

# Journal Pre-proof

Machine Learning-Based Radiomics for Predicting Outcomes in Cervical Cancer Patients Undergoing Concurrent Chemoradiotherapy

Wang Xin, Su Rixin, Li Linrui, Qin Zhihui, Liu Long, Zhang Yu



PII: S0010-4825(24)00678-4

DOI: <https://doi.org/10.1016/j.compbiomed.2024.108593>

Reference: CBM 108593

To appear in: *Computers in Biology and Medicine*

Received Date: 22 November 2023

Revised Date: 9 May 2024

Accepted Date: 10 May 2024

Please cite this article as: W. Xin, S. Rixin, L. Linrui, Q. Zhihui, L. Long, Z. Yu, Machine Learning-Based Radiomics for Predicting Outcomes in Cervical Cancer Patients Undergoing Concurrent Chemoradiotherapy, *Computers in Biology and Medicine*, <https://doi.org/10.1016/j.compbiomed.2024.108593>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2024 Published by Elsevier Ltd.

# **Machine Learning-Based Radiomics for Predicting Outcomes in Cervical Cancer Patients Undergoing Concurrent Chemoradiotherapy**

Wang Xin<sup>1</sup>, Su Rixin<sup>1</sup>, Li Linrui<sup>1</sup>, Qin Zhihui<sup>1</sup>, Liu Long<sup>2\*</sup>, Zhang Yu<sup>1\*</sup>

<sup>1</sup>Department of Radiology, The First Affiliated Hospital of Anhui Medical University, Hefei 230022, Anhui, China

<sup>2</sup>Department of Hepatobiliary and Pancreatic Surgery, The Second Hospital of Zhejiang University, Hangzhou 310000, Zhejiang, China

\*Corresponding authors: Liu Long, liulong6179@163.com; Zhang Yu, zhangyu199831@163.com

Wang Xin and Su Rixin contributed equally to the work.

# Machine Learning-Based Radiomics for Predicting Outcomes in Cervical Cancer Patients Undergoing Concurrent Chemoradiotherapy

## Abstract

**Purposes** To investigate the value of machine learning-based radiomics for predicting disease-free survival (DFS) and overall survival (OS) undergoing concurrent chemoradiotherapy (CCRT) for patients with locally advanced cervical cancer (LACC).

**Materials and methods** In this multicentre study, 700 patients with IB2-IVA cervical cancer who underwent CCRT with ongoing follow-up were retrospectively analyzed. Three-dimensional radiomics features of primary lesions and its surrounding 5 cm region in T2WI sequences were collected. Six machine learning methods were used to construct the optimal radiomics model for accurate prediction of DFS and OS after CCRT in LACC patients. Eventually, TCGA and GEO databases were used to explore the mechanisms of radiomics in predicting the progression and survival of cervical cancer. This study adhered CLEAR for reporting and its quality was assessed using RQS and METRICS.

**Results** In the prediction of DFS, the RSF model combined tumor and peritumor radiomics demonstrated the best predictive efficacy, with the AUC for predicting 1-year, 3-year, and 5-year DFS in the training, validation, and test sets of 0.986, 0.989, 0.990, and 0.884, 0.838, 0.823, and 0.829, 0.809, 0.841, respectively. In the prediction of OS, the GBM model best performer, with AUC of 0.999, 0.995, 0.978, and 0.981, 0.975, 0.837, and 0.904, 0.860, and 0.905. Differential genes in TCGA and GEO suggest that the prediction of radiomics model may be associated with KDELR2 and HK2.

**Conclusion** Machine learning-based radiomics models help to predict DFS and OS after CCRT in LACC patients, and the combination of tumor and peritumor information has higher predictive efficacy, which can provide a reliable basis for therapeutic decision-making in cervical cancer patients.

**Keywords** Machine learning; Radiomics; Cervical cancer; Concurrent chemoradiotherapy; Outcome

## Abbreviations

ADC	Apparent diffusion coefficient
CCRT	Concurrent chemoradiotherapy
DFS	Disease-free survival
DGE	Differential gene expression
EBRT	External beam radiotherapy
EQD	2 2-Gy equivalent dose
FIGO	International federation of gynecology and obstetrics
GBM	Gradient boosting machine
GEO	Gene expression omnibus
ICBT	Intracavitary brachytherapy
ICC	Intraclass correlation coefficient
LACC	Locally advanced cervical cancer
LASSO	Least absolute shrinkage and selection operator
NCCN	National comprehensive cancer network
OS	Overall survival
ROI	Region of interest
RSF	Random survival forest
SCC-Ag	Squamous cell carcinoma antigen
SVM	support vector machine
TCGA	The cancer genome atlas
VOI	Volume of interest

## Introduction

Cervical cancer is the most common malignant tumor of the female reproductive system, and most patients are detected at locally advanced stages with a poor prognosis[1-3]. Concurrent chemoradiotherapy (CCRT) is the treatment of choice for patients with locally advanced cervical cancer (LACC)[1-4]. The current refinement of comprehensive treatment centers on CCRT has resulted in a 5-year survival rate of 70%-90% for patients with LACC[1,5,6]. However, the prognosis of different cervical cancer patients after undergoing CCRT varies greatly due to the heterogeneity of patient's constitution and gene expression differences in tumor tissues[7,8]. About one-third of patients with LACC will experience recurrence after receiving CCRT, and the 5-year survival rate of patients with recurrent cervical cancer will be less than 50%[8-10]. Meanwhile, some patients may experience disease progression and death in a short period due to mutations in key genes, which leads to high invasiveness and metastatic ability of the tumor[11]. If there are reliable prognostic markers that can predict the sensitivity of cervical cancer patients to CCRT and key prognostic genes at an early stage, clinicians can adjust the treatment plan timely and give patients more accurate and individualized treatment, while researchers can explore new therapeutic targets and ideas, which may even be expected to change the existing diagnostic and therapeutic modes of cervical cancer. However, there is currently no effective means to predict the outcome of LACC patients after CCRT, and relying only on conventional clinical parameters such as The FIGO stage cannot achieve satisfactory prediction accuracy[9]. Therefore, finding accurate predictive indicators for recurrence and survival after CCRT in cervical cancer is an urgent clinical problem.

Radiomics is a technique for non-invasive quantitative analysis and extraction of high-throughput medical imaging[12,13]. By deconstructing and analyzing the three-dimensional texture features of the lesion, it can extract a large number of radiomics parameters, which are potentially relevant to the intrinsic heterogeneity, aggressiveness, and immunoreactivity of the tumor[13-15]. In addition, it has been pointed out that the texture features of the peritumor area also contain important information, which not only reveals the information of subclinical lesions but also reflects the microenvironment of the tumor growth and molecular regulation information[14,15]. Although there is a wealth of biological information in radiomics parameters, how to efficiently find, screen, identify, and utilise these massive features has become one of the major conflicts in the application of radiomics[13]. The application of machine learning effectively solves the problem of screening and combining massive

micro features. By applying multiple machine learning methods to achieve the maximum utilization and combination of radiomics features to quantify the intrinsic biological properties of tumors and the sensitivity of tumor-related treatments, it is expected to achieve accurate prediction of disease progression and patient prognosis[14-16]. Currently, machine learning-based radiomics has been successfully applied to the precise staging, immune assessment, and efficacy prediction of various tumors such as esophageal cancer, breast cancer, and lung cancer[14-18]. Therefore, we sought to explore the value of machine learning models based on tumor and peritumor radiomics in predicting DFS and OS after CCRT in patients with LACC and tried to reveal the correlation between radiomics features and gene expression to provide new ideas and efficient tools for clinical decision-making in cervical cancer.

## Materials and methods

### 1. Patients:

The data of 845 patients with LACC who attended the First Affiliated Hospital of University of Science and Technology of China (Center 1) or the First Affiliated Hospital of Anhui Medical University (Center 2) from January 1, 2013, to January 1, 2021 were retrospectively collected. Patient informed consent was waived by the Ethics Committee, Ethics No: 2023-YXK-12 and PJ2024-03-49. Based on the inclusion and exclusion criteria 700 patients from 845 were included in the study, of which 570 patients from Center 1, were randomly divided into training and validation sets in a 2:1 manner, and 130 patients from Center 2 served as the test set (Figure 1). All enrolled patients were in locally advanced stage (IB2-IVA) of the 2018 International Federation of Gynecology and Obstetrics (FIGO) staging criteria, receiving standard CCRT and conventional MRI 2 weeks before CCRT, and were followed up until October 1, 2023, or death. General clinicopathological data including age, pre-treatment squamous cell carcinoma antigen (SCC-Ag), maximum tumor diameter, lymph node metastasis (LNM), histological types, internal and external irradiation dose, total dose of chemotherapy, total duration of treatment and radiotherapy complications were collected from hospital medical record system. The 2-Gy equivalent dose (EQD2) was calculated based on linear-quadratic formalism using an alpha-beta for tumor tissue of 10 Gy. This study adhered to CLEAR standards[19], as detailed in the *Supplementary Material*.

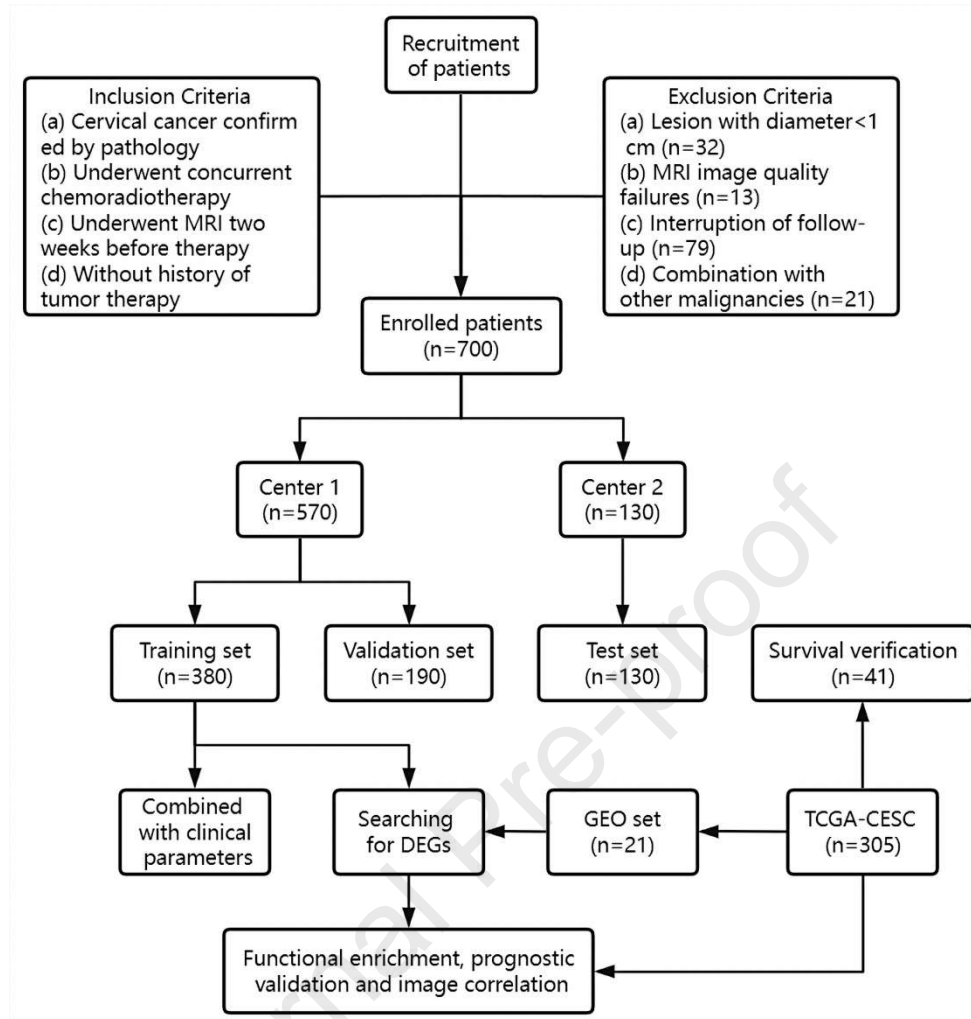


Figure 1 Flowchart of patient recruitment and study overview

## 2. Protocol:

**2.1 Scanning protocol:** MRI scans were performed using a 3.0T system (Signa Excite HD, GE, USA) with an 8-channel Torsor phased array coil. The scanning sequence included: axial T1WI, axial T2WI, axial T2WI pressure lipids, sagittal T2WI, and axial DWI (Table S1).

**2.2 Treatment protocol:** The treatment for LACC refers to the National Comprehensive Cancer Network (NCCN)[2] including external beam radiotherapy (EBRT), intracavitary brachytherapy (ICBT), and concurrent platinum-containing chemotherapy.

**2.3 Follow-up protocols:** All patients are followed up on a long-term basis through outpatient, inpatient reviews, or telephone contact. Every 3-6 months for 2 years, every 6-12 months for the 3rd-5th year, and annually thereafter. Follow-up included clinical manifestations, laboratory and

imaging examinations.

The endpoints in this research include disease-free survival (DFS): the period from the start of treatment to tumor recurrence, secondary malignancy, death, or last follow-up; and overall survival (OS): the period from the start of treatment to death or last follow-up[9,16]. More details on scanning, treatment, and follow-up protocols are provided in *Supplementary Material*.

### 3. Image analysis:

Firstly pretreatment axial T2WI images (DICOM format) were imported from the PACS system into the 3D slicer software and the N4ITK bias field correction module was used to obtain a standard normal distribution of image intensities. The tumor region of interest (ROI) was then delineated by a senior radiologist layer by layer, and the region 5 cm outside the tumor ROI was the peritumor ROI[19]. Each level of ROI was fused into a volume of interest (VOI) to extract three-dimensional texture features. Finally, 1037 radiomics parameters were extracted from each image by PyRadiomics software. The radiomics parameters were standardized and normalized by Z-score, and ComBat was used to reduce the batch effect of the parameters. The flowchart of segmentation delineation and model construction is shown in Figure 2.

One hundred patients were randomly selected after 4 weeks to assess intra- and inter-observer reproducibility of ROI for radiomics. Intra-group reproducibility analysis was performed for the two images delineated by the same physician, and inter-group reproducibility analysis was performed for the first image delineated by two different physicians. An intraclass correlation coefficient (ICC) > 0.75 indicated good reproducibility[20]. We also collected the apparent diffusion coefficient (ADC) value of all patients to assist in model construction. The specific process of ADC value and radiomics parameters extraction and screening is described in *Supplementary Material*.



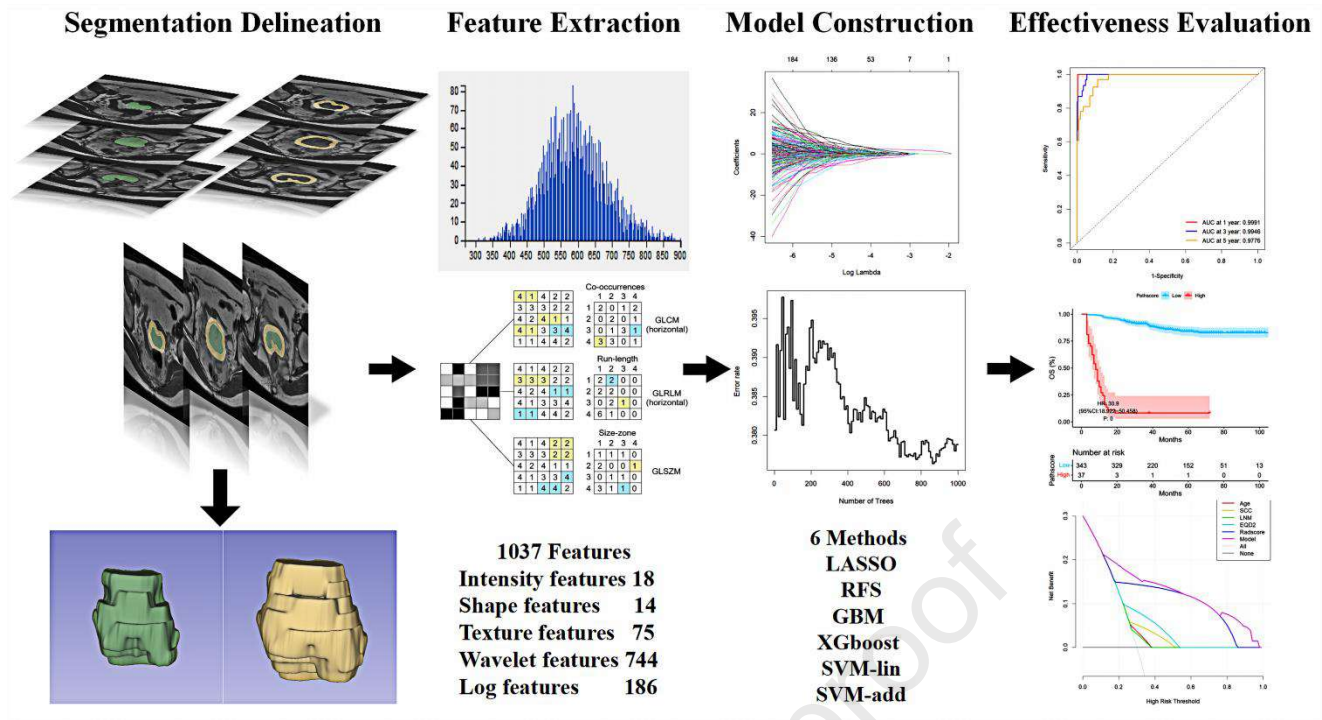


Figure 2 Flowchart for segmentation delineation and model construction

#### 4. Construction of radiomics model:

Spearman correlation analysis was performed on the radiomics parameters with  $ICC > 0.75$  (intra and inter), and redundant parameters with correlation  $> 0.80$  were removed. Afterwards, six machine learning methods, the least absolute shrinkage and selection operator (LASSO), Random Survival Forest (RSF), Gradient Boosting Machine (GBM), XGboost, Linear Kernel Support Vector Machine (SVM-lin) and Polynomial Kernel Support Vector Machine (SVM-add), were used to screen the radiomics features. The optimal combination model (highest sum of 1, 3, and 5 year AUC values in validation and test sets) was named as the radiomics score (Radscore). More details are described in *Supplementary Material*.

#### 5. Union of radiomics and clinical indicators:

To analyze the complementarity between the radiomics and clinical parameters and to further perfect the prediction model, we united multiple clinical indicators with the best Radscore. The Thresholds for each parameter are derived from the `surv_cutpoint` function (Table S2). A comprehensive nomogram model was constructed based on the independent predictors in multivariate

Cox regression analysis to reveal the outcomes of cervical cancer from more dimensions.

## 6. Finding functional genes for radiomics models:

Radiomics parameters were extracted from 53 cervical cancer patients with qualifying MR images in TCIA. Stratification of patients based on their radiomics score (thresholds is the best cut-off value of the model in the training group). DFS and OS-related differential gene expression (DGE) were derived and enriched for analysis.

The GSE56363 dataset contains whole human genome sequencing of 21 patients with LACC, all of whom underwent CCRT. Patients were grouped according to efficacy at 6 months after CCRT to look for genes associated with CCRT susceptibility. The DGE in GSE56363 was taken to intersect with the DFS-related and OS-related DGE in TCGA, respectively, to obtain genes both associated with Radscore (prognosis) and CCRT efficacy (sensitivity).

To further clarify the association between DGE and radiomics features, Spearman/pearson analysis was used to analyze the correlation between the DGE and the top 10 sorted radiomics features in the machine learning model.  $P < 0.05$  and correlation coefficients  $> 0.6$  are considered as a strong correlation[21].

## 7. Statistical analysis:

The software of SPSS27.0 and R4.2.0 were used for statistical analysis. The Z-score function of the scale package was first used to standardize and normalize the radiomics parameters. The ComBat function of the SVA package was used to reduce the batch effect caused by different acquisition times. The ICC function of the psych package was used to test the robustness of radiomics parameters. The glmnet package, randomForest package, gbm package, xgboost package, and survivalsvm package were used to construct LASSO, RSF, GBM, XGboost, and SVM, respectively. The rms package was used to construct the nomogram model. The timeROC, rmda, and ggDCA packages were used to plot the ROC curves, calibration curve, decision curve, and clinical impact curves of models.

## Result

### 1. Baseline characteristics:

A total of 700 patients were included in this study, 570 in Center 1 and 130 in Center 2. The median age was 55 years, range 23-89 years. The median follow-up was 49 months, range 2-131 months. 176 patients (25.1%) experienced recurrence, including 42 local recurrence, 104 distant metastasis, and 30 both recurrence and metastasis. 155 patients (22.1%) died, of which 127 were cervical cancer-specific deaths and 28 non-cancer deaths. The general clinical data were not statistically different between Center 1 and Center 2 (Table 1).

Table 1 Baseline data for 700 patients with cervical cancer

Parameters	Center 1 (n=570)	Center 2 (n=130)	<i>P</i> value
Recurrence	140 (24.6%)	36 (27.7%)	0.479
Death	125 (21.9%)	30 (23.1%)	0.802
Age	56.60±10.16	55.52±10.70	0.520
SCC-Ag (ng/ml)	2.69±8.82	2.73±6.76	0.094
Max tumor diameter (cm)	4.34±1.54	4.32±1.53	0.613
Lymph node metastasis	264 (46.3%)	60 (46.2%)	0.931
Histology			0.775
SCC	485 (85.1%)	112 (86.2%)	
AC	74 (13.0%)	16 (12.3%)	
ASC	7 (1.2%)	1 (0.8%)	
NEC	4 (0.7%)	1 (0.8%)	
FIGO stage			0.943
IB	29 (5.1%)	7 (5.4%)	
IIA	75 (13.2%)	18 (13.8%)	
IIB	161 (28.2%)	34 (26.2%)	
IIIA	10 (1.8%)	2 (1.5%)	
IIIB	21 (3.7%)	5 (3.8%)	
IIIC1	213 (37.4%)	50 (38.5%)	
IIIC2	10 (1.8%)	3 (2.3%)	
IVA	51 (8.9%)	11 (8.5%)	
EBRT (Gy)	47.89±2.55	47.93±2.89	0.887
ICBT (Gy)	29.33±2.56	29.24±2.37	0.462
EQD2 (Gy)	86.10±4.26	85.78±4.39	0.422
Total chemotherapy dose (mg)	195.65±45.75	193.26±45.58	0.361
Duration of CCRT (day)	50.07±2.59	50.08±2.54	0.811
Radiotherapy sensitizer	69 (12.1%)	27 (20.8%)	0.098

SCC-Ag: squamous cell carcinoma antigen; SCC: squamous cell carcinoma; AC: adenocarcinoma;

ASC: adenosquamous carcinoma; NEC: neuroendocrine carcinoma; FIGO: international federation of gynecology and obstetrics; EBRT: external beam radiotherapy; ICBT: intracavitary brachytherapy; EQD2: 2-Gy equivalent dose.

## 2.Screening of radiomics parameters and construction of the best combination model:

A total of 1037 features were extracted from each lesion of the tumor and peritumor, including 917 highly robust parameters with intra- and inter-group ICC>0.75 in tumor lesions. After removing redundant parameters with a correlation greater than 0.9, the remaining 240 radiomics features were used for machine learning modeling. There were 925 high robustness parameters in the peritumor lesion, and 247 features remained after redundancy removal.

In the study of DFS, the RSF-based radiomics model demonstrated the best predictive efficacy (Figure 3). The peritumor radiomics features were complementary to the prediction of cervical cancer recurrence. The 1-, 3-, and 5-year AUC of the combined model in the training, validation, and test sets were 0.986, 0.989, 0.990 and 0.884, 0.838, 0.823 and 0.829, 0.809, 0.841, respectively (Table 2).

The RQS[22] and METRICS[23] were used to assess the research methodology, where the RQS score was 22 out of 36 (Table S3) and the METRICS score was 80.3% (Table S4). More details can be found in the *Supplementary Material*.

Table 2 Comparison of efficacy of different classifiers for predicting DFS

Group		AUC	LASSO	RSF	GBM	XGboost	SVM-lin	SVM-add
Tumor	Training set	1-year	0.801	0.979	0.885	0.763	0.752	0.800
		3-year	0.751	0.974	0.870	0.764	0.680	0.755
		5-year	0.754	0.977	0.869	0.794	0.753	0.853
	Validation set	1-year	0.654	<b>0.809</b>	0.765	0.662	0.697	0.689
		3-year	0.643	<b>0.753</b>	0.685	0.635	0.658	0.649
		5-year	0.645	<b>0.797</b>	0.649	0.702	0.748	0.773
	Test set	1-year	0.637	<b>0.799</b>	0.734	0.603	0.699	0.701
		3-year	0.592	<b>0.765</b>	0.699	0.582	0.632	0.733
		5-year	0.581	<b>0.752</b>	0.631	0.690	0.705	0.716
	Peritumor	1-year	0.732	0.936	0.921	0.780	0.716	0.793
		3-year	0.694	0.968	0.929	0.803	0.633	0.724
		5-year	0.665	0.928	0.914	0.820	0.748	0.853
	Validation set	1-year	0.701	<b>0.743</b>	0.665	0.622	0.594	0.598

Combine	Test set	3-year	0.654	<b>0.703</b>	0.670	0.592	0.568	0.583
		5-year	0.606	<b>0.670</b>	0.516	0.644	0.686	0.676
		1-year	0.613	<b>0.733</b>	0.656	0.601	0.541	0.610
		3-year	0.583	<b>0.711</b>	0.551	0.578	0.589	0.644
		5-year	0.531	<b>0.782</b>	0.513	0.622	0.611	0.703
	Training set	1-year	0.806	0.986	0.938	0.740	0.779	0.868
		3-year	0.779	0.989	0.936	0.748	0.690	0.798
		5-year	0.763	0.990	0.925	0.770	0.794	0.920
	Validation set	1-year	0.752	<b>0.884</b>	0.760	0.574	0.633	0.636
		3-year	0.718	<b>0.838</b>	0.683	0.556	0.618	0.628
		5-year	0.658	<b>0.823</b>	0.582	0.609	0.718	0.743
	Test set	1-year	0.776	<b>0.829</b>	0.784	0.610	0.754	0.738
		3-year	0.679	<b>0.809</b>	0.769	0.590	0.702	0.795
		5-year	0.614	<b>0.841</b>	0.819	0.656	0.738	0.794

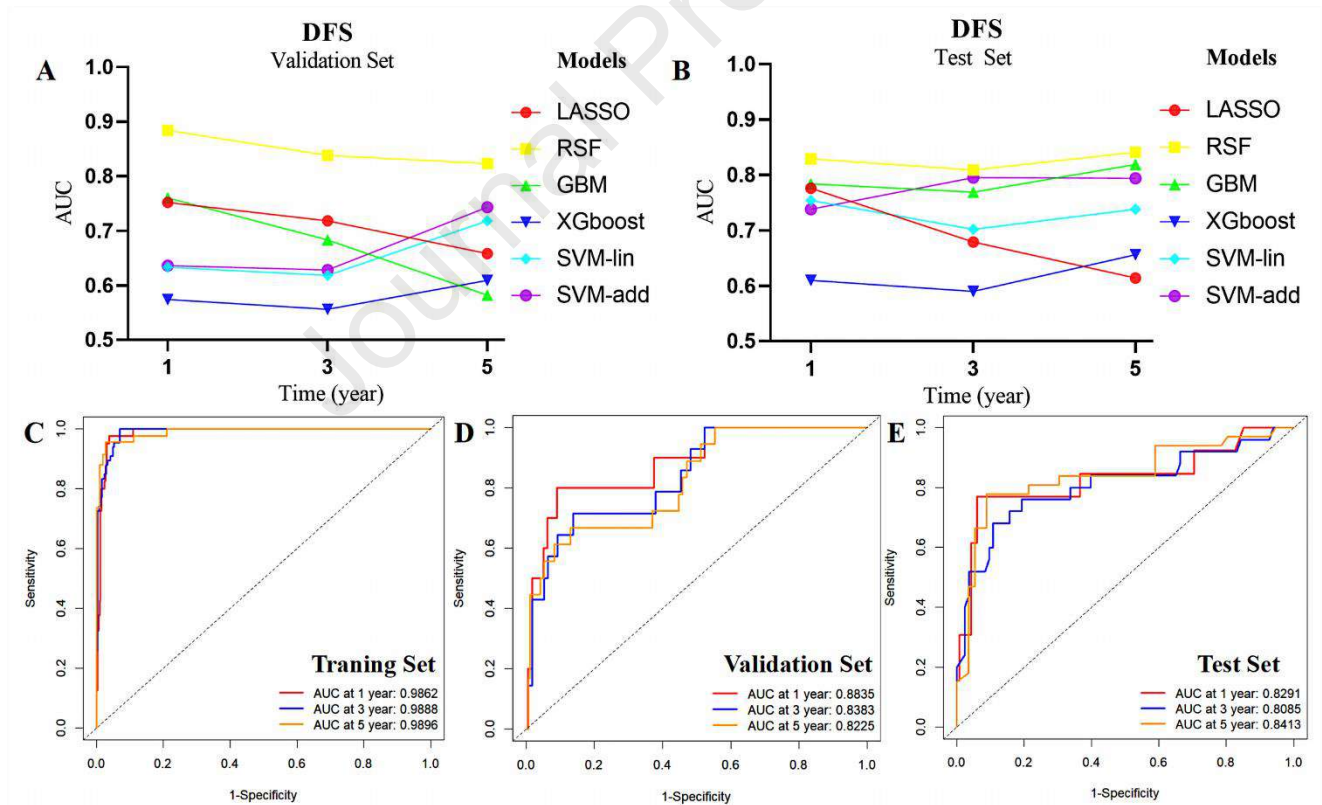


Figure 3 Predictive efficacy of machine learning-based radiomics models in predicting DFS. Figures A and B show the comparison of the AUC values in each model for predicting 1-year, 3-year, and 5-year DFS after CCRT in cervical cancer patients in the validation and test sets, in which the

RSF model has the best overall performance; Figures C, D, and E show the ROC curves of the RSF model for predicting DFS in the training, validation, and test sets.

In the study of OS, the GBM-based radiomics model demonstrated the best predictive efficacy (Figure 4). The peritumor information was complementary for the prediction of cervical cancer survival. The 1-, 3-, and 5-year AUC of the combined model were 0.999, 0.995, 0.978 and 0.981, 0.975, 0.837 and 0.904, 0.860, 0.905 in the training, validation, and test sets, respectively (Table 3). Additional assessment of model efficacy using 41 patients from the TCGA cohort showed AUC of 0.869, 0.839, and 0.951 for 1-, 3-, and 5-year OS. The Logrank test showed a P-value of 0.018 and an HR of 11.608 for stratification according to Radscore (Figure S1).

Table 3 Comparison of efficacy of different classifiers for predicting OS

Group		AUC	LASSO	RSF	GBM	XGboost	SVM-lin	SVM-add
Tumor	Training set	1-year	0.884	0.966	0.842	0.907	0.720	0.735
		3-year	0.909	0.992	0.936	0.869	0.692	0.740
		5-year	0.652	0.990	0.917	0.880	0.751	0.821
	Validation set	1-year	0.662	0.957	<b>0.976</b>	0.538	0.935	0.873
		3-year	0.762	0.753	<b>0.916</b>	0.644	0.775	0.785
		5-year	0.741	0.663	<b>0.722</b>	0.526	0.773	0.805
	Test set	1-year	0.776	0.852	<b>0.855</b>	0.677	0.533	0.611
		3-year	0.802	0.735	<b>0.783</b>	0.695	0.680	0.746
		5-year	0.733	0.663	<b>0.804</b>	0.600	0.663	0.781
Peritumor	Training set	1-year	0.884	0.991	0.997	1.000	0.872	0.899
		3-year	0.879	0.996	0.992	0.985	0.869	0.912
		5-year	0.667	0.996	0.965	0.944	0.814	0.874
	Validation set	1-year	0.925	0.941	<b>0.952</b>	0.955	0.880	0.864
		3-year	0.873	0.984	<b>0.972</b>	0.806	0.870	0.892
		5-year	0.783	0.837	<b>0.868</b>	0.603	0.847	0.868
	Test set	1-year	0.861	0.811	<b>0.862</b>	0.890	0.656	0.623
		3-year	0.800	0.765	<b>0.807</b>	0.715	0.701	0.688
		5-year	0.706	0.868	<b>0.885</b>	0.699	0.690	0.704
Combine	Training set	1-year	0.977	0.999	0.999	0.852	0.960	0.968
		3-year	0.939	0.995	0.995	0.921	0.866	0.921
		5-year	0.733	0.999	0.978	0.930	0.841	0.917
	Validation set	1-year	0.976	0.987	<b>0.981</b>	0.848	0.949	0.960
		3-year	0.965	0.973	<b>0.975</b>	0.668	0.917	0.938



Test set	5-year	0.784	0.826	<b>0.837</b>	0.539	0.878	0.800
	1-year	0.885	0.897	<b>0.904</b>	0.908	0.650	0.674
	3-year	0.857	0.802	<b>0.860</b>	0.738	0.742	0.790
	5-year	0.744	0.906	<b>0.905</b>	0.724	0.752	0.826

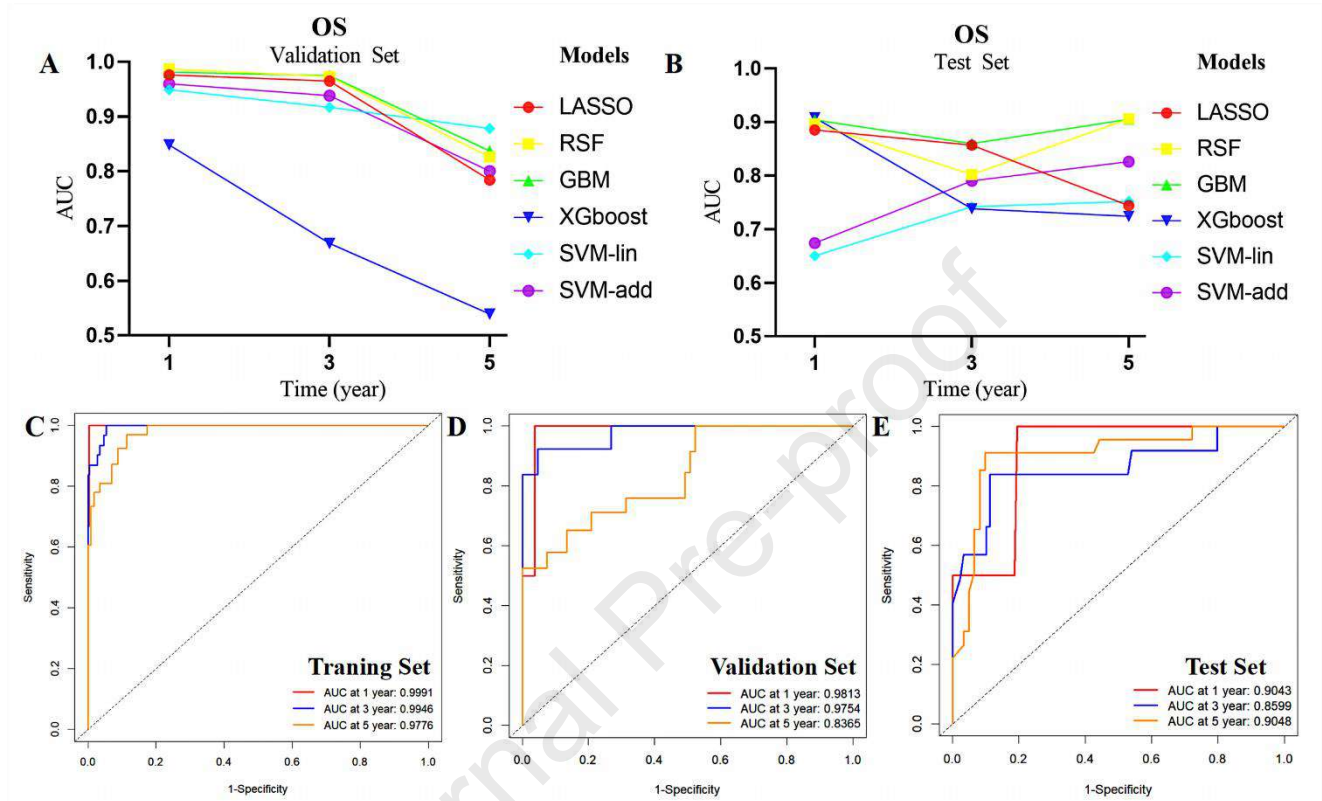


Figure 4 Predictive efficacy of machine learning-based radiomics models in predicting OS. Figures A and B show the comparison of the AUC values in each model for predicting 1-year, 3-year, and 5-year OS after CCRT in cervical cancer patients in the validation and test sets, in which the GBM model has the best overall performance; Figures C, D, and E show the ROC curves of the GBM model for predicting OS in the training, validation, and test sets.

### 3. The complementary role of clinical indicators to radiomics models:

To further analyze the complementarity of radiomics models with clinical parameters and to perfect the prediction models, we combined multiple clinical indicators with the best radiomics score. The results showed that in the prediction of DFS, SCC-Ag, FIGO stage, EBRT, EQD2, ADC value, and Radscore-DFS were significantly correlated with DFS in cervical cancer patients, in which SCC, ADC, EQD2, and Rad-score were independent prognostic factors for DFS after CCRT in LACC patients. The nomogram was constructed based on the independent predictors, and its AUC for

predicting 1-year, 3-year, and 5-year DFS was 0.976, 0.931, 0.947, and 0.914, 0.898, 0.903, and 0.883, 0.832, 0.798 in the training, validation, and test sets, respectively (Figure 5). The calibration curves, decision curves, and clinical impact curves indicate high model fit, stability, and clinical benefit. In the prediction of OS, age, SCC-Ag, LNM, FIGO stage, EBRT, EQD2, and Radscore-OS were significantly associated with OS in cervical cancer patients, where age, SCC, LNM, EQD2 and Rad-score were independent prognostic factors for OS. The AUC of the model for predicting OS at 1, 3, and 5 years were 0.991, 0.994, 0.973, and 0.989, 0.974, 0.916, and 0.930, 0.921, 0.855 in the training, validation, and test sets, respectively, with high stability and clinical benefit (Figure 6).

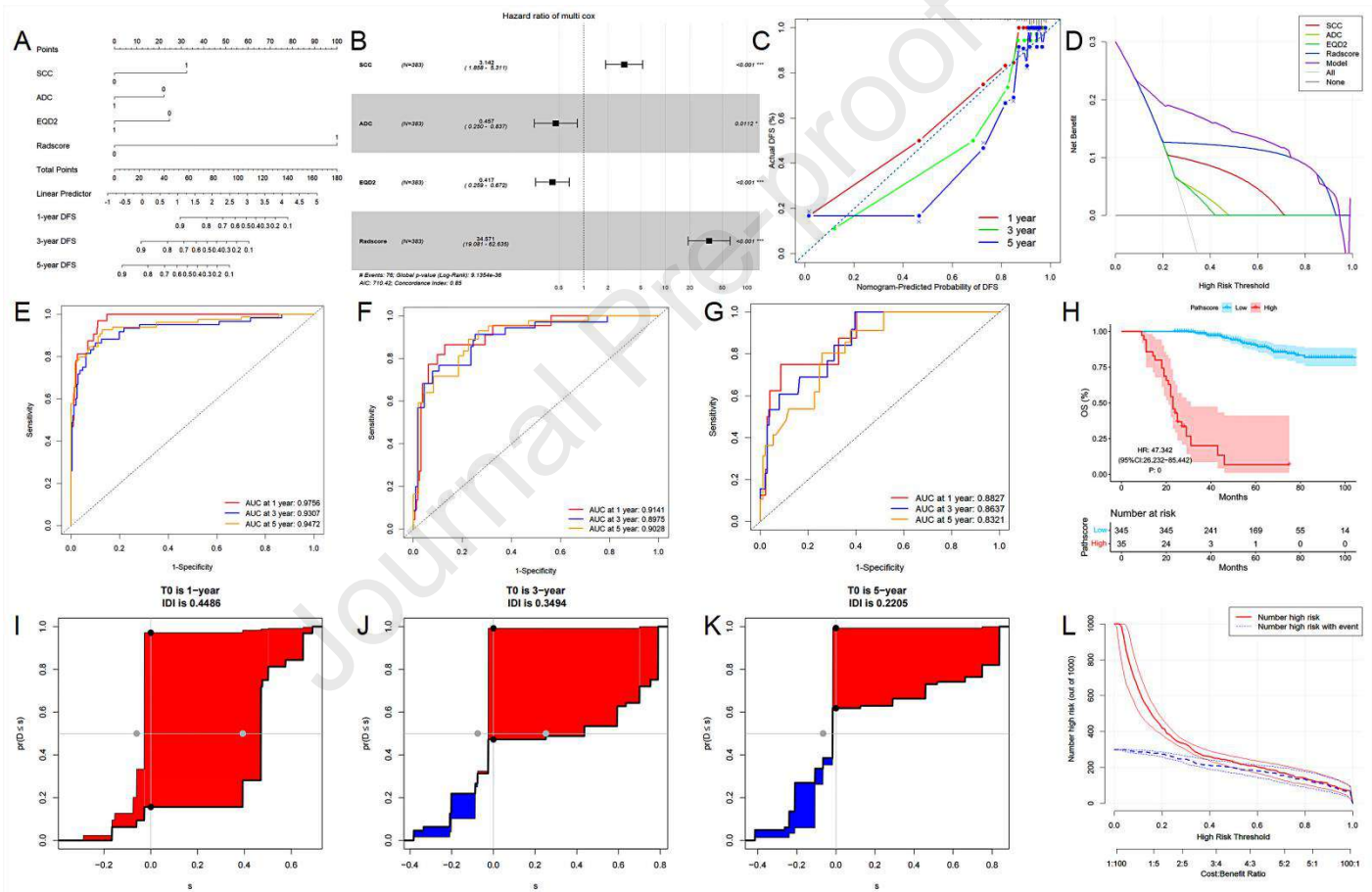


Figure 5 Predictive efficacy of the nomogram for predicting DFS with the combination of radiomics score and clinical parameters. Figure A shows the visualization of the nomogram; Figure B shows the forest plot of the nomogram; Figures C and D show the calibration curves and decision curves of the model, which demonstrates that the net clinical benefit of the joint model outperforms the single parameter at most thresholds (except 0.73-0.88); Figures E, F, and G show the ROC curves of the model in the training, validation, and test sets; Figures H shows the LogRank test of the model;



Figures I, J, and K show the IDI curves of the model; Figure L shows the clinical impact curves of the model.

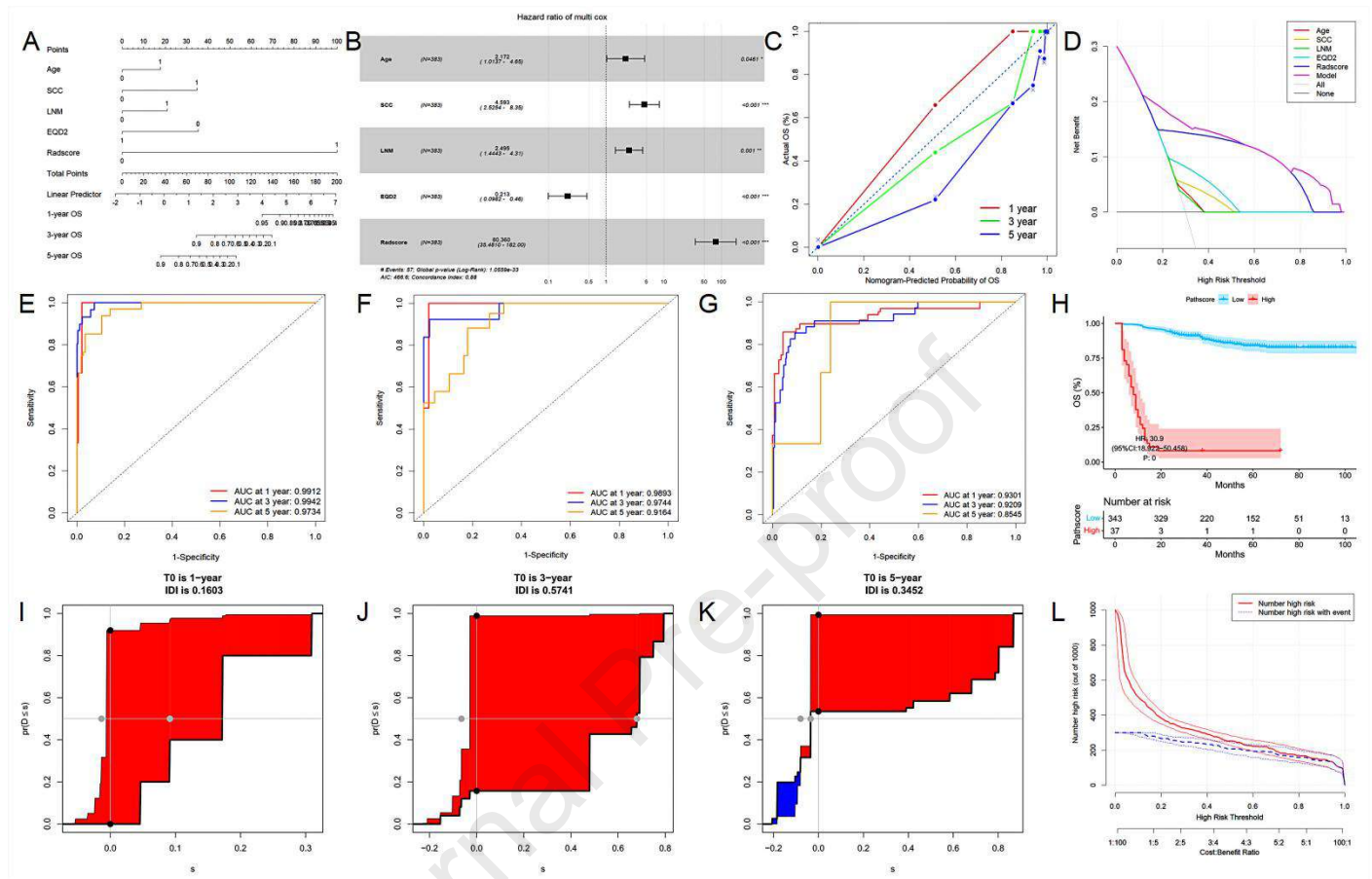


Figure 6 Predictive efficacy of the nomogram for predicting OS with the combination of radiomics score and clinical parameters. Figure A shows the visualization of the nomogram; Figure B shows the forest plot of the nomogram; Figures C and D show the calibration curves and decision curves of the model, which demonstrates that the net clinical benefit of the joint model outperforms the single parameter at all number thresholds; Figures E, F, and G show the ROC curves of the model in the training, validation, and test sets; Figure H shows the LogRank test of the model; Figures I, J, and K show the IDI curves of the model; Figure L shows the clinical impact curves of the model.

#### 4. Radiomics related differential genes:

Radiomics parameters of 53 patients with qualifying MR images of cervical cancer in TCIA were used to calculate Radscore. Patients were grouped according to the best cut-off value (DFS of 14.38 and OS of 2.79), while DFS and OS-related differential genes were derived. Using  $P < 0.05$  and LogFC

absolute value  $>1$  as the cut-off basis, there were 438 DGE for DFS (up-regulated 377, down-regulated 61), and the enrichment analysis showed that the functions were clustered in lesion adhesion, actin filament binding, regulation of epithelial-to-mesenchymal transition, and tumor necrosis factor-alpha signaling pathway (Figure 7). There were 232 DGE for OS (up-regulated 130, down-regulated 102), and the functions were clustered in response to hypoxia, cellular carbohydrate metabolic processes, regulation of lipid metabolic processes, and central carbon metabolism in cancer.

The DGE in GSE56363 (644 up-regulated and 342 down-regulated) were taken to be intersected with the DFS- and OS-related DGE in TCGA, respectively, yielding 43 DFS intersected genes and 26 OS intersected genes.

In the analysis of the correlation of DGE with the top 10 ranked radiomics features (Figure S2), KDELR2 showed a significant correlation with the tumor region wavelet-HHH-glcmm-DifferenceAverage ( $P<0.001$ ,  $r=0.618$ ). HK2 was significantly correlated with the tumor region wavelet-HHH-glszm-SmallAreaLowGrayLevelEmphasis ( $P<0.001$ ,  $r=-0.649$ ) and the peritumor region log-sigma-glszm-HighGrayLevelZoneEmphasis ( $P<0.001$ ,  $r=0.601$ ). In addition, co-expression network analysis and immune correlation analysis of KDELR2 and HK2 were executed to assist in clarifying their functions (Figure S3). The expression validation and prognostic analysis in the TCGA database also demonstrated that KDELR2 and HK2 were highly expressed in cervical cancer and significantly correlated with high-stage and poor prognosis (Figure S4).

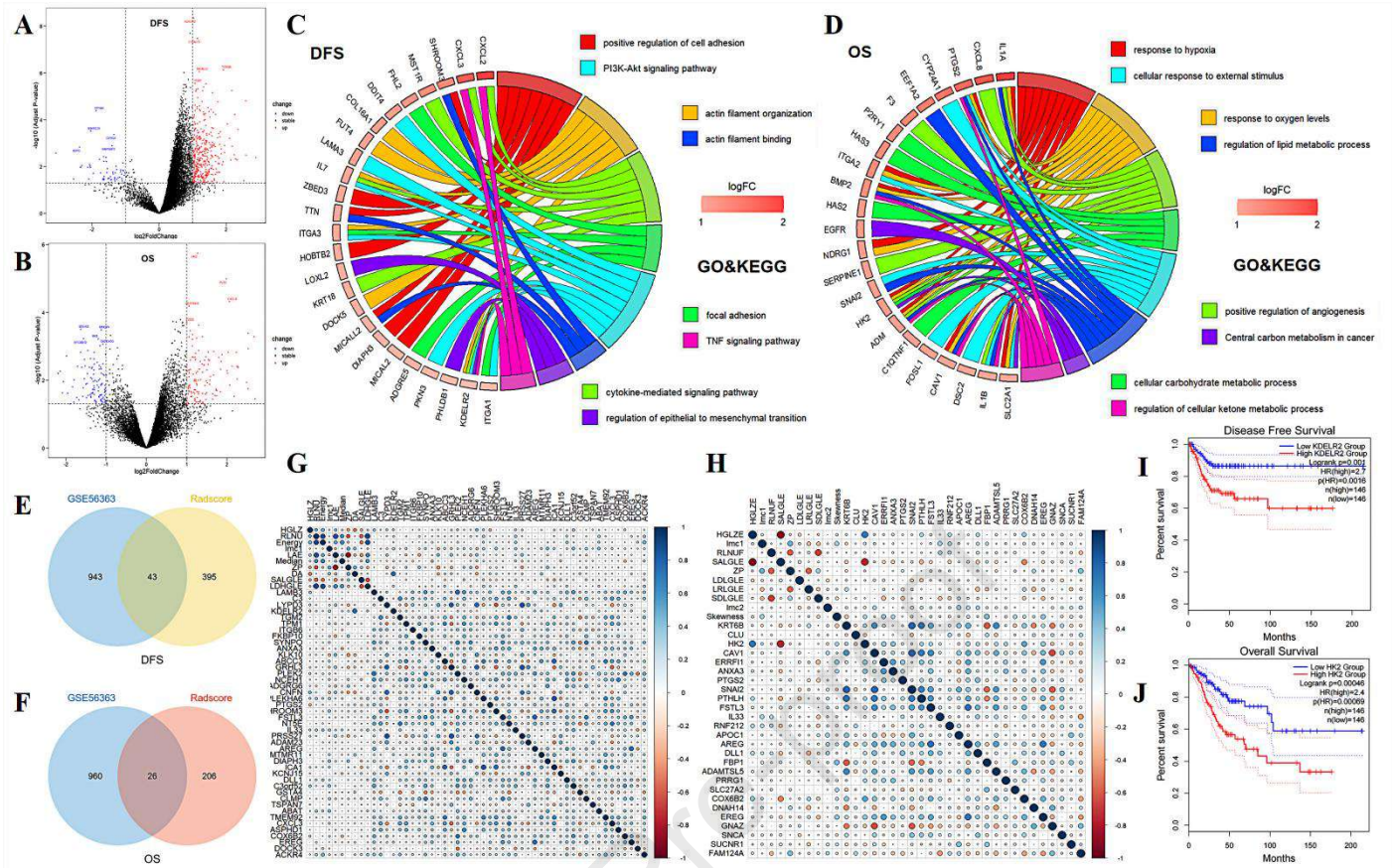


Figure 7 Radiomics-related differential genes and their functional predictions. Figures A and B show the differential gene volcano maps stratified according to the Radscore of DFS and OS; Figures C and D show the GO and KEGG enrichment analysis of DFS and OS differential genes; Figures E and F show the Venn diagrams of the intersection of DFS and OS differential genes in TCGA with the CCRT sensitivity differential genes in GSE56363; Figures G and H show heatmaps of the correlation of DFS and OS differential genes with the top 10 texture features in radiomics models; Figures I, J show the recurrence and survival validation of KDELR2 and HK2 in TCGA.

## Discussion

In this study, 700 patients with LACC from two Centers were followed up for a long period and closely analyzed. In our results, MRI-based radiomics demonstrated the ability to predict the evolution and outcome after CCRT in LACC patients. Meanwhile, the radiomics features of peritumor are complementary and enhanced with the tumor radiomics features. The screening and combination of prognostically relevant radiomics parameters for cervical cancer were further optimized by employing multiple machine learning methods, which improved the accuracy of predicting outcomes. In addition,

radiomics modeling demonstrates incremental value to existing clinical markers, and the combination of the two can effectively strengthen the stability of the model, allowing the model to maintain a high degree of accuracy while maintaining a higher degree of stability and generalizability. Eventually, by analyzing the differential genes in the TCGA and GEO databases, the possible molecular mechanisms by which the radiomics model predicts the outcomes after CCRT in patients with cervical cancer, as well as the associations between specific radiomics features and genes, were revealed to a certain extent. The optimal radiomics model achieves optimization of existing diagnostic modalities and accurate prediction of DFS and OS after CCRT for cervical cancer with LACC. The application of combined model can help early and accurate clinical prediction of the prognostic level as well as the risk and nodes of tumor progression after receiving first-line CCRT in patients with LACC, thus determining whether to use CCRT as the preferred treatment option and whether to intensify the treatment and review, to improve clinical outcomes.

Radiomics is a technique for quantitatively describing tissue texture differences at the microscopic level and from a structural perspective, and the automation and intelligence of its clinical applications have developed rapidly[13]. By objectively analyzing and deconstructing the pixel differences between lesions, it can reveal the micropathological and biological information behind the macroscopic imaging[13,16]. At the same time, the ability of radiomics to reveal the biological information of tumors and their microenvironments, and to further explore their potential correlation with clinical events, is expected to assist in disease prevention and clinical treatment[24]. In this study, we extracted "three-dimensional" texture features of the primary lesion and non-invasively achieved a "virtual biopsy" of the whole tumor[25]. The radiomics GBM model based on primary lesions showed a very high prediction level in predicting 1-year DFS and OS after CCRT, with AUC of 0.979, 0.842 (training set), and 0.809, 0.976 (validation set) and 0.799, 0.855 (test set), respectively. This suggests that the radiomics features of tumors have the potential to reveal the recurrence and survival after CCRT in patients with LACC.

In addition, subclinical lesions around the tumor are thought to contain a great deal of biological information, including local inflammation levels, tumor micro infiltration status, tumor regulatory factors, and other important information[15,26]. It has been suggested that the 5 cm region around cervical cancer lesions contains almost all subclinical lesions, and its radiomics features have proved to be of value in predicting lymph node metastasis[27-29]. Therefore, in this study, we also collected and analyzed the radiomics of the 5 cm region around the primary lesion of cervical cancer. The



results showed that although the overall predictive efficacy of the peritumor radiomics features was slightly worse than that of tumors, the peritumor radiomics also had its advantages and values in some models. Meanwhile, in the combined model, the peritumor radiomics showed its independence and complementarity with the tumor radiomics. The combined model showed more efficient and stable predictive efficacy in the prediction of DFS and OS after CCRT in LACC patients, with 1-, 3- and 5-year AUC greater than 0.800. It suggests that there exists a large amount of critical information and strong clinical prognostic correlation in the area around the primary lesion, and the combination of tumor and the peritumor radiomics features will further improve the model's connotation and efficacy in cervical cancer[28].

The reproducibility of radiomics parameters is one of the main contradictions in its clinical translation[30]. In this study, the N4ITK module in 3D Slicer software was used to correct the images at each Center to obtain a standard normal distribution of the image intensities and to minimize the heterogeneity that exists in the images before feature extraction. Also, after feature extraction ComBat was used to reduce the batch effect of the parameters. All extracted radiomics parameters were standardized and normalized by Z-score to enable comparability between radiomics features. Moreover, only stable parameters with intra- and inter-group correlation coefficients  $> 0.75$  in this study will be included in the subsequent statistical analysis to ensure the reproducibility and reliability of the model. Optimization of radiomics models is one of the main contradictions to improve their clinical value[31]. Six machine learning methods were used to select the best combination model. Also, the high RQS and METRICS scores suggest that this study has sufficient scientific and rigorous methodological design. Compared with previous radiomics studies on cervical cancer, the RQS score of this study was much higher than the median level of 13.5 (5.5-15.75)[32], and the METRICS report was rated as excellent, which indicates that the results of this study have a high degree of reliability. In the prediction of DFS, the RSF model showed the best predictive efficacy, with an AUC of 0.809-0.884. The current state-of-the-art (OSTA) model[7], is LASSO-based, with an AUC of 0.860-0.916. Its lack of multiple model comparisons and independent validation sets may be with model overfitting. The GBM model performed best in the prediction of OS. The OSTA model[33] is a deep learning-based radiomics model with an AUC of 0.713-0.871. It suffers from fewer sample samples and mixed radiomics features. The performance and ease of use are worse than our model.

To further analyze the complementarity between radiomics and clinical parameters and to refine the model, we combined multiple clinical markers with the best radiomics score. The results showed

that SCC-Ag, ADC, and EQD2 had an enhancing effect on the radiomics model in the prediction of DFS. The combined model showed better efficacy, and more importantly, the model became more stable. In the prediction of OS, Age, SCC-Ag, LNM, and EQD2 complemented Rad-score. The overall improvement in model predictive efficacy was modest but more stable. It is worth mentioning among the clinical parameters that EQD2 was an independent prognostic factor for DFS and OS[7]. This suggests that radiotherapy is an important factor influencing the evolution and outcome after CCRT in patients with LACC. Unlike objective clinical and imaging parameters, radiotherapy dose is a parameter that can be subject to human control. This means that it can be used not only for prediction but also to improve clinical outcomes by controlling the dose. When a patient receives a radiotherapy dose higher than the cut-off value (DFS of 83.75 Gy and OS of 83 Gy), one will tend to have a better prognosis.

Data-driven radiomics studies have the nature of post generation of scientific hypotheses, and therefore the biological significance of the results cannot be directly known[13]. However, clarifying the radiomics features and the underlying mechanisms behind the models is important for understanding disease development, exploring novel diagnostic and therapeutic modalities, and promoting the clinical application of the models. In this study, we collected radiomics parameters and gene sequencing results from 53 cervical cancer patients in the TCGA cohort. Patients were grouped by optimal radiomics scores and differential genes for DFS and OS were derived. The differential genes of DFS were enriched to lesion adhesion, actin filament binding, regulation of epithelial-to-mesenchymal transition, and TNF signaling pathway. This indicated that the DFS radiomics model may enable the prediction of recurrence by revealing adhesion, migration, and EMT in cervical cancer. The differential genes of OS were enriched to responses to hypoxia, cellular carbohydrate metabolic processes, regulation of lipid metabolic processes, and central carbon metabolism in cancer. This suggests that OS radiomics models may indirectly reveal hypoxia and metabolism in cervical cancer to achieve prediction of patient outcomes. The intersection of the differential genes with DGE in GSE56363 was then taken to further screen for high-quality genes associated with radiotherapy efficacy and prognosis. The associations between the differential genes and the 10 most important radiomics features in the model were analyzed, where KDELR2 was positively correlated with the DifferenceAverage of the tumor in the DFS model; HK2 was negatively correlated with the SmallAreaLowGrayLevelEmphasis of the tumor in the OS model, and positively

correlated with the HighGrayLevelZoneEmphasis of the peritumor. Meanwhile, the predictive value of KDELR2 and HK2 for DFS and OS in cervical cancer was further validated in the TCGA database.

Meng[34] found that KDELR2 increases Golgi-mediated matrix metalloproteinase secretion to drive tumor progression by regulating KIF20A. Wei[35] found that the KDELR-Gao-Rab1/3 signaling axis controls vesicular transport and material delivery to drive breast cancer invasion and metastasis. Rakhee[36] discovered that KDELR2 does not alter cell proliferation or tumor growth, but can strongly promote EMT, invasion, and metastasis. The association of Difference Average (Difference Average measures the relationship between occurrences of pairs with similar intensity values and occurrences of pairs with differing intensity values) with KDELR2 may reveal the association of grey-scale differences in tumor with tumor invasion and EMT, thus enabling the prediction of recurrence and metastasis after CCRT in cervical cancer patients. HK2 is a key glycolytic enzyme that catalyzes the conversion of glucose to glucose-6-phosphate. Jiao[37] found that HK2 regulates glycolysis through the ubiquitination system mediated by TRAF6 and SQSTM1. In addition, HK2 not only improves the energy supply of malignant cells by making them more dependent on glycolytic metabolic patterns and better adapted to survive in hypoxic environments but also immortalizes malignant cells and protects them from apoptosis through direct interactions with mitochondria[38]. Yang[39] found that HK2 promotes both spheroid-forming ability and cisplatin resistance of cervical cancer cells, which correlates with poor prognosis. Therefore, HK2, which is negatively correlated with SALGLE (SALGLE measures the proportion in the image of the joint distribution of smaller size zones with lower gray-level values) and positively correlated with HGLZE (HGLZE measures the distribution of the higher gray-level values, with a higher value indicating a greater proportion of higher gray-level values and size zones in the image), may indirectly reveal the metabolic status and hypoxia level of the tumor by reflecting the distribution and ratio of high and low grey levels in the tumor, thus predicting the resistance to chemoradiotherapy and the prognosis of cervical cancer patients. The proposal of KDELR2 and HK2 is the first step in exploring the mechanism of radiomics to predict the outcome after CCRT in cervical cancer patients, and more experiments are needed to validate their working mechanism as well as diagnostic value in cervical cancer in the future.

This paper also has many limitations. Firstly, radiomics parameters, gene expression, and patient prognosis vary between pathology types[2], and different pathology types of cervical cancer were not analyzed independently in this study, which may have an impact on the analysis of the results. Second, the TCGA cohort used in this study was small in number and lacked recurrence-related follow-up,

which made the external validation of the model as well as the analysis of genetic correlation less convincing. Eventually, the genetic mechanism part of this study is still in the conjecture stage, and there is no clear evidence to prove the inevitability of the association between radiomics features and genes, which needs to be verified by a large number of basic experiments in the future.

In conclusion, the radiomics features of tumor and peritumor in cervical cancer can reveal the biological information and the tumor microenvironment to a certain extent, which has a great potential for application in the prediction of gene expression level and clinical prognosis. The joint model based on machine learning was outstanding in the prediction of DFS and OS after CCRT in patients with LACC, which can provide accurate and noninvasive evaluation indexes for individualized treatment and prognosis prediction in cervical cancer patients.



**Ethics approval and consent to participate**

Institutional Review Board approval from Ethics Committee of Anhui Provincial Cancer Hospital (Reference No: 2023-YXK-12) and The First Affiliated Hospital of Anhui Medical University (Reference No: PJ2024-03-49) was obtained.

**Consent for publication**

All authors have read and approved the content and agree to submit for consideration for publication in the journal.

**Availability of data and material**

All data generated or analysed during this study are included in this published article.

**Competing interests**

The authors declare that they have no competing interests.

**Funding**

The article is supported by the Mechanism of Orai3/Claudin-5 involvement in the formation of radiological brain oedema and its clinical application clinical application (2022AH051175)

**Authors' contributions**

All authors have had access to the data and all drafts of the manuscript.

Specific contributions are as follows: study design: LL, ZZ; data collection: WX, SRX, QZH, LLR; data management and analysis: WX, SRX, QZH, LLR; manuscript drafting: WX, SRX; manuscript review: all. All authors read and approved the final manuscript.

**Acknowledgements**

All authors sincerely thank two senior radiologists, Master Fei Gao and Dr. Chao Wei, for delineating lesions and image analysis for this study.

## Reference

- [1] Cohen PA, Jhingran A, Oaknin A, et al. Cervical cancer. *Lancet*. 2019, 393: 169-182. DOI: 10.1016/S0140-6736(18)32470-X.
- [2] Koh WJ, Abu-Rustum NR, Bean S, et al. Cervical Cancer, Version 3.2019, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw*. 2019, 17: 64-84. DOI:10.6004/jnccn.2019.0001.
- [3] Wenzel HHB, Olthof EP, Bekkers RLM, et al. Primary or adjuvant chemoradiotherapy for cervical cancer with intraoperative lymph node metastasis - A review. *Cancer Treat Rev*. 2022, 102: 102311. DOI:10.1016/j.ctrv.2021.102311.
- [4] Zheng X, Li C, Zhang L, et al. Combining Intravoxel Incoherent Motion Diffusion Weighted Imaging and Texture analysis for a Nomogram to Predict Early Treatment Response to Concurrent Chemoradiotherapy in Cervical Cancer Patients. *J Oncol*. 2021, 2021: 9345353. DOI:10.1155/2021/9345353.
- [5] Weng D, Xiong H, Zhu C, et al. Adjuvant chemotherapy versus adjuvant concurrent chemoradiotherapy after radical surgery for early-stage cervical cancer: a randomized, non-inferiority, multicenter trial. *Front Med*. 2023, 17(1): 93-104. DOI:10.1007/s11684-021-0892-z.
- [6] Yang J, Mead-Harvey C, Polen-De C, et al. Survival outcomes in patients with cervical cancer treated with open versus robotic radical hysterectomy: Our surgical pathology interrogation. *Gynecol Oncol*. 2020, 159(2): 373-380. DOI:10.1016/j.ygyno.2020.08.031.
- [7] Zhang Y, Liu L, Zhang K, et al. Nomograms Combining Clinical and Imaging Parameters to Predict Recurrence and Disease-free Survival After Concurrent Chemoradiotherapy in Patients With Locally Advanced Cervical Cancer. *Acad Radiol*. 2023, 30(3): 499-508. DOI:10.1016/j.acra.2022.08.002.
- [8] Gao S, Du S, Lu Z, et al. Multiparametric PET/MR (PET and MR-IVIM) for the evaluation of early treatment response and prediction of tumor recurrence in patients with locally advanced cervical cancer. *Eur Radiol*. 2020, 30(2): 1191-1201. DOI:10.1007/s00330-019-06428-w.
- [9] Tseng JY, Yen MS, Twu NF, et al. Prognostic nomogram for overall survival in stage IIB-IVA cervical cancer patients treated with concurrent chemoradiotherapy[J]. *Am J Obstet Gynecol*, 2010, 202(2):174.e1-7. DOI:10.1016/j.ajog.2009.09.028.
- [10] Paskeh MDA, Mirzaei S, Gholami MH, et al. Cervical cancer progression is regulated by SOX transcription factors: Revealing signaling networks and therapeutic strategies. *Biomed Pharmacother*. 2021, 144: 112335. DOI:10.1016/j.biopha.2021.112335.

- [11] Hu C, Liu T, Zhang W, et al. miR-145 inhibits aerobic glycolysis and cell proliferation of cervical cancer by acting on MYC. *FASEB J.* 2023, 37(4): e22839. DOI:10.1096/fj.202201189RR.
- [12] Shih IL, Yen RF, Chen CA, et al. PET/MRI in Cervical Cancer: Associations Between Imaging Biomarkers and Tumor Stage, Disease Progression, and Overall Survival. *J Magn Reson Imaging.* 2021, 53(1): 305-318. DOI:10.1002/jmri.27311.
- [13] Hatt M, Le Rest CC, Tixier F, et al. Radiomics: Data Are Also Images. *J Nucl Med.* 2019, 60(Suppl 2): 38S-44S. DOI:10.2967/jnumed.118.220582.
- [14] Hu Y, Xie C, Yang H, et al. Assessment of Intratumoral and Peritumoral Computed Tomography Radiomics for Predicting Pathological Complete Response to Neoadjuvant Chemoradiation in Patients With Esophageal Squamous Cell Carcinoma. *JAMA Netw Open.* 2020, 3(9): e2015927. DOI:10.1001/jamanetworkopen.2020.15927.
- [15] Xia TY, Zhou ZH, Meng XP, et al. Predicting Microvascular Invasion in Hepatocellular Carcinoma Using CT-based Radiomics Model. *Radiology.* 2023, 307(4): e222729. DOI:10.1148/radiol.222729.
- [16] Xu X, Zhang HL, Liu QP, et al. Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma. *J Hepatol.* 2019, 70(6): 1133-1144. DOI:10.1016/j.jhep.2019.02.023.
- [17] Yu Y, He Z, Ouyang J, et al. Magnetic resonance imaging radiomics predicts preoperative axillary lymph node metastasis to support surgical decisions and is associated with tumor microenvironment in invasive breast cancer: A machine learning, multicenter study. *EBioMedicine.* 2021, 69: 103460. DOI:10.1016/j.ebiom.2021.103460.
- [18] Tong H, Sun J, Fang J, et al. A Machine Learning Model Based on PET/CT Radiomics and Clinical Characteristics Predicts Tumor Immune Profiles in Non-Small Cell Lung Cancer: A Retrospective Multicohort Study. *Front Immunol.* 2022, 13: 859323. DOI:10.3389/fimmu.2022.859323.
- [19] Kocak B, Baessler B, Bakas S, et al. CheckList for Evaluation of Radiomics research (CLEAR): a step-by-step reporting guideline for authors and reviewers endorsed by ESR and EuSoMII. *Insights Imaging.* 2023, 14(1):75. DOI: 10.1186/s13244-023-01415-8.
- [20] Yang J, Shi W, Yang Z, et al. Establishing a predictive model for tumor mutation burden status based on CT radiomics and clinical features of non-small cell lung cancer patients. *Transl Lung Cancer Res.* 2023, 12(4):808-823. DOI: 10.21037/tlcr-23-171.
- [21] Mossel E, Delli K, van Nimwegen JF, et al. Ultrasonography of major salivary glands compared with parotid and labial gland biopsy and classification criteria in patients with clinically suspected primary Sjögren's syndrome. *Ann Rheum Dis.* 2017, 76(11): 1883-1889. DOI:10.1136/annrheumdis-2017-211250.

- [22] Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017, 14(12):749-762. DOI: 10.1038/nrclinonc.2017.141.
- [23] Kocak B, Akinci D'Antonoli T, Mercaldo N, et al. METHodological RadiomICs Score (METRICS): a quality scoring tool for radiomics research endorsed by EuSoMII. *Insights Imaging*. 2024, 15(1):8. DOI: 10.1186/s13244-023-01572-w.
- [24] Bera K, Braman N, Gupta A, et al. Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nat Rev Clin Oncol*. 2022, 19(2): 132-146. DOI:10.1038/s41571-021-00560-7.
- [25] Costa G, Cavinato L, Masci C, et al. Virtual Biopsy for Diagnosis of Chemotherapy-Associated Liver Injuries and Steatohepatitis: A Combined Radiomic and Clinical Model in Patients with Colorectal Liver Metastases. *Cancers (Basel)*. 2021, 13(12): 3077. DOI:10.3390/cancers13123077.
- [26] Jiang Y, Wang H, Wu J, et al. Noninvasive imaging evaluation of tumor immune microenvironment to predict outcomes in gastric cancer. *Ann Oncol*. 2020, 31(6): 760-768. DOI:10.1016/j.annonc.2020.03.295.
- [27] Shi J, Dong Y, Jiang W, et al. MRI-based peritumoral radiomics analyzis for preoperative prediction of lymph node metastasis in early-stage cervical cancer: A multi-center study. *Magn Reson Imaging*. 2022, 88: 1-8. DOI:10.1016/j.mri.2021.12.008.
- [28] Zhang Z, Wan X, Lei X, et al. Intra- and peri-tumoral MRI radiomics features for preoperative lymph node metastasis prediction in early-stage cervical cancer. *Insights Imaging*. 2023, 14(1): 65. DOI:10.1186/s13244-023-01405-w.
- [29] Yu F, Huang X, Zhou D, et al. Genetic, DNA methylation, and immune profile discrepancies between early-stage single primary lung cancer and synchronous multiple primary lung cancer. *Clin Epigenetics*. 2023, 15(1): 4. DOI:10.1186/s13148-023-01422-y.
- [30] Park JE, Park SY, Kim HJ, et al. Reproducibility and Generalizability in Radiomics Modeling: Possible Strategies in Radiologic and Statistical Perspectives. *Korean J Radiol*. 2019, 20(7): 1124-1137. DOI:10.3348/kjr.2018.0070.
- [31] Pszczolkowski S, Manzano-Patrón JP, Law ZK, et al. Quantitative CT radiomics-based models for prediction of haematoma expansion and poor functional outcome in primary intracerebral haemorrhage. *Eur Radiol*. 2021, 31(10): 7945-7959. DOI:10.1007/s00330-021-07826-9.
- [32] Ren J, Li Y, Liu XY, et al. Diagnostic performance of ADC values and MRI-based radiomics analysis for detecting lymph node metastasis in patients with cervical cancer: A systematic review and meta-analysis. *Eur J Radiol*. 2022, 156:110504. DOI: 10.1016/j.ejrad.2022.110504.

- [33] Xu C, Liu W, Zhao Q, et al. CT-based radiomics nomogram for overall survival prediction in patients with cervical cancer treated with concurrent chemoradiotherapy. *Front Oncol.* 2023, 13:1287121. DOI: 10.3389/fonc.2023.1287121.
- [34] Meng X, Li W, Yuan H, et al. KDELR2-KIF20A axis facilitates bladder cancer growth and metastasis by enhancing Golgi-mediated secretion. *Biol Proced Online.* 2022, 24(1): 12. DOI:10.1186/s12575-022-00174-y.
- [35] Wei H, Ma W, Lu X, et al. KDELR2 promotes breast cancer proliferation via HDAC3-mediated cell cycle progression. *Cancer Commun (Lond).* 2021, 41(9): 904-920. DOI:10.1002/cac2.12180.
- [36] Bajaj R, Kundu ST, Grzeskowiak CL, et al. IMPAD1 and KDELR2 drive invasion and metastasis by enhancing Golgi-mediated secretion. *Oncogene.* 2020, 39(37): 5979-5994. DOI:10.1038/s41388-020-01410-z.
- [37] Jiao L, Zhang HL, Li DD, et al. Regulation of glycolytic metabolism by autophagy in liver cancer involves selective autophagic degradation of HK2 (hexokinase 2). *Autophagy.* 2018, 14(4): 671-684. DOI:10.1080/15548627.2017.1381804.
- [38] Liu Y, Murray-Stewart T, Casero RA Jr, et al. Targeting hexokinase 2 inhibition promotes radiosensitization in HPV16 E7-induced cervical cancer and suppresses tumor growth. *Int J Oncol.* 2017, 50(6): 2011-2023. DOI:10.3892/ijo.2017.3979.
- [39] Yang H, Hou H, Zhao H, et al. HK2 Is a Crucial Downstream Regulator of miR-148a for the Maintenance of Sphere-Forming Property and Cisplatin Resistance in Cervical Cancer Cells. *Front Oncol.* 2021, 11: 794015. DOI:10.3389/fonc.2021.794015.

- A multicentre large sample study has built a stable and reliable predictive tool that can efficiently predict the progression and prognosis of cervical cancer patients after concurrent chemoradiotherapy.
- Complementary of radiomics features in tumors and their surrounding 5 cm were analysed, and a prognostic model for cervical cancer was optimised in combination with clinical parameters.
- Multiple machine learning methods were used to screen the radiomics features and find the best combination of models.
- Possible molecular mechanisms of radiomics model to predict the prognosis of cervical cancer patients and the association between specific radiomics features and genes were revealed.

**Declaration of interests**

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: