European Journal of Radiology 132 (2020) 109277

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

CT-based deep learning radiomics analysis for evaluation of serosa invasion in advanced gastric cancer

Rui-Jia Sun [a,1], Meng-Jie Fang [b,c,1], Lei Tang [a,1], Xiao-Ting Li [a], Qiao-Yuan Lu [a], Di Dong [b,c,**], Jie Tian [b,d,e,**], Ying-Shi Sun [a,*]

[a] Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Radiology, Peking University Cancer Hospital & Institute, Beijing, 100142, China
[b] CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China
[c] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, 100049, China
[d] Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine, Beihang University, Beijing, 100191, China
[e] Engineering Research Center of Molecular and Neuro Imaging of Ministry of Education, School of Life Science and Technology, Xidian University, Xi'an, Shaanxi, 710126, China

ARTICLE INFO

Keywords:
Stomach neoplasms
Multi-detector computed tomography
Radiomics
Deep learning

ABSTRACT

Purpose: This work aimed to develop and validate a deep learning radiomics model for evaluating serosa invasion in gastric cancer.

Materials and Methods: A total of 572 gastric cancer patients were included in this study. Firstly, we retrospectively enrolled 428 consecutive patients (252 in the training set and 176 in the test set I) with pathological confirmed T3 or T4a. Subsequently, 144 patients who were clinically diagnosed cT3 or cT4a were prospectively allocated to the test set II. Histological verification was based on the surgical specimens. CT findings were determined by a panel of three radiologists. Conventional hand-crafted features and deep learning features were extracted from three phases CT images and were utilized to build radiomics signatures via machine learning methods. Incorporating the radiomics signatures and CT findings, a radiomics nomogram was developed via multivariable logistic regression. Its diagnostic ability was measured using receiver operating characteristiccurve analysis.

Results: The radiomics signatures, built with support vector machine or artificial neural network, showed good performance for discriminating T4a in the test I and II sets with area under curves (AUCs) of 0.76−0.78 and 0.79−0.84. The nomogram had powerful diagnostic ability in all training, test I and II sets with AUCs of 0.90 (95 % CI, 0.86−0.94), 0.87 (95 % CI, 0.82−0.92) and 0.90 (95 % CI, 0.85−0.96) respectively. The net reclassification index revealed that the radiomics nomogram had significantly better performance than the clinical model (p-values < 0.05).

Conclusions: The deep learning radiomics model based on CT images is effective at discriminating serosa invasion in gastric cancer.

1. Introduction

Gastric cancer, a common malignancy with a poor prognosis, is listed as the second leading cause of cancer mortality worldwide [1]. Accurate staging plays an important role in determining the appropriate management of patients and the prediction of postoperative survival [2–4].

Abbreviation: MDCT, multidetector computed tomography; AUC, areaunder curve; ROC, receiver operating characteristic curve; SVM, support vector machine; ANN, artificial neural network; NRI, Netre classification index; ROI, regions of interest; AIC, akaike information criterion; DCNN, deep convolutional neural networks; ICC, intra-/inter-class correlation coefficients; RBF, radial basis function.

* Corresponding author.
** Corresponding authors at: CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China.
E-mail addresses: sunruijia328@163.com (R.-J. Sun), fangmengjie2015@ia.ac.cn (M.-J. Fang), terrytang78@163.com (L. Tang), 13520120308@163.com (X.-T. Li), luqiaoyuan85@163.com (Q.-Y. Lu), di.dong@ia.ac.cn (D. Dong), jie.tian@ia.ac.cn (J. Tian), sys27@163.com (Y.-S. Sun).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ejrad.2020.109277
Received 14 April 2020; Received in revised form 14 July 2020; Accepted 7 September 2020
Available online 11 September 2020

For the pretreatment baseline cases, serosa positive is considered to increase the risk of micrometastases and peritoneal dissemination [3]. In the present study, the good control of micrometastases is considered to be one of the effects of neoadjuvant chemotherapy which could contribute to the decrease of peritoneal recurrence and improve long-term survival of local advanced gastric cancer [4,5]. Furthermore, the accurate assessment of pretreatment serosa invasion is a prerequisite to select patients in terms of avoiding overtreatment [6].

In clinical practice, noninvasive imaging modalities such as multidetector computed tomography (MDCT) play an important role in the diagnosis and staging of gastric malignancies [7,8]. Previous study compared the overall accuracy of T staging via endoscopic ultrasonography (EUS) and MDCT, which was 75 % and 77 %, respectively [9]. However, EUS is only used to detect intracavitary lesions. For clinical T4a (cT4a) gastric cancer staging, the reported diagnostic sensitivity of EUS was unacquirable [9]. MDCT is the most effective way to detect lesions that break through the serosa [7,9,10].

However, the differentiation of serosa invasion remains a diagnostic dilemma for radiologists. CT accuracy for preoperative T staging was reported to be 75 %–94 % [7,10,11]. But for cT4a gastric cancer staging, the reported diagnostic accuracy was about 77 % [9]. The individual perspective and clinical experience of radiologists could interpret the difference.

Radiomics analysis of large imaging datasets has been successfully employed in the field of oncology for noninvasively profiling tumor heterogeneity and there is a growing interest within the field in exploring the associations between tumor heterogeneity and imaging features [12–16]. Meanwhile, the novel deep learning techniques have shown the promising capabilities in principal representation learning and extraction of correlative features without human intervention [17]. To our knowledge, few studies [18,19] have explored a radiomics-based classification of different gastric tumors. We hypothesized that deep learning could potentially add valuable information to diagnosis by capturing more features beyond a visual interpretation. In this present work, we investigate the value of deep learning radiomics analysis for differentiating T3 and T4a stage gastric cancers.

2. Materials and methods

2.1. Patients

This study contained two parts, including phase I retrospective study and phase II validation study. The flowchart is shown in Fig. 1.

The phase I retrospective study was approved by the institutional review board of our institution, and the requirement of informed consent was waived. A total of 428 consecutive patients who underwent MDCT scan prior to surgical treatment from January 2009 to December 2015 were enrolled retrospectively according to the following inclusion and exclusion criteria. The inclusion criteria were patients who had pathological proven pT3 or pT4 gastric adenocarcinoma. The exclusion criteria were patients who (a) had history of previous chemoradiation; (b) had history of previous gastrectomy or endoscopic surgery; (c) had not undergone gastrectomy within 2 weeks after preoperative CT.

To make the sample size of training set enough for a relative sufficient model development, while to ensure high power and generalisability of test results, we allocated the patients to the training set and test set I according to the time of preoperative CT in a 2:1 ratio by January 1 st 2014. The first 252 patients were allocated to the training set and the subsequent 176 patients were allocated to the test set I.

The phase II validation study was approved by our institution review board, and informed consent was obtained from all patients. Patients who were scheduled to undergo preoperative CT and were clinical diagnosed cT3 or cT4a between January 2016 and January 2017 were eligible to be prospectively included in this study. The exclusion criteria were the same as phase I. A total of 144 consecutive patients who met the study criteria were recruited and allocated to the test set II. The pathological results were not known until the gastrectomy was undertaken. The patient enrolment flowchart of the phase II validation study is shown in Fig. 2. The clinical characteristics of all patients are shown in Table 1.

2.2. Histological evaluation

The pathological features of gastric cancer were officially reported by a specialist pathologist with 15 years of experience in gastrointestinal pathology. Pathological T stage was determined using the pathological findings of surgical specimens as reference standards according to the

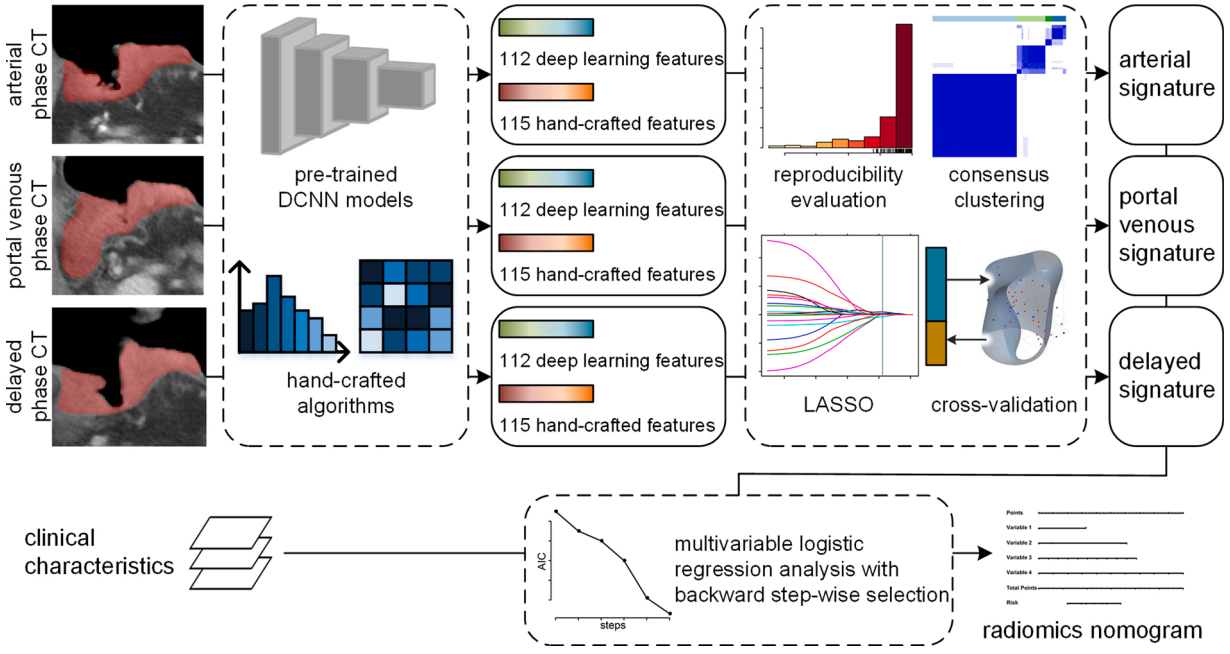

Fig. 1. Radiomics modeling and analysis workflow in this study.

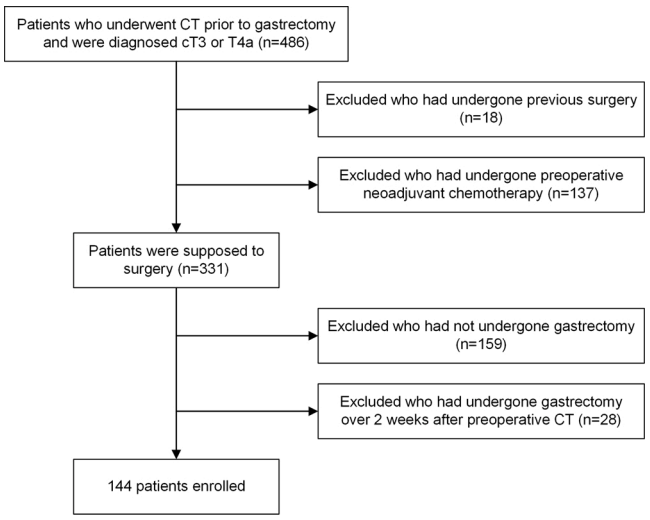

Fig. 2. Flowchart of the phase II validation study profile based on inclusion criteria.

Table 1
Clinicopathological characteristics of patients in the three sets.

	Training set	Test set I	Test set II
Age (years)	64.4 ± 12.1	61.5 ± 10.7	59.2 ± 10.7
Gender			
Male	176	118	103
Female	76	58	41
Pathological T stage			
T2	0	0	8
T3	89	94	73
T4	163	82	63
Location			
Fundus	118	52	37
Body	70	62	69
Antrum	56	51	30
Multiple	8	11	8

eighth edition of the AJCC and the UICC TNM classification [20]. The tumor locations were recorded as fundus, body and antrum. When the tumor occupied more than two areas, the pathological T stage was determined according to the deepest invasion of the tumor.

2.3. CT protocol

Each patient had been fasting prior to the CT examination for more than 8 h. To reduce gastric motility and enable gastric distention, the patients were given an intramuscular injection of 10 mg Anisodamine 10–15 min before the examination and received 8 g gas-producing crystals orally with 10 mL of water orally shortly before CT scanning. Upper abdominal unenhanced CT scans from the diaphragmatic domes to 2 cm below the lower margin of the air-distended gastric body were acquired. Subsequently, the scans were started after intravenous injection of non-ionic contrast material (1.5 ml/kg body weight; Omnipaque 300, GE Healthcare) at a rate of 3 ml/s by a high pressure injector via the antecubital vein. The following scan parameters were used: tube peak voltage 120–140 KV, tube current 300 mAs, collimation thickness 1.25 mm, helical pitch 1.5:1. Contrast-enhanced CT scans were performed in the arterial phase (start delay, 30 s), in the portal venous phase (70 s), in the delayed phase (180 s).

2.4. CT finding evaluation

CT images were analyzed by 2 radiologists (with 6 and 15 years of experience in abdominal imaging, respectively), both were blinded to the histopathology results. For the cases (n = 93, **Table S1**) with discrepancy in the subjective evaluation between the two readers, these were reviewed by a third expert (with over 20 years of experience in abdominal imaging) to make an arbitration for further analysis.

In the paper written by Kim et al., the association between the high density outer layer of the gastric wall and the involvement of the serosa was hypothesized and defined as hyperattenuating serosa sign [21]. The high enhanced serosa sign on CT was defined as a focal or diffuse thickened hyperattenuating outer layer of the gastric wall. For all patients, the clinical stages of T3 and T4, as well as the following conventional CT features were recorded: the extraluminal extension, including nodular extension and cord irregular outer layer, and the blurring and obliteration of the fat plane [2]. The extraluminal extension was defined as the irregular outer layer of the gastric wall. The blurring and obliteration of the fat plane was defined as linear or reticular structures in the fatty layer surrounding the cancerous lesion exhibiting a lack of a clear perigastric fat plane. Examples of these CT findings were shown in Fig. 3**a-c**.

2.5. Image segmentation

CT images were retrieved from the picture archiving and communication system (PACS) and then loaded into ITK-SNAP software (version 3.6.0; www.itksnap.org) for manual segmentation. A radiologist (reader 1) outlined the regions of interest (ROIs) on CT images of the three phases respectively. For each CT phase, the radiologist selected the slice with the largest tumor area and a two-dimensional ROI was then delineated using a margin that included all tumor area (Fig. 4**a-f**). Therefore, there were three different ROIs being segmented for each patient in this study.

ROIs were drawn along the contour of the tumor carefully to avoid involving adjacent fluid or air. If a high mixed-intensity cord like extragastric extension was found, the ROIs were drawn with contouring of the abnormal density. If low mixed-density was found (<20HU) such as fat blurring and exhibiting a lack of a clear perigastric fat plane, the ROIs were drawn avoiding fat extension(Fig. 4f). When the tumor occupied more than two areas, all the region were selected and segmented. After over three months, 50 patients in the training set were selected randomly. Then, the reader 1 and another radiologist (reader 2) segmented their ROIs and evaluated their CT findings again to assess intra-/inter-reader agreement of the radiomics analysis.

2.6. Radiomics feature extraction

Both deep learning features and hand-crafted features were extracted to quantify the tumor phenotype (Fig. 1). We referred to the Image Biomarker Standardisation Initiative (IBSI) and Radiomic Ontology [22, 23] to standardize the algorithms of feature extraction. For each CT phase, 112 deep learning features and 115 hand-crafted features were extracted.

We constructed and trained three deep convolutional neural networks (DCNNs) which all contained 6 weighted layers to extract deep learning features on the three groups of ROIs (**Fig. S1**). Note that, in order to avoid over-fitting of our model, we trained the DCNN in an unsupervised manner, and used the hierarchical convolutional structure as feature extraction network. The architecture and implementation of our DCNNs and the feature extraction pipeline were described in the **Supplemental material** in detail.

2.7. Feature selection and radiomics signature building

We built three radiomics signatures reflecting the phenotypic characteristics of the tumor in arterial phase, portal venous phase and delayed phase CT images respectively as independent predictors of serosa invasion, i.e. the arterial signature, the portal venous signature and the delayed signature.

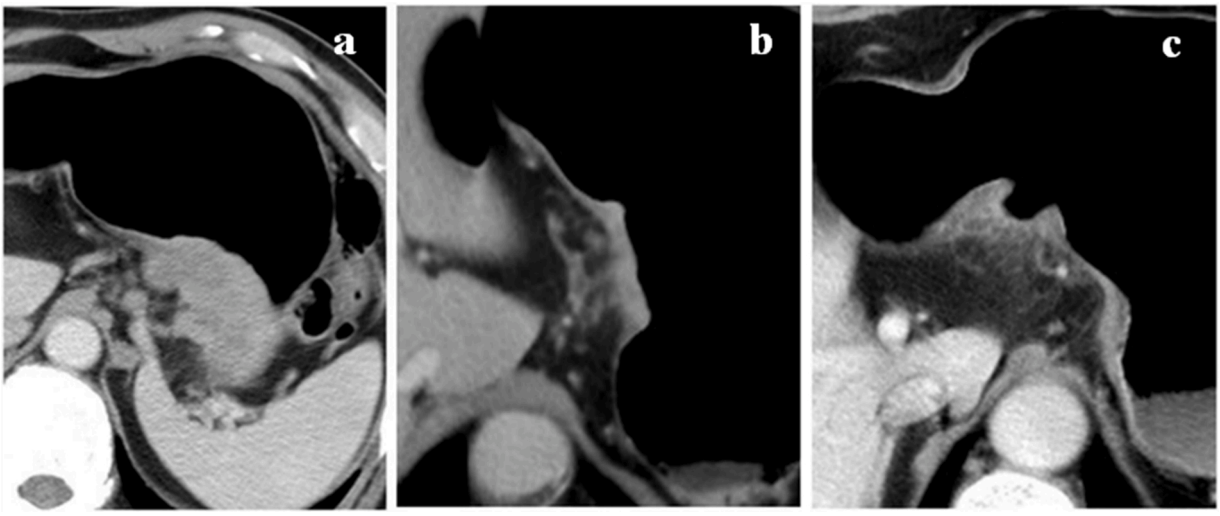

Fig. 3. Conventional CT features of serosa invasion and example of the manual segmentation. (a) Nodular extension; (b) Cord out layer; (c) Perigastric fat infiltration.

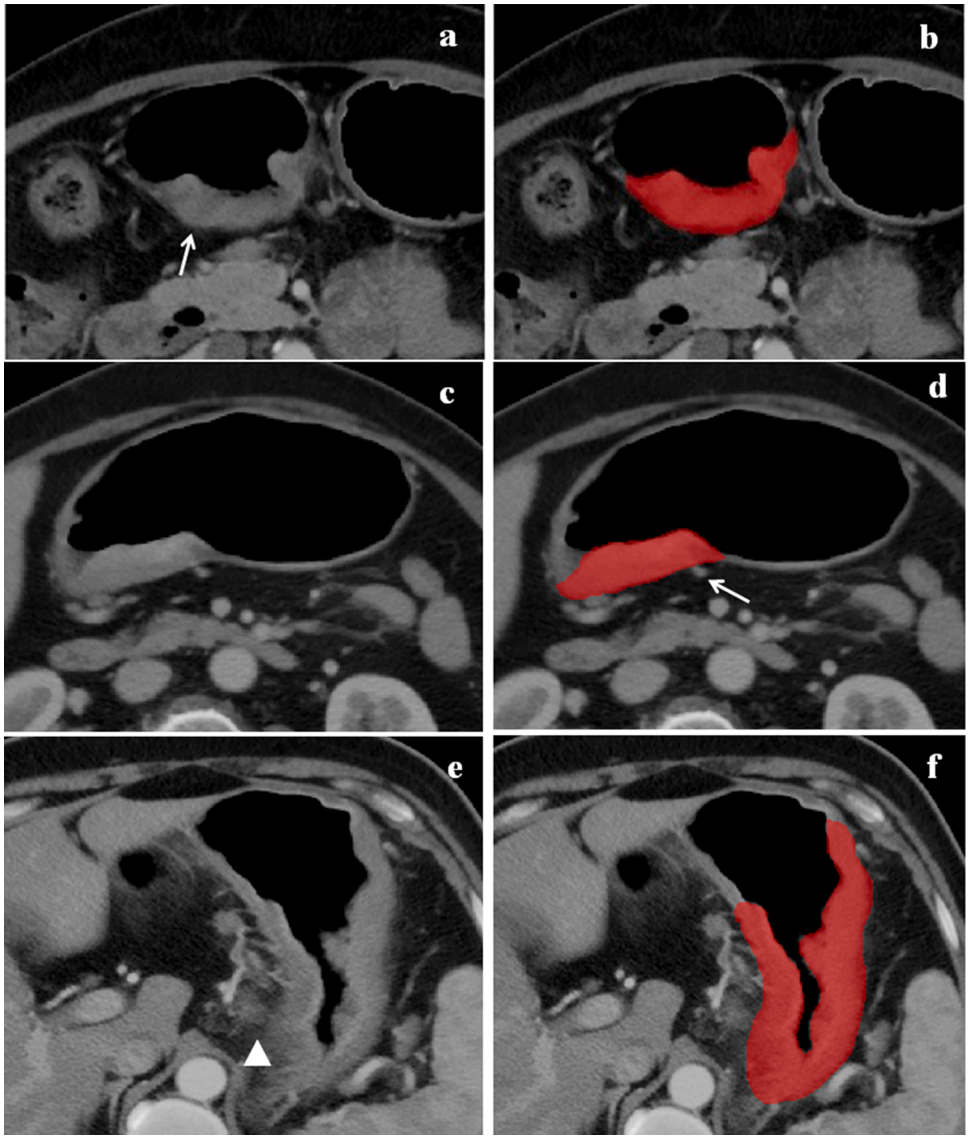

Fig. 4. Examples of the manual segmentation. (a) (c) (e) Diffusely infiltrating mass with heterogeneous enhancement; (b) (d) (f) Manual segmentations on the portal venous phase CT. ROIs were drawn along the contour of the tumor, avoiding the blood vessel (arrow) and fat infiltration (triangle).

Based on the training set, we implemented multiple processes of feature selection and signature building to avoid model over-fitting and to improve the performance (Fig. 1), as details described in **Supplemental material**. We firstly utilized the intra-/inter-class correlation coefficients (ICC) to assess feature reproducibility. Then, the consensus clustering [24,25] and the least absolute shrinkage and selection operator (LASSO) method were successively conducted to reduce the features' redundancy and to determine the candidate set of predictive features. Finally, five machine learning models were compared via cross-validation. The best performing model, as well as the best combination of features, was selected to build the signature of each CT phase.

2.8. Statistical analysis

We conducted univariate analysis to assess the differences in patient variables between the different groups using the independent *t*-test or Mann-Whitney *U* test for continuous variables, and Fisher's exact test or chi-square test for categorical variables. Multivariable logistic regression analysis was used to identify independent predictors. Backward, step-wise selection was applied by using the likelihood-ratio test with Akaike's information criterion (AIC) as the stopping rule. We built a quantitative radiomics nomogram to predict the probability of serosa invasion with the clinical characteristics and radiomics signatures, as well as a clinical model containing only the clinical characteristics. We also built two phenotypic models, which contained only the signatures incorporating hand-crafted features (phenotypic model 1) and only the signatures incorporating deep learning features (phenotypic model 2) respectively, for the purpose of comparison. The collinearity of each variable in the regression model was checked using variance inflation factor (VIF). The variables with VIF > 5 were excluded.

The radiomics nomogram was assessed via the area under curve (AUC) of the receiver operating characteristic (ROC) curve, the calibration curve with the Hosmer-Lemeshow test and accuracy. Delong test was used to compared different ROC curves. Net reclassification index (NRI) was used to evaluate the improvement of predictive performances.

To assess the association of radiomics nomogram with serosa invasion in different clinical subgroups, a stratified analysis was presented by the gender, age, tumor location and CT system on the patients of test sets. The 50 selected patients, with the re-segmented ROIs and the re-evaluated CT findings, were utilized to determine the intra-/inter-reader agreement of radiomics nomogram via the Kappa test. Furthermore, to evaluate the reproducibility of our model's diagnostic performance, we repeated the randomized assignment of training/test sets 10 times. Subsequently, the model was re-trained and validated repeatedly.

The softwares used in this study are reported in the **Supplemental material**.

3. Results

3.1. Clinical characteristics of the patients

Demographic data and CT findings in the training set (n = 252), test I set (n = 176) and test II set (n = 144) are listed in Table 2. The median (range) ages of the three sets were 64 (29–90) years, 62 (30–86) years and 60 (29–85) years respectively, and the proportion of females were 30.2 %, 33.0 % and 28.5 % respectively. cT, nodular extension, cord out layer, perigastric fat infiltration and the high enhanced serosa sign differed significantly between T3 group and T4 group in the training set (p-values < 0.05).

3.2. Radiomics signature building and validation

As presented in **Table S2-S4**, most of radiomics features (132 arterial features, 152 portal venous features and 113 delayed features) demonstrated to have a good intra-reader agreement. In the consensus clustering step, 40, 58 and 33 distinct clusters were obtained respectively (**Fig. S2**) and only the medoid features were retained. The feautres were further reduced to 18, 15 and 14 predictive candidate features, which had non-zero coefficients in the LASSO logistic regression (**Fig. S3**). Based on these features, we built signatures using five kinds of models and validated them through cross-validation in the training set. For arterial phase CT, the best model was artificial neural network (ANN) with five hidden nodes, and it yielded a cross-validation accuracy of 0.706. For portal venous phase CT, radial basis function (RBF) kernel support vector machine (SVM) with the penalty parameter of 1.41 and the kernel coefficient of 0.22 yielded the highest cross-validation accuracy of 0.725. For delayed phase CT, yielding a cross-validation accuracy of 0.726, the best model was ANN with seven hidden nodes. The detailed results of cross-validation are listed in **Table S5**. Therefore, three radiomics signatures were built using the two different kinds of models, and as listed in **Table S6**, they had six, four and three input features respectively. The experimental results demonstrated the three

Table 2
Univariate analysis of clinical characteristics and CT findings.

Characteristic	Training set				Test set I				Test set II			
	Total	T3	T4	p-value	Total	T3	T4	p-value	Total	T2/T3	T4	p-value
Age(years)	64.4 ± 12.1	64.9 ± 10.6	64.1 ± 12.9	0.618	61.5 ± 10.7	61.0 ± 10.6	62.0 ± 10.8	0.742	59.2 ± 10.7	58.8 ± 10.2	59.7 ± 11.4	0.634
Gender				0.232				0.201				0.254
Male	176	58	118		118	67	51		103	61	42	
Female	76	31	45		58	27	31		41	20	21	
cT stage				<0.001				<0.001				0.001
3	69	43	26		72	52	20		72	52	20	
4	183	46	137		104	42	62		72	29	43	
Nodular extension				<0.001				<0.001				0.001
negative	126	62	64		119	77	42		104	67	37	
positive	126	27	99		57	17	40		40	14	26	
Cord out layer				<0.001				0.001				0.001
negative	66	41	25		52	38	14		65	46	19	
positive	186	48	138		124	56	68		79	35	44	
Fat infiltration				<0.001				<0.001				<0.001
negative	105	59	46		103	74	29		68	51	17	
positive	147	30	117		73	20	53		76	30	46	
High enhanced serosa sign				<0.001				<0.001				<0.001
negative	128	68	60		112	78	34		102	68	34	
positive	124	21	103		64	16	48		42	13	29	

signatures yielded good performance for discriminating T4 stage in the test set I and II with AUCs of 0.76−0.78 and 0.79−0.84 (**Table S6**).

3.3. Development and validation of radiomics nomogram

Beginning with the three radiomics signatures and the clinical characteristics, multivariate logistic regression analysis with stepwise backward elimination was applied. The results determined that the radiomics signatures, nodular extension, fat infiltration and high enhanced serosa sign remained significant after adjustment for cofactors (**Table S7**). A radiomics nomogram was built by using the above regression coefficients as a quantitative method for noninvasive T4 stage prediction (Fig. 5). Meanwhile, the clinical model incorporated cT stage, nodular extension, fat infiltration and high enhanced serosa sign (**Table S8**). In regression analysis, the radiomics nomogram yielded a relative low AIC (Table 3), suggesting it had a good global model fit.

The radiomics nomogram showed good discriminationon all training, test I and II sets with AUCs of 0.90 (95 % CI, 0.86−0.94), 0.87 (95 % CI, 0.82−0.92) and 0.90 (95 % CI, 0.85−0.96) respectively (Fig. 6**a-c**). The Delong test was implemented on the ROC curves to assess possible overfitting and revealed that the differences were not statistically significant between AUCs on the three sets with p-values > 0.408. The radiomics nomogram yielded the highest accuracies (range 0.80−0.85, Table 3) and the best calibration to the overall risk (**Fig. S4**) among the four comparative models on the test sets. Furthermore, the NRI revealed that it had significantly better performance than the clinical model and the phenotypic model 1 (NRI, range 0.140−0.240; p-values < 0.008). Built based on the deep learning features, the phenotypic model 2 achieved somewhat weaker diagnostic ability than the radiomics nomogram, and yielded a relative low accuracy on the test set I. The nomogram calibration curve demonstrated good agreement between prediction and observation in the three sets (Fig. 6d). The Hosmer-Lemeshow test was not significant (p-values = 0.827, 0.119 and 0.198), which suggested that there was no departure from a perfect fit. The stratified analysis is shown in **Supplemental material**, which indicates the performance of radiomics nomogram was not affected by patient age, gender, tumor location and CT system (Delong test p-values > 0.05) (**Fig. S5**), and therefore implies its generalization on various kinds of cases. The Kappa test showed good intra-/inter-reader agreement for our radiomics nomogram with values of 0.720 and 0.688, respectively. Furthermore, we split the whole data set into paired training (50 %) and test (50 %) sets 10 times, followed by the repeating construction and validation of the classification model. In this experiment, there was no significant difference found between the resulted AUCs which ranged from 0.847 to 0.910 in the holdout test sets.

4. Discussion

Our study demonstrated the potential use for CT-based radiomics signature to act as a reproducible imaging marker for an accurate differentiation of serosa invasion. In order to facilitate further research and develpment, we made the developed models open access in our website (www.radiomics.net.cn/platform.html).

Most of the CT criteria for T staging of gastric cancer were based on a multilayered pattern of the gastric wall. The ability of CT to differentiate between the layers of the gastric wall is limited, especially when the tumor invades serosa. A previous study indicated that the CT diagnostic accuracy for serosa positive gastric cancer was about 77 % [9], which was a lower value than the test results (test set I: 80 %; test set II: 85 %) obtained in our study.

This study applies not only traditional radiomics methods, but also incorporates features extracted from deep learning neural networks. The method of deep learning in the field of computer vision has become increasingly mature, providing a new opportunity for automated analysis of medical images and assisting doctors in realizing high-accuracy intelligent diagnosis of diseases [26,27]. Most of them use DCNN models or other commonly used deep learning algorithms, or several merging methods for image classification detection. The use of deep learning methods can provide a lot of convenience for image classification.

The limited sample sizes in medical image analysis studies could introduce bias in the results of feature selection, model building and model evaluation and therefore affect the performance and generalizability of the developed systems [28,29]. In this study, we designed and conducted the strategies of multiple feature selection and cross-validation to construct the relative optimal predictive models based on our training set. Although the satisfactory accuracies were obtained on the two independent test sets, it should be noted that the power of ROC curve and calibration curve analyses would be improved if further larger multicenter datasets could be used validate the model.

According to the literature [30,31], the growth, invasion and metastasis of gastric cancer mainly depend on angiogenesis and could be evaluated using the perfusion CT. Different from these previous researches, in which the radiologists measured the image characteristics (e.g., blood flow and blood volume) from perfusion CT and directly evaluated their correlation with other clinical markers, we extracted the features from the contrast-enhanced CT images and combined them with CT findings to develop a predictive model for serosa invasion. The contrast-enhanced CT could achieve a good visualization of the tumor vascularity and reveal enhancement heterogeneity. In the arterial phase, the mucosa at the lesion presented as a focal enhanced line. In the portal

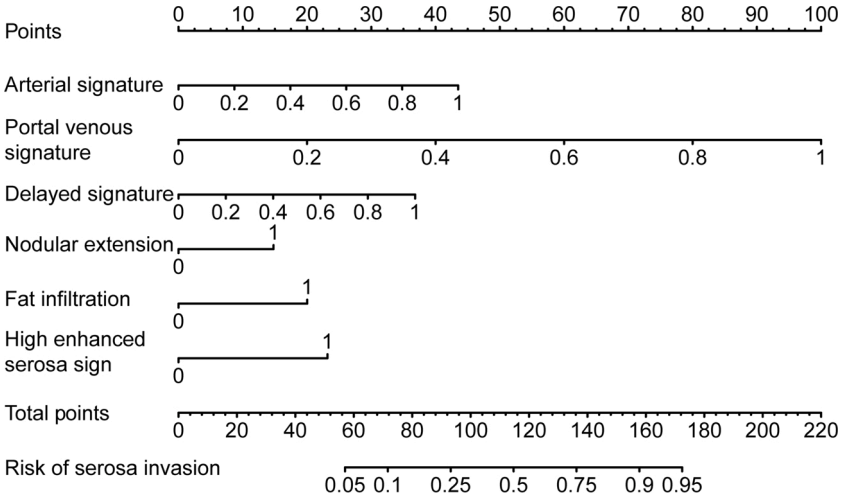

Fig. 5. Radiomics nomogram based on radiomics signatures and clinical characteristics.

Table 3
Predictive performances of each model on the training and test sets.

Models	AIC	Training set				Test set I				Test set II			
		AUC (95 % CI)	Acc (95 %CI)	Sen	Spe	AUC (95 % CI)	Acc (95 %CI)	Sen	Spe	AUC (95 % CI)	Acc (95 %CI)	Sen	Spe
Radiomics nomogram	288.7	0.90 (0.86−0.94)	0.81 (0.76−0.86)	0.80	0.84	0.87 (0.82−0.92)	0.80 (0.73−0.85)	0.73	0.85	0.90 (0.85−0.96)	0.85 (0.78−0.90)	0.75	0.93
Clinical model	388.1	0.79 (0.73−0.84)	0.72 (0.66−0.77)	0.74	0.69	0.80 (0.73−0.86)	0.73 (0.66−0.79)	0.63	0.81	0.78 (0.70−0.85)	0.73 (0.65−0.80)	0.62	0.81
Phenotypic model 1	392.6	0.78 (0.72−0.84)	0.69 (0.63−0.75)	0.66	0.75	0.76 (0.68−0.83)	0.68 (0.60−0.74)	0.63	0.71	0.80 (0.72−0.87)	0.72 (0.64−0.79)	0.67	0.77
Phenotypic model 2	364.4	0.82 (0.77−0.88)	0.74 (0.68−0.79)	0.71	0.80	0.84 (0.78−0.90)	0.78 (0.71−0.84)	0.76	0.80	0.88 (0.82−0.93)	0.81 (0.73−0.87)	0.65	0.93

NOTE. Acc, accuracy; Sen, sensitivity; Spe, specificity.

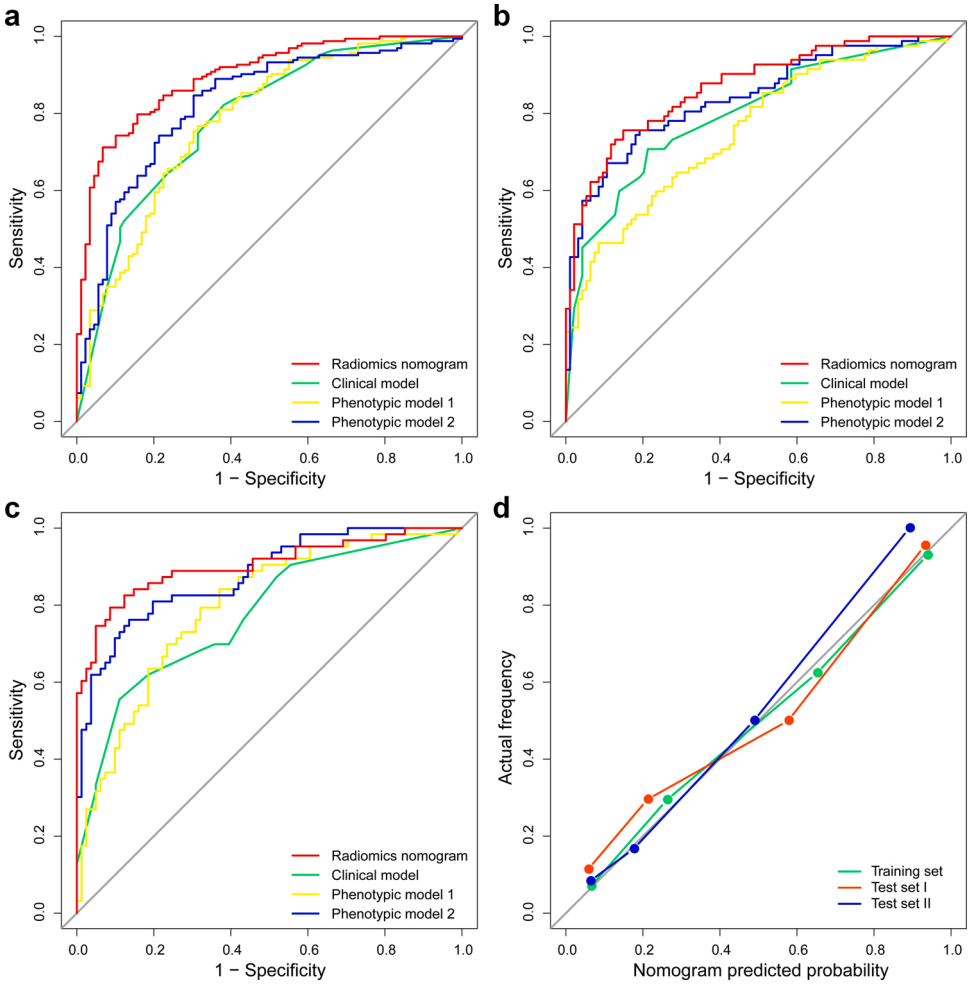

Fig. 6. ROC curves of the four predicted models on (a) the training set, (b) the test set I and (c) the test set II, and (d) calibration curve of the radiomics nomogram.

venous phase, the markedly increased interstitial fibrous tissues reduced the flow-out speed of the contrast media, resulting in well-proportioned enhancement of the lesion. Thus, the arterial phase can reflect the blood supply and functional capillary density of gastric cancer and the portal venous phase may reflect more dysfunctional neo-vessels and represent the distribution of contrast media in interstitial spaces [32], which might account for different diagnosis performances between different phases. Advanced gastric cancer is characterized by gradual enhancement from the mucosa to the serosa. The enhancement peak is achieved in the delayed phase. If no enhancement is observed in the serosa in the delayed phase, it cannot be considered that the tumor has extramural infiltration even if the fat around the gastric wall is blurred or disappeared. Therefore, multiphase enhancement is beneficial to the identification of tumor wall infiltration.

In the previous studies, most of the models had not been applied to the clinic. If the prediction performance of the model on the test set is better, it shows that it has the potential to be applied in clinical diagnosis. If the model is even significant in the prospective data, it indicate clinical predictive efficacy. Blind to the postoperative pathological results, we prospectively collected the test set II, and validated the performance of our models on it.

There were some limitations of the study. First, we did not perform a 3D image segmentation. In our study, ROIs of gastric cancers were drawn on the largest axial section and were drawn along the contour of

the tumor as visualized on each phase (arterial, portal venous, and delayed phase) of CT. The influence of 2D vs. 3D segmentation to the model should be further studied. Second, our method needed manually segmentation of the tumor, which was time consuming. The automated segmentation method, as well as the fully automated model, is valuable to be developed in the future. Third, all the patients were from a single center, the model may perform differently if multicenter datasets with different parameters are used. A much larger dataset from multiple centers, with the prognosis data ought to be investigated to validate the reproducibility of our radiomics model.

In conclusion, we presented a CT-based deep learning radiomics model for discriminating serosa invasion of advanced gastric cancer, which is discriminative to yield good diagnostic accuracy.

Author contribution

Conception and design: Ying-Shi Sun, Jie Tian, Di Dong
Collection and assembly of data: Rui-Jia Sun, Lei Tang, Qiao-Yuan Lu
Data analysis and interpretation: Meng-Jie Fang, Rui-Jia Sun, Xiao-Ting Li
Manuscript writing: Meng-Jie Fang, Rui-Jia Sun, Lei Tang
Final approval of manuscript: All authors.

Funding

This study was supported by Beijing Municipal Administration of Hospitals Clinical Medicine Development of Special Funding Support (No.ZYLX201803), 'Beijing Hospitals Authority' Ascent Plan, (Code: DFL20191103), National Natural Science Foundation of China (81971584, 91959116, 91959130, 81971776, 81771924, 81227901, 82022036), National Key R&D Program of China (2019YFC0117705, 2017YFC1309101, 2017YFC1309104), the Bureau of International Cooperation of Chinese Academy of Sciences (173211KYSB20160053), and the Youth Innovation Promotion Association CAS (2017175).

Declaration of Competing Interest

The authors declare no conflicts of interest.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ejrad.2020.109277.

References

[1] A. Jemal, F. Bray, M.M. Center, J. Ferlay, E. Ward, D. Forman, Global cancer statistics, CA Cancer J. Clin. 61 (2011) 69–90, https://doi.org/10.3322/caac.20107.
[2] T. Yoshikawa, M. Sasako, S. Yamamoto, et al., Phase II study of neoadjuvant chemotherapy and extended surgery for locally advanced gastric cancer, Br. J. Surg. 96 (2009) 1015–1022, https://doi.org/10.1002/bjs.6665.
[3] M. Yano, T. Yasuda, Y. Fujiwara, S. Takiguchi, H. Miyata, M. Monden, Preoperative intraperitoneal chemotherapy for patients with serosa-infiltrating gastric cancer, J. Surg. Oncol. 88 (2004) 39–43, https://doi.org/10.1002/jso.20133.
[4] Y. Yonemura, I. Ninomiya, M. Kaji, et al., Prophylaxis with intraoperative chemohyperthermia against peritoneal recurrence of serosal invasion-positive gastric cancer, World J. Surg. 19 (1995) 450–454, https://doi.org/10.1007/bf00299188.
[5] K. Ott, F. Lordick, K. Herrmann, B.J. Krause, C. Schuhmacher, J.R. Siewert, The new credo: induction chemotherapy in locally advanced gastric cancer: consequences for surgical strategies, Gastric Cancer 11 (2008) 1–9, https://doi.org/10.1007/s10120-007-0448-1.
[6] N. Boku, A. Ohtsu, Y. Shimada, et al., Phase II study of a combination of irinotecan and cisplatin against metastatic gastric cancer, J. Clin. Oncol. 17 (1) (1999) 319–323, https://doi.org/10.1200/JCO.1999.17.1.319.
[7] A. Ba-Ssalamah, M. Prokop, M. Uffmann, P. Pokieser, B. Teleky, G. Lechner, Dedicated multidetector CT of the stomach: spectrum of diseases, Radiographics 23 (3) (2003) 625–644, https://doi.org/10.1148/rg.233025127.
[8] J.Y. Kim, S.H. Kim, J.M. Lee, J.K. Han, J.M.Y. Lee, B.I. Choi, Differentiating malignant from benign wall thickening in postoperative stomach using helical computed tomography: results of multivariate analysis, J. Comput. Assist. Tomogr. 31 (3) (2007) 455–462, https://doi.org/10.1097/01.rct.0000243454.15684.e2.
[9] S.W. Hwang, D.H. Lee, S.H. Lee, et al., Preoperative staging of gastric cancer by endoscopic ultrasonography and multidetector-row computed tomography, J. Gastroenterol. Hepatol. 25 (2010) 512–518, https://doi.org/10.1111/j.1440-1746.2009.06106.x.
[10] D. Tsurumaru, M. Miyasaka, Y. Nishimuta, et al., Differentiation of early gastric cancer with ulceration and resectable advanced gastric cancer using multiphasic dynamic multidetector CT, Eur. Radiol. 26 (2016) 1330–1337, https://doi.org/10.1007/s00330-015-3938-2.
[11] J.W. Kim, S.S. Shin, S.H. Heo, et al., Diagnostic performance of 64-section CT using CT gastrography in preoperative T staging of gastric cancer according to 7th edition of AJCC cancer staging manual, Eur. Radiol. 22 (2012) 654–662, https://doi.org/10.1007/s00330-011-2283-3.
[12] H.J. Aerts, E.R. Velazquez, R.T. Leijenaar, et al., Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach, Nat Commun Jun 3 (5) (2014) 4006, https://doi.org/10.1038/ncomms5006.
[13] P. Lambin, R.T. Leijenaar, T.M. Deist, et al., Radiomics: the bridge between medical imaging and personalized medicine, Nat. Rev. Clin. Oncol. 14 (12) (2017) 749–762, https://doi.org/10.1038/nrclinonc.2017.141.
[14] J. Song, J. Shi, D. Dong, et al., A new approach to predict progression-free survival in stage IV EGFR-mutant NSCLC patients with EGFR-TKI therapy, Clin. Cancer Res. 24 (15) (2018) 3583–3592, https://doi.org/10.1158/1078-0432.CCR-17-2507.
[15] Y.Q. Huang, C.H. Liang, L. He, et al., Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer, J. Clin. Oncol. 34 (18) (2016) 2157–2164, https://doi.org/10.1200/JCO.2015.65.9128.
[16] D. Dong, F. Zhang, L.Z. Zhong, et al., Development and validation of a novel MR imaging predictor of response to induction chemotherapy in locoregionally advanced nasopharyngeal cancer: a randomized controlled trial substudy (NCT01245959), BMC Med. 23 (1) (2019) 190, https://doi.org/10.1186/s12916-019-1422-6.
[17] S. Wang, J. Shi, Z. Ye, et al., Predicting EGFR mutation status in lung adenocarcinoma on CT image using deep learning, Eur. Respir. J. 53 (3) (2019) 1800986, https://doi.org/10.1183/13993003.00986-2018.
[18] A. Ba-Ssalamah, D. Muin, R. Schernthaner, et al., Texture-based classification of different gastric tumors at contrast-enhanced CT, Eur. J. Radiol. 82 (10) (2013) 537–543, https://doi.org/10.1016/j.ejrad.2013.06.024.
[19] D. Dong, L. Tang, Z.Y. Li, et al., Development and validation of an individualiZ.Y. eD nomogram to identify occult peritoneal metastasis in patients with advanced gastric cancer, Ann. Oncol. 30 (3) (2019) 431–438, https://doi.org/10.1093/annonc/mdz001.
[20] T. Jäger, D. Neureiter, R. Urbas, et al., Applicability of American joint committee on cancer and college of american pathologists regression grading system in rectal cancer, Dis. Colon Rectum 60 (8) (2017) 815–826, https://doi.org/10.1097/DCR.0000000000000806.
[21] T.U. Kim, S. Kim, J.W. Lee, et al., MDCT features in the differentiation of T4a gastric cancer from less-advanced gastric cancer: significance of the hyperattenuating serosa sign, Br. J. Radiol. 86 (1029) (2013) 20130290, https://doi.org/10.1259/bjr.20130290.
[22] A. Zwanenburg, S. Leger, M. Vallières, S. Löck, Results from the image biomarker standardisation initiative, Radiother. Oncol. 127 (2018) 543–544, https://doi.org/10.1016/S0167-8140(18)31291-X.
[23] A. Traverso, Radiomics Ontology, 2017. https://bioportal.bioontology.org/ontologies/RO.
[24] M.D. Wilkerson, D.N. Hayes, Consensus cluster plus: a class discovery tool with confidence assessments and item tracking, Bioinformatics 26 (12) (2010) 1572–1573, https://doi.org/10.1093/bioinformatics/btq170.
[25] C. Parmar, R.T. Leijenaar, P. Grossmann, et al., Radiomic feature clusters and prognostic signatures specific for lung and head & neck cancer, Sci. Rep. 5 (2015) 11044, https://doi.org/10.1038/srep11044.
[26] T.C. Chiang, Y.S. Huang, R.T. Chen, C.S. Huang, R.F. Chang, Tumor detection in automated breast ultrasound using 3-D CNN and prioritized candidate aggregation, IEEE Trans. Med. Imaging 38 (1) (2019) 240–249, https://doi.org/10.1109/TMI.2018.2860257.
[27] D. Dong, M. Fang, L. Tang, et al., Deep learning radiomic nomogram can predict the number of lymph node metastasis in locally advanced gastric cancer: an international multicenter study, Ann. Oncol. 31 (7) (2020) 912–920, https://doi.org/10.1016/j.annonc.2020.04.003.
[28] K. Nie, L. Shi, Q. Chen, et al., Rectal cancer: assessment of neoadjuvant chemoradiation outcome based on radiomics of multiparametric MRI, Clin. Cancer Res. 22 (21) (2016) 5256–5264, https://doi.org/10.1158/1078-0432.CCR-15-2997.
[29] C. Parmar, D. Barry, J. Hosny, et al., Data analysis strategies in medical imaging, Clin. Cancer Res. 24 (15) (2018) 3492–3499, https://doi.org/10.1158/1078-0432.CCR-18-0385.
[30] A. Satoh, K. Shuto, S. Okazumi, et al., Role of perfusion CT in assessing tumor blood flow and malignancy level of gastric cancer, Dig. Surg. 27 (4) (2010) 253–260, https://doi.org/10.1159/000288703.
[31] J. Yao, Z.G. Yang, H.J. Chen, T.W. Chen, J. Huang, Gastric adenocarcinoma: can perfusion CT help to noninvasively evaluate tumor angiogenesis? Abdom. Imaging 36 (1) (2011) 15–21, https://doi.org/10.1007/s00261-010-9609-5.
[32] X.H. Chen, K. Ren, P. Liang, Y.R. Chai, K.S. Chen, J.B. Gao, Spectral computed tomography in advanced gastric cancer: can iodine concentration non-invasively assess angiogenesis? World J. Gastroenterol. 23 (2017) 1666–1675, https://doi.org/10.3748/wjg.v23.i9.1666.