

Deep learning integrated clinical-radiomics model for prognostic prediction in locally advanced cervical cancer through 18 F-FDG PET/CT and MRI combination

Jinling Song

Zhejiang Cancer Hospital

Kun Chen

Zhejiang Cancer Hospital

Yuhuang Cai

Wenzhou Medical University

Heqing Yi

yihq@zjcc.org.cn

Zhejiang Cancer Hospital

Research Article

Keywords: Cervical cancer, Radiomics, Deep learning, 18F-FDG PET/CT, MRI

Posted Date: October 18th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10116621/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Background

Precise pretreatment risk stratification is critical for guiding treatment intensification and personalized follow-up in patients with locally advanced cervical cancer (LACC) after concurrent chemoradiotherapy (CCRT). This study developed and validated a deep learning integrated clinical-radiomics model for prognosis prediction in patients with LACC treated with CCRT, using dual-modal pretreatment imaging from ^{18}F -FDG PET/CT and contrast-enhanced MRI.

Methods

We retrospectively collected 346 patients (276 for training and 70 for testing) with LACC between June 2016 and October 2021. Radiomics and deep learning signatures were extracted from PET/CT and MRI. These signatures were integrated with the clinicopathological factors into a nomogram for clinical applications. Models were constructed using eleven machine learning algorithms, and a clinical-radiomics-deep learning model (CRDLM) was developed. The evaluation of the model's performance was conducted through the metrics of the area under the curve (AUC).

Results

The CRDLM demonstrated superior predictive performance (AUCs of 0.998 (95% CI: 1.00–1.00) for training cohort and 0.901 (95% CI: 0.83–0.97) for testing cohort), which is significantly superior to conventional clinical factors (FIGO stage and SCC-Ag, AUC = 0.605–0.658). In addition, the CRDLM which combined radiomics with deep learning significantly outperformed the radiomics-only model (AUC = 0.793, $p = 0.006$). Sensitivity, specificity, and accuracy of CRDLM reached 0.947, 0.725, and 0.786, respectively.

Conclusions

The CRDLM provides robust prediction of disease progression in LACC patients who received CCRT, which may facilitate personalized treatment strategies and improve clinical outcomes. Further validation in multicenter prospective studies is warranted.

Introduction

Cervical cancer is one of the most common malignancies and the fourth leading cause of cancer-related death in women worldwide. Approximately 37%–60% of patients are diagnosed at locally advanced stages, resulting in a poor prognosis [1, 2]. The standard treatment for patients with locally advanced cervical cancer (LACC), as defined by the International Federation of Gynecology and Obstetrics (FIGO) staging IIB–IVA, is concomitant chemoradiation therapy (CCRT) using a cisplatin-based regimen, combined with pelvic external beam radiotherapy (EBRT) and subsequent brachytherapy (BT) [3, 4]. However, approximately one-third of LACC patients experience disease recurrence within two years of completing treatment, and the five-year overall survival rate for patients with recurrence is less than 50% [4–6]. To optimize treatment strategies and improve risk-adapted care, it is crucial to determine whether patients can achieve long-term disease-free survival with early-stage treatment. Currently, the outcomes of LACC patients after CCRT relied solely on conventional clinical parameters, such as FIGO stage. No effective methods provide satisfactory prediction accuracy [7]. Therefore, it is urgent to identify accurate predictive indicators for recurrence and survival after CCRT in LACC.

Fluorine-18 fluorodeoxyglucose (^{18}F -FDG) positron emission tomography/computed tomography (PET/CT) and magnetic resonance imaging (MRI) have well-established roles in managing LACC patients, particularly in clinical staging, monitoring treatment efficacy, and assessing prognosis [4]. Our previous study also demonstrated that metabolic parameters of cervical lesions and lymph nodes on ^{18}F -FDG PET/CT were associated with disease progression in LACC [8]. Recent attention has focused on the extraction and analysis of a variety of quantitative features from ^{18}F -FDG PET/CT through radiomics [9, 10]. This approach involves deconstructing three-dimensional (3D) texture features of lesions to derive numerous radiomic parameters, potentially related to tumor heterogeneity, aggressiveness, and immunoreactivity [11]. Previous studies have showed the potential of radiomics derived from pretreatment PET/CT or MRI for predicting survival in LACC patients treated with chemoradiotherapy despite the limited sample sizes [12–14]. Although radiomic parameters contain extensive biological information, efficiently identifying, screening, and utilizing these features remains a significant challenge in clinic [9, 15].

Deep learning (DL), an artificial intelligence technique for autonomously deriving quantitative representations from medical images, is an evolving field in radiomics research. Without explicit feature predefinition or selection, DL surpasses conventional radiomics by automatically extracting hierarchical, high-dimensional features directly from raw medical images, enabling it to navigate the data space and achieve superior problem-solving capabilities for more comprehensive tumor characterization [16]. Although radiomics remains the dominant approach due to its interpretability [17], integrating it with DL may offer exceptional performance in predicting disease progression.

In this study, we aim to develop and validate a clinical-radiomics-deep learning model (CRDLM) based on pretreatment ^{18}F -FDG PET/CT and contrast-enhanced MRI to predict disease progression in LACC patients treated with CCRT.

Materials and Methods

Patient inclusion

This study was approved by the institutional review board of our hospital (IRB-2022-423), and informed consent was waived due to its retrospective nature. Consecutive patients from June 2016 to October 2021 were retrospectively enrolled based on the following inclusion criteria: 1) histologically confirmed cervical cancer based on biopsy specimens, staged IIB-IVA according to the 2019 FIGO classification; 2) chemoradiotherapy as the primary treatment modality; 3) no prior treatments before ^{18}F -FDG PET/CT and enhanced MRI scans; and 4) availability of complete medical history and clinicopathological data. The exclusion criteria included: 1) secondary malignancies; 2) serious infections or inflammation; and 3) hepatic or renal dysfunction. A total of 346 patients were included in the final clinical database. Figure 1 illustrates the flowchart of the selection process.

Image acquisition and segmentation

PET/CT and MRI imaging protocol

^{18}F -FDG PET/CT scans were performed using a dedicated PET/CT system (Discovery 710; GE Medical Systems) in accordance with PET/CT tumor imaging guidelines [18]. All MRI studies were conducted using a 1.5-T unit (GE HealthCare, SIGNA Voyager, Waukesha, WI, USA) with a phased-array body coil, two weeks prior to the start of CCRT using standardized imaging protocols. Pelvic MRI was performed at least 10 days after cone biopsy to avoid false-positive findings due to post-biopsy inflammation. No patients had contraindications to MRI. Detailed information on the PET/CT and MRI protocols can be found in *Supplementary Methods*.

Imaging region segmentation

The PET and MRI images were processed independently by a single expert radiation oncologist (Reader 1). Pretreatment images (in DICOM format) were imported from the PACS system into ITK-SNAP software (version 3.6.0; <http://www.itksnap.org>). The tumor region of interest (ROI) was delineated layer-by-layer on (i) T2, (ii) CE-T1, (iii) DWI, and (iv) the ADC map derived from DWI-MRI. Each sequence was segmented independently due to anatomical variations across imaging modalities. Metabolically active volumes were delineated separately on PET and CT images using the same method. A month later, 30 patients in the cohort were randomly selected, and their ROIs were segmented again by reader 1 and a nuclear physician (reader 2) to construct two re-segmentation datasets for the assessment of intrareader/interreader reproducibility of radiomic features. An intraclass correlation coefficient (ICC) > 0.75 indicated good reproducibility.

Radiomics analysis

The radiomics model was constructed using multi-sequence images, as shown in Fig. 2 (orange dotted box). All analyses were performed using the uAI Research Portal (uRP, United Imaging Intelligence) [19], which included feature extraction, feature selection, and model construction. Each sequence was used to construct a separate model.

The detailed modeling process for each sequence was as follows. To minimize heterogeneity, images were standardized to similar resolution, and intensities were normalized with a fixed bin width. Initially, 2264 features were extracted from the image ROI and the corresponding ROI mask, including first-order statistics, shape-based features, and higher-order texture characteristics. Z-score normalization was applied to mitigate dimensional effects. Feature selection was conducted in two stages. First, the Minimum Redundancy Maximum Relevance (mRMR) method was used to identify and eliminate redundant features, retaining those with the highest relevance to the outcome variable. After mRMR, Least Absolute Shrinkage and Selection Operator (LASSO) regression was employed to further refine the feature set.

The refined feature set was then used to construct models using eleven commonly used machine learning algorithms: Support Vector Machine (SVM), Logistic Regression (LR), Gaussian Process (GP), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbors (KNN), Gradient Boosting Decision Trees (GBDT), AdaBoost, Bagging Decision Tree, Quadratic Discriminant Analysis (QDA), and Stochastic Gradient Descent (SGD). Each algorithm was trained and tested using the selected features, and the optimal algorithm for each sequence was identified. The corresponding radiomics score (Radscore) was derived from the predicted probabilities. To further refine the model, different multi-sequence combinations, such as PET/CT, MR (CE-T1, T2, DWI, ADC), and PET/CT + MR, were explored.

Deep learning analysis

A multi-sequence-based deep learning model was developed to predict the prognosis of cervical cancer patients, as shown in Fig. 2 (green dotted box). Each sequence of medical images (e.g., PET, CT, CE-T1, T2, DWI, ADC) was processed independently using a 2.5D ConvNeXt model [20], which is an extension of the 2D ConvNeXt architecture, designed to process 3D data while capturing spatial context. The model utilized pretrained ConvNeXt-Tiny parameters to enhance performance and ensure faster convergence.

For each sequence, the 3D image data was cropped based on the annotated ROI to extract 3D patches containing the ROI. These patches were then split into 2D slices, and each 2D slice, along with the preceding and following two slices, was concatenated to form 5-channel 2D slices for input into the

ConvNeXt blocks. After training, deep features were extracted using global average pooling (GAP), reducing spatial dimensions to produce a 1D feature vector for each 5-channel 2D slice. These feature vectors were then averaged to create a comprehensive feature representation for each patient.

To improve computational efficiency and avoid overfitting, dimensionality reduction was performed using principal component analysis (PCA), factor analysis, or independent component analysis. The reduced feature vectors were input into a machine learning algorithm (as described in Radiomics Analysis) to generate a prognosis prediction score (DLscore) for each sequence. Multi-sequence combinations, such as PET/CT, MR (CE-T1, T2, DWI, ADC), and PET/CT + MR, were also explored to refine the deep learning-based prediction model. The deep learning model was implemented with Python 3.8.13 and PyTorch 1.10.1, and dimensionality reduction and model construction were performed in uRP. The study design and pipeline are shown in Fig. 2.

Model comparison and evaluation

Univariate ($p < 0.1$) and multivariate ($p < 0.05$) logistic regression analyses were used to identify independent clinical risk factors, forming the basis for constructing the Clinic model. Two integrated models (Rad-Clin score and DL-Clin score) were developed by combining predictive probabilities from both imaging and clinical models. The Rad-DL model was constructed by integrating Radscore and DLscore, and a comprehensive clinical-radiomics-deep learning model (CRDLM) was created incorporating all three components.

Receiver operating characteristic (ROC) curves were used to assess model performance, with calculations for the area under the ROC curve (AUC), sensitivity (SEN), specificity (SPEC), accuracy (ACC), positive predictive value (PPV), and F1-Score. The DeLong test was employed to compare predictive efficiency between models. Additionally, net reclassification improvement (NRI) and integrated discriminatory improvement (IDI) were calculated. Calibration curve analysis was conducted, and the Brier loss was computed to assess consistency between actual recurrence status and predicted probabilities. Decision curve analysis (DCA) was performed to evaluate the clinical applicability of the models by quantifying net benefit at different risk thresholds. A p -value of less than 0.05 was considered statistically significant.

Treatment and Follow-up

All LACC patients underwent pelvic $\pm$ lombo-aortic external beam radiotherapy (EBRT) according to the 2015 NCCN guidelines. EBRT was delivered using intensity-modulated radiation therapy (IMRT), volumetric-modulated arc therapy (VMAT), or helical tomotherapy (HT) in 25–28 fractions of 1.8 Gray (Gy), for a total dose of 45–50.4 Gy. For patients with para-aortic nodal involvement, the superior border extended to the renal vessels or the upper margin of T12. In cases with significantly enlarged lymph nodes, additional radiotherapy (10–15 Gy) was applied with highly conformal EBRT. Concomitant chemotherapy with cisplatin (40 mg/m²) was administered weekly for 3–6 courses during radiotherapy, followed by pulse-dose rate brachytherapy (BT) with an Ir-192 source, delivering 24–36 Gy (biologically effective dose 38.4–57.6 Gy) in four to six fractions to point A. A small number of patients received radical radiotherapy alone.

Reviews were conducted 1 month after radical chemoradiotherapy, every 3–6 months for 2 years, and every 6–12 months for 3–5 years after treatment. Disease progression was determined via gynecologic examination, serum tumor markers, and enhanced abdominopelvic CT with or without follow-up ¹⁸F-FDG PET/CT. When clinical assessment, serum tumor markers, or imaging studies indicated abnormalities, additional diagnostic studies or pathological confirmation were performed to evaluate cancer progression. Progression-free survival (PFS) was calculated from the date of biopsy to disease progression, relapse, death from any cause, or the most recent follow-up.

Statistical analysis

The mean and standard deviation (SD) were utilized to describe normally distributed continuous data, and non-normally distributed continuous data were expressed as the median (min-max). Continuous data were compared using the independent sample T-test or Mann-Whitney U test as appropriate. Categorical data were expressed as frequencies and percentages, and were compared with the Chi-square test. Univariate ($p < 0.1$) and multivariate ($p < 0.05$) analyses of clinical characteristics were conducted to identify the independent clinical risk factors via the logistic regression method. A p -value less than 0.05 was indicative of statistical significance unless specified. All the statistical analyses were completed with Python software (version 3.8.13) and R software (version 4.4.3).

Results

Clinical characteristics

A total of 346 patients were included in this study (276 in the training cohort and 70 in the testing cohort). The mean patient age was 61.36 ± 10.35 years. The median follow-up was 35.82 months (range, 1.17–67.92 months). During follow-up, 96 patients (27.75%) experienced disease progression. No statistically significant differences were observed in baseline clinical data between the training and testing cohorts ($p > 0.005$). The clinical characteristics of the patients are summarized in Table 1.

Table 1
Clinical characteristics

Age (years, mean $\pm$ SD)	Training set(n = 276)	Testing set(n = 70)	p
	61.37 $\pm$ 10.49	61.36 $\pm$ 9.87	0.731
FIGO Staging			0.427
Ia-Ib	47(17.0%)	14(20%)	
Ia-IIIb	64(23.2%)	20(28.6%)	
IIc-IIIa	165(59.8%)	36(51.5%)	
Histological type			0.29
SCC	262(94.9%)	66(94.3%)	
AC/ASC	14(5.1%)	4(5.7%)	
Lombo-aortic EBRT			0.126
Yes	75(27.2%)	20(28.6%)	
No	201(72.8%)	50(71.4%)	
CEA (ng/mL, mean $\pm$ SD)	17.14 $\pm$ 19.23	14.90 $\pm$ 18.22	0.376
SCC-Ag (ng/mL, mean $\pm$ SD)	5.11 $\pm$ 14.69	4.39 $\pm$ 9.85	0.644
FIGO: international federation of gynecology and obstetrics; SCC: squamous cell carcinoma; AC: adenocarcinoma; ASC: adenosquamous carcinoma; EBRT: external beam radiation therapy; CEA: carcinoembryonic antigen; SCC-Ag: squamous cell carcinoma antigen			

Six candidate clinicopathological variables with potential prognostic value were evaluated to construct the clinicopathological model: age, FIGO stage, histological type, lombo-aortic EBRT, CEA, and SCC-Ag levels. Univariate and multivariate Cox regression analyses were performed to compare prognostic significance. Only FIGO stage and serum SCC-Ag level were retained for model construction ($p = 0.019$ and $p = 0.001$, respectively; Table S1). This model demonstrated limited predictive performance for progression-free survival (PFS), with an AUC of 0.658 (95% CI: 0.59–0.73) in the training cohort and 0.605 (95% CI: 0.46–0.75) in the testing cohort (Table 2).

Table 2
Performance of Clinic, Radiomics, DL, and Combined Models in training and testing cohorts

	Training set						Testing set					
	AUC (95% CI)	SEN	SPEC	ACC	PPV	F1-Score	AUC (95% CI)	SEN	SPEC	ACC	PPV	F1-Score
Clinicopathological model	0.658[0.59, 0.73]	0.589	0.611	0.605	0.352	0.441	0.605[0.46, 0.75]	0.684	0.588	0.614	0.382	0.491
Radiomics-based model												
PET	0.780[0.72, 0.84]	0.836	0.571	0.641	0.412	0.552	0.706[0.58, 0.83]	0.684	0.647	0.657	0.419	0.520
CT	0.661[0.59, 0.73]	0.575	0.655	0.634	0.375	0.454	0.656[0.52, 0.79]	0.737	0.627	0.657	0.424	0.538
CE-T1	0.708[0.64, 0.77]	0.890	0.404	0.533	0.349	0.502	0.652[0.52, 0.79]	0.947	0.412	0.557	0.375	0.537
T2	0.792[0.74, 0.84]	0.877	0.433	0.551	0.358	0.508	0.593[0.45, 0.74]	0.737	0.471	0.543	0.341	0.467
DWI	0.793[0.74, 0.85]	0.986	0.360	0.525	0.356	0.524	0.665[0.53, 0.80]	0.842	0.431	0.543	0.356	0.500
ADC	0.898[0.86, 0.93]	0.932	0.547	0.649	0.425	0.584	0.603[0.46, 0.75]	0.737	0.490	0.557	0.350	0.475
PET/CT	0.783[0.73, 0.84]	0.863	0.562	0.641	0.414	0.560	0.747[0.63, 0.87]	0.895	0.510	0.614	0.405	0.557
MR	0.999[1.00, 1.00]	1.000	0.764	0.826	0.603	0.753	0.764[0.65, 0.88]	0.737	0.667	0.686	0.452	0.560
Radscore	0.996[0.99, 1.00]	1.000	0.729	0.801	0.570	0.726	0.791[0.68, 0.90]	0.789	0.686	0.714	0.484	0.600
Rad-Clin score	0.997[0.99, 1.00]	1.000	0.877	0.909	0.745	0.854	0.793[0.69, 0.90]	0.684	0.784	0.757	0.542	0.605
DL-based model												
PET	0.851[0.81, 0.90]	0.932	0.567	0.663	0.436	0.594	0.746[0.63, 0.86]	0.842	0.529	0.614	0.400	0.542
CT	0.739[0.68, 0.80]	0.767	0.567	0.620	0.389	0.516	0.702[0.57, 0.83]	0.789	0.569	0.629	0.405	0.536
CE-T1	0.787[0.73, 0.84]	0.781	0.626	0.667	0.429	0.553	0.677[0.54, 0.81]	0.737	0.569	0.614	0.389	0.509
T2	0.796[0.74, 0.85]	0.712	0.734	0.728	0.491	0.581	0.676[0.54, 0.81]	0.684	0.569	0.600	0.371	0.481
DWI	0.743[0.68, 0.80]	0.836	0.542	0.620	0.396	0.537	0.723[0.60, 0.85]	0.895	0.510	0.614	0.405	0.557
ADC	0.785[0.73, 0.84]	0.822	0.626	0.678	0.441	0.574	0.685[0.55, 0.82]	0.737	0.549	0.600	0.378	0.500
PET/CT	0.910[0.88, 0.94]	0.781	0.857	0.837	0.663	0.717	0.793[0.69, 0.90]	0.684	0.804	0.771	0.565	0.619
MR	0.887[0.85, 0.93]	0.712	0.837	0.804	0.612	0.658	0.817[0.72, 0.92]	0.737	0.784	0.771	0.560	0.636
DLscore	0.968[0.95, 0.99]	0.753	0.956	0.902	0.859	0.803	0.872[0.79, 0.95]	0.789	0.863	0.843	0.682	0.732
DL-Clin score	0.973[0.96, 0.99]	0.918	0.892	0.899	0.753	0.827	0.877[0.80, 0.96]	0.789	0.882	0.857	0.714	0.750
Combined model												
Rad-DL score	0.988[0.98,1.00]	1.000	0.719	0.793	0.562	0.719	0.898[0.83,0.97]	1.000	0.706	0.786	0.559	0.717
CRDLM	0.998[1.00–1.00]	1.000	0.946	0.960	0.869	0.930	0.901[0.83,0.97]	0.947	0.725	0.786	0.563	0.706
AUC: area under the curve; DL: deep learning; SEN: sensitivity; SPEC: specificity; ACC: accuracy; PPV: positive predictive value												

The feature extraction and selection

A total of 34 radiomics features were selected from six multi-sequence datasets: six from CE-T1, seven from T2, five from DWI, two from ADC, seven from PET, and seven from CT. Details of the selected features are provided in the Table S2 and Supplementary Fig S1-S3.

Radiomics-based model performance

The radiomics models were constructed using multi-sequence images. The MRI-based radiomics model achieved strong predictive performance, with an AUC of 0.999 (95% CI: 1.00–1.00) in the training cohort and 0.764 (95% CI: 0.65–0.88) in the testing cohort (Table 2, Fig. 3). However, incorporating PET/CT radiomics did not significantly improve performance [Radscore: AUC, 0.996 (95% CI: 0.99–1.00) in training; 0.791 (95% CI: 0.68–0.90) in testing] nor did adding clinicopathological features [Rad-Clin score: AUC, 0.997 (95% CI: 0.99–1.00) in training; 0.793 (95% CI: 0.69–0.90) in testing; $p = 0.81$]. Radscore did not outperform either PET/CT- or MRI-based radiomics alone ($p = 0.46$ and $p = 0.09$, respectively; Tables S3) according to the DeLong test.

Deep learning-based model performance

The DL models demonstrated similar predictive performance: AUCs of 0.910 (95% CI: 0.88–0.94) and 0.887 (95% CI: 0.85–0.93) in the training cohort, and 0.793 (95% CI: 0.69–0.90) and 0.817 (95% CI: 0.72–0.92) in the testing cohort, respectively. The DLscore model, which combined PET/CT and MRI multi-sequence images, achieved higher predictive performance, with an AUC of 0.968 (95% CI: 0.95–0.99) in training and 0.872 (95% CI: 0.79–0.95) in testing (Table 2, Fig. 3). Incorporating clinicopathological features further improved performance [DL-Clin score: AUC, 0.973 (95% CI: 0.96–0.99) in training; 0.877 (95% CI: 0.80–0.96) in testing]. However, no statistically significant differences were found between DLscore, DL-Clin score, and the radiomics models in predicting PFS (DeLong test, $p > 0.05$) (Tables S4).

Clinical-radiomics DL model

Finally, a combined model integrating radiomics and DL features (Rad-DL score) was developed. This model achieved superior predictive performance in both cohorts: AUCs of 0.988 (95% CI: 0.98–1.00) in training and 0.898 (95% CI: 0.83–0.97) in testing (Table 2, Fig. 4). The CRDLM yielded the best results, with AUCs of 0.998 (95% CI: 1.00–1.00) in training and 0.901 (95% CI: 0.83–0.97) in testing, significantly outperforming Rad-Clin score (DeLong test, $p = 0.006$).

Comparison the performance of models to predict progression

When comparing Radscore and DLscore models using PET/CT and MRI data, neither the DeLong test ($p = 0.270$) nor NRI/IDI analyses showed significant differences (NRI = 1.176, $p = 0.119$; IDI = 0.023, $p = 0.192$). Incorporating clinical factors modestly improved performance: DL-Clin score achieved higher accuracy than Rad-Clin score (AUC 0.877 vs. 0.793; DeLong test, $p = 0.246$; NRI = 0.203, $p = 0.109$; IDI = 0.340, $p < 0.001$). The combined model (Rad-DL score) demonstrated improved predictive performance compared with Radscore (AUC 0.898 vs. 0.791; DeLong test, $p = 0.074$; NRI = 0.230, $p = 0.035$; IDI = 0.061, $p < 0.001$) but not DLscore (AUC 0.898 vs. 0.872; DeLong test, $p = 0.190$; NRI = 0.054, $p = 0.326$; IDI = 0.083, $p < 0.001$). Adding clinical factors produced similar results: the CRDLM outperformed Rad-Clin score (AUC 0.901 vs. 0.790; DeLong test, $p = 0.006$), with NRI and IDI analyses confirming significant improvement (NRI = 0.204, $p = 0.060$; IDI = 0.309, $p < 0.001$). However, no significant difference was found between the CRDLM and DL-Clin score (AUC 0.901 vs. 0.877; DeLong test, $p = 0.573$; NRI = 0.001, $p = 0.496$; IDI = 0.401, $p = 0.344$) (Tables S5-6). The CRDLM achieved the highest sensitivity (0.947 vs. 0.684) and greater accuracy (0.786 vs. 0.757) than Rad-Clin score (Table 2, Fig. 5).

Discussion

In this study, we applied deep learning methods to develop a new predictive model integrated with radiomics (CRDLM) for prognosis prediction in LACC patients ($n = 346$) treated with CCRT. To our knowledge, this represents the largest cohort of LACC patients with longitudinal follow-up reported to date. Our findings demonstrated that the CRDLM achieved optimal predictive performance for disease evolution and outcomes after CCRT in LACC patients. Specifically, the CRDLM based on multiparametric MRI and ^{18}F -FDG PET/CT, achieved excellent performance with AUCs of 0.998 in the training cohort and 0.901 in the testing cohort. These preliminary results suggest that the model enables early and accurate prognostic prediction, including the risk and sites of tumor progression following first-line CCRT in LACC patients. Such prediction is expected to guide treatment selection, decisions regarding therapy intensification, and follow-up strategies, ultimately improving clinical outcomes.

This study employed a combined PET/CT and MRI multimodal imaging strategy, primarily leveraging the complementary advantages of both modalities in revealing tumor biological characteristics. ^{18}F -FDG PET/CT is a widely used diagnostic tool in oncology, and several studies have reported associations between conventional PET parameters at the primary cervical tumor and adverse outcomes[21, 22]. FDG uptake intensity is influenced by factors such as glucose transporter and hexokinase activity, neoangiogenesis, and proliferative activity, which collectively reflect tumor aggressiveness[23, 24]. MRI, with its superior soft tissue resolution and multiparametric capability, has also shown promise for survival prediction in LACC. Multi-sequence MRI radiomics, particularly from T2WI, contrast-enhanced T1WI (CE-T1), DWI, and ADC maps, has been reported to outperform single-sequence approaches [25–27]. Compared with previous studies of the similar type, this study not only expanded the sample size but also integrated radiomics with deep learning methods. The constructed model demonstrated superior predictive performance[12, 13].

In terms of model construction strategy, we integrated radiomics with deep learning and demonstrated its superiority. Radiomics enables noninvasive extraction of diagnostic, therapeutic, and prognostic features from medical images [28]. Prior PET radiomics studies in LACC demonstrated improved prognostic accuracy compared with conventional PET parameters[29–33]. Deep learning, particularly convolutional neural networks, enables automated extraction of hierarchical image features, capturing subtle patterns beyond handcrafted [34]. By integrating DL with radiomics, we developed a clinical–

radiomics DL model that outperformed traditional models. Incorporating clinical-pathological information further improved prediction of progression-free survival. Compared with the radiomics-only model (Rad-Clin score), the DL-enhanced model showed significant improvement in prognostic accuracy (AUC 0.901 vs. 0.793). This performance improvement may be attributed to the ability of radiomics features to stably quantify traditional imaging phenotypes, and deep learning features automatically extract information about tumor heterogeneity. The complementary advantages of both approaches at the feature level collectively enhance the model's predictive capability.

In the development of high-dimensional small-sample data models, overfitting risks and feature selection instability are common issues that may compromise research conclusions. To address this, we adopted a phased, multi-strategy approach for feature screening and modeling. First, we employed the mRMR algorithm for preliminary screening. This algorithm aims to eliminate redundant features while retaining those most strongly associated with the outcome variable, thereby reducing multicollinearity among features and enhancing model generalization[35–37]. Subsequently, we introduced LASSO regression for secondary feature selection. Its L1 regularization property compresses unimportant feature coefficients to zero, achieving dual objectives of feature selection and model regularization—a classic approach for addressing overfitting in high-dimensional data[38, 39]. During model construction, we systematically compared 11 commonly used machine learning algorithms and selected the optimal algorithm for each sequence, ensuring objectivity and peak performance in model building. Furthermore, in deep learning, we employed the pre-trained ConvNeXt-Tiny model for transfer learning and applied dimensionality reduction techniques like PCA to extracted deep features[40, 41]. These measures collectively aim to reduce model parameter complexity, enhance computational efficiency, and effectively suppress overfitting.

However, this study still has some limitations. First, the relatively limited number of events in the test set and the lack of external validation may affect the model's generalization ability to some extent. Second, although the fusion model demonstrated superior predictive performance, interpreting its internal mechanisms remains challenging. Specifically, the relationship between the high-dimensional features extracted by the deep learning module and tumor biological behavior remains unclear. Future studies could explore the biological basis underlying the model's features using approaches such as image genomics to enhance interpretability and clinical acceptance. Third, ROIs in this study were manually delineated. Future research may consider adopting deep learning-based automated segmentation methods to improve consistency and reproducibility. Finally, our follow-up revealed a relatively limited number of mortality events, insufficient for robust OS analysis. Therefore, larger-scale, multicenter prospective studies with extended follow-up periods remain necessary to further validate and refine the findings of this research.

Conclusion

In summary, our clinical-radiomics DL model, based on pretreatment multi-sequence MRI and ¹⁸F-FDG PET/CT images, shows strong potential for predicting disease progression in LACC patients following CCRT. Future studies with larger, prospectively validated datasets are warranted to further support its clinical utility and to assist clinicians in developing personalized treatment strategies for LACC patients.

Abbreviations

Abbreviations

Definitions

LACC: Locally advanced cervical cancer
Locally advanced cervical cancer

ADC

Apparent diffusion coefficient

AUC

Area under the curve

CCRT

Concurrent chemoradiotherapy

CE-T1

Contrast-enhanced T1-weighted imaging

CRDLM

Clinical-radiomics-deep learning model

CT

Computed tomography

DL

Deep learning

DWI

Diffusion-weighted imaging

LASSO

Least Absolute Shrinkage and Selection Operator

MRI: Magnetic resonance imaging

Magnetic resonance imaging

PET

Positron emission tomography

PFS

Progression-free survival

ROC: Receiver operating characteristic

Receiver operating characteristic

¹⁸F-FDG

Fluorine-18 fluorodeoxyglucose

Declarations

Funding:

This work was supported by Zhejiang Province Natural Science Foundation of China (Grant No. LTGY24H180013).

Conflicts of interest/Competing interests:

The authors declare that they have no conflict of interest.

Availability of data and material:

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

Authors' contributions:

JLS, the first author, collecting, analyzed and interpreted image data; drafted the manuscript. KC the co-first author, revised the manuscript. YHC collected clinical data and follow-up LACC patients. HQY, the corresponding authors, responsible for the study designing, and obtaining grants and funding to support the trial, enhanced the intellectual content of the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate:

This retrospective chart review study involving human participants was in accordance with the ethical standards of the institutional and national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The Human Investigation Committee (IRB number: IRB-2022-423) of Zhejiang Cancer Hospital approved this study. The institutional review board granted a waiver of informed consent for this retrospective analysis.

Declaration of generative AI use:

The authors did not use any generative AI tools or AI-assisted technologies in the design, analysis, or writing of this manuscript.

Data Availability:

The datasets generated during the current study are available.

Consent for publication:

Not Applicable.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* 2021;71:209–49. 10.3322/caac.21660.
2. Cohen PA, Jhingran A, Oaknin A, Denny L. Cervical cancer. *Lancet.* 2019;393:169–82. 10.1016/s0140-6736(18)32470-x.
3. Cibula D, Raspollini MR, Planchamp F, Centeno C, Chargari C, Felix A, et al. ESGO/ESTRO/ESP Guidelines for the management of patients with cervical cancer - Update 2023. *Virchows Arch.* 2023;482:935–66. 10.1007/s00428-023-03552-3.
4. Koh WJ, Abu-Rustum NR, Bean S, Bradley K, Campos SM, Cho KR, et al. Cervical Cancer, Version 3.2019, NCCN Clinical Practice Guidelines in Oncology. *J Natl Compr Canc Netw.* 2019;17:64–84. 10.6004/jnccn.2019.0001.
5. Reducing uncertainties about the effects of chemoradiotherapy for cervical cancer: a systematic review and meta-analysis of individual patient data from 18 randomized trials. *J Clin Oncol.* 2008;26:5802–12. 10.1200/jco.2008.16.4368.
6. Tseng JY, Yen MS, Twu NF, Lai CR, Horng HC, Tseng CC, et al. Prognostic nomogram for overall survival in stage IIB-IVA cervical cancer patients treated with concurrent chemoradiotherapy. *Am J Obstet Gynecol.* 2010;202:e1741–7. 10.1016/j.ajog.2009.09.028.
7. Rose PG, Java J, Whitney CW, Stehman FB, Lanciano R, Thomas GM, et al. Nomograms Predicting Progression-Free Survival, Overall Survival, and Pelvic Recurrence in Locally Advanced Cervical Cancer Developed From an Analysis of Identifiable Prognostic Factors in Patients From NRG Oncology/Gynecologic Oncology Group Randomized Trials of Chemoradiotherapy. *J Clin Oncol.* 2015;33:2136–42. 10.1200/jco.2014.57.7122.

8. Song J, Pang W, Yi H, Ji J, Ye X, Li L. Tumor and metastatic lymph nodes metabolic activity on (18)F-FDG-PET/CT to predict progression-free survival in locally advanced cervical cancer. *Abdom Radiol (NY)*. 2024;49:975–84. 10.1007/s00261-023-04158-8.
9. Hatt M, Le Rest CC, Tixier F, Badic B, Schick U, Visvikis D. Radiomics: Data Are Also Images. *J Nucl Med*. 2019;60:s38–44. 10.2967/jnumed.118.220582.
10. Mayerhoefer ME, Materka A, Langs G, Häggström I, Szczypiński P, Gibbs P, et al. Introduction to Radiomics. *J Nucl Med*. 2020;61:488–95. 10.2967/jnumed.118.222893.
11. Xia TY, Zhou ZH, Meng XP, Zha JH, Yu Q, Wang WL, et al. Predicting Microvascular Invasion in Hepatocellular Carcinoma Using CT-based Radiomics Model. *Radiology*. 2023;307:e222729. 10.1148/radiol.222729.
12. Lucia F, Visvikis D, Desseroit MC, Miranda O, Malhaire JP, Robin P, et al. Prediction of outcome using pretreatment (18)F-FDG PET/CT and MRI radiomics in locally advanced cervical cancer treated with chemoradiotherapy. *Eur J Nucl Med Mol Imaging*. 2018;45:768–86. 10.1007/s00259-017-3898-7.
13. Lucia F, Visvikis D, Vallières M, Desseroit MC, Miranda O, Robin P, et al. External validation of a combined PET and MRI radiomics model for prediction of recurrence in cervical cancer patients treated with chemoradiotherapy. *Eur J Nucl Med Mol Imaging*. 2019;46:864–77. 10.1007/s00259-018-4231-9.
14. Ferreira M, Lovinfosse P, Hermesse J, Decuypere M, Rousseau C, Lucia F, et al. [(18)F]FDG PET radiomics to predict disease-free survival in cervical cancer: a multi-scanner/center study with external validation. *Eur J Nucl Med Mol Imaging*. 2021;48:3432–43. 10.1007/s00259-021-05303-5.
15. Collarino A, Feudo V, Pasciuto T, Florit A, Pfaehler E, de Summa M, et al. Is PET Radiomics Useful to Predict Pathologic Tumor Response and Prognosis in Locally Advanced Cervical Cancer? *J Nucl Med*. 2024;65:962–70. 10.2967/jnumed.123.267044.
16. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018;18:500–10. 10.1038/s41568-018-0016-5.
17. Gu W, Chen Y, Zhu H, Chen H, Yang Z, Mo S, et al. Development and validation of CT-based radiomics deep learning signatures to predict lymph node metastasis in non-functional pancreatic neuroendocrine tumors: a multicohort study. *EClinicalMedicine*. 2023;65:102269. 10.1016/j.eclinm.2023.102269.
18. Delbeke D, Coleman RE, Guiberteau MJ, Brown ML, Royal HD, Siegel BA, et al. Procedure guideline for tumor imaging with 18F-FDG PET/CT 1.0. *J Nucl Med*. 2006;47:885–95.
19. Wu J, Xia Y, Wang X, Wei Y, Liu A, Innanje A, et al. uRP: An integrated research platform for one-stop analysis of medical images. *Front Radiol*. 2023;3:1153784. 10.3389/fradi.2023.1153784.
20. Liu Z, Mao H, Wu CY, Feichtenhofer C, Darrell T, Xie S. A ConvNet for the 2020s. 2022 IEEE/CVF Conference on Computer Vision and, Recognition P. (CVPR); 2022. pp. 11966–76.
21. Han L, Wang Q, Zhao L, Feng X, Wang Y, Zou Y, et al. A Systematic Review and Meta-Analysis of the Prognostic Impact of Pretreatment Fluorodeoxyglucose Positron Emission Tomography/Computed Tomography Parameters in Patients with Locally Advanced Cervical Cancer Treated with Concomitant Chemoradiotherapy. *Diagnostics (Basel)*. 2021;11. 10.3390/diagnostics11071258.
22. Han S, Kim H, Kim YJ, Suh CH, Woo S. Prognostic Value of Volume-Based Metabolic Parameters of (18)F-FDG PET/CT in Uterine Cervical Cancer: A Systematic Review and Meta-Analysis. *AJR Am J Roentgenol*. 2018;211:1112–21. 10.2214/ajr.18.19734.
23. Rufini V, Collarino A, Calcagni ML, Meduri GM, Fuoco V, Pasciuto T, et al. The role of (18)F-FDG-PET/CT in predicting the histopathological response in locally advanced cervical carcinoma treated by chemo-radiotherapy followed by radical surgery: a prospective study. *Eur J Nucl Med Mol Imaging*. 2020;47:1228–38. 10.1007/s00259-019-04436-y.
24. Bos R, van Der Hoeven JJ, van Der Wall E, van Der Groep P, van Diest PJ, Comans EF, et al. Biologic correlates of (18)fluorodeoxyglucose uptake in human breast cancer measured by positron emission tomography. *J Clin Oncol*. 2002;20:379–87. 10.1200/jco.2002.20.2.379.
25. Zhang X, Zhao J, Zhang Q, Wang S, Zhang J, An J, et al. MRI-based radiomics value for predicting the survival of patients with locally advanced cervical squamous cell cancer treated with concurrent chemoradiotherapy. *Cancer Imaging*. 2022;22:35. 10.1186/s40644-022-00474-2.
26. Autorino R, Gui B, Panza G, Boldrini L, Cusumano D, Russo L, et al. Radiomics-based prediction of two-year clinical outcome in locally advanced cervical cancer patients undergoing neoadjuvant chemoradiotherapy. *Radiol Med*. 2022;127:498–506. 10.1007/s11547-022-01482-9.
27. Yan Q, Wu M, Zhang J, Yang J, Lv G, Qu B, et al. MRI radiomics and nutritional-inflammatory biomarkers: a powerful combination for predicting progression-free survival in cervical cancer patients undergoing concurrent chemoradiotherapy. *Cancer Imaging*. 2024;24:144. 10.1186/s40644-024-00789-2.
28. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology*. 2016;278:563–77. 10.1148/radiol.2015151169.
29. Chen SW, Shen WC, Hsieh TC, Liang JA, Hung YC, Yeh LS, et al. Textural features of cervical cancers on FDG-PET/CT associate with survival and local relapse in patients treated with definitive chemoradiotherapy. *Sci Rep*. 2018;8:11859. 10.1038/s41598-018-30336-6.
30. Schernberg A, Reuze S, Orhac F, Buvat I, Dercle L, Sun R, et al. A score combining baseline neutrophilia and primary tumor SUV(peak) measured from FDG PET is associated with outcome in locally advanced cervical cancer. *Eur J Nucl Med Mol Imaging*. 2018;45:187–95. 10.1007/s00259-017-3824-z.
31. Mu W, Liang Y, Hall LO, Tan Y, Balagurunathan Y, Wenham R, et al. (18)F-FDG PET/CT Habitat Radiomics Predicts Outcome of Patients with Cervical Cancer Treated with Chemoradiotherapy. *Radiol Artif Intell*. 2020;2:e190218. 10.1148/ryai.2020190218.

32. de Alencar NRG, Machado MAD, Mourato FA, de Oliveira ML, Moraes TF, Mattos Junior LAR, et al. Exploratory analysis of radiomic as prognostic biomarkers in (18)F-FDG PET/CT scan in uterine cervical cancer. *Front Med (Lausanne)*. 2022;9:1046551. 10.3389/fmed.2022.1046551.
33. Cho HW, Lee ES, Lee JK, Eo JS, Kim S, Hong JH. Prognostic value of textural features obtained from F-fluorodeoxyglucose (F-18 FDG) positron emission tomography/computed tomography (PET/CT) in patients with locally advanced cervical cancer undergoing concurrent chemoradiotherapy. *Ann Nucl Med*. 2023;37:44–51. 10.1007/s12149-022-01802-z.
34. Jiang X, Li J, Kan Y, Yu T, Chang S, Sha X, et al. MRI Based Radiomics Approach With Deep Learning for Prediction of Vessel Invasion in Early-Stage Cervical Cancer. *IEEE/ACM Trans Comput Biol Bioinform*. 2021;18:995–1002. 10.1109/tcbb.2019.2963867.
35. Chen D, Zhang R, Huang X, Ji C, Xia W, Qi Y, et al. MRI-derived radiomics assessing tumor-infiltrating macrophages enable prediction of immune-phenotype, immunotherapy response and survival in glioma. *Biomark Res*. 2024;12:14. 10.1186/s40364-024-00560-6.
36. Gao X, Yang L, She T, Wang F, Ding H, Lu Y, et al. Habitat-aware radiomics and adaptive 2.5D deep learning predict treatment response and long-term survival in ESCC patients undergoing neoadjuvant chemoimmunotherapy. *Eur J Nucl Med Mol Imaging*. 2026;53:1651–70. 10.1007/s00259-025-07522-6.
37. Zhang H, Jiang X, Yang G, Tang Y, Qi L, Chen M, et al. Machine learning-based construction and validation of an radiomics model for predicting ISUP grading in prostate cancer: a multicenter radiomics study based on [68Ga]Ga-PSMA PET/CT. *Eur J Nucl Med Mol Imaging*. 2025;53:306–19. 10.1007/s00259-025-07412-x.
38. Warkentin MT, Al-Sawaihey H, Lam S, Liu G, Diergaarde B, Yuan JM, et al. Radiomics analysis to predict pulmonary nodule malignancy using machine learning approaches. *Thorax*. 2024;79:307–15. 10.1136/thorax-2023-220226.
39. Guo F, Sun S, Deng X, Wang Y, Yao W, Yue P, et al. Predicting axillary lymph node metastasis in breast cancer using a multimodal radiomics and deep learning model. *Front Immunol*. 2024;15:1482020. 10.3389/fimmu.2024.1482020.
40. Kamsari M, Sadeghi S, de Oliveira GG, Alves AM, Sarshar NT, Anari S, et al. The role of data augmentation and attention mechanisms in UNet and ConvNeXt architectures for optimizing breast tumor segmentation. *Sci Rep*. 2025;15:45268. 10.1038/s41598-025-29841-2.
41. Ashraf M, Aslam MZ, Saeed N, Hussain SJ. PalmNeXt: a ConvNeXt-based deep learning model for pest detection in date palm leaves. *Front Plant Sci*. 2025;16:1738129. 10.3389/fpls.2025.1738129.

Figures

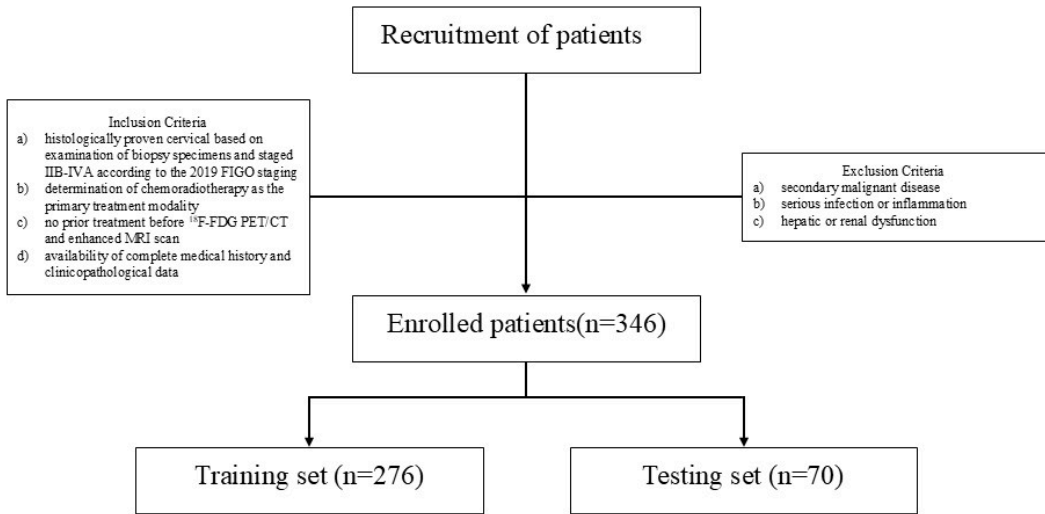


Figure 1

The flowchart of the selection process.

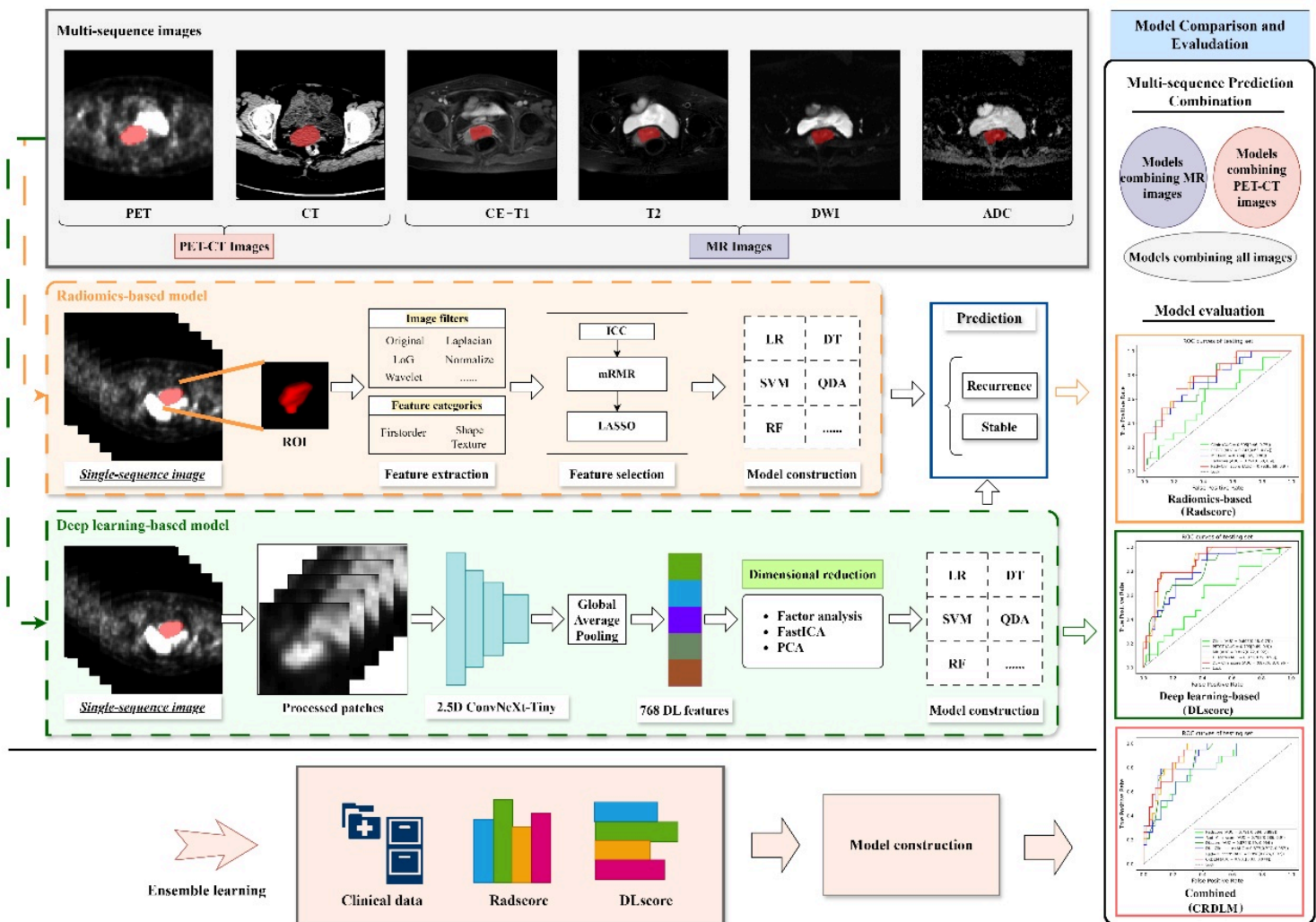


Figure 2

Schematic overview of the study design and workflow. Using multi-sequence images, radiomics-based (orange dotted box), deep learning-based (green dotted box) and combined models (bottom section) were developed, compared and evaluated for predicting progression in locally advanced cervical cancer.

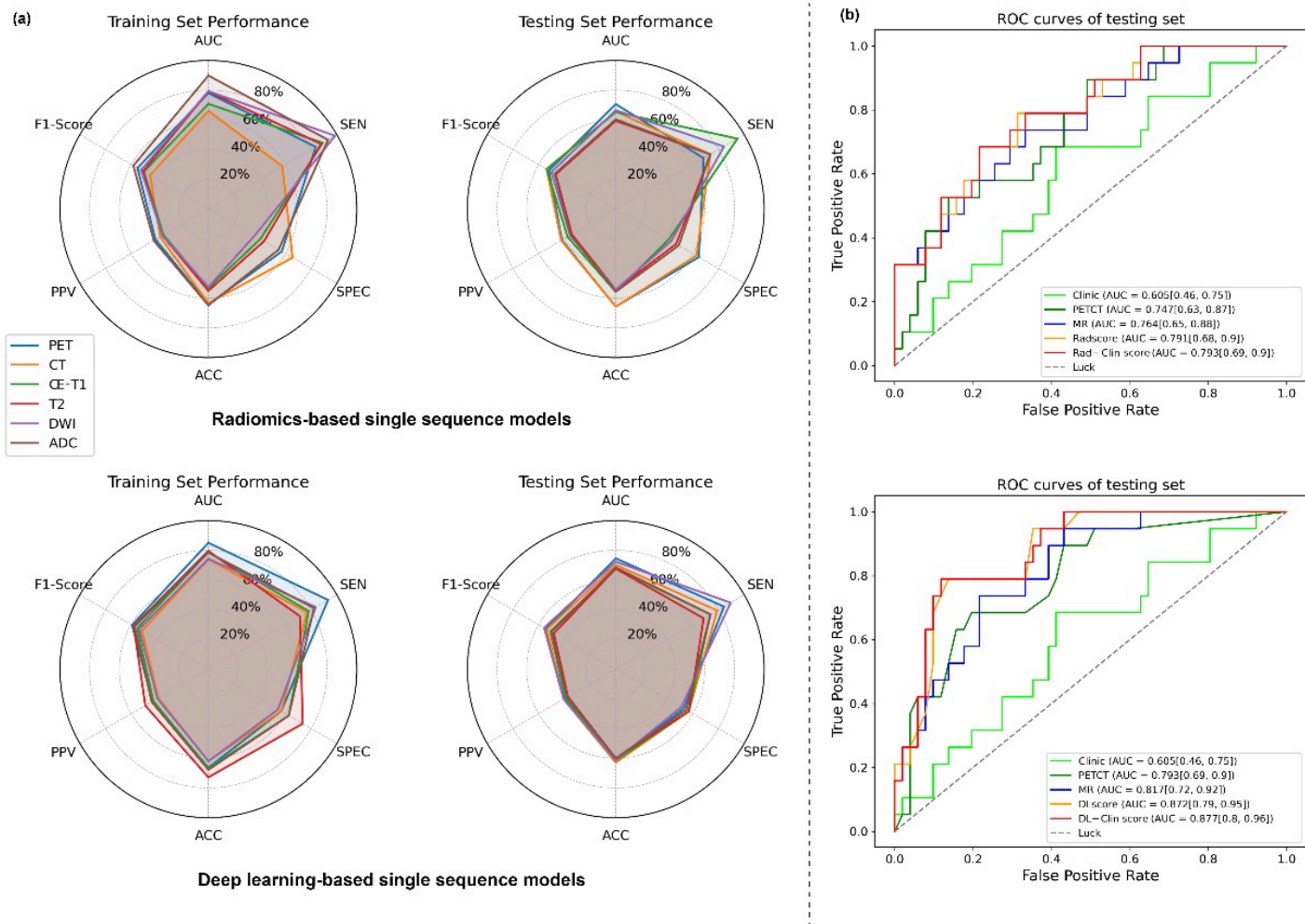


Figure 3

Performance of radiomics-based (top row) and deep learning-based (bottom row) models using single imaging sequences and integrated data. (a) Radar chart illustrating the performance of models trained on individual sequences: PET, CT, CE-T1, T2, DWI and ADC. (b) ROC curves comparing the performance of clinical (Clinic), PETCT, MR, and integrated models in the testing set.

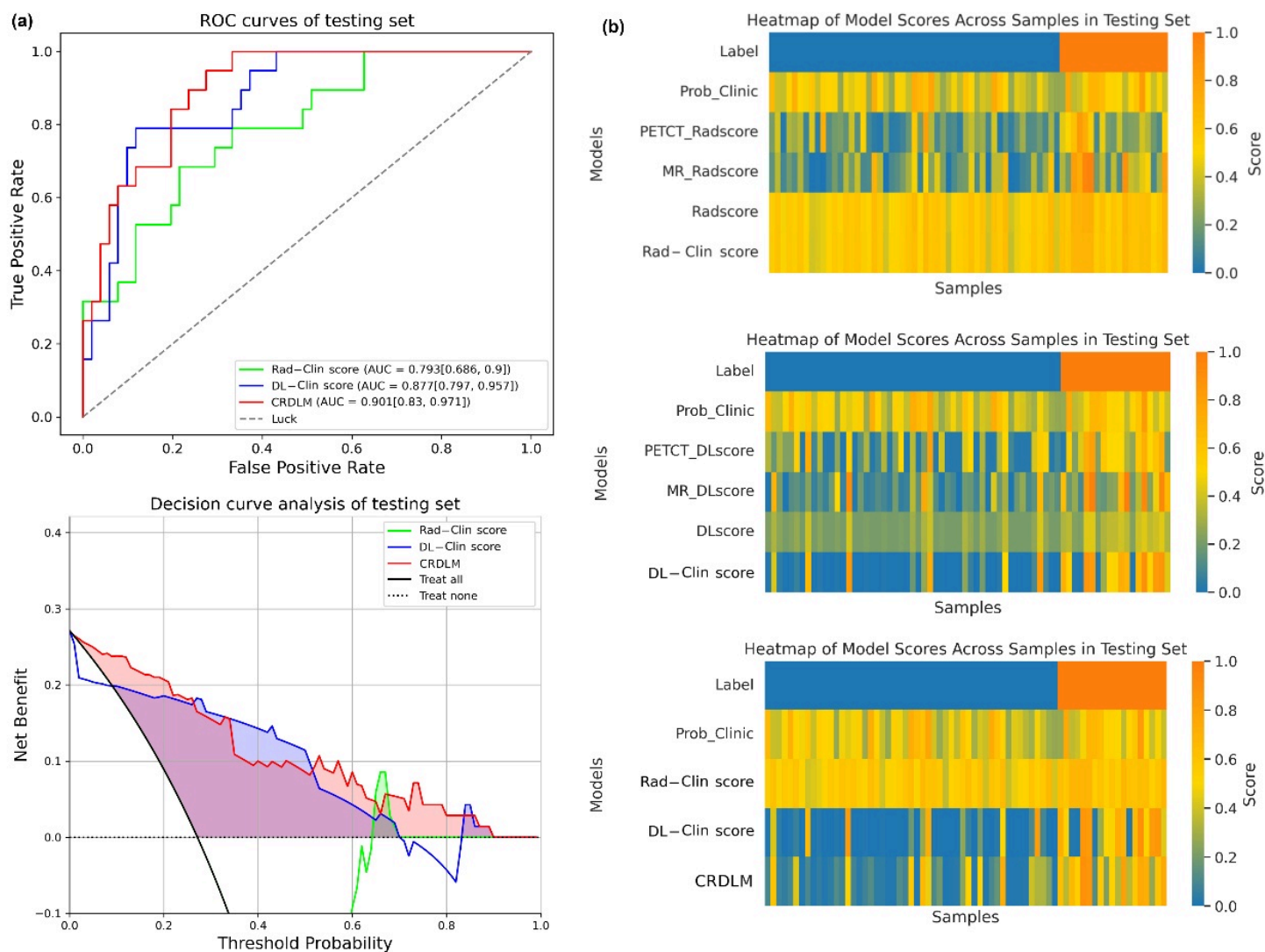


Figure 4

Performance comparison of integrated and combined models. (a) ROC curves and decision curve analysis for the Rad-Clin score, DL-Clin score and CRDLM models in the testing set. (b) Distribution of predictive scores from the Radscore, DLscore, and combined models.

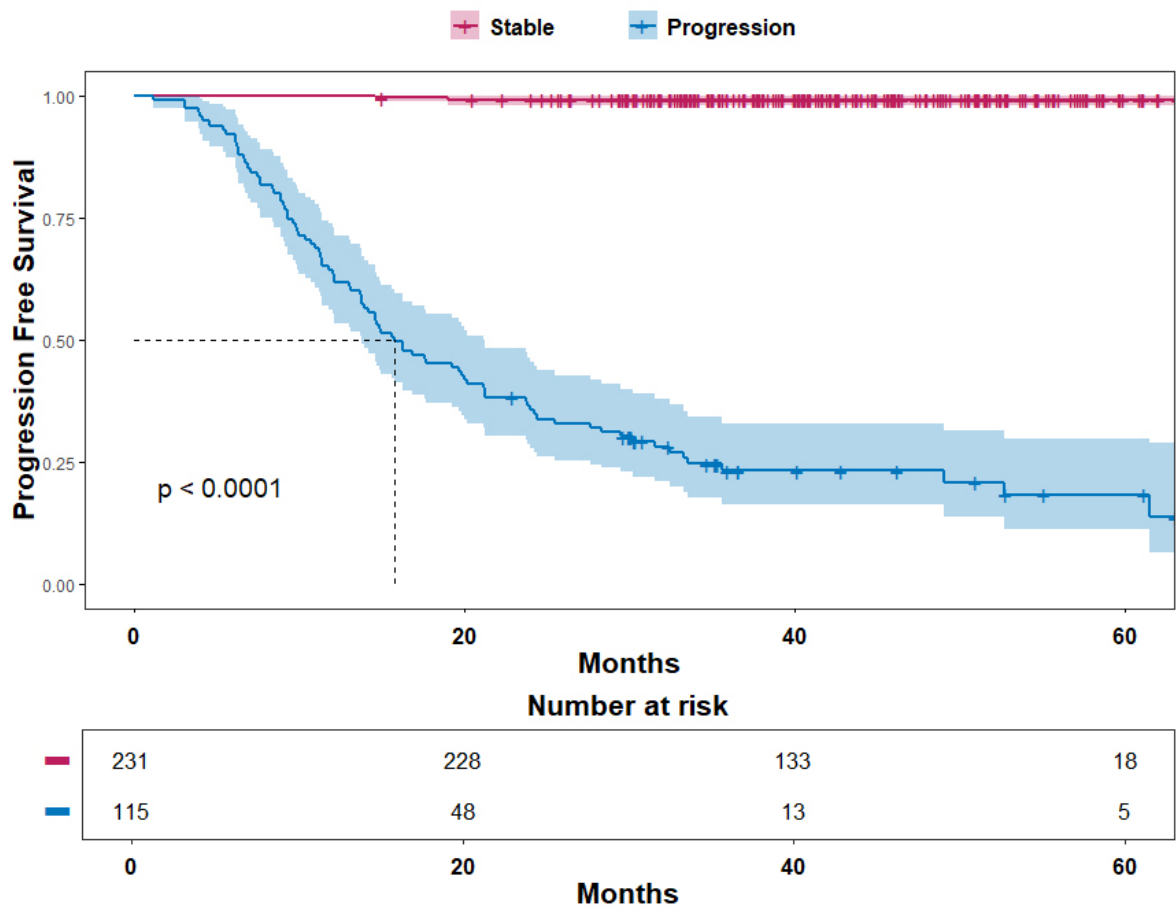


Figure 5

Clinical application of the CRDLM. Kaplan–Meier curves demonstrate a significant difference in progression-free survival (PFS) between the predicted progression and stable groups.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [supplementary.zip](#)