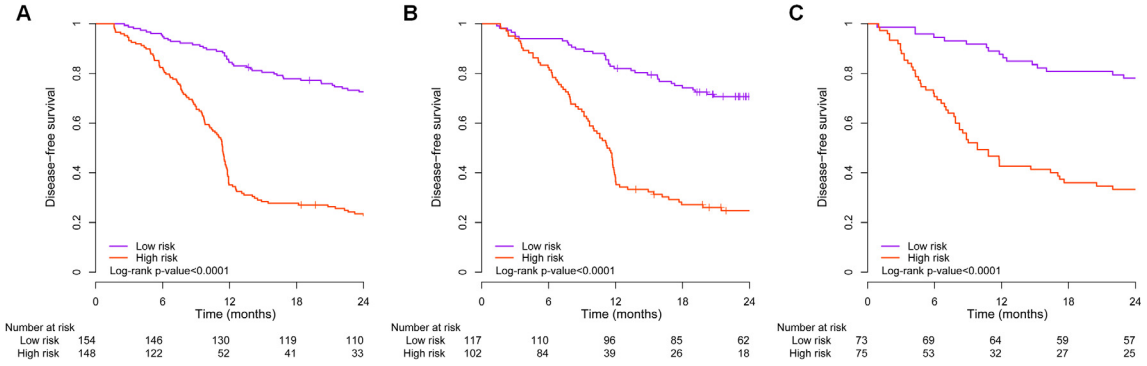

Fig. 4. Two-year Kaplan-Meier curves for the radiomic nomogram in the training (A), internal test (B) and external test set (C).

radiomic features are susceptive to the parameters of imaging and interpolation, more advanced isotropic interpolation and feature selection should be designed and evaluated in the future.

In this study, CEA, CA199, cN stage and Borrmann type were demonstrated as important risk predictors and incorporated into our radiomic nomogram. Previous studies have indicated that TNM staging is the most commonly used system to predict outcome for AGC patients [13,33]. The cumulative recurrence rates at 5 years for peritoneal, hematogenous, and lymphatic metastases of T2/3 and T4 are 8.0% and 53.6%, 13.0% and 33.1%, 7.7% and 31.2%, respectively [34]. In our study, patients with cT4a/T4b and/or cN2/N3 were at greater risk of experiencing early recurrence compared with whose with T2/T3 and/or cN0/N1 in univariate analysis, all of which is consistent with that of previous studies. However, cT stage was not included in our radiomic nomogram, which might be because it was associated with other risk factors.

Pre-operative serum CEA, CA199 and CA724 levels were reported to be significantly associated with the invasiveness of GC and survival [35]. CEA is known as an independent risk factor for predicting hematogenous recurrences of GC, especially liver metastasis relapse. CA199 was evaluated significantly in the gastric patients with later N stage, the prognosis of N1 stage gastric patients with high level of CA199 was significantly poorer than that of patients with lower level [36]. Macroscopic type (Borrmann classification) is also a simple and valuable predictor for lymph node metastasis and survival in AGC patients, and it is a known risk factor for peritoneal dissemination [37]. Borrmann types III and IV GC were common when tumors were found to invade deep into the gastric wall. In this study, all the above mentioned factors were associated with early recurrence.

As a preoperative tool, our nomogram included only preoperative factors and did not use postoperative factors such as operative type, surgical approach, or postoperative chemotherapy. However, the surgical approach and postoperative therapy may influence the prognosis in AGC patients. The Japanese gastric cancer treatment guidelines [38] recommend distal gastrectomy with D2 lymph node dissection as the standard surgical treatment for locally AGC because of the confirmed feasibility and safety. One randomized controlled trial showed that the compliance rates of D2 lymphadenectomy, postoperative morbidity, and mortality rate were similar between laparoscopic distal gastrectomy and open gastrectomy [39]; moreover, no significant statistical difference existed in the overall survival for patients in either group [40]. Besides surgical management, postoperative chemotherapy was found to be an independent predictor of early recurrence. Adjuvant chemotherapy with capecitabine plus oxaliplatin, fluorouracil, or leucovorin after D2 gastrectomy improved the outcome of the patients with resectable GC and should be considered as a treatment option [7,9]. In this study, to further validate the consistency of our nomogram, we performed stratified analysis on operative type, surgical approach, and postoperative chemotherapy. The results showed that the nomogram was not affected by these factors and it remained consistent, thus implying that our method can be generalized.

Our study has several limitations. Selection bias occurred when strict criteria were used (randomization hypothesis is compromised), which may affect the model training. In our study, the criteria introduced selection bias by removing patients with the best prognosis (e.g., those of early gastric cancer) as well as the worst prognosis (e.g., those were lost in follow up at all and those of IVb stage). The selection bias thus limits our model only accurate in AGC patients in a common condition. Therefore, while our results clearly showed the potential of radiomics approach in the prognosis of AGC patients, it is needed to increase the patient sample size and include other patient population in the future. Moreover, our method was developed based on the patients of Asian race/ethnicity. As the epidemiology of gastric cancer in Eastern countries is very different from that in Western countries, further validation on patients of different ethnicity should be studied. While the prognostic ability and generalization of the two-dimensional feature-based models were assessed in this study, the performance of the three-dimensional features remains to be further investigated. Finally, although satisfactory independent validation results were obtained in this study, many key techniques in radomics pipeline (such as automated segmentation, advanced isotropic interpolation, stable feature selection) deserve further research to improve the robustness and performance of radiomic model.

In conclusion, we were able to develop a radiomic nomogram that incorporates a radiomic signature and clinical risk factors to effectively predict early recurrence in patients with AGC following curative resection; the nomogram was validated through diverse methods showing powerful prognostic ability. We expect that the radiomic nomogram may serve as a potential tool for guiding individual care for these patients.

Competing interests

The authors have declared that no competing interest exists.

Acknowledgements

This work was supported by the National Key R&D Program of China (2017YFC1308700, 2017YFA0205200, 2017YFC1309100, 2017YFA0700401), National Natural Science Foundation of China (81772006, 91959130, 81971776, 81771924, 81501616, 81227901), Jiangsu Provincial Research Foundation for Basic Research of China (BK20151334), Zhenjiang Innovation Capacity Building Program (technological infrastructure) – R & D project of China (SS2015023), Jiangsu Provincial Key R & D Special Fund (BE2015666), the Beijing Natural Science Foundation (L182061), Key Research Program of the Chinese Academy of Sciences (No.-KGZD-EW-T03) and the Youth Innovation Promotion Association of the Chinese Academy of Sciences (2017175).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.radonc.2019.11.023.

References

[1] Fitzmaurice C, Dicker D, Pain A, Hamavid H, Moradi-Lakeh M, MacIntyre MF, et al. The global burden of cancer 2013. JAMA Oncol 2015;1:505–27.
[2] Zhang ZY, Dai ZL, Yin XW, Li SH, Li SP, Ge HY. Meta-analysis shows that circulating tumor cells including circulating microRNAs are useful to predict the survival of patients with gastric cancer. BMC Cancer 2014;14:773.
[3] Tegels JJ, De Maat MF, Hulsewe KW, Hoofwijk AG, Stoot JH. Improving the outcomes in gastric cancer surgery. World J Gastroenterol 2014;20:13692–704.
[4] Deng J, Liang H, Wang D, Sun D, Pan Y, Liu Y. Investigation of the recurrence patterns of gastric cancer following a curative resection. Surg Today 2011;41:210–5.
[5] Rohatgi PR, Yao JC, Hess K, Schnirer I, Rashid A, Mansfield PF, et al. Outcome of gastric cancer patients after successful gastrectomy: influence of the type of recurrence and histology on survival. Cancer 2006;107:2576–80.
[6] Spolverato G, Ejaz A, Kim Y, Squires MH, Poultsides GA, Fields RC, et al. Rates and patterns of recurrence after curative intent resection for gastric cancer: a United States multi-institutional analysis. J Am Coll Surg 2014;219:664–75.
[7] Bang YJ, Kim YW, Yang HK, Chung HC, Park YK, Lee KH, et al. Adjuvant capecitabine and oxaliplatin for gastric cancer after D2 gastrectomy (CLASSIC): a phase 3 open-label, randomised controlled trial. Lancet (London, England) 2012;379:315–21.
[8] Leong T, Smithers BM, Haustermans K, Michael M, Gebski V, Miller D, et al. TOPGEAR: a randomized, phase III trial of perioperative ecf chemotherapy with or without preoperative chemoradiation for resectable gastric cancer: interim

results from an international, intergroup trial of the AGITG, TROG, EORTC and CCTG. Ann Surg Oncol 2017;24:2252–8.
[9] Smalley SR, Benedetti JK, Haller DG, Hundahl SA, Estes NC, Ajani JA, et al. Updated analysis of SWOG-directed intergroup study 0116: a phase III trial of adjuvant radiochemotherapy versus observation after curative gastric cancer resection. J Clin Oncol 2012;30:2327–33.
[10] Kang YK, Chang HM, Yook JH, Ryu MH, Park I, Min YJ, et al. Adjuvant chemotherapy for gastric cancer: a randomised phase 3 trial of mitomycin-C plus either short-term doxifluridine or long-term doxifluridine plus cisplatin after curative D2 gastrectomy (AMC0201). Br J Cancer 2013;108:1245–51.
[11] Lee J, Lim DH, Kim S, Park SH, Park JO, Park YS, et al. Phase III trial comparing capecitabine plus cisplatin versus capecitabine plus cisplatin with concurrent capecitabine radiotherapy in completely resected gastric cancer with D2 lymph node dissection: the ARTIST trial. J Clin Oncol 2012;30:268–73.
[12] Badgwell B, Blum M, Elimova E, Estrella J, Chiang YJ, Das P, et al. Frequency of resection after preoperative chemotherapy or chemoradiotherapy for gastric adenocarcinoma. Ann Surg Oncol 2016;23:1948–55.
[13] Bando E, Makuuchi R, Tokunaga M, Tanizawa Y, Kawamura T, Terashima M. Impact of clinical tumor-node-metastasis staging on survival in gastric carcinoma patients receiving surgery. Gastric Cancer 2017;20:448–56.
[14] Tan P, Yeoh KG. Genetics and molecular pathogenesis of gastric adenocarcinoma. Gastroenterology 2015;149. 1153-62.e3.
[15] Sasako M, Inoue M, Lin JT, Khor C, Yang HK, Ohtsu A. Gastric cancer working group report. Jpn J Clin Oncol 2010;40:i28–37.
[16] Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.
[17] Park SR, Kim MJ, Ryu KW, Lee JH, Lee JS, Nam BH, et al. Prognostic value of preoperative clinical staging assessed by computed tomography in resectable gastric cancer patients: a viewpoint in the era of preoperative treatment. Ann Surg 2010;251:428–35.
[18] Tsurumaru D, Miyasaka M, Nishimuta Y, Asayama Y, Nishie A, Kawanami S, et al. Differentiation of early gastric cancer with ulceration and resectable advanced gastric cancer using multiphasic dynamic multidetector CT. Eur Radiol 2016;26:1330–7.
[19] Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016;278:563–77.
[20] Dong D, Tang L, Li ZY, Fang MJ, Gao JB, Shan XH, et al. Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced gastric cancer. Annals Oncol 2019.
[21] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer (Oxford, England) 1990;2012:441–6.
[22] Giganti F, Antunes S, Salerno A, Ambrosi A, Marra P, Nicoletti R, et al. Gastric cancer: texture analysis from multidetector computed tomography as a potential preoperative prognostic biomarker. Eur Radiol 2017;27:1831–9.
[23] Li W, Zhang L, Tian C, Song H, Fang M, Hu C, et al. Prognostic value of computed tomography radiomics features in patients with gastric cancer following curative resection. Eur Radiol 2018.
[24] Jiang Y, Chen C, Xie J, Wang W, Zha X, Lv W, et al. Radiomics signature of computed tomography imaging for prediction of survival and chemotherapeutic benefits in gastric cancer. EBioMedicine 2018;36:171–82.
[25] Zwanenburg A, Leger S, Vallières M, Löck S. Image biomarker standardisation initiative. Radiother Oncol 2016;127:S543–4.
[26] Berenguer R, Pastor-Juan MDR, Canales-Vazquez J, Castro-Garcia M, Villas MV, Mansilla Legorburo F, et al. Radiomics of CT features may be nonreproducible and redundant: influence of CT acquisition parameters. Radiology 2018;288:407–15.
[27] Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. Machine Learn 2003;52:91–118.
[28] Parmar C, Leijenaar RT, Grossmann P, Rios Velazquez E, Bussink J, Rietveld D, et al. Radiomic feature clusters and prognostic signatures specific for Lung and Head & Neck cancer. Sci Rep 2015;5:11044.
[29] Del Poggio P, Olmi S, Ciccarese F, Di Marco M, Rapaccini GL, Benvegnu L, et al. Factors that affect efficacy of ultrasound surveillance for early stage hepatocellular carcinoma in patients with cirrhosis. Clin Gastroenterol Hepatol 2014;12. 1927-33.e2.
[30] Larue R, Van De Voorde L, van Timmeren JE, Leijenaar RTH, Berbee M, Sosef MN, et al. 4DCT imaging to assess radiomics feature stability: an investigation for thoracic cancers. Radiother Oncol 2017;125:147–53.
[31] Whybra P, Parkinson C, Foley K, Staffurth J, Spezi E. Assessing radiomic feature robustness to interpolation in (18)F-FDG PET imaging. Sci Rep 2019;9:9649.
[32] Mackin D, Fave X, Zhang L, Yang J, Jones AK, Ng CS, et al. Harmonizing the pixel size in retrospective computed tomography radiomics studies. PLoS ONE 2017;12:e0178524.
[33] Bando E, Makuuchi R, Irino T, Tanizawa Y, Kawamura T, Terashima M. Validation of the prognostic impact of the new tumor-node-metastasis clinical staging in patients with gastric cancer. Gastric Cancer 2018.
[34] Nakanishi Y, Ohara M, Domen H, Shichinohe T, Hirano S, Ishizaka M. Differences in risk factors between patterns of recurrence in patients after curative resection for advanced gastric carcinoma. World J Surg Oncol 2013;11:98.
[35] Shimada H, Noie T, Ohashi M, Oba K, Takahashi Y. Clinical significance of serum tumor markers for gastric cancer: a systematic review of literature by the Task Force of the Japanese Gastric Cancer Association. Gastric Cancer 2014;17:26–33.
[36] Li F, Li S, Wei L, Liang X, Zhang H, Liu J. The correlation between pre-operative serum tumor markers and lymph node metastasis in gastric cancer patients undergoing curative treatment. Biomarkers 2013;18:632–7.
[37] Kanda M, Mizuno A, Fujii T, Shimoyama Y, Yamada S, Tanaka C, et al. Tumor infiltrative pattern predicts sites of recurrence after curative gastrectomy for stages 2 and 3 gastric cancer. Ann Surg Oncol 2016;23:1934–40.
[38] Japanese gastric cancer treatment guidelines 2010 (ver. 3). Gastric cancer 2011;14:113–123.
[39] Hu Y, Huang C, Sun Y, Su X, Cao H, Hu J, et al. Morbidity and mortality of laparoscopic versus open D2 distal gastrectomy for advanced gastric cancer: a randomized controlled trial. J Clin Oncol 2016;34:1350–7.
[40] Park YK, Yoon HM, Kim YW, Park JY, Ryu KW, Lee YJ, et al. Laparoscopy-assisted versus Open D2 distal gastrectomy for advanced gastric cancer: results from a randomized phase II multicenter clinical trial (COACT 1001). Ann Surg 2018;267:638–45.