



# A Radiomics Nomogram for Non-Invasive Prediction of Progression-Free Survival in Esophageal Squamous Cell Carcinoma

Ting Yan<sup>1</sup>, Lili Liu<sup>1</sup>, Zhenpeng Yan<sup>1</sup>, Meilan Peng<sup>1</sup>, Qingyu Wang<sup>2</sup>, Shan Zhang<sup>2</sup>, Lu Wang<sup>1</sup>, Xiaofei Zhuang<sup>3</sup>, Huijuan Liu<sup>1</sup>, Yanchun Ma<sup>1</sup>, Bin Wang<sup>2\*</sup> and Yongping Cui<sup>1\*</sup>

<sup>1</sup> Key Laboratory of Cellular Physiology of the Ministry of Education, Department of Pathology, Shanxi Medical University, Taiyuan, China, <sup>2</sup> College of Information and Computer, Taiyuan University of Technology, Taiyuan, China, <sup>3</sup> Department of Thoracic Surgery, Shanxi Cancer Hospital, Taiyuan, China

## OPEN ACCESS

### Edited by:

Tianyi Yan,  
Beijing Institute of Technology, China

### Reviewed by:

Ling Chen,  
Tianjin Central Hospital of Gynecology  
Obstetrics, China  
Liqun Kuang,  
North University of China, China

### \*Correspondence:

Yongping Cui  
cuiyp@sxmu.edu.cn  
Bin Wang  
wangbin01@tyut.edu.cn

Received: 27 February 2022

Accepted: 11 April 2022

Published: 16 May 2022

### Citation:

Yan T, Liu L, Yan Z, Peng M, Wang Q, Zhang S, Wang L, Zhuang X, Liu H, Ma Y, Wang B and Cui Y (2022) A Radiomics Nomogram for Non-Invasive Prediction of Progression-Free Survival in Esophageal Squamous Cell Carcinoma. *Front. Comput. Neurosci.* 16:885091. doi: 10.3389/fncom.2022.885091

To construct a prognostic model for preoperative prediction on computed tomography (CT) images of esophageal squamous cell carcinoma (ESCC), we created radiomics signature with high throughput radiomics features extracted from CT images of 272 patients (204 in training and 68 in validation cohort). Multivariable logistic regression was applied to build the radiomics signature and the predictive nomogram model, which was composed of radiomics signature, traditional TNM stage, and clinical features. A total of 21 radiomics features were selected from 954 to build a radiomics signature which was significantly associated with progression-free survival ( $p < 0.001$ ). The area under the curve of performance was 0.878 (95% CI: 0.831–0.924) for the training cohort and 0.857 (95% CI: 0.767–0.947) for the validation cohort. The radscore of signatures' combination showed significant discrimination for survival status. Radiomics nomogram combined radscore with TNM staging and showed considerable improvement over TNM staging alone in the training cohort (C-index, 0.770 vs. 0.603;  $p < 0.05$ ), and it is the same with clinical data (C-index, 0.792 vs. 0.680;  $p < 0.05$ ), which were confirmed in the validation cohort. Decision curve analysis showed that the model would receive a benefit when the threshold probability was between 0 and 0.9. Collectively, multiparametric CT-based radiomics nomograms provided improved prognostic ability in ESCC.

**Keywords:** esophageal squamous cell carcinoma, computed tomography, progression-free survival, radiomics, nomogram

## INTRODUCTION

Esophageal cancer (EC) remains the seventh most frequently occurring cancer and the sixth most prevalent cause of cancer deaths globally (Bray et al., 2018). An estimated 477,900 new cases and 375,000 annual deaths have been reported in China, and most of them were esophageal squamous cell carcinoma (ESCC) (Chen W. et al., 2016). Most patients with ESCC are diagnosed at an advanced stage due to the vague symptoms in the early stage and have a meager 5-year survival rate (<20%) (Pennathur et al., 2013;

Chen et al., 2017). However, surgery is still the most sanative treatment, and the 5-year survival rate of resectable EC treated with surgery alone is only 34–36% (Omloo et al., 2007). Hence, effective means to preoperatively predict the prognosis of patients with ESCC is necessary.

Prognosis survival evaluation of EC mainly depends on traditional Tumor Node Metastasis (TNM) staging. However, the TNM system only considers anatomical features and neglects the intrinsic factors of the tumor, resulting in an inaccurate prognosis (Wang et al., 2011). Then scholars started to collect clinical data, such as age, gender, body mass index (BMI), and the quality of life (Tang et al., 2013; Ng et al., 2014; Zeng et al., 2015; Zhang et al., 2016). However, the performance is still weak, for they failed to reflect the internals of tumors. Furthermore, prognostic evaluation by multi-omics approaches is based on molecular features of a small portion of tumor tissue, which limits the understanding of the heterogeneous tumor.

Radiomics, as a non-invasive, quantitative, and low-cost approach, can objectively and comprehensively evaluate tumor heterogeneity by converting medical images into high-dimensional, mineable, and quantitative imaging features via high-throughput extraction of data-characterization algorithms (Aerts et al., 2014; Gillies et al., 2016). These features can reveal disease progression, providing valuable information for personalized therapy and decision-support (Chicklore et al., 2013; Cameron et al., 2016; Huynh et al., 2016; Jin and Kong, 2016; Kotrotsou et al., 2016; Parekh and Jacobs, 2016; Ginsburg et al., 2017; Lee et al., 2017; Marin et al., 2017; Scalco and Rizzo, 2017; Shafiq-Ul-Hassan et al., 2017). Previous studies have shown that the radiomics signature alone or merged with clinical parameters could enhance predictive accuracy in cancers (Huang Y. et al., 2016; Huang Y. Q. et al., 2016; Zhang et al., 2017). Recently, the most widely-used imaging modality in radiomics is computed tomography (CT), which is universally used for preoperative diagnostics of ESCC. Due to the poor contrast resolution, it is not easy to distinguish the different histologic layers of the esophageal wall. However, it is believed that a lot of digital information could be deeply excavated through radiomics approaches.

In the present study, we developed CT-based radiomics as a novel approach for individualized, pretreatment evaluation of progression-free survival (PFS) in ESCC patients (stage I–III). Additionally, we sought to reveal the association between radiomics and clinical information.

## MATERIALS AND METHODS

### Patients and Clinical Characteristics

Shanxi Medical University Review Board approved this retrospective study. The entire cohort was obtained from the Institutional Picture Archiving and Communication System (PACS) at Shanxi Cancer Hospital from February 2016 to October 2018. The patients who had histologically confirmed ESCC (TNM stage: I–III) and underwent surgery after diagnosis, underwent pretreatment CT scans from neck to abdomen and signed informed consent. All methods were carried out following the relevant guidelines and regulations.

To determine the patients that could be included, we developed the following criteria: (1) pathologically confirmed ESCC; (2) underwent surgery for ESCC; (3) standard contrast-enhanced CT was performed preoperatively; and (4) complete clinical and follow-up information was available. We randomly divided the patients into training and validation cohorts by a ratio of about 3:1. We trained models in the training cohort and validated them in the validation cohort.

Clinical characteristics including age, gender, tumor location (upper, middle, lower), drinking history, smoking history, genetic alterations, and pathologic features including depth of invasion, TNM stage, and lymph node metastasis information were collected from patient records. These clinicopathologic characteristics are presented in **Table 1**.

### Follow Up and Clinical Endpoint

All patients were followed up every 1–3 months during the first 2 years, every 6 months in years 2–5, and annually after

**TABLE 1** | Patient and tumor characteristics in the training and validation cohorts.

	Training (N = 204)	Validation (N = 68)	P-value
Gender			0.539
Male	146 (71.6%)	46 (67.6%)	
Female	58 (28.4%)	22 (32.4%)	
Age			0.398
Median (interquartile range)	60.22	60.44	
≤56	63 (30.9%)	19 (27.9%)	
56–66	92 (45.1%)	27 (39.7%)	
≥66	49 (24.0%)	22 (32.4%)	
Location			0.452
Up	10 (4.9%)	5 (7.4%)	
Mid	135 (66.2%)	48 (70.6%)	
Down	59 (28.9%)	15 (22.1%)	
Drinking			0.662
Yes	75 (36.8%)	23 (33.8%)	
No	129 (63.2%)	45 (66.2%)	
Smoking			0.569
Yes	118 (57.8%)	42 (61.8%)	
No	86 (42.2%)	26 (38.2%)	
Genetic history			0.880
Yes	64 (31.4%)	22 (32.4%)	
No	140 (68.6%)	46 (67.6%)	
Invasion degree			0.887
Full layer	121 (59.3%)	41 (60.3%)	
Non-full layer	83 (40.7%)	27 (39.7%)	
TNM			0.556
I	19 (9.3%)	9 (13.2%)	
II	105 (51.5%)	36 (52.9%)	
III	80 (39.2%)	23 (33.8%)	
Lymph node metastasis			0.255
Yes	88 (43.1%)	24 (35.3%)	
No	116 (56.9%)	44 (64.7%)	

that. To provide an efficient tool, which would allow earlier personalized treatment, we chose PFS as the endpoint (Sargent et al., 2005). We defined PFS from the first day of treatment to the date of disease progression (locoregional recurrences or distant metastases), death from any cause, or the date of the last follow-up visit (censored). The minimum follow-up time to ascertain the PFS was 2 months.

### CT Acquisition and Segmentation

All patients were underwent the contrast-enhanced CT using a 64-channel multi-detector CT scanner (LightSpeed VCT, GE Medical Systems, Milwaukee, Wis, USA). The acquisition parameters were: 120 kV; 160 mA; 0.5-s rotation time; detector collimation: 64 × 0.625 mm; field of view: 350 × 350 mm; and matrix: 512 × 512. After routine non-enhanced CT, contrast-enhanced CT was performed after a 25-s delay following intravenous administration of 85 mL of iodinated contrast material (Ultravist 370; Bayer Schering Pharma, Berlin, Germany) at a rate of 3.0 mL/s with a pump injector (Ulrich CT Plus 150, Ulrich Medical, Ulm, Germany). All images were reconstructed with a thick slice of 5.0 mm. We converted the image format from DICOM to NII for feature selection without any preprocessing.

Note that segmentation is required before the extraction of quantitative radiomics features; we performed three-dimensional manual segmentation using 3D-Slicer software (<https://www.slicer.org/>), an open platform for medical image processing. The chief physician of Shanxi Cancer Hospital, with more than 5 years' experience in interpreting chest radiology, outlined the tumor regions for each CT image layer. The tumor segmentation was guided and verified by the specialist. The region of interest (ROI) covered the whole tumor mass, was delineated on each CT slice, and used in subsequent feature extraction.

### Selection of Radiomics Features and Building of Radiomics Signature

We performed the calculation through our homemade Python scripts (Python3.6, <https://www.python.org>) for radiomics feature extraction based on the segmentation results. A total of 954 features were obtained by calling feature calculation in pyradiomics package (open-source python package; <https://pyradiomics.readthedocs.io/en/latest/>), which included the following four categories: (1) first-order statistics features; (2) size- and shape-based features; (3) texture features; and 4) wavelet features; and five typical matrixes: Gray-Level Co-occurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM), Gray Level Size Zone Matrix (GLSZM), Gray Level

**TABLE 3 |** Radiomics signature selection results with descriptions.

Feature name	Feature coefficient
HLL_firstorder_Skewness	0.066
HLH_firstorder_Median	-1.812
HLH_glszm_SmallAreaEmphasis	-13.697
HLH_glszm_ZoneEntropy	0.092
HLL_glcm_ClusterShade	-0.004
HLL_glcm_InverseVariance	-6.470
HLL_glszm_GrayLevelNonUniformityNormalized	-0.612
HLL_glszm_SizeZoneNonUniformityNormalized	15.084
HLL_gldm_SmallDependenceHighGrayLevelEmphasis	-0.0008
HLL_ngtdm_Complexity	-0.0007
LHH_glszm_LargeAreaLowGrayLevelEmphasis	1.02e-06
LHH_gldm_DependenceNonUniformityNormalized	31.635
LHH_ngtdm_Busyness	0.001
LHL_glcm_Idn	12.445
LHL_glszm_LargeAreaHighGrayLevelEmphasis	-1.42e-10
LHL_gldm_SmallDependenceLowGrayLevelEmphasis	-82.462
LLH_firstorder_Energy	9.77e-10
LLH_glcm_Contrast	0.029
LLH_glszm_SizeZoneNonUniformity	1.37e-05
LLH_ngtdm_Complexity	5.66e-05
LLL_gldm_LargeDependenceHighGrayLevelEmphasis	2.44e-06

*Skewness:*The asymmetric distribution of the Mean value. Depending on where the tail is elongated and the mass of the distribution is concentrated, it can be positive or negative.  
*Median:*The median gray level intensity within ROI.  
*Small Area Emphasis (SAE):* A measure of the distribution of small size zones, with a greater value indicative of more smaller size zones and more fine textures.  
*Zone Entropy (ZE):* The degree of instability and variation in spatial and regional differences of the image distribution range.  
*Cluster Shade:* A measure of skewness and uniformity of the GLCM. A higher cluster shade implies greater asymmetry about the mean.  
*Size Zone Non-Uniformity Normalized:* The variability of size zone volumes throughout images, with a lower value indicating more homogeneity among zone size volumes in images. it's the normalized version of the SZN formula.  
*Small Dependence High Gray Level Emphasis (SDHGLE):* Measures the joint distribution of small dependence with higher gray-level values.  
*Large Area Low Gray Level Emphasis (LALGLE):* The proportion in images of the joint distribution of larger size zones with lower gray-level values.  
*Dependence Non-Uniformity Normalized (DNN):* Measures the similarity of dependence throughout images, with a lower value indicating more homogeneity among dependencies in images. This is the normalized version of the DLN formula.  
*Busyness:* A measure of the change from a pixel to its neighbor. A high value for busyness indicates a 'busy' image, with rapid changes of intensity between pixels and their neighborhood.  
*IDN (inverse difference normalized):* Another measure of a local homogeneity of images. Unlike Homogeneity1, IDN normalizes the difference between neighboring intensity values by dividing over the total number of discrete intensity values.  
*Large Area High Gray Level Emphasis (LAHGLE):* The proportion in images of the joint distribution of larger size zones with higher gray-level values.  
*Small Dependence Low Gray Level Emphasis (SDLGLE):* Measures the joint distribution of small dependence with lower gray-level values.  
*Energy:* The sum of squares of gray level intensity within ROI.  
*Contrast:* A measure of local intensity variation, favoring values away from the diagonal (i=j). A larger value correlates with a greater disparity in intensity values among neighboring voxels.  
*Large Dependence High Gray Level Emphasis (LDHGLE):* Measures the joint distribution of large dependence with higher gray-level values.

**TABLE 2 |** Radiomics features selection results based on the ANOVA.

Result category	CT
Number of selected features	221
The best-performance feature	HLL-original_glcm_InverseVariance ( $\rho = 2.316589e-04$ )

Dependence Matrix (GLDM) and Neighboring Gray Tone Difference Matrix (NGTDM).

We built the radiomics signature with selected features in the training cohort. To reduce over-fitting or any types of bias, we applied the following two steps: First, the best features based on univariate statistical tests (2-sample *t*-test) between death and censoring groups in the primary cohort were selected and executed by using Matlab 2016b. Second, we used our homemade R scripts to select features that were most significant by using the least absolute shrinkage and selection operator (LASSO) method, which would be a suitable methodology for the feature selection through regression of high-dimensional data. Additionally, the accuracy of the prediction model could be improved by regularizing the features through penalized estimation. We added the L1 penalty term to the normal linear model, and the parameter lambda controls the complexity of regression. When the  $\lambda$  is large, it indicates no effect on the estimated regression parameters, while the  $\lambda$  gets smaller, most covariate coefficients were shrunk to zero. Then the remaining variables with non-zero coefficients were selected by the  $\lambda$  that the 10-fold cross-validation error was the most minor (Kumamaru et al., 2016; Vasquez et al., 2016). When performing 10-fold cross-validation, the training cohort was divided into 10 equal parts; each called a fold. Next, it will train a series of models. The first model was trained using the first fold as the test set and the other folds (2–10) as the training set. Then another model was constructed using 2nd fold as the test set and the 1st, 3th–10th folds as the training set. The process is repeated with 3th–10th folds as a test set.

Finally, the radiomics signature was built by combining those variables in the primary cohort and validated in the validation cohort. The radiomics signature is a linear combination of selected features with respective weights, which would be calculated as a factor (Radiomics score, Rad-score) for the further prediction model. The Rad-score calculated by using the following formula:

$$\text{Rad-score} = c_1F_1 + c_2F_2 + \dots + c_iF_i \quad (1)$$

Where  $F_i$  is the selected radiomics feature,  $c_i$  is the LASSO coefficient of  $F_i$ . Then, the assessment method of the logistic regression model is the receiver operating characteristic (ROC) curve and its area under the curve (AUC).

## Prognostic Validation of Radiomics Signature

We calculated Rad-score for each patient with ESCC and grouped them according to two rules. (1) The patients were divided into high-risk and low-risk groups based on the median Rad-score. (2) Patients with median scores were placed in high-risk groups. The radiomics signature discriminative performance of the survival status was assessed according to the overall distribution of ESCC patients. And then, the potential association of radiomics signature and clinical feature with PFS was assessed in the training cohort and validated in the validation cohort. Kaplan–Meier survival analysis was used in these two cohorts. Stratified

analyses were implemented to determine the PFS in high-risk and low-risk patient subgroups. Univariate Cox Proportional Hazards Models were performed to explore the C-index of the radiomics signature.

## Performance of TNM Staging and Clinical Nomograms in the Training Cohort Before and After Addition of Rad-Score

The nomogram with the predicting model was based on the multivariable logistic regression analysis. The following candidate factors: TNM stage (dummy variable: “0” for I, “1” for II, “2” for III), the status of clinical features, and Rad-scores were involved in a diagnostic model for preoperative prediction of ESCC. The nomogram is a graphical representation of this prediction model in the training cohort. The prognostic performance of TNM staging and clinical nomograms in the training cohort before and after the addition of the Rad-score was quantitatively measured using Harrell’s concordance index (C-Index), which is commonly used for the evaluation of the discriminative power of prognostic models (Harrell, 2018). The value of the C-index could range from 0.5, which indicated no discriminative ability, to 1.0, which showed perfect ability to distinguish between the patients who suffered disease progression or death and those who did not. Bootstrap analyses with 1,000 resample were applied to obtain a C-index with a 95% confidence interval (CI) (Canty and Ripley, 2021) corrected for potential overfitting. The calibration curves were drawn for assessing the agreement between the predicted probability of 3-year PFS and actual 3-year PFS (Pencina et al., 2011).

## Nomogram Validation in the Validation Cohort

The prognostic performance of TNM staging and clinical nomograms in the validation cohort before and after the addition of the Rad-score was tested by the above method. The calibration curve and C-index were calculated through multivariable Cox proportional hazard regression analyses. The decision curve analysis (DCA) was introduced to evaluate the quantified net benefit of our prediction model in the validation cohort (Vickers et al., 2008; Shen et al., 2018).

## Association of Radiomics Features With Clinical Data

A heat map analysis was used to evaluate the associations between clinical data and radiomics features.

## Statistical Analysis

All the statistical analyses were performed using IBM SPSS software (version 26; IBM Corp, Armonk, NY, USA), Matlab 2016b (Mathworks, Natick, USA), and R software (version 4.1.1, Boston, MA, USA). In this study, 2-sample *t*-test was applied to confirm whether differences between death and censoring groups in the primary cohort. The differences in gender, age, TNM stage, smoking status, drinking status, location, genetic history, invasion degree and metastasis between the training and validation data sets were assessed by using the  $\chi^2$ -test. The

following R packages were used: the glmnet package was used for the LASSO logistic regression model analysis, the pROC package was used for the ROC curves, the Hmisc package was used for the comparisons between the C-indices, the survival package was used for Kaplan–Meier survival analyses, the rms package was used for the nomograms and calibration curves, the rmda package was used to implement the DCA, and the gplots and pheatmap packages were used for heat maps.

## RESULTS

### Clinical Characteristics of All the Patients

A total of 272 consecutive patients who met the criteria (192 men and 80 women; mean age, 60.27 years  $\pm$  7.43) were included and divided into two cohorts by a ratio of 3:1 using computer-generated random numbers. A total of 204 patients were enrolled in the training cohort (146 men and 58 women; mean age, 60.22 years  $\pm$  7.28), while 68 patients were enrolled in the independent validation cohort (46 men and 22 women; mean age, 60.44 years  $\pm$  7.90). The clinical characteristics with statistics of the training and validation cohorts are summarized in **Table 1**. No significant differences were found between these two cohorts in terms of gender, age, history of smoking and drinking, location, genetic history, invasion degree, lymph node metastasis, and overall TNM Stage ( $p = 0.255$ – $0.887$ ). The median PFS was 35.35 months (range, 2–75 months).

### Radiomics Feature Selection and Radiomics Signature Building

A total of 954 features were extracted from CT images and might contain many redundant and highly correlated features. To find out robust and valuable features, we performed the following steps: First, 221 features were selected by univariate statistical tests ( $p < 0.05$ ) (**Table 2**). Then, based on the LASSO logistic regression algorithm approach in the training cohort, we selected the features with non-zero coefficients. As a result, 21 radiomics features were screened out from 221 features (**Table 3**). The procedures of parameter tuning and feature space reduction of the regression model are illustrated in **Figure 1**. Then the 21 features were selected to build the radiomics signature and involved in the Rad-score-based prognostic model. The discriminative ability of the survival status based on radiomics signatures was assessed by ROC in both cohorts, respectively (**Figure 2A**).

### Prognostic Validation of Radiomics Signature

Rad-score for each patient in the training cohort and the validation cohort correspondingly showed that the higher the Rad-score, the greater the probability of death (**Figures 2B,C**). Besides, in the training cohort, the radiomics signature from CT images yielded the highest C-index, which was 0.758 (95% CI: 0.708–0.808). In the validation cohort, the radiomics signature from CT images yielded a C-index of 0.748 (95% CI: 0.656–0.840). It showed a significant discrimination between the PFS of high-risk and low-risk patients (**Figure 3**).

### Performance of TNM Staging and Clinical Nomograms in the Training Cohort Before and After the Addition of Rad-Score

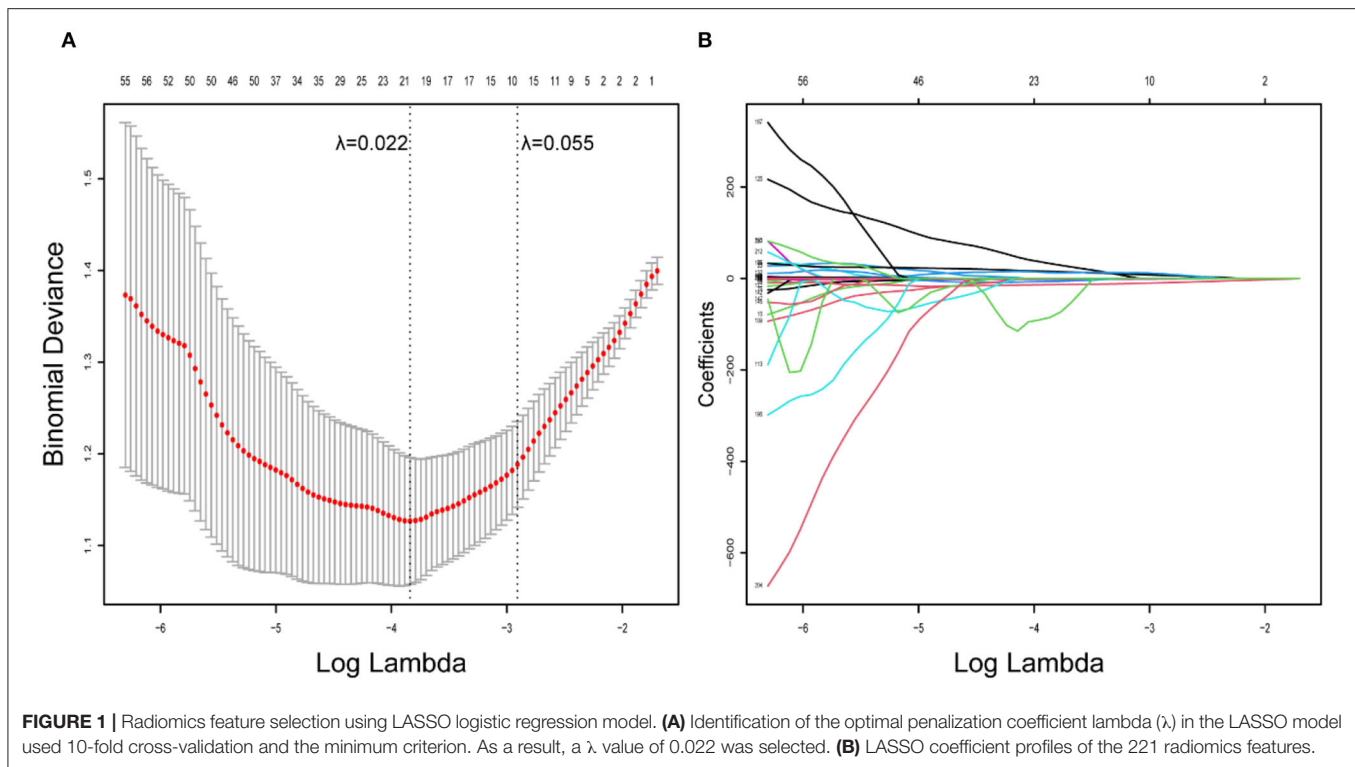
We developed a radiomics nomogram that integrated the radiomics signature from the CT images with the traditional TNM staging system, which yielded a C-index of 0.603 (95% CI: 0.549–0.657). This nomogram significantly improved the discrimination ability in evaluating PFS (C-index: 0.770; 95% CI: 0.721–0.819) than TNM staging system ( $p < 0.05$ ; **Figure 4A**), and showed good calibration as well (**Figure 4B**). Moreover, a radiomics nomogram was created by integrating the radiomics signature from the CT images with all clinical data, whose nomogram yielded a C-index of 0.680 (95% CI: 0.626–0.734). We found that the radiomics nomogram possessed good calibration and seemed to be more accurate than the clinical nomogram for evaluating PFS (C-index: 0.792; 95% CI: 0.748–0.836) with a  $p < 0.05$  (**Figures 4C,D**).

### The Validation of Nomograms in the Validation Cohort

In the validation cohort, the C-index of the traditional TNM staging system is 0.572 (95% CI: 0.478–0.666). We integrated the radiomics signature with the TNM staging system to produce a radiomics nomogram, which showed an improvement over the TNM staging system alone (C-index: 0.760; 95% CI: 0.673–0.847). The calibration curve of probability in PFS evaluation showed good agreement between nomogram-evaluated and actual observation (Figure not shown). While the clinical nomogram yielded a C-index of 0.605 (95% CI: 0.501–0.709) in the validation cohort and was advanced by combining with radiomics signature (C-index: 0.779; 95% CI: 0.697–0.861). The calibration curves of this nomogram showed good agreement between nomogram-evaluated and actual survival (Figure not shown). The DCA for the prediction model derived from the addition of Rad-score before and after is presented in **Figure 5A**. It showed that the predictive model collaborated with Rad-score had a better net benefit than that with only traditional TNM staging combined with clinical features.

### Association of Radiomics Features With Clinical Data

The ESCC patients with similar patterns of radiomics expression were clustered through unsupervised clustering (**Figure 5B**). Then we organized a heat map to determine the association between radiomics features and clinical data (**Figure 5B**). The results showed significant correlations between signature features LHL\_glcm\_Idn, LHL\_glszm\_Large Area High Gray Level Emphasis with drinking ( $p < 0.001$ ) as well as gender ( $p < 0.001$ ). LHL\_glcm\_Idn was associated with smoking ( $p < 0.01$ ). LHL\_gldm\_Small Dependence Low Gray Level Emphasis was associated with location ( $p < 0.01$ ). LHL\_glcm\_Idn and LHL\_glszm\_Large Area High Gray Level Emphasis were associated with invasion degree ( $p < 0.05$ ). HLH\_glszm\_Zone Entropy, LHL\_glcm\_Idn and LLH\_firstorder\_Energy were associated with overall stage ( $p < 0.05$ ). In contrast, no radiomics feature was significantly associated with age, genetic history



and metastasis (for all,  $p > 0.05$ ). These  $p$ -values for all the correlations of radiomics features with clinical data are shown in **Supplementary Table 1**.

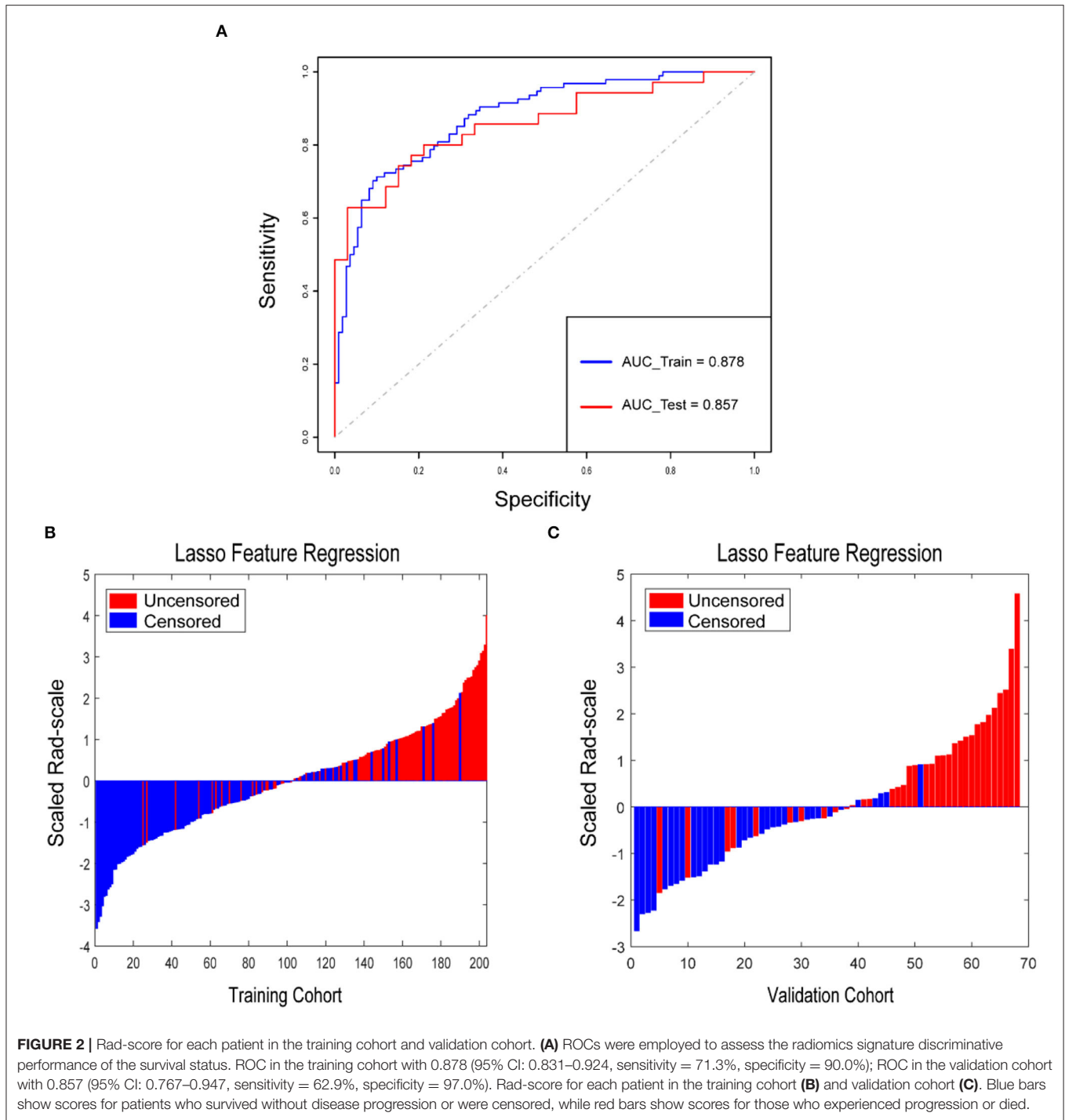
## DISCUSSION

Here, we first developed and validated a new approach based on CT radiomics to evaluate PFS before treatment in ESCC (stage I-III). The radiomics signature from CT images demonstrated better prognostic performance than traditional clinical information alone. It could be competently differentiated between patients with high-risk and low-risk, who had significantly different 3-year PFS, and were defined according to the median Rad-score. The developed radiomics nomogram transcended the traditional TNM staging system and clinical nomogram alone.

In clinical practice, CT, magnetic resonance imaging (MRI), positron emission tomography (PET), and endoscopic ultrasound (EUS) have their own advantages and disadvantages in the staging of esophageal cancer or even cancer. But the use of these modalities is limited to their cost in both time and money. CT owns the highest cost performance of its high availability and non-invasive process. However, the traditional prognosis is dependent on the doctors' observation, which differs greatly according to the experience. Moreover, the evaluation from traditional clinical information is even more inadequate. It is believed that there is still a lot of digital data that can be deeply excavated through the radiomics methodology and used for judgment conversely. Therefore, we analyzed all acquired CT

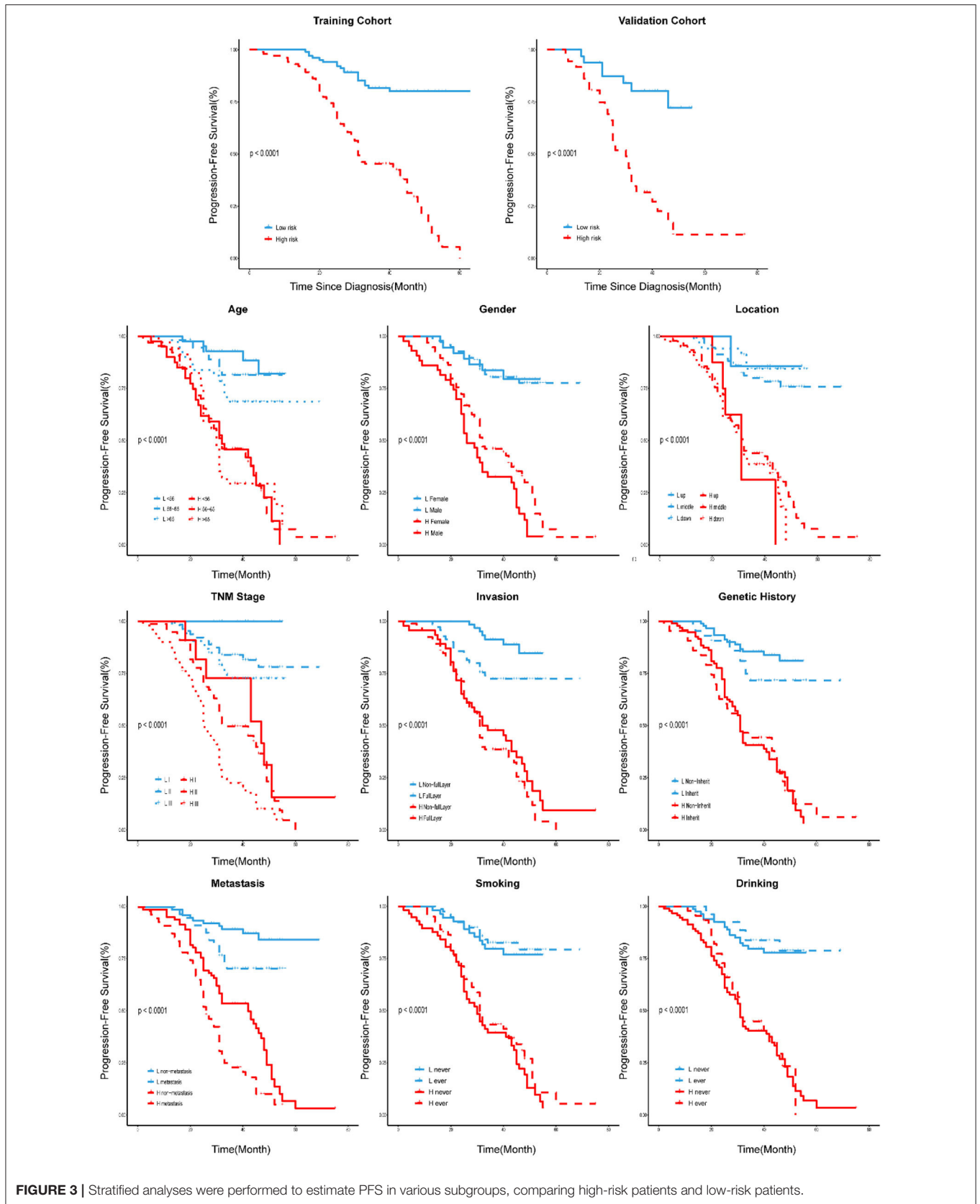
images and constructed a CT-based radiomics signature. And the results confirmed our expectations that the radiomics signatures have the potential for evaluating prognosis in ESCC.

To build the radiomics signature, we selected 21 potential predictors from 954 candidate features by selecting highly correlated features with event outcomes and LASSO logistic regression. The radiomics features obtained are generally accurate. Because the regression coefficients of most features have shrunk toward zero during model fitting. It allowed the identification of features that had the strongest association with PFS (Ndhlovu et al., 2013) and avoided overfitting (Hepp et al., 2016). The radiomics signatures could reveal adequate discrimination both in the training cohort (C-index, 0.758) and the validation cohort (C-index, 0.748). Additionally, the selected features were used to improve radiomics signature and Rad-scores. We sorted the Rad-scores of all the patients with the labeled living status in **Figure 2B**, suggesting that the Rad-score could potentially differentiate the two types of patients. Other related statistical analyses also supported that the radiomics signature could be used as a biomarker in the prognosis of ESCC. We found that compared to the traditional TNM staging system and clinical nomogram, the radiomics signature dominated our nomogram in the training and validation cohorts. It means the radiomics signature has better discrimination and prognosis than classical radiologists, indicating the clinical importance of our findings due to the traditional clinical information and TNM staging are routinely used in clinical practice (Li et al., 2015; Wu et al., 2015).



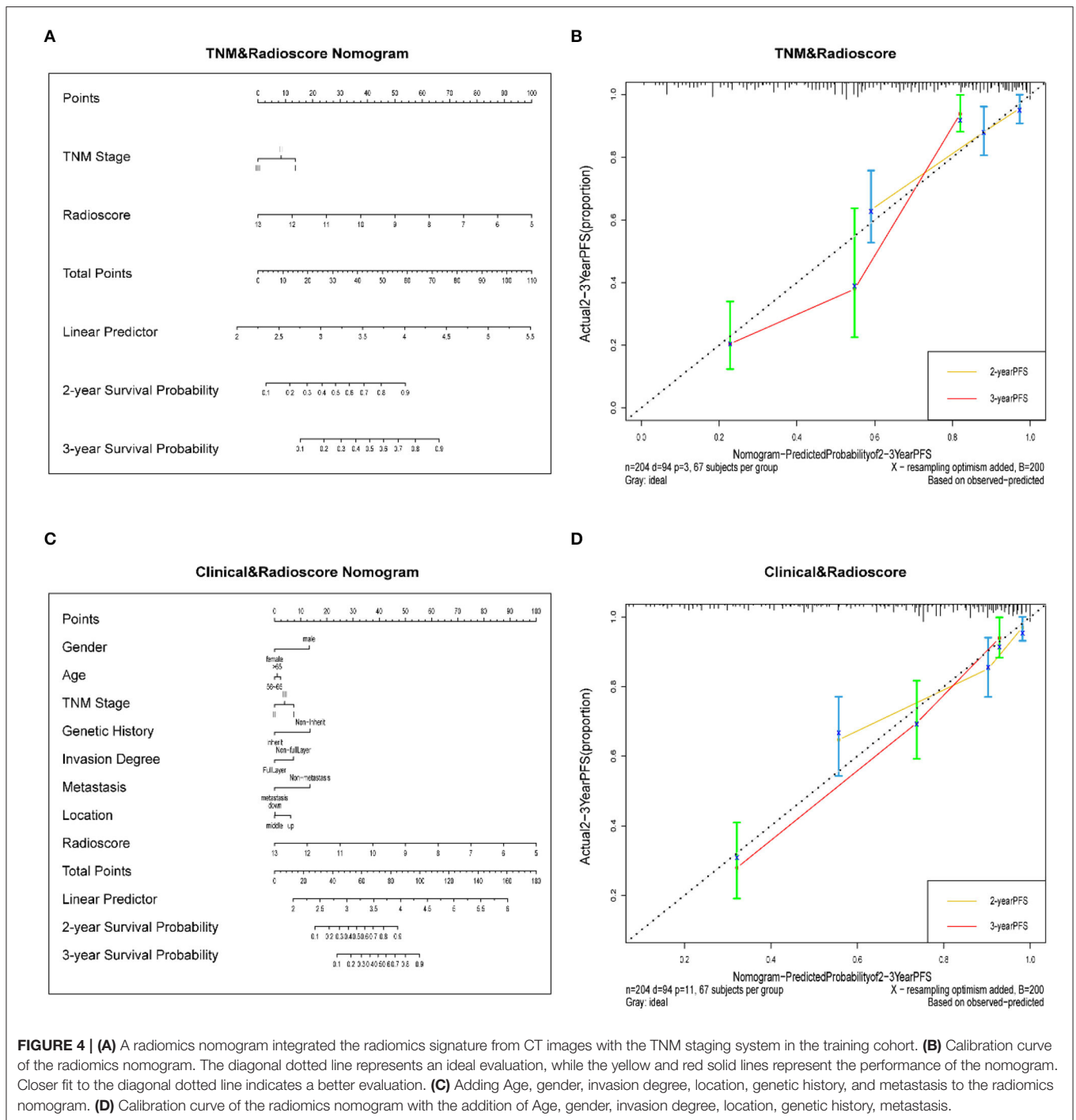
Generally, doctors use the traditional TNM staging system for risk prediction and treatment planning. However, there were apparent differences in PFS with the same clinical identified disease stage, indicating that tumor heterogeneity would affect the survival outcomes. Patients with ESCC (stage I-III) with shorter PFS may benefit from the prognostic model because they may give up aggressive treatments to avoid suffering and overspending. Here, we developed the radiomics features

possessing better prognostic ability than the traditional TNM staging system for pretreatment of PFS in the validation and training cohorts. Our study focused on the patients with stage I-III tumors (Table 1), and the patients with stage I accounted for a small proportion (9.3% in the training cohort, 13.2% in the validation cohort). Consequently, it might be not easy to stratify PFS accurately since the similar information of clinical stage. Additionally, the traditional TNM stage mainly reflects



**FIGURE 3 |** Stratified analyses were performed to estimate PFS in various subgroups, comparing high-risk patients and low-risk patients.

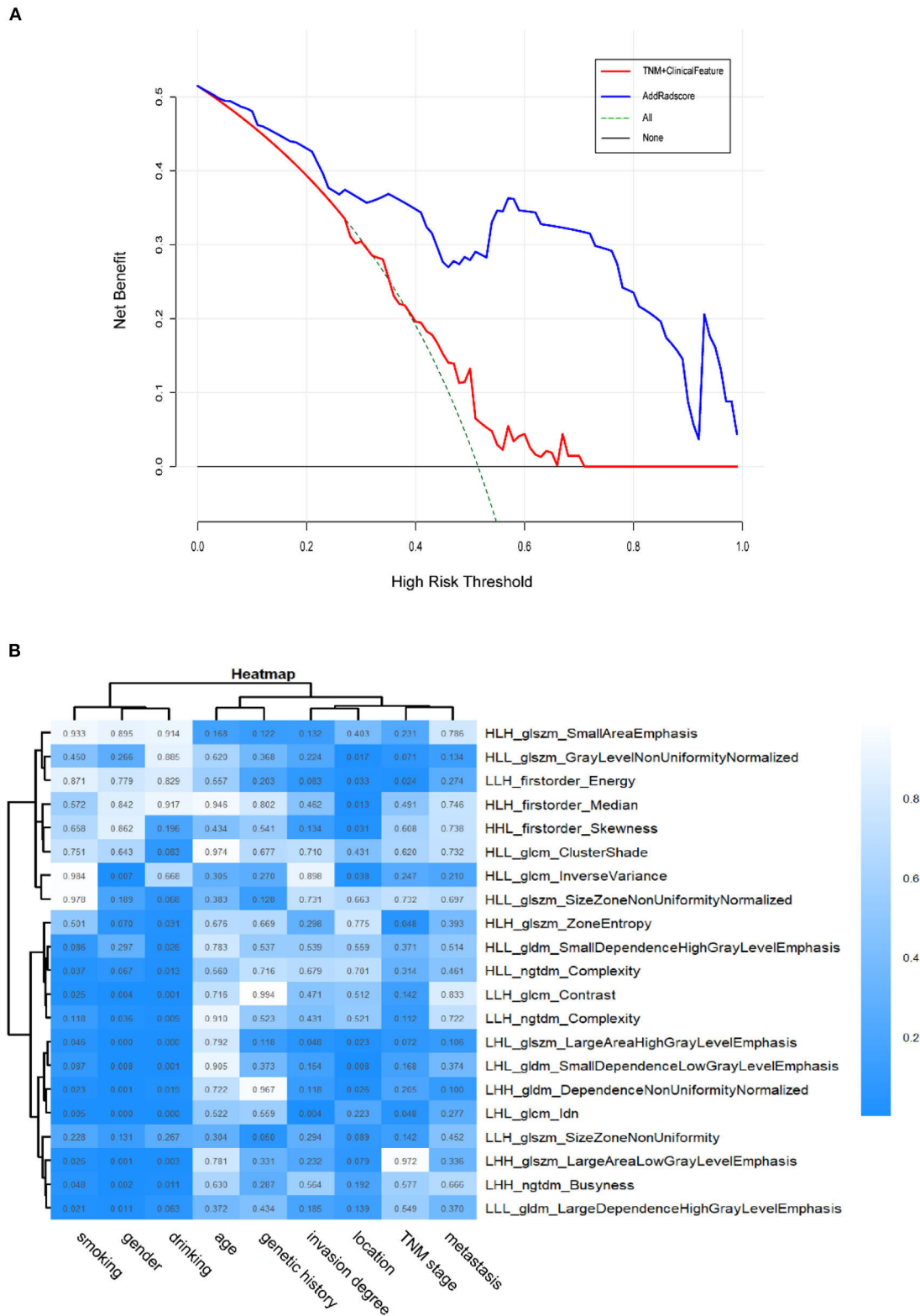




**FIGURE 4 | (A)** A radiomics nomogram integrated the radiomics signature from CT images with the TNM staging system in the training cohort. **(B)** Calibration curve of the radiomics nomogram. The diagonal dotted line represents an ideal evaluation, while the yellow and red solid lines represent the performance of the nomogram. Closer fit to the diagonal dotted line indicates a better evaluation. **(C)** Adding Age, gender, invasion degree, location, genetic history, and metastasis to the radiomics nomogram. **(D)** Calibration curve of the radiomics nomogram with the addition of Age, gender, invasion degree, location, genetic history, metastasis.

cancer patients’ clinicopathologic features, such as tumor size, lymph node involvement, and distant metastasis status. They do have prognostic value in tumor treatment but neglected the intratumor heterogeneity, which was deemed as a crucial factor for tumor progression and prognosis (Yan et al., 2019). As a result, it provided an inefficient nomogram performance in both the training cohort (C-index, 0.603) and the validation cohort (C-index, 0.572). While the radiomics approach extracted

the features of the entire tumor from medical images, which produced a more comprehensive way to involve the intratumor heterogeneity non-invasively. It might be why the combination of radiomics signatures and traditional TNM staging could provide a better nomogram performance in both the training cohort (C-index, 0.770) and validation cohort (C-index, 0.760). Hence, the radiomics signatures could assist the prognosis for ESCC complementarily to the traditional TNM staging.



**FIGURE 5 | (A)** The DCA of the radiomics-comparison-based nomogram. The black dotted line describes the scheme of no treatment. The green dotted line describes the scheme of treatment. The red line represents our predictive model with only traditional TNM staging combined with clinical features. And the blue line (Continued)

**FIGURE 5** | represents our personalized prediction model that added Rad-score. The x-axis is the threshold probability and the y-axis is the net benefit. It can be seen the personalized prediction model with Rad-score added had a better net benefit than the traditional predictive model when the threshold is in the range of 0–0.9. Hence, the patient with ESCC would receive benefit from taking our CT-based radiomics nomogram guidance. **(B)** Heatmap of associations between selected radiomics features and clinical data.  $p < 0.05$  indicates statistical associations, as determined using *t*-tests.

Previous studies reported that clinical information including gender, pathological type, tumor differentiation, depth of invasion, and regional lymph node metastasis was associated with overall survival (OS) outcomes through univariate analysis. While multivariate analysis showed that pathological type, depth of invasion, and regional lymph node metastasis were the independent predictors of OS (Liu et al., 2019). Besides, the tumor volume of ESCC could be used as an important prognostic factor for radiotherapy and chemotherapy assessment (Chen et al., 2013; Li et al., 2013; Chen Y. et al., 2016). Therefore, we exploited a clinical nomogram that combined available risk factors (age, gender, invasion degree, location, genetic history, metastasis) with the overall stage, but it doesn't exhibit well (C-index of training cohort, 0.680; C-index of validation cohort, 0.605). Then, we developed the nomogram by combining the radiomics signature in both the training cohort (C-index, 0.792) and the validation cohort (C-index, 0.779). This process suggested that radiomics signatures have important prognostic value for patients with ESCC.

Unlike the traditional methods, radiomics system is a non-invasive and low-spending approach, which could provide new insights into the associations between intrinsic tumor properties and biological behaviors. We analyzed the relationship between radiomics features and tumor-associated characteristics and observed some radiomics features were related to the general information of patients (gender, drinking, or smoking information, **Figure 5B**). Additionally, our radiomics system showed some radiomics features associated with invasion degree as well (**Figure 5B**). As a result, the present study may provide some different insights into the mechanisms of lymphatic metastasis of ESCC, which require future investigation.

There were several limitations in our study. First, we used thick-slice CT images rather than thin-slice images to extract radiomics signatures. Zhao et al. (2016) found that thin-slice images could reflect texture features of tumors more complete than thick-slice images. For the measurement of tumor volumes, thin-slice images had less measurement variability. We will further study the effect of thin-slice CT images for the staging of ESCC and confirm whether the performance is comparable with thick-slice images. Second, all data involved in this study are derived from the same hospital, resulting in the lack of multi-center validation. Further investigations on the applicability to the patients of other institutions are still required. Third, the analysis did not cover two-way or higher-order interactions of the radiomics features. If the interaction(s) strongly associated with

the outcomes were applied, the prognostic performance of our nomogram might be significantly improved. However, revealing the interactions of multiple factors is challenging. In brief, our study clearly showed that the radiomics approach is potential for the prognosis of ESCC patients.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Shanxi Medical University Review Board. Written informed consent to participate in this study was provided by the patient/participants or patient/participants' legal guardian/next of kin.

## AUTHOR CONTRIBUTIONS

TY conceived the study, designed the experiments, analyzed the data, and wrote the manuscript. BW and YC edited the manuscript. LL, MP, ZY, QW, and SZ supervised data analysis. XZ provided clinical information and coordinated and performed segmentation of CT images. LW, HL, and YM performed the statistical analyses. All authors accessed the study data and reviewed and approved the final manuscript.

## FUNDING

This work was supported by funding from the National Natural Science Foundation of China (62176177, 81702449); the Fundamental Research Program of Shanxi Province (20210302123292, 20210302123112), the Research Project Supported by Shanxi Scholarship Council of China (2021-039), the Central Guidance on Local Science and Technology Development Fund of Shanxi Province (YDZJSX2021A018), the Shenzhen Project of Science and Technology (JCYJ20190813094203600).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncom.2022.885091/full#supplementary-material>

## REFERENCES

- Aerts, H. J., Velazquez, E. R., Leijenaar, R. T., Parmar, C., Grossmann, P., and Carvalho, S. (2014). Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.* 5, 4006. doi: 10.1038/ncomms5006
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Cameron, A., Khalvati, F., Haider, M. A., and Wong, A. (2016). MAPS: a quantitative radiomics approach for prostate cancer detection. *IEEE Trans. Biomed. Eng.* 63, 1145–1156. doi: 10.1109/TBME.2015.2485779
- Canty, A., and Ripley, B. D. (2021). *Bootstrap R (S-Plus) Functions. R package Version.* Hamilton.
- Chen, C. Z., Chen, J. Z., Li, D. R., Lin, Z. X., Zhou, M. Z., and Li, D. S. (2013). Long-term outcomes and prognostic factors for patients with esophageal cancer following radiotherapy. *World J. Gastroenterol.* 19, 1639–1644. doi: 10.3748/wjg.v19.i10.1639
- Chen, W., Zeng, H., Chen, R., Xia, R., Yang, Z., and Xia, C. (2017). Evaluating efficacy of screening for upper gastrointestinal cancer in China: a study protocol for a randomized controlled trial. *Chin. J. Cancer Res.* 29, 294–302. doi: 10.21147/j.issn.1000-9604.2017.04.02
- Chen, W., Zheng, R., Baade, P. D., Zhang, S., Zeng, H., and Bray, F. (2016). Cancer statistics in China, 2015. *CA Cancer J. Clin.* 66, 115–132. doi: 10.3322/caac.21338
- Chen, Y., Zhang, Z., Jiang, G., and Zhao, K. (2016). Gross tumor volume is the prognostic factor for squamous cell esophageal cancer patients treated with definitive radiotherapy. *J. Thorac. Dis.* 8, 1155–1161. doi: 10.21037/jtd.2016.04.08
- Chicklore, S., Goh, V., Siddique, M., Roy, A., Marsden, P. K., and Cook, G. J. (2013). Quantifying tumour heterogeneity in 18F-FDG PET/CT imaging by texture analysis. *Eur. J. Nucl. Med. Mol. Imaging* 40, 133–140. doi: 10.1007/s00259-012-2247-0
- Gillies, R. J., Kinahan, P. E., and Hricak, H. (2016). Radiomics: images are more than pictures, they are data. *Radiology* 278, 563–577. doi: 10.1148/radiol.2015151169
- Ginsburg, S. B., Alghohary, A., Pahwa, S., Gulani, V., Ponsky, L., and Aronen, H. J. (2017). Radiomic features for prostate cancer detection on MRI differ between the transition and peripheral zones: preliminary findings from a multi-institutional study. *J. Magn. Reson. Imaging* 46, 184–193. doi: 10.1002/jmri.25562
- Harrell, F. E. (2018). *Hmisc: Harrell Miscellaneous. R Package Version.* Nashville.
- Hepp, T., Schmid, M., Gefeller, O., Waldmann, E., and Mayr, A. (2016). Approaches to regularized regression - a comparison between gradient boosting and the lasso. *Methods Inf. Med.* 55, 422–430. doi: 10.3414/ME16-01-0033
- Huang, Y., Liu, Z., He, L., Chen, X., Pan, D., and Ma, Z. (2016). Radiomics signature: a potential biomarker for the prediction of disease-free survival in early-stage (I or II) non-small cell lung cancer. *Radiology* 281, 947–957. doi: 10.1148/radiol.2016152234
- Huang, Y. Q., Liang, C. H., He, L., Tian, J., Liang, C. S., and Chen, X. (2016). Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J. Clin. Oncol.* 34, 2157–2164. doi: 10.1200/JCO.2015.65.9128
- Huynh, E., Coroller, T. P., Narayan, V., Agrawal, V., Hou, Y., and Romano, J. (2016). CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. *Radiother. Oncol.* 120, 258–266. doi: 10.1016/j.radonc.2016.05.024
- Jin, J. Y., and Kong, F. M. (2016). Personalized Radiation Therapy (PRT) for Lung Cancer. *Adv. Exp. Med. Biol.* 890, 175–202. doi: 10.1007/978-3-319-24932-2\_10
- Kotrotsou, A., Zinn, P. O., and Colen, R. R. (2016). Radiomics in brain tumors: an emerging technique for characterization of tumor environment. *Magn. Reson. Imaging Clin. N. Am.* 24, 719–729. doi: 10.1016/j.mric.2016.06.006
- Kumamaru, K. K., Saboo, S. S., Aghayev, A., Cai, P., Quesada, C. G., and George, E. (2016). CT pulmonary angiography-based scoring system to predict the prognosis of acute pulmonary embolism. *J. Cardiovasc. Comput. Tomogr.* 10, 473–479. doi: 10.1016/j.jcct.2016.08.007
- Lee, G., Lee, H. Y., Park, H., Schiebler, M. L., van Beek, E. J. R., and Ohno, Y. (2017). Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: State of the art. *Eur. J. Radiol.* 86, 297–307. doi: 10.1016/j.ejrad.2016.09.005
- Li, A. C., Xiao, W. W., Wang, L., Shen, G. Z., Xu, A. A., and Cao, Y. Q. (2015). Risk factors and prediction-score model for distant metastasis in nasopharyngeal carcinoma treated with intensity-modulated radiotherapy. *Tumour Biol.* 36, 8349–8357. doi: 10.1007/s13277-015-3574-0
- Li, R., Chen, T. W., Hu, J., Guo, D. D., Zhang, X. M., and Deng, D. (2013). Tumor volume of resectable adenocarcinoma of the esophagogastric junction at multidetector CT: association with regional lymph node metastasis and N stage. *Radiology* 269, 130–138. doi: 10.1148/radiol.13122269
- Liu, Y., Chen, X., Wang, Y., Wang, F., Gong, Y., and Zhang, J. (2019). Clinical features and prognostic factors for surgical treatment of esophageal squamous cell carcinoma in elderly patients. *J. Buon.* 24, 1240–1244.
- Marin, Z., Batchelder, K. A., Toner, B. C., Guimond, L., Gerasimova-Chechkina, E., and Harrow, A. R. (2017). Mammographic evidence of microenvironment changes in tumorous breasts. *Med. Phys.* 44, 1324–1336. doi: 10.1002/mp.12120
- Ndhlovu, Z. M., Chibnik, L. B., Proudfoot, J., Vine, S., McMullen, A., Cesa, K., et al. (2013). High-dimensional immunomonitoring models of HIV-1-specific CD8 T-cell responses accurately identify subjects achieving spontaneous viral control. *Blood* 121, 801–11. doi: 10.1182/blood-2012-06-436295
- Ng, S. H., Lin, C. Y., Chan, S. C., Lin, Y. C., Yen, T. C., and Liao, C. T. (2014). Clinical utility of multimodality imaging with dynamic contrast-enhanced MRI, diffusion-weighted MRI, and 18F-FDG PET/CT for the prediction of neck control in oropharyngeal or hypopharyngeal squamous cell carcinoma treated with chemoradiation. *PLoS ONE* 9, e115933. doi: 10.1371/journal.pone.0115933
- Omlou, J. M., Lagarde, S. M., Hulscher, J. B., Reitsma, J. B., Fockens, P., and van Dekken, H. (2007). Extended transthoracic resection compared with limited transhiatal resection for adenocarcinoma of the mid/distal esophagus: five-year survival of a randomized clinical trial. *Ann. Surg.* 246, 992–1000. discussion 1000-1. doi: 10.1097/SLA.0b013e31815c4037
- Parekh, V., and Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert. Rev. Precis. Med. Drug. Dev.* 1, 207–226. doi: 10.1080/23808993.2016.1164013
- Pencina, M. J., D'Agostino, R. B. Sr., and Steyerberg, E. W. (2011). Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat. Med.* 30, 11–21. doi: 10.1002/sim.4085
- Pennathur, A., Gibson, M. K., Jobe, B. A., and Luketich, J. D. (2013). Oesophageal carcinoma. *Lancet* 381, 400–412. doi: 10.1016/S0140-6736(12)60643-6
- Sargent, D. J., Wieand, H. S., Haller, D. G., Gray, R., Benedetti, J. K., and Buysse, M. (2005). Disease-free survival versus overall survival as a primary end point for adjuvant colon cancer studies: individual patient data from 20,898 patients on 18 randomized trials. *J. Clin. Oncol.* 23, 8664–8670. doi: 10.1200/JCO.2005.01.6071
- Scalco, E., and Rizzo, G. (2017). Texture analysis of medical images for radiotherapy applications. *Br. J. Radiol.* 90, 20160642. doi: 10.1259/bjr.20160642
- Shafiq-Ul-Hassan, M., Zhang, G. G., Latifi, K., Ullah, G., Hunt, D. C., and Balagurunathan, Y. (2017). Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. *Med. Phys.* 44, 1050–1062. doi: 10.1002/mp.12123
- Shen, C., Liu, Z., Wang, Z., Guo, J., Zhang, H., and Wang, Y. (2018). Building CT radiomics based nomogram for preoperative esophageal cancer patients lymph node metastasis prediction. *Transl. Oncol.* 11, 815–824. doi: 10.1016/j.tranon.2018.04.005
- Tang, L. Q., Chen, Q. Y., Fan, W., Liu, H., Zhang, L., and Guo, L. (2013). Prospective study of tailoring whole-body dual-modality [18F]fluorodeoxyglucose positron emission tomography/computed tomography with plasma Epstein-Barr virus DNA for detecting distant metastasis in endemic nasopharyngeal carcinoma at initial staging. *J. Clin. Oncol.* 31, 2861–2869. doi: 10.1200/JCO.2012.46.0816
- Vasquez, M. M., Hu, C., Roe, D. J., Chen, Z., Halonen, M., and Guerra, S. (2016). Least absolute shrinkage and selection operator type methods for the identification of serum biomarkers of overweight and obesity: simulation and application. *BMC Med. Res. Methodol.* 16, 154. doi: 10.1186/s12874-016-0254-8

- Vickers, A. J., Cronin, A. M., Elkin, E. B., and Gonen, M. (2008). Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Med. Inform. Decis. Mak.* 8, 53. doi: 10.1186/1472-6947-8-53
- Wang, H. Y., Sun, B. Y., Zhu, Z. H., Chang, E. T., To, K. F., and Hwang, J. S. (2011). Eight-signature classifier for prediction of nasopharyngeal [corrected] carcinoma survival. *J. Clin. Oncol.* 29, 4516–4525. doi: 10.1200/JCO.2010.33.7741
- Wu, S., Xia, B., Han, F., Xie, R., Song, T., and Lu, L. (2015). Prognostic Nomogram for Patients with Nasopharyngeal Carcinoma after Intensity-Modulated Radiotherapy. *PLoS ONE* 10, e0134491. doi: 10.1371/journal.pone.0134491
- Yan, T., Cui, H., Zhou, Y., Yang, B., Kong, P., and Zhang, Y. (2019). Multi-region sequencing unveils novel actionable targets and spatial heterogeneity in esophageal squamous cell carcinoma. *Nat. Commun.* 10, 1670. doi: 10.1038/s41467-019-09255-1
- Zeng, L., Guo, P., Li, J. G., Han, F., Li, Q., and Lu, Y. (2015). Prognostic score models for survival of nasopharyngeal carcinoma patients treated with intensity-modulated radiotherapy and chemotherapy. *Oncotarget* 6, 39373–39383. doi: 10.18632/oncotarget.5781
- Zhang, B., Tian, J., Dong, D., Gu, D., Dong, Y., and Zhang, L. (2017). Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma. *Clin. Cancer Res.* 23, 4259–4269. doi: 10.1158/1078-0432.CCR-16-2910
- Zhang, M., Wei, S., Su, L., Lv, W., and Hong, J. (2016). Prognostic significance of pretreated serum lactate dehydrogenase level in nasopharyngeal carcinoma among Chinese population: a meta-analysis. *Medicine* 95, e4494. doi: 10.1097/MD.0000000000004494
- Zhao, B., Tan, Y., Tsai, W. Y., Qi, J., Xie, C., and Lu, L. (2016). Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci. Rep.* 6, 23428. doi: 10.1038/srep23428

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Yan, Liu, Yan, Peng, Wang, Zhang, Wang, Zhuang, Liu, Ma, Wang and Cui. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.