



Machine learning analysis for the noninvasive prediction of lymphovascular invasion in gastric cancer using PET/CT and enhanced CT-based radiomics and clinical variables

Lijing Fan¹ · Jing Li² · Huiling Zhang³ · Hongkun Yin³ · Rongguo Zhang³ · Jibin Zhang¹ · Xuejun Chen²

Received: 24 July 2021 / Revised: 6 October 2021 / Accepted: 7 October 2021 / Published online: 28 January 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Purpose Lymphovascular invasion (LVI) is associated with metastasis and poor survival in patients with gastric cancer, yet the noninvasive diagnosis of LVI is difficult. This study aims to develop predictive models using different machine learning (ML) classifiers based on both enhanced CT and PET/CT images and clinical variables for preoperatively predicting lymphovascular invasion (LVI) status of gastric cancer.

Methods A total of 101 patients with gastric cancer who underwent surgery were retrospectively recruited, and the LVI status was confirmed by pathological analysis. Patients were randomly divided into a training dataset ($n = 76$) and a validation dataset ($n = 25$). By 3D manual segmentation, radiomics features were extracted from the PET and venous phase CT images. Image models, clinical models, and combined models were constructed by selected enhanced CT-based and PET-based radiomics features, clinical factors, and a combination of both, respectively. Three ML classifiers including adaptive boosting (AdaBoost), linear discriminant analysis (LDA), and logistic regression (LR) were used for model development. The performance of these predictive models was evaluated with respect to discrimination, calibration, and clinical usefulness.

Results Ten radiomics features and eight clinical factors were selected for the development of predictive models. In the validation dataset, the area under curve (AUC) values of clinical models using AdaBoost, LDA, and LR classifiers were 0.742, 0.706, and 0.690, respectively. The image models using AdaBoost, LDA, and LR classifiers achieved an AUC of 0.849, 0.778, and 0.810, respectively. The combined models showed improved performance than the image models and the clinical models, with the AUC values of AdaBoost, LDA, and LR classifier yielding 0.944, 0.929, and 0.921, respectively. The combined models also showed good calibration and clinical usefulness for LVI prediction.

Conclusion ML-based models integrating PET/CT and enhanced CT radiomics features and clinical factors have good discrimination capability, which could serve as a noninvasive, preoperative tool for the prediction of LVI and assist surgical treatment decisions in patients with gastric cancer.

Jibin Zhang and Xuejun Chen have contributed equally to this work and should be considered co-correspondent authors.

✉ Jibin Zhang
zjb12042@126.com

✉ Xuejun Chen
chenxuejun1967@163.com

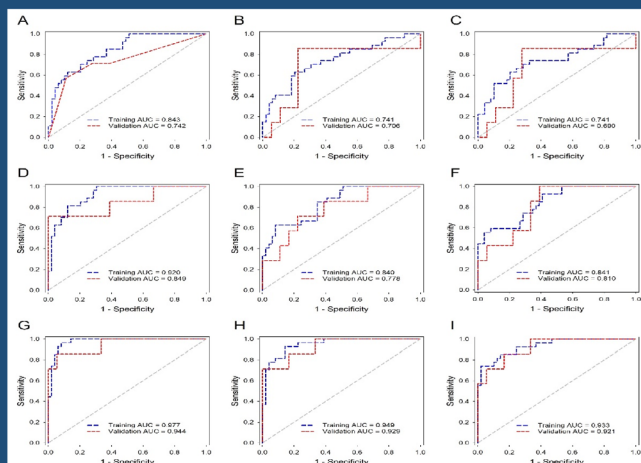
¹ Department of Radiology, Affiliated Suzhou Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Suzhou 215002, China

² Department of Radiology, Affiliated Tumor Hospital of Zhengzhou University, Henan Cancer Hospital, Zhengzhou 450008, China

³ Institute of Advanced Research, Beijing Infervision Technology Co., Ltd, Beijing, China

Graphic abstract

Machine learning analysis for the noninvasive prediction of lymphovascular invasion in gastric cancer using PET/CT and enhanced CT-based radiomics and clinical variables



Combination of enhanced CT and PET-based radiomics feature and clinical variables facilitated the preoperative prediction of LVI (Validation AUC 0.690~0.742/0.778~0.849 increased to 0.921~0.944)

Abdominal Radiology
The Official Journal of the Society of Abdominal Radiology www.abdominalradiology.org

Fan et al; 2021 DOI.....

Keywords Lymphovascular invasion · Gastric cancer · Radiomics · Machine learning · Enhanced CT · PET/CT

Abbreviations

| | |
|----------|---|
| LVI | Lymphovascular invasion |
| AUC | Area under curve |
| ROC | Receiver operating characteristics |
| LASSO | Least absolute shrinkage and selection operator |
| CTV-P | mean CT values of the venous period (CTV-P) |
| ML | Machine learning |
| AdaBoost | Adaptive boosting |
| LDA | Linear discriminant analysis |
| LR | Logistic regression |

Introduction

Gastric cancer is the fifth most common malignant tumor in the world after lung cancer, breast cancer, colorectal cancer, and prostate cancer and the third leading cause of cancer mortality in men and women [1]. Recent studies have indicated that gastric cancer is 10.79% of new cancer cases each year, and 12.80% of all cancer deaths are attributed to gastric cancer [2]. Surgery is considered the best treatment for early gastric cancer, such as endoscopic resection, laparoscopic,

or open gastrectomy with D1 or D2 lymph node resection [3], [4]. For advanced gastric cancer, surgery remains the preferred approach for resectable gastric cancer, with preoperatively and perioperatively adjuvant chemotherapy improving the efficacy [5]. Gastric cancer is characterized by high postoperative recurrence, and early gastric cancer has a recurrence rate of 1.6% [6]. Lymphovascular invasion (LVI) is defined as the lymphatic vessels and/or blood vessel invasion of malignant tumor cells within the primary tumor and surrounding tissues [7]. LVI in gastric cancer is an independent risk factor for patient prognosis, especially in advanced cancers with lymph node metastases [8]. Even after surgical resection, patients with early gastric cancer who were positive for LVI had a higher rate of recurrence and a lower overall survival rate of 3–5 years than those with a negative LVI status [9]. Moreover, LVI is also associated with lymph node metastasis and is an independent predictor of poor prognosis in patients with gastric cancer [10].

Therefore, accurate assessment of LVI was important in predicting prognosis of gastric cancer and if the status of LVI can be detected and predicted noninvasively before operation, there will aid more optimal clinical decision-making and personalized treatment for patients [11]. LVI is detected after surgery by pathological examination, and

preoperative noninvasive diagnosis of LVI is limitation relying solely on traditional clinical examinations (such as dynamic enhanced CT). Dynamic enhanced CT has been reported to be an effective tool for assessing tumor angiogenesis by quantitative enhancement measurement [12]. Fluoro18-fluorodeoxyglucose positron emission tomography/computed tomography (PET/CT), as a new quantitative and non-invasive imaging technology, uses 18F-FDG as a tracer to reflect the differences in metabolic levels of tissues in vivo through the differences in uptake distribution of tracers in different tissues [13]. Enhanced CT or PET/CT can be further quantitatively analyzed by radiomics with machine-learning approaches [14]. Recently, several applications of radiomics in gastric cancer have been reported in the prediction of histopathological characteristics [15], prediction of lymph node metastasis [16], and evaluation of patients' prognosis [17]. However, it remains unclear whether radiomic models based on both enhanced CT and PET/CT can be used as preoperative prediction tools for LVI in gastric cancer patients, and there are few relevant literatures. By contrast, Li et al. reported that clinicopathological characteristics including the level of CA19-9, Lauren classification, tumor differentiation, TNM stage, and gastric wall invasive depth are associated with the presence of LVI in patients with gastric cancer [18], and Shen et al. found that LVI is significantly correlated with tumor size, age, status of ALN, and histological grade in breast cancer [19], indicating that clinical variables might also help predict LVI.

Therefore, this study aimed to investigate the value of PET/CT radiomics and clinical variables in preoperative prediction of LVI status in gastric cancer, as well as develop and validate a combined model that incorporates radiomics features and clinical characteristics to improve model performance.

Materials and methods

Patient enrollment

The study was approved by the ethical committee of the Affiliated Cancer Hospital of Zhengzhou University, Henan Cancer Hospital (Zhengzhou, China), and the need for written informed consent form was waived. Patients with gastric cancer with histologically confirmed LVI status were retrospectively recruited from February 2014 to June 2019. The exclusive criteria were as follows: (1) abdominal enhanced CT and PET/CT examination was performed more than 2 weeks before surgery or biopsy, (2) tumor lesions could not be identified on the enhanced CT or PET/CT images, (3) patients with insufficient clinicopathological characteristics information, and (4) patients had received any anticancer

therapy before scan. Finally, 34 gastric patients with LVI and 67 patients with non-LVI were selected for this study.

The patients were randomly divided into a training dataset (27 LVI and 49 non-LVI, $n = 76$) and a validation dataset (7 LVI and 18 non-LVI, $n = 25$). The predictive models were developed in the training dataset, and the validation dataset was a held-out dataset and never used before the evaluation and comparison of model performance (Fig. 1).

Collection and pre-processing of clinical variables

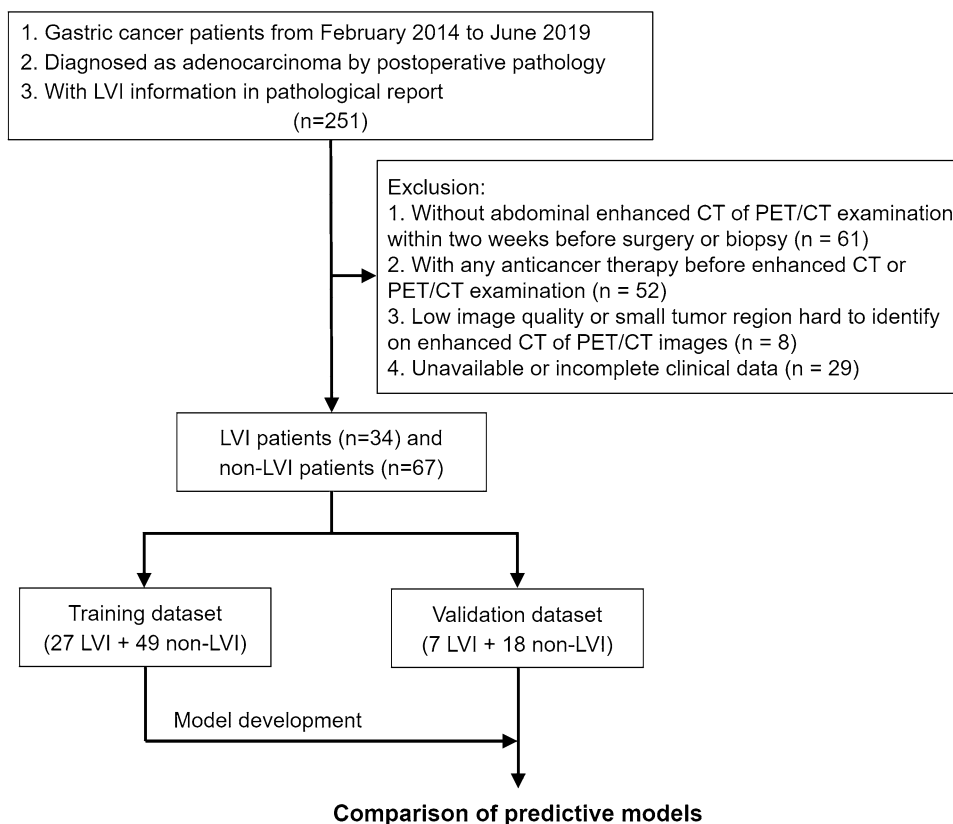
Basic demographic information including age and gender were directly collected from the electronic medical record system. By applying fixed staging criteria, the depth of tumor invasion, clinical T stage, and clinical N stage on CT images were prospectively diagnosed by an experienced radiologist with 9 years of experience and confirmed by a senior expert with more than 20 years of experience [20]. The clinical T stage and clinical N stage were assessed based on the AJCC/UICC 8th edition staging system in gastric cancer, and the local regional lymph nodes with shorter diameter greater than 10 mm were considered as suspected metastasis. The two radiologists showed excellent inter-rater agreement in assessing clinical T stage (Weighted kappa = 0.936). The exemplars of different cT staging signs in gastric cancer are shown in Fig. 2. Borrmann classification was based on the results of endoscopy examination. The mean CT values of the venous period (CTV-P value) and thickness of tumor were measured by an experienced radiologist using a radiology workstation. Before being used as input information for model development, clinical variables were pre-processed and recorded in relation to the diagnosis as follows:

- (1) Age, an actual variable.
- (2) Gender, a dichotomous variable (male = 1, female = 0).
- (3) Clinical T stage, a dichotomous variable (T1 = 0, T2 = 1, T3 = 2, T4 = 3).
- (4) Clinical N stage, a dichotomous variable (N0 = 1, N1 = 2, N2 = 3, N3 = 4).
- (5) Borrmann type, a dichotomous variable (type I = 1, type II = 2, type III = 3, type IV = 4).
- (6) Tumor location, a dichotomous variable (cardiac fundus and lower esophagus = 1, gastric curvature = 2, large curvature of the stomach = 3, gastric antrum and pylorus = 4, full stomach = 5).
- (7) CTV-P value, an actual variable.
- (8) Thickness of tumor, an actual variable.

Image acquisition

Enhanced CT images

Abdominal CT examinations were performed by Siemens Somatom Perspective CT Scanner (Siemens, Forchheim,

Fig. 1 Flowchart of the study

Germany). The scan covered the upper or the entire abdomen. Twenty patients with gastric cancer were examined with a single-phase CT-enhanced scan (venous phase), and 81 patients were examined with triple-phase spiral CT-enhanced scan. For the enhanced CT scan, patients were infused with 1.5 ml/kg of iodine (370 mg I/ml, 50 ml) with a pump injector, at a rate of 3.0 ml/s into an antecubital vein. Imaging was obtained with a post-injection delay of 30 s, 70 s, and 300 s after initiation of contrast material injection, corresponding to the arterial, venous, and delayed phases, respectively. The CT scanning parameters were as follows: 120 kV tube voltage, 200–250 mA tube current, 0.7 s tube rotation time, 40–50 cm field of view, 512×512 matrix, and 5 mm section thickness. All CT images were reconstructed with 5 mm slice thickness.

PET/CT images

CT plain scan was performed first, with tube voltage 120 kV and tube current 80mAs, reconstruction spacing 3.75 mm and pitch 1.25, and then 3D PET image was collected every 3 min, with axial field of 15.7 cm and fault resolution of 4 mm. After the display, Fourier reconstruction (FORE) was used to reconstruct the image, and CT data were used to attenuate PET image twice. Finally, PET and CT data were

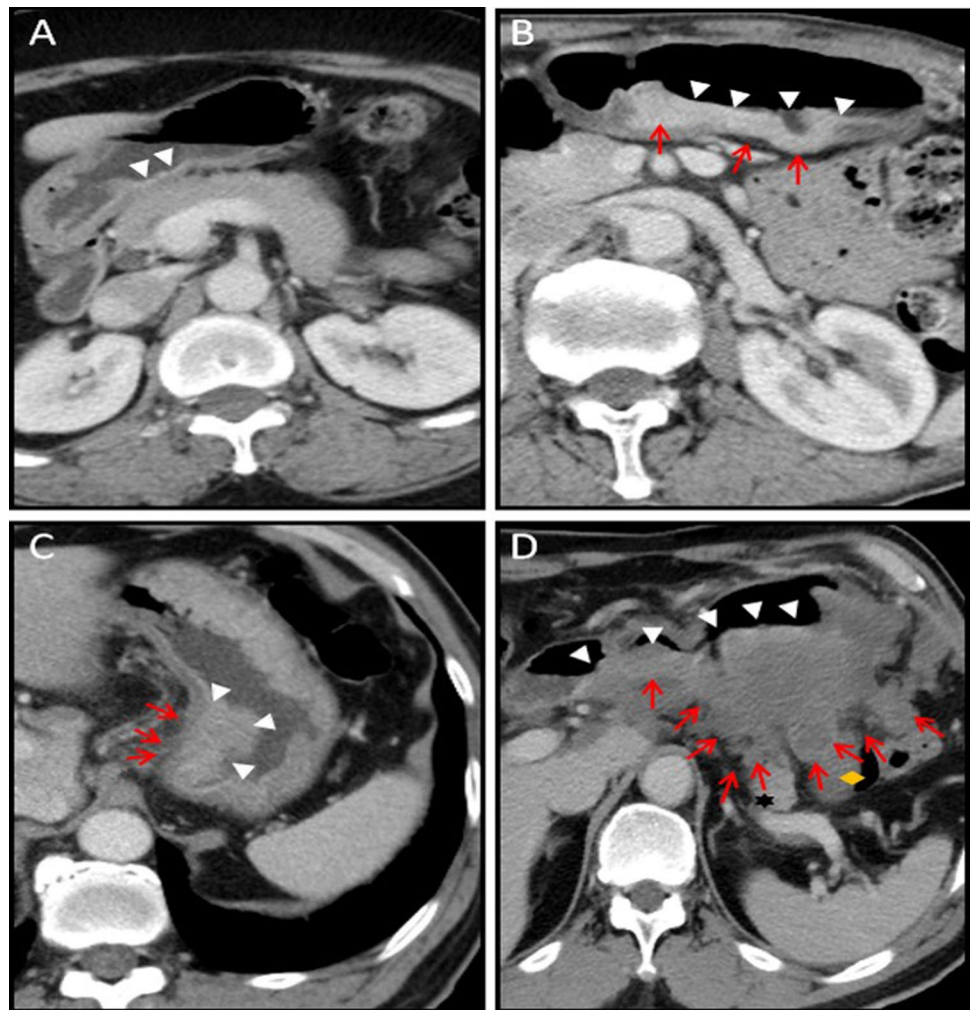
transmitted to Xeleris workstation to obtain multi-plane image and fusion image.

Tumor region segmentation and radiomics features extraction

All the CT images and PET images were retrieved from the Picture Archiving and Communication System to a local workstation for the regions of interests (ROIs) segmentation and analysis. A gastrointestinal radiologist with 9 years of experience manually labeled the 3D tumor regions on each slice of the CT images and PET images by ITK-SNAP software (v3.8.0, <http://www.itksnap.org>). The segmentation results were then reviewed and modified by a senior gastrointestinal radiologist with more than 20 years of experience. Both radiologists were blinded to the pathologic results.

We used an open-source python package (PyRadiomics version 2.1.2, <https://github.com/Radiomics/pyradiomics>) to automatically extract radiomics features from the manually segmented ROIs in the enhanced CT and PET/CT images. A total of 1454 radiomics features, including shape features ($n = 14$), first-order intensity statistics features ($n = 288$) and texture features [Gray Level Dependence Matrix (GLDM, $n = 224$), Gray Level Co-occurrence Matrix (GLCM, $n = 336$), Gray Level Size Zone Matrix (GLSZM, $n = 256$), Gray Level Run Length Matrix (GLRLM, $n = 256$),

Fig. 2 CT signs of gastric cancer by T1–4 staging. **a** The mucosa of the gastric antrum significantly enhanced (Short white arrow), not exceeding 50% of the thickness of the gastric wall. **b** Thickened locally of the gastric antrum wall, high enhancement of lesion involving thickness over 50% (Short white arrow), smooth serosal layer (red arrow). **c** Gastric cardia mass, slightly uneven enhancement (Short white arrow), slightly blurred fat space on serosal surface with short and thin cords (red arrow). **d** Stiff and non-uniform thickening gastric antrum and gastric body wall with multiple lymph nodes metastasis around the stomach (Short white arrow), nodular changes in serosal surface (red arrow), indistinguishable from the pancreatic tail (hexagon) and intestine (rhombus) (Color figure online)



Gray Neighboring Gray Tone Difference Matrix (NGTDM, $n=80$), were extracted from the manually labeled ROIs, and all these features had been used in the literature [21].

Evaluation of inter- and intra-observer reproducibility and selection of radiomics features

In order to ensure the reproducibility and accuracy of the radiomics features, 30 CT scans and 30 PET scans (each from 20 non-LVI and 10 LVI patients) were randomly selected, and the ROIs were manually labeled by radiologist 1 and radiologist 2. Then, radiologist 1 repeated the same procedure 2 weeks later. The radiomics features from the paired ROIs were automatically extracted and the inter- and intra-observer agreement was assessed by intra-class correlation coefficients (ICC). The radiomics features which met the criteria of having an ICC greater than 0.75 were considered as good agreement and used for further analysis.

To avoid overfitting problems and reduce computation complexity, the Least absolute shrinkage and selection operator (LASSO) algorithm was applied to select the key

radiomics features most closely associated with the determination of gastric cancer LVI status [22], with penalty parameter tuning conducted by ten-fold cross-validation. Based on the optimal log(λ) sequence (-2.695 for enhanced CT-based features and -2.434 for PET/CT-based features), six enhanced CT-based key features and four PET/CT-based key features were selected for further analysis, including shape ($n=2$), first-order ($n=3$), GLCM ($n=1$), GLDM ($n=1$), and GLSZM ($n=3$) features. The feature extraction and selection process were implemented by Python 3.6.0 (www.python.org).

Development of predictive models

To investigate the capability of radiomics features and clinical variables in the preoperative differentiation of the LVI from non-LVI patients with gastric cancer, we developed 3 types of models with different machine-learning (ML) classifiers by different input data, respectively. The clinical variables were used to conduct the clinical model, the selected radiomics features were used to develop the image model, and the combined

model was constructed with the clinical variables and selected radiomics features.

Logistic regression (LR), adaptive boosting (AdaBoost), and linear discriminant analysis (LDA) classifiers were used to discriminate LVI groups from non-LVI groups of patients with gastric cancer. Moreover, LVI groups and non-LVI groups were defined as positive and negative in the classification, respectively.

LR was a statistical modeling technique where the probability of a category was related to a set of explanatory variables [23]. The logistic model was defined by the following equations:

$$z = a_0 + \sum_{i=1}^n a_i x_i, \quad (1)$$

$$P(z) = \frac{e^z}{1 + e^z}, \quad (2)$$

where Z was a measure of the contribution of the explanatory variables x_i ($i = 1, \dots, n$), a_i were the regression coefficients obtained by maximum likelihood in conjunction with their standard errors Δa_i , and $P(z)$ was the categorical response of variables.

AdaBoost classifier was an ensemble classifier that included various weak classifiers [24]. In each iteration during the learning process, a component classifier C_k was trained according to the weights.

If the training pattern was classified correctly, its chance of being used again in the next component classifier decreased; conversely, if the pattern was classified mistakenly, its chance of being used again increased. The ensemble classifier could be expressed as follows:

$$C(x) = \text{sign} \left\{ \sum_{k=1}^N \alpha_k c_k(x) \right\}, \quad (3)$$

where x denoted the input vector. N denoted the number of classifiers. α_k ($k = 1, 2, \dots, N$) represented the weight of each weak classifier.

LDA was a classical linear classifier we used to solve the problem [25]:

$$y = \beta_0 + \beta^T x, \quad (4)$$

where x denoted the feature vector. β denoted the weight, and β_0 was the constant, which were determined by maximizing the distance between the 2 classes' means and minimizing the interclass variance.

The development and validation of all models were performed with InferScholar platform version 3.5 (InferVision).

Evaluation of model performance

The performance of the predictive models was evaluated by the area under the curve (AUC) and the sensitivity and specificity in the validation dataset. Sensitivity denoted the number of correctly predicted vascular invasion samples divided by the total number of vascular invasion samples, and specificity denoted the number of correctly predicted vascular non-invasion samples to the total number of vascular non-invasion samples. AUC was defined as the area under curves, which could evaluate the classifiers' performance for the identification of LVI and non-LVI samples across the entire range independent of class distributions. An $\text{AUC} > 0.7$ was considered good classification performance [26].

Calibration analysis and decision curve analysis

We used 1,000 bootstrapping resamples for the evaluation of calibration, and the consistency between the actual and predicted LVI probability was represented graphically and assessed by Hosmer–Lemeshow test [27]. The potential net benefit of the predictive models at different threshold probabilities was quantified and the clinical usefulness was evaluated by the decision curve analysis (DCA) [28].

Statistical analysis

The differences between continuous clinical variables were evaluated by the Mann–Whitney U test, and the chi-squared test or Fisher's exact test was used to compare dichotomous clinical characteristics between groups. Receiver operating characteristic (ROC) curve analysis was performed, and the discrimination performance of different models was quantified by the AUC, sensitivity, and specificity. Delong's test was used to compare the difference between 2 AUCs of different models [29]. A heatmap of selected radiomic features was generated by HemI v1.0 software [30]. The calibration curve was plotted using the "rms" package (version 6.2), and the decision curve was plotted using the "rmda" package (version 1.6). A two-sided p value less than 0.05 was considered statistically significant. All statistical analyses were performed with SPSS (version 21.0) and R software (version 3.6.3).

Results

Study design and patient characteristics

Two hundred and fifty-one patients were diagnosed with gastric cancer from February 2014 to June 2019 in our hospital and were initially recruited. According to the inclusion and exclusion criteria, 74 men (median age, 64 years;

Table 1 Comparison of clinical variables between the LVI group and non-LVI group

| Characteristic | Training dataset (<i>n</i> = 76) | | | Validation dataset (<i>n</i> = 25) | | |
|------------------------------------|-----------------------------------|--------------------------|----------------|-------------------------------------|--------------------------|----------------|
| | LVI (<i>n</i> = 27) | non-LVI (<i>n</i> = 49) | <i>p</i> value | LVI (<i>n</i> = 7) | non-LVI (<i>n</i> = 18) | <i>p</i> value |
| Gender, male/female | 20/7 | 35/14 | 0.81 | 6/1 | 13/5 | 0.48 |
| Age, mean ± SD, years | 60.3 ± 14.3 | 60.4 ± 10.3 | 0.95 | 61.4 ± 9.0 | 61.4 ± 9.5 | 0.99 |
| cT category | | | 0.96 | | | 0.47 |
| T1 | 2 | 4 | | 0 | 0 | |
| T2 | 2 | 3 | | 2 | 2 | |
| T3 | 9 | 14 | | 2 | 4 | |
| T4 | 14 | 28 | | 3 | 12 | |
| cN category | | | 0.81 | | | 0.77 |
| N0 | 6 | 8 | | 2 | 3 | |
| N1 | 11 | 24 | | 3 | 10 | |
| N2 | 8 | 15 | | 2 | 5 | |
| N3 | 2 | 2 | | 0 | 0 | |
| Borrmann type | | | 0.10 | | | 0.72 |
| I | 2 | 6 | | 0 | 1 | |
| II | 4 | 9 | | 1 | 5 | |
| III | 15 | 32 | | 5 | 11 | |
| IV | 6 | 2 | | 1 | 1 | |
| Tumor location | | | 0.12 | | | 0.22 |
| Cardiac fundus and lower esophagus | 7 | 22 | | 0 | 8 | |
| Gastric curvature | 5 | 10 | | 3 | 3 | |
| Large curvature of the stomach | 2 | 0 | | 0 | 1 | |
| Gastric antrum and pylorus | 5 | 10 | | 3 | 5 | |
| Full stomach | 8 | 7 | | 1 | 1 | |
| CTV-P, mean ± SD | 95.2 ± 26.8 | 80.3 ± 24.4 | 0.02* | 77.6 ± 30.6 | 74.4 ± 18.6 | 0.75 |
| Thickness of tumor, mean ± SD, mm | 16.5 ± 6.4 | 22.1 ± 12.4 | 0.03* | 17.7 ± 6.6 | 17.1 ± 4.7 | 0.82 |

Differences in characteristic dichotomous variables between the 2 datasets were calculated with the chi-square test or Fisher's exact test; the Mann–Whitney *U* test was conducted to compare differences in actual variables

* $p < 0.05$

CTV-P CT values of the venous period

range, 24–78 years) and 27 women (median age, 55 years; range, 24–84 years) were enrolled in the study. Based on pathological test results, our classification splits the patients into 2 groups: LVI ($n = 34$) and non-LVI ($n = 67$) group. The prevalence of LVI was 33.7% (34/101).

No significant differences were observed between the training and validation datasets in LVI prevalence (35.5% and 32.0% in the training and validation datasets, respectively. $p = 0.75$). No significant differences were observed in age, gender, cT category, cN category, Borrmann type, and tumor location between LVI and non-LVI groups in the training and validation datasets (all $ps > 0.05$). The LVI group showed higher CTV-P value and lower tumor thickness than the non-LVI group did in the training dataset; however, their differences were not significant in the validation dataset. The demographic and clinicopathological characteristics of all patients are summarized in Table 1.

Radiomics feature selection

There were 1105 and 726 radiomics features extracted from the enhanced CT and PET images showed good agreement with ICC > 0.75 after inter- and intra-observer agreement analysis. Six CT-based radiomics features and 4 PET-based radiomics features with non-zero coefficients remained after LASSO penalty was applied (Fig. 3). The heatmap of these selected key radiomics features in the training and validation dataset was plotted according to the normalized radiomics feature values (Fig. 4).

Development and validation of the predictive models

The performance of predictive models was evaluated by ROC analysis with respect to the AUC in the

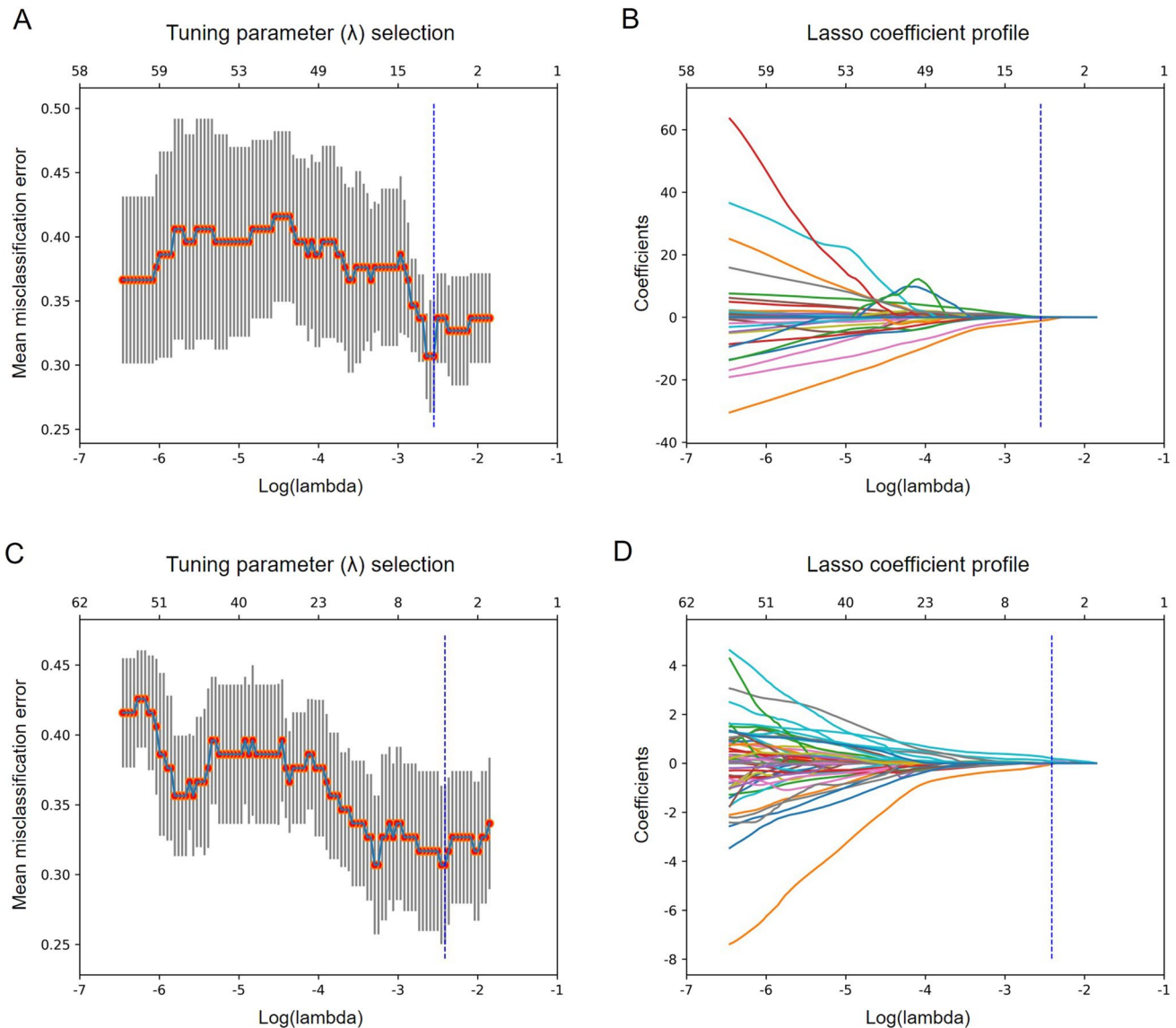


Fig. 3 Selection of radiomics features by the variance threshold method and LASSO regression. **A** The variance threshold approach-based dimensionality reduction profile of the CT-based radiomics features. **B** The coefficient profile plot of 6 CT-based non-zero coefficients against the optimal $\log(\lambda)$ sequence. **C** The variance

threshold approach-based dimensionality reduction profile of the PET-based radiomics features. **D** The coefficient profile plot of 4 PET-based non-zero coefficients against the optimal $\log(\lambda)$ sequence

validation dataset. The clinical models using AdaBoost, LDA, and LR classifiers had reached AUCs of 0.742 (95% CI, 0.529–0.894), 0.706 (95% CI, 0.492–0.870), and 0.690 (95% CI, 0.476–0.858), respectively (Fig. 5a–c). The AUCs of the image models using AdaBoost, LDA, and LR classifiers were 0.849 (95% CI, 0.650–0.960), 0.778 (95% CI, 0.568–0.918), and 0.810 (95% CI, 0.604–0.937), respectively (Figs. 5d–f). The combined models achieved an improvement in discrimination efficacy compared with the other models, with AUCs of AdaBoost, LDA, and LR classifiers achieving 0.944 (95% CI, 0.774–0.997), 0.929 (95% CI, 0.751–0.993), and 0.921 (95% CI, 0.741–0.990),

respectively (Fig. 5g–i). The detailed performance of each model is presented in Table 2.

Model comparison

For the clinical, image, or combined models, the AdaBoost, LDA, and LR classifiers showed no significant difference in performance in the validation dataset (DeLong's test, all $p > 0.05$). By contrast, although not statistically significant, the combined models showed better discrimination capability than the clinical models (DeLong's test, AdaBoost classifier $p = 0.177$, LDA classifier $p = 0.063$, LR classifier

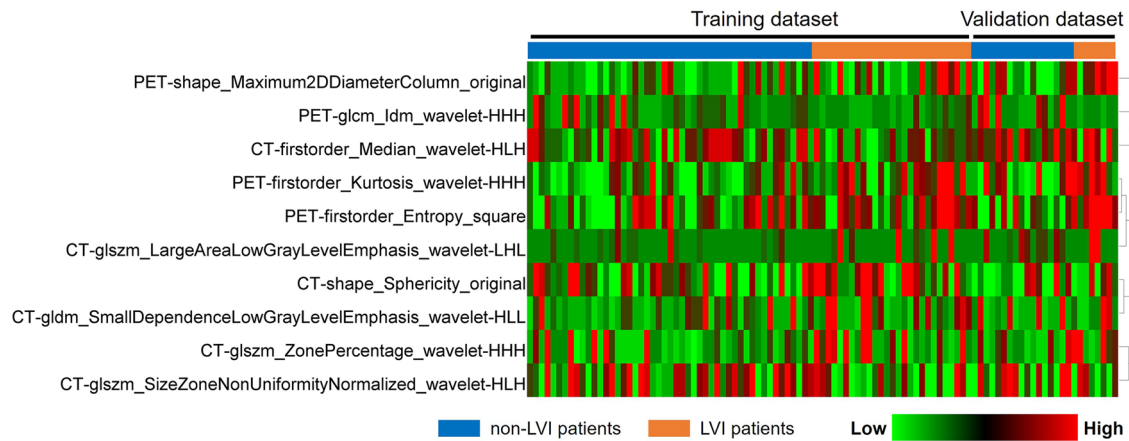


Fig. 4 Heatmap of selected CT and PET-based radiomics features in the training and validation datasets. Each row represented a radiomic feature, and each column corresponded to 1 patient (separately grouped for LVI and non-LVI patients)

$p=0.132$, respectively) or the image models (DeLong's test, AdaBoost classifier $p=0.136$, LDA classifier $p=0.164$, LR classifier $p=0.146$, respectively), indicating that incorporating radiomics features and clinical variables would benefit the prediction of LVI status in patients with gastric cancer.

Calibration and clinical usefulness analysis

The calibration curve analysis demonstrated good agreement between prediction and observation for the combined models, with the non-significant statistic of the Hosmer–Lemeshow test suggested no significant deviation from an ideal fitting ($p=0.926$ for AdaBoost classifier, Fig. 6a; $p=0.744$ for LDA classifier, Fig. 6B; $p=0.907$ for LR classifier, Fig. 6c). The DCA for the combined models using different classifier is presented in Fig. 7, and the results indicating that these combined models were clinically useful.

Discussion

As one of the most common independent predictors of prognosis of gastric cancer, LVI refers to the invasion of lymphatic space, blood vessels, or both in the peritumor region by tumor embolus. The positive LVI status in gastric cancer increases the incidence rate of lymph node metastasis and distant metastasis and reduces the survival rate of patients [18]. TNM is the most used tool for the assessment of cancer prognosis [31], but the combination of LVI status and a TNM staging system can improve the accuracy of predicting the prognosis of patients with N0 stage gastric cancer [32]. Because there are no significant image features, it is difficult for radiologists to differentiate between LVI and non-LVI on preoperative CT images. Li et al. found that the level of CA19-9, tumor size, Lauren grading, tumor

differentiation, depth of gastric wall infiltration, involvement of lymph nodes, distant metastasis, and advanced TNM stage were significantly correlated with the presence of vascular invasion in gastric cancer [18]. However, Lauren grading, tumor differentiation, depth of gastric wall infiltration, and involvement of lymph nodes are confirmed by postoperative pathology diagnosis, and the TNM stage is also difficult to classify before biopsy or surgery [33]. Therefore, it remains challenging to preoperatively predict the presence of LVI by clinicopathological characteristics. According to our review of the literature, no studies have used both enhanced CT and PET/CT images incorporating clinical variables for the prediction of LVI in gastric cancer.

In this study, the cT stage and cN stage, which can be classified on preoperative CT images, were used for model development. Other clinical variables, including age, gender, Borrmann type, tumor location, thickness of tumor, and CTV-P value, were used as risk factor candidates in this study. The Borrmann type was associated with LVI in advanced proximal gastric cancer [34]. Tumor location (the upper and middle third of stomach) is a significant risk factor for submucosal or lymphovascular invasion [35]. Ma et al. found that ΔAP (Tumor CT attenuation difference between non-contrast and portal) could be an independent predictor of LVI because of the increased microvascular permeability by lymphatic vascular structure destruction, suggesting that the CT values of the venous period might be correlated with LVI in gastric cancer [36]. Univariate analysis showed significant differences in thickness of tumor and CTV-P value between the LVI group and non-LVI group in the training dataset, which was consistent with another study [8].

Gastric cancer is a clinically heterogeneous disease, and CT images and clinical variables would enable different insights into different tumor biology characteristics on multiple levels. We developed 3 types of predictive

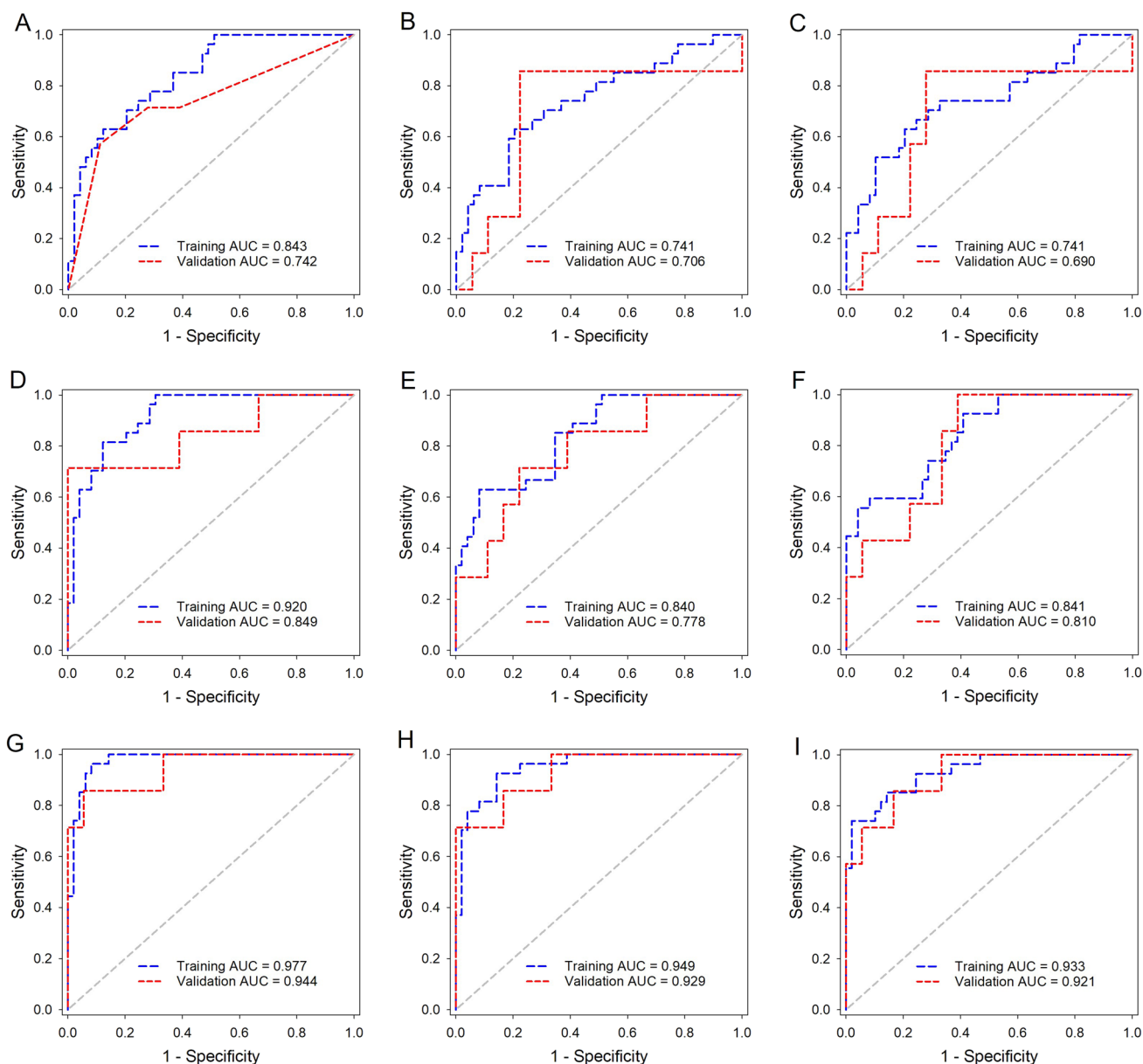


Fig. 5 Performance evaluation of the predictive models. ROC curve of the clinical models using **A** AdaBoost, **B** LDA, and **C** LR classifiers in the training and validation datasets, respectively. ROC curve of the image models using **D** AdaBoost, **E** LDA, and **F** LR classifiers

in the training and validation datasets, respectively. ROC curve of the combined models using **G** AdaBoost, **H** LDA, and **i** LR classifiers in the training and validation datasets, respectively

models by radiomics features from both CT images, clinical variables, and the combination of both, respectively in the study. The combined model has a better performance than simply clinical model and image model in AdaBoost (AUC = 0.944, 0.742, 0.849), LDA (AUC = 0.929, 0.706, 0.778), and LR (AUC = 0.921, 0.690, 0.810) classifier, respectively. The results indicated that the combined model could reflect the tumor heterogeneity more accurately and be more reliable. We reasoned that both radiomics features selected from both enhanced CT and PET/CT

images and preoperative clinical variables of patients with gastric cancer could be used for prediction of LVI status by machine-learning-based methods, and the predictive performance could be improved by their combination. In other studies, researchers found that incorporating radiomics features could lead to improved performance in evaluating the survival and chemotherapeutic benefits of patients with gastric cancer than TNM staging system and clinicopathologic nomogram [37], and the combined model could also add more efficiency than either feature alone for

Table 2 Performance comparison of the predictive models in the validation dataset

| Models | Clinical model | Image model | Combined model |
|----------------------------|------------------------|------------------------|------------------------|
| AdaBoost classifier | | | |
| Cut-off threshold | 0.364 | 0.535 | 0.409 |
| Sensitivity | 57.1% (4/7) | 71.4% (5/7) | 85.7% (6/7) |
| Specificity | 88.9% (16/18) | 100.0% (18/18) | 94.4% (17/18) |
| *AUC (95% CI) | 0.742 (0.529–0.894) | 0.849 (0.650–0.960) | 0.944 (0.774–0.997) |
| LDA classifier | | | |
| Cut-off threshold | 0.391 | 0.352 | 0.534 |
| Sensitivity | 85.7% (6/7) | 71.4% (5/7) | 71.4% (5/7) |
| Specificity | 77.8% (14/18) | 77.8% (14/18) | 100.0% (18/18) |
| #AUC (95% CI) | 0.706 (0.492–0.870) | 0.778 (0.568–0.918) | 0.929 (0.751–0.993) |
| LR classifier | | | |
| Cut-off threshold | 0.384 | 0.296 | 0.299 |
| Sensitivity | 85.7% (6/7) | 100.0% (7/7) | 85.7% (6/7) |
| Specificity | 72.2% (13/18) | 61.1% (11/18) | 83.3% (15/18) |
| §AUC (95% CI) | 0.690 (0.476–0.858) | 0.810 (0.604–0.937) | 0.921 (0.741–0.990) |

AUC area under the receiver operating characteristic curve, *AdaBoost* adaptive boosting, *LDA* linear discriminant analysis, *LR* logistic regression

*ROC comparison of AdaBoost classifier-based predictive models by Delong's test: Combined model *versus* Image model, $p=0.177$, z statistic = 1.350; Combined model *versus* Clinical model, $p=0.136$, z statistic = 1.492; Image model *versus* Clinical model, $p=0.557$, z statistic = 0.587

#ROC comparison of LDA classifier-based predictive models by Delong's test: Combined model *versus* Image model, $p=0.063$, z statistic = 1.861; Combined model *versus* Clinical model, $p=0.164$, z statistic = 1.391; Image model *versus* Clinical model, $p=0.708$, z statistic = 0.375

§ROC comparison of LR classifier-based predictive models by Delong's test: Combined model *versus* Image model, $p=0.132$, z statistic = 1.507; Combined model *versus* Clinical model, $p=0.146$, z statistic = 1.454; Image model *versus* Clinical model, $p=0.527$, z statistic = 0.633

predicting preoperatively the LVI status of patients with breast cancer [38].

We also compared the performance of 3 ML classifiers—AdaBoost, LDA, and LR—and found no significant difference. It may be that the sample size is relatively small. In clinical practice, appropriate methods should be used according to different needs. Considering the stability and prognostic properties, a combined model using the AdaBoost classifier might be the best choice.

Several limitations in this study deserve consideration. First, as a retrospective study, the patients were enrolled from one research institute and the sample size was limited; thus, prospective multicenter research with a larger sample is necessary. Second, the value of radiomics features extracted from multiphase contrast-enhanced CT images and other clinical variables in the prediction of LVI also requires investigation. Third, because three-dimensional manual tumor segmentation was time-consuming and complicated, a deep-learning method for the automatic segmentation of gastric cancer should be developed.

In conclusion, the results of this study showed that the LVI status of gastric cancer could be predicted by radiomics features from enhanced CT and PET/CT images, as well as incorporating with clinical variables. The combined model incorporating radiomics features and clinical variables had a better performance and is a potential non-invasive tool for detecting LVI in patients with gastric cancer preoperatively. Further research is required before application in clinical practice.

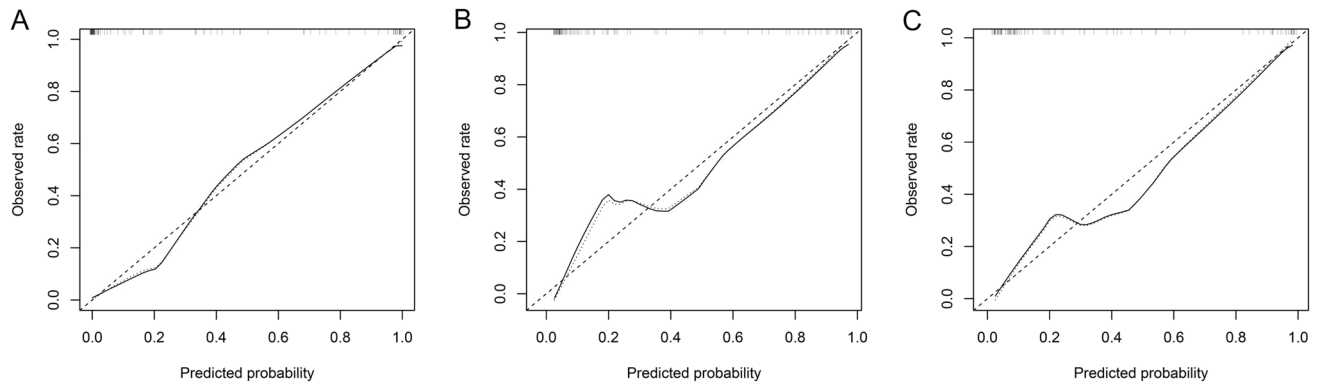


Fig. 6 Calibration curves of the combined models using **a** AdaBoost, **b** LDA, and **c** LR classifiers, respectively. The solid line represented the performance of the model without correction for overfit. The dot-

ted line was the bootstrap-corrected performance of the model with a scatter estimate for future accuracy

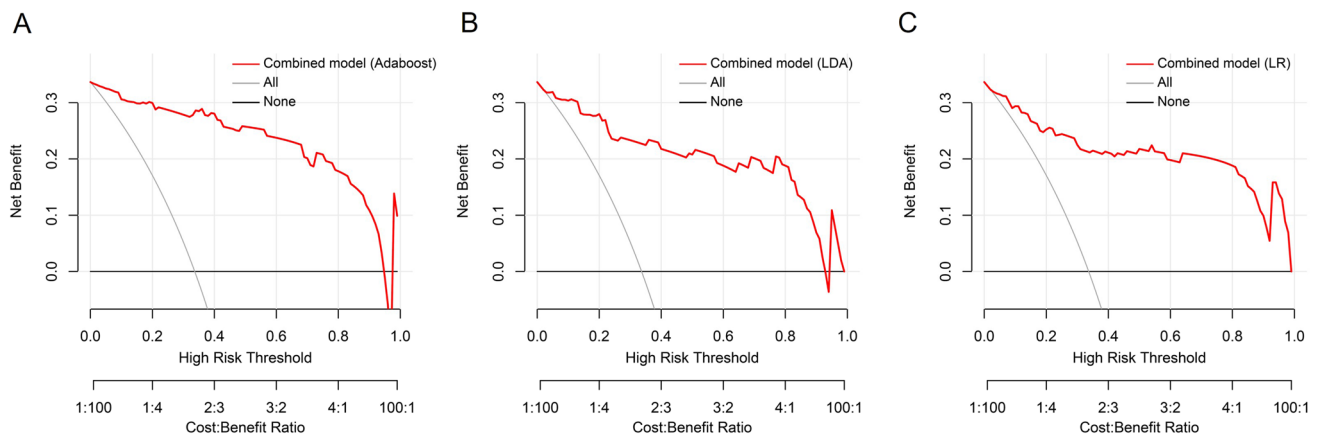


Fig. 7 Decision curve analysis for the combined models using **A** AdaBoost, **B** LDA, and **C** LR classifiers, respectively. The red line represented the net benefit of the model across the full range of

threshold probabilities. Gray line: all positive, assuming all patients should undergo surgery to confirm the presence of LVI. Black line: all negative, assuming no possibility of LVI in patients

Authors' contributions Study concept and design: XC and JL. Data collection and annotation: LF and JL. Model development and data analysis: HZ, RZ, and HY. Literature search and manuscript writing: JL and HY. Manuscript editing: CX and JZ. Contributed equally to this work and should be considered co-correspondent authors: XC and JZ.

Funding The authors state that this work has not received any funding.

Data availability All data generated or analyzed during this study are included in this published article.

Code availability No complex statistical methods were necessary for this paper.

Declarations

Conflict of interest The authors of this manuscript declare relationships with the following companies.

Ethical approval Institutional Review Board approval was obtained.

Consent to participate Not applicable.

Consent for publication Not applicable.

References

1. J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D.M. Parkin, D. Forman, F. Bray, Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012, *Int. J. Cancer*. (2015). <https://doi.org/10.1002/ijc.29210>.
2. L. Yang, R. Zheng, N. Wang, Y. Yuan, S. Liu, H. Li, S. Zhang, H. Zeng, W. Chen, Incidence and mortality of stomach cancer in China, 2014, *Chinese J. Cancer Res.* (2018). <https://doi.org/10.21147/j.issn.1000-9604.2018.03.01>.

3. H. Li, L.Q. Feng, Y.Y. Bian, L.L. Yang, D.X. Liu, Z. Bin Huo, L. Zeng, Comparison of endoscopic submucosal dissection with surgical gastrectomy for early gastric cancer: An updated meta-analysis, *World J. Gastrointest. Oncol.* (2019). <https://doi.org/10.4251/wjgo.v11.i2.161>.
4. K. Sun, S. Chen, J. Ye, H. Wu, J. Peng, Y. He, J. Xu, Endoscopic resection versus surgery for early gastric cancer: a systematic review and meta-analysis, *Dig. Endosc.* (2016). <https://doi.org/10.1111/den.12596>.
5. L. Marano, K. Polom, A. Patrìti, G. Roviello, G. Falco, A. Stracqualursi, R. De Luca, R. Petrioli, M. Martinotti, D. Generali, D. Marrelli, N. Di Martino, F. Roviello, Surgical management of advanced gastric cancer: An evolving issue, *Eur. J. Surg. Oncol.* (2016). <https://doi.org/10.1016/j.ejso.2015.10.016>.
6. S. J., L. H.-J., C. G.S., H. S.-U., K. M.-C., R. S.W., K. W., S. K.Y., K. H.-H., H. W.J., Recurrence following laparoscopy-assisted gastrectomy for gastric cancer: A multicenter retrospective analysis of 1,417 patients, *Ann. Surg. Oncol.* (2010).
7. M.E. Orme, M.A.J. Chaplain, A mathematical model of vascular tumour growth and invasion, *Math. Comput. Model.* (1996). [https://doi.org/10.1016/0895-7177\(96\)00053-2](https://doi.org/10.1016/0895-7177(96)00053-2).
8. H. FUJIKAWA, K. KOUMORI, H. WATANABE, K. KANO, Y. SHIMODA, T. AOYAMA, T. YAMADA, T. HIROSHI, N. YAMAMOTO, H. CHO, M. SHIOZAWA, T. YOSHIKAWA, S. MORINAGA, Y. RINO, M. MASUDA, T. OGATA, T. OSHIMA, The Clinical Significance of Lymphovascular Invasion in Gastric Cancer, *In Vivo* (Brooklyn). (2020). <https://doi.org/10.21873/invivo.11942>.
9. Z. C.-D., N. F.-L., Z. X.-T., D. D.-Q., Lymphovascular invasion as a predictor for lymph node metastasis and a prognostic factor in gastric cancer patients under 70 years of age: A retrospective analysis, *Int. J. Surg.* (2018). <https://doi.org/10.1016/j.ijsu.2018.03.073> LK - <http://rug.on.worldcat.org/atoztitles/link/?sid=EMBASE&issn=17439159&id=doi:10.1016%2Fj.ijsu.2018.03.073&atitle=Lymphovascular+invasion+as+a+predictor+for+lymph+node+metastasis+and+a+prognostic+factor+in+gastric+cancer+patients+under+70+years+of+age%3A+A+retrospective+analysis&title=Int.+J.+Surg.&title=International+Journal+of+Surgery&volume=53&issue=&page=214&epage=220&aulast=Zhang&aufirst=Chun-Dong&aunit=C.-D.&aufull=Zhang+C.-D.&coden=&isbn=&pages=214-220&da>.
10. J.H. Pyo, H. Lee, Y.W. Min, B.H. Min, J.H. Lee, K.M. Kim, H. Kim, K. Kim, J.J. Kim, Feasibility of Endoscopic Resection in Early Gastric Cancer with Lymphovascular Invasion, *Ann. Surg. Oncol.* (2019). <https://doi.org/10.1245/s10434-018-07119-4>.
11. W. Du, Y. Wang, D. Li, X. Xia, Q. Tan, X. Xiong, Z. Li, Preoperative Prediction of Lymphovascular Space Invasion in Cervical Cancer With Radiomics –Based Nomogram, *Front. Oncol.* (2021). <https://doi.org/10.3389/fonc.2021.637794>.
12. M. Komori, Y. Asayama, N. Fujita, K. Hiraka, D. Tsurumaru, Y. Kakeji, H. Honda, Extent of arterial tumor enhancement measured with preoperative mdct gastrography is a prognostic factor in advanced gastric cancer after curative resection, *Am. J. Roentgenol.* (2013). <https://doi.org/10.2214/AJR.12.9206>.
13. L.M. Fanchon, S. Dogan, A.L. Moreira, S.A. Carlin, C.R. Schmittlein, E. Yorke, A.P. Apte, I.A. Burger, J.C. Durack, J.P. Erinjeri, M. Maybody, H. Schöder, R.H. Siegelbaum, C.T. Sofocleous, J.O. Deasy, S.B. Solomon, J.L. Humm, A.S. Kirov, Feasibility of in situ, high-resolution correlation of tracer uptake with histopathology by quantitative autoradiography of biopsy specimens obtained under 18F-FDG PET/CT guidance, *J. Nucl. Med.* (2015). <https://doi.org/10.2967/jnumed.114.148668>.
14. P. Lambin, E. Rios-Velazquez, R. Leijenaar, S. Carvalho, R.G.P.M. Van Stiphout, P. Granton, C.M.L. Zegers, R. Gillies, R. Boellard, A. Dekker, H.J.W.L. Aerts, Radiomics: Extracting more information from medical images using advanced feature analysis, *Eur. J. Cancer.* (2012). <https://doi.org/10.1016/j.ejca.2011.11.036>.
15. S. Liu, S. Liu, C. Ji, H. Zheng, X. Pan, Y. Zhang, W. Guan, L. Chen, Y. Guan, W. Li, J. He, Y. Ge, Z. Zhou, Application of CT texture analysis in predicting histopathological characteristics of gastric cancers, *Eur. Radiol.* (2017). <https://doi.org/10.1007/s00330-017-4881-1>.
16. Y. Wang, W. Liu, Y. Yu, J. Juan Liu, H. dan Xue, Y. fei Qi, J. Lei, J. chun Yu, Z. yu Jin, CT radiomics nomogram for the preoperative prediction of lymph node metastasis in gastric cancer, *Eur. Radiol.* (2020). <https://doi.org/10.1007/s00330-019-06398-z>.
17. F. Giganti, S. Antunes, A. Salerno, A. Ambrosi, P. Marra, R. Nicoletti, E. Orsenigo, D. Chiari, L. Albarello, C. Staudacher, A. Esposito, A. Del Maschio, F. De Cobelli, Gastric cancer: texture analysis from multidetector computed tomography as a potential preoperative prognostic biomarker, *Eur. Radiol.* (2017). <https://doi.org/10.1007/s00330-016-4540-y>.
18. P. Li, H.Q. He, C.M. Zhu, Y.H. Ling, W.M. Hu, X.K. Zhang, R.Z. Luo, J.P. Yun, D. Xie, Y.F. Li, M.Y. Cai, The prognostic significance of lymphovascular invasion in patients with resectable gastric cancer: A large retrospective study from Southern China, *BMC Cancer.* (2015). <https://doi.org/10.1186/s12885-015-1370-2>.
19. S. Di Shen, S.Z. Zhong, C.Z. Wang, W.H. Huang, Correlation of lymphovascular invasion with clinicopathological factors in invasive breast cancer: A meta-analysis, *Int. J. Clin. Exp. Med.* (2015).
20. T. Fukagawa, H. Katai, J. Mizusawa, K. Nakamura, T. Sano, M. Terashima, S. Ito, T. Yoshikawa, N. Fukushima, Y. Kawachi, T. Kinoshita, Y. Kimura, H. Yabusaki, Y. Nishida, Y. Iwasaki, S.W. Lee, T. Yasuda, M. Sasako, A prospective multi-institutional validity study to evaluate the accuracy of clinical diagnosis of pathological stage III gastric cancer (JCOG1302A), *Gastric Cancer.* (2018). <https://doi.org/10.1007/s10120-017-0701-1>.
21. H. Sun, H. Qu, L. Chen, W. Wang, Y. Liao, L. Zou, Z. Zhou, X. Wang, S. Zhou, Identification of suspicious invasive placentation based on clinical MRI data using textural features and automated machine learning, *Eur. Radiol.* (2019). <https://doi.org/10.1007/s00330-019-06372-9>.
22. W. Sauerbrei, P. Royston, H. Binder, Selection of important variables and determination of functional form for continuous predictors in multivariable model building, in: *Stat. Med.*, 2007. <https://doi.org/10.1002/sim.3148>.
23. L. Dong, J. Wesseloo, Y. Potvin, X. Li, Discrimination of Mine Seismic Events and Blasts Using the Fisher Classifier, Naive Bayesian Classifier and Logistic Regression, *Rock Mech. Rock Eng.* (2016). <https://doi.org/10.1007/s00603-015-0733-y>.
24. Y. Freund, R.E. Schapire, 1996 Experiments with a New Boosting Algorithm, *Proc. 13th Int. Conf. Mach. Learn.* Doi:10.1.1.133.1040.
25. T. V. Bandos, L. Bruzzone, G. Camps-Valls, Classification of hyperspectral images with regularized linear discriminant analysis, *IEEE Trans. Geosci. Remote Sens.* (2009). <https://doi.org/10.1109/TGRS.2008.2005729>.
26. W. Liang, L. Zhang, G. Jiang, Q. Wang, L. Liu, D. Liu, Z. Wang, Z. Zhu, Q. Deng, X. Xiong, W. Shao, X. Shi, J. He, Development and validation of a nomogram for predicting survival in patients with resected non-small-cell lung cancer, *J. Clin. Oncol.* (2015). <https://doi.org/10.1200/JCO.2014.56.6661>.
27. S. Finazzi, D. Poole, D. Luciani, P.E. Cogo, G. Bertolini, Calibration belt for quality-of-care assessment based on dichotomous outcomes, *PLoS One.* (2011). <https://doi.org/10.1371/journal.pone.0016110>.
28. A.J. Vickers, E.B. Elkin, Decision curve analysis: A novel method for evaluating prediction models, *Med. Decis. Mak.* (2006). <https://doi.org/10.1177/0272989X06295361>.

29. E.R. DeLong, D.M. DeLong, D.L. Clarke-Pearson, Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach, *Biometrics*. (1988). <https://doi.org/10.2307/2531595>.
30. W. Deng, Y. Wang, Z. Liu, H. Cheng, Y. Xue, HemI: A toolkit for illustrating heatmaps, *PLoS One*. (2014). <https://doi.org/10.1371/journal.pone.0111988>.
31. H. In, I. Solsky, B. Palis, M. Langdon-Embry, J. Ajani, T. Sano, Validation of the 8th Edition of the AJCC TNM Staging System for Gastric Cancer using the National Cancer Database, *Ann. Surg. Oncol.* (2017). <https://doi.org/10.1245/s10434-017-6078-x>.
32. L. J., D. Y., X. J.-W., W. J.-B., L. J.-X., C. Q.-Y., C. L.-L., L. M., T. R.-H., Z. C.-H., L. P., H. C.-M., Combination of lymphovascular invasion and the AJCC TNM staging system improves prediction of prognosis in N0 stage gastric cancer: Results from a high-volume institution, *BMC Cancer*. (2019). <https://doi.org/10.1186/s12885-019-5416-8> LK - <http://rug.on.worldcat.org/atoztitles/link/?sid=EMBASE&issn=14712407&id=doi:10.1186%2Fs12885-019-5416-8&atitle=Combination+of+lymphovascular+invasion+and+the+AJCC+TNM+staging+system+improves+prediction+of+prognosis+in+N0+stage+gastric+cancer%3A+Results+from+a+high-volume+institution&stitle=BMC+Cancer&title=BMC+Cancer&volume=19&issue=1&spage=&epage=&aulast=Lu&aufirst=Jun&aunit=J.&aufull=Lu+J.&coden=BCMAC&isbn=&pages=-&date=2019&aunit1=J&aunitm=>
33. M. Ohashi, S. Morita, T. Fukagawa, T. Wada, R. Kushima, H. Onaya, H. Katai, Evaluation of 64-Channel Contrast-Enhanced Multi-detector Row Computed Tomography for Preoperative N Staging in cT2-4 Gastric Carcinoma, *World J. Surg.* (2016). <https://doi.org/10.1007/s00268-015-3318-8>.
34. S. Gao, G.H. Cao, P. Ding, Y.Y. Zhao, P. Deng, B. Hou, K. Li, X.F. Liu, Retrospective evaluation of lymphatic and blood vessel invasion and Borrmann types in advanced proximal gastric cancer, *World J. Gastrointest. Oncol.* (2019). <https://doi.org/10.4251/wjgo.v11.i8.642>.
35. Y.Y. Choi, S.J. Kim, C.W. Choi, D.H. Kang, H.W. Kim, S.B. Park, H.S. Nam, Risk factors of submucosal or lymphovascular invasion in early gastric cancer <2cm, *Med. (United States)*. (2016). <https://doi.org/10.1097/MD.0000000000003822>.
36. Z. Ma, C. Liang, Y. Huang, L. He, C. Liang, X. Chen, X. Huang, Y. Xiong, Z. Liu, Can lymphovascular invasion be predicted by preoperative multiphasic dynamic CT in patients with advanced gastric cancer?, *Eur. Radiol.* (2017). <https://doi.org/10.1007/s00330-016-4695-6>.
37. Y. Jiang, Q. Yuan, W. Lv, S. Xi, W. Huang, Z. Sun, H. Chen, L. Zhao, W. Liu, Y. Hu, L. Lu, J. Ma, T. Li, J. Yu, Q. Wang, G. Li, Radiomic signature of 18F fluorodeoxyglucose PET/CT for prediction of gastric cancer survival and chemotherapeutic benefits, *Theranostics*. (2018). <https://doi.org/10.7150/thno.28018>.
38. Z. Liu, B. Feng, C. Li, Y. Chen, Q. Chen, X. Li, J. Guan, X. Chen, E. Cui, R. Li, Z. Li, W. Long, Preoperative prediction of lymphovascular invasion in invasive breast cancer with dynamic contrast-enhanced-MRI-based radiomics, *J. Magn. Reson. Imaging*. (2019). <https://doi.org/10.1002/jmri.26688>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.