

Research Paper

Association of MRI-derived radiomic biomarker with disease-free survival in patients with early-stage cervical cancer

Jin Fang^{1*}, Bin Zhang^{1*}, Shuo Wang^{2,3*}, Yan Jin^{4*}, Fei Wang^{1*}, Yingying Ding⁴, Qiuying Chen¹, Liting Chen¹, Yueyue Li¹, Minmin Li¹, Zhuozhi Chen¹, Lizhi Liu⁵✉, Zhenyu Liu^{3,6}✉, Jie Tian^{2,3,6}✉, Shuixing Zhang¹✉

1. Medical Imaging Center, The First Affiliated Hospital of Jinan University, Guangzhou, 510630, China;
2. Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine, Beihang University, Beijing, 100191, China;
3. CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China;
4. Department of Radiology, the Third Affiliated Hospital of Kunming Medical University (Yunnan Cancer Hospital), Kunming, Yunnan, 650031, China;
5. Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Guangzhou, 510060, China;
6. University of Chinese Academy of Sciences, Beijing, 100080, China.

*Jin Fang, Bin Zhang, Shuo Wang, Yan Jin and Fei Wang contributed equally to this work.

✉ Corresponding authors: Lizhi Liu, M.D., Ph.D., No.651 Dongfeng East Road, Yuexiu District, Guangzhou, Guangdong, 510060, China. E-mail: liulizh@sysucc.org.com; Tel: +86-20-87343217; Fax: +86-20-87342125. Zhenyu Liu, Ph.D., CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, China. No.95 Zhongguancun east road, Beijing, 100190, China. Email: zhenyu.liu@ia.ac.cn; Tel: +86-10-82618465; Fax: +86-10-62527995. Jie Tian, Ph.D., Fellow of AIMBE, IAMBE, IEEE, SPIE, OSA, IAPR, ISMRM, Director of the CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, China. No.95 Zhongguancun east road, Beijing, 100190, China. Email: jie.tian@ia.ac.cn; Tel: +86-10-82618465; Fax: +86-10-62527995. Shuixing Zhang, M.D., Ph.D., No. 613 Huangpu West Road, Tianhe District, Guangzhou, Guangdong 510630, China. E-mail: shui7515@126.com; Tel: +86-20-38688417; Fax: +86-20-38688888.

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). See <http://ivyspring.com/terms> for full terms and conditions.

Received: 2019.06.09; Accepted: 2019.12.12; Published: 2020.01.16

Abstract

Pre-treatment survival prediction plays a key role in many diseases. We aimed to determine the prognostic value of pre-treatment Magnetic Resonance Imaging (MRI) based radiomic score for disease-free survival (DFS) in patients with early-stage (IB-IIA) cervical cancer.

Methods: A total of 248 patients with early-stage cervical cancer underwent radical hysterectomy were included from two institutions between January 1, 2011 and December 31, 2017, whose MR imaging data, clinicopathological data and DFS data were collected. Patients data were randomly divided into the training cohort (n = 166) and the validation cohort (n=82). Radiomic features were extracted from the pre-treatment T2-weighted (T2w) and contrast-enhanced T1-weighted (CET1w) MR imagings for each patient. Least absolute shrinkage and selection operator (LASSO) regression and Cox proportional hazard model were applied to construct radiomic score (Rad-score). According to the cutoff of Rad-score, patients were divided into low- and high- risk groups. Pearson's correlation and Kaplan-Meier analysis were used to evaluate the association of Rad-score with DFS. A combined model incorporating Rad-score, lymph node metastasis (LNM) and lymphovascular space invasion (LVI) by multivariate Cox proportional hazard model was constructed to estimate DFS individually.

Results: Higher Rad-scores were significantly associated with worse DFS in the training and validation cohorts ($P<0.001$ and $P=0.011$, respectively). The Rad-score demonstrated better prognostic performance in estimating DFS (C-index, 0.753; 95% CI: 0.696-0.805) than the clinicopathological features (C-index, 0.632; 95% CI: 0.567-0.700). However, the combined model showed no significant improvement (C-index, 0.714; 95%CI: 0.642-0.784).

Conclusion: The results demonstrated that MRI-derived Rad-score can be used as a prognostic biomarker for patients with early-stage (IB-IIA) cervical cancer, which can facilitate clinical decision-making.

Key words: cervical cancer; magnetic resonance imaging; radiomics; disease-free survival

Introduction

As the fourth leading cause of cancer-derived death in women, cervical cancer is a global health problem [1], with an estimated 570,000 cases and 311,000 deaths worldwide in 2018 [2]. During the past decades, increasing early-stage cervical cancer has been detected largely due to the popularity of cancer screening [3, 4]. For early-stage cervical cancer, the standard treatment is radical hysterectomy with pelvic lymph node dissection. However, locoregional recurrence or distant metastasis is not rare in patients with early-stage cervical cancer, only about 70% patients have a 5-year disease-free survival (DFS) [5]. Prediction of patients' survival can help to determine whether more intensive observation and aggressive treatment regimens should be administered, which might improve clinical outcome.

Previous studies have identified the depth of invasion, lymph node metastasis (LNM) and lymphovascular invasion (LVI) as risk factors for recurrence and metastasis in cervical cancer patients [6, 7]. These risk factors are determined by random sampling biopsy or surgery, which have limitations including procedure-related complications, sampling error, and interobserver variability [8]. In this scenario, non-invasively prognostic biomarkers that allow assessment of tumor heterogeneity are warranted.

Radiomics has emerged as a promising method to evaluate tumor heterogeneity by extracting large set of high-dimensional features from a series of medical images, such as computed tomography (CT), positron emission tomography-computed tomography (PET/CT) and magnetic resonance imaging (MRI) data [9]. It has been conclusively shown that radiomic features could be used to diagnose precisely, evaluate treatment response, and predict survival in various types of cancers [10-15].

MRI is routinely used in clinical workup to diagnose, stage and monitor cervical cancer, with the advantages of lower cost and higher spatial and contrast resolution of pelvic tissues and organs, as well as no radiation [16-18]. Nevertheless, whether the radiomic features extracted from MRI could be used to predict survival in patients with early-stage cervical cancer remains unclear.

Therefore, this study aimed to develop a radiomic score by pre-treatment MRI to estimate 3-year DFS in patients with early-stage cervical cancer, and to further construct a combined model incorporating the radiomic score and the clinicopathological features for the individual prediction of DFS.

Methods

Patients

Institutional ethics review board approval was acquired for this study, and written informed consent was not required for this retrospective study. This study was conducted in agreement with the Declaration Helsinki.

This retrospective study included patients with confirmed early-stage cervical cancer from two tertiary centers in a large metropolitan setting of China (Yunnan Cancer Hospital and Sun Yat-sen University Cancer Center) between January 1, 2011 and December 31, 2017. The patient demographics, laboratory test results, pretreatment MRI imaging data, pathologic results and survival outcome data were reviewed. All the patients met the following inclusion criteria: (a) patients with pathologically confirmed early-stage cervical cancer (Federation of Gynecology and Obstetrics [FIGO] stages IB-IIA); (b) patients underwent contrast-enhanced pelvic MRI scans within two-week period before surgery; and (c) patients underwent radical hysterectomy and bilateral pelvic lymph node dissection. The exclusion criteria of this study were as following: (1) patients treated with neoadjuvant chemotherapy or radiotherapy pre-operatively; (2) patients with diagnosis of other cancers meanwhile; and (3) patients without clinical data including age, neutrophils, lymphocyte, platelet, squamous cell carcinoma antigen (SCCA) and human papillomavirus (HPV), pretreatment T2-weighted (T2w) and contrast-enhanced T1-weighted (CET1w) images. Finally, 248 patients (mean age of 47.77 ± 9.89 years) were included in this study. Supplementary Methods 2 showed the patient selection flowchart from the two centers. Eligible patients' data were randomly divided into a training cohort ($n = 166$) and an independent validation cohort ($n = 82$) at a ratio of 2:1.

Treatments and follow-up

Radical hysterectomy and bilateral pelvic lymph node dissection were conducted for all patients. Adjuvant chemoradiotherapy after radical surgery was administered in 181 patients. Regular follow-up was conducted every 3-6 months during the first two years after operation, 2 times annually for 3-5 years, and then once a year thereafter. The endpoint of our study was DFS, which is defined as the period from the date of surgery to the date of the first locoregional recurrence, distant metastasis, death, or the last visit in follow up. Locoregional recurrences and distant metastasis were confirmed by gynecological examination, imaging modalities such as CT, MRI and positron emission tomography-computed tomography

(PET/CT), or biopsy-proven. The available information was collected from patients' medical records.

MR Image acquisitions

Abdomen and pelvic MRI examinations were conducted at least two weeks after biopsy to avoid the impact of post-biopsy inflammation and within two weeks before surgery. The MRI images were obtained by different MR devices at two institutions. To avoid the possibility of image information loss, we acquired the Digital Imaging and Communications in Medicine (DICOM) images from Picture Archiving and Communication System (PACS) directly without any compression or down sampling. Details regarding the acquisition parameters and MRI retrieval procedure were presented in Supplementary Methods 1.

Radiomic analysis

The radiomic analysis workflow included four steps as illustrated in Figure 1: tumor image segmentation, radiomic feature extraction, feature selection and model building.

Tumor image segmentation

We used open-source ITK-SNAP software (www.itksnap.org) for three-dimensional manual segmentation in axial T2w and sagittal CET1w images. Regions-of-interest (ROIs) were manually delineated by a radiologist who had 8 years of experience in gynecological MR imaging interpretation, and validated by a senior radiologist with 10 years of experience in segmentation results validation.

Radiomic feature extraction

After manually segmenting the tumor ROI, we standardized the T2w and CET1w images by z-score normalization to obtain a standard normal distribution of the image intensities. This procedure was aimed at reducing the image intensity shift caused by different equipments and scanning parameters. Afterward, we resampled CET1w and T2w images into $1 \text{ mm} \times 1 \text{ mm}$ and $0.65 \text{ mm} \times 0.65 \text{ mm}$, respectively. Here, $1 \text{ mm} \times 1 \text{ mm}$ and $0.65 \text{ mm} \times 0.65 \text{ mm}$ were the median voxel spacing of CET1w and T2w images in this dataset. Then we extracted radiomic features from T2w and CET1w images respectively through an open-source software PyRadiomics [19]. Considering the relatively low resolution and large variance between different equipments for axial MR images, we extracted two dimensional (2D) features in axial T2w images and sagittal CET1w images instead of three dimensional (3D) features to increase the robustness of features. Specifically, we extracted 2D features from all the 2D slices of tumor ROI, and averaged the feature values

in all the image slices as the final results for each patient [20-21].

For each slice, we extracted 1299 2D radiomic features including: i) intensity features ($n = 19$): these features were the first-order statistics calculated from the tumor intensities such as entropy, reflecting the signal intensity of different tumors, ii) shape features ($n = 16$): these features represented the size and shape information of tumors, which showed prognostic value in cervical cancer, iii) texture features ($n = 74$): these features measured the relationship between each tumor voxel and its surrounding environments, which can quantify intra-tumor heterogeneity and the use of complex tumor patterns such as size-zone matrix, iv) wavelet features ($n = 736$): we decomposed MR images into low and high frequencies and extracted the features in group i and ii from each frequency range. The wavelet transformation enabled us to quantify high-dimensional multi-frequency tumor information that is difficult to be visually interpreted, and v) Laplacian of Gaussian (LoG) features ($n = 454$): these features were textural features extracted through a Laplacian of Gaussian spatial band-pass filter. These features described the tumor information from multi-scale space that combined both the very detailed and macroscopic tumor texture patterns.

Radiomic feature selection

Although the radiomic features reflected tumor information from various perspectives, not all of them were associated with the DFS in cervical cancer. Consequently, we used a two-step feature selection method to retain only the most strong features that are significantly associated with DFS. First, we evaluated the predictive performance of each radiomic feature by univariate Cox regression in the training cohort. Specifically, we built a univariate Cox proportional hazards model (Cox model) for each radiomic feature in the training cohort, and calculated the P value of the feature in predicting DFS. Afterward, the features with P value less than 0.05 were treated as significant prognostic factors and selected as candidate features. Second, we used regularized multivariate logistic regression with the LASSO penalty for multivariate feature selection. The LASSO regularization involved a parameter λ to control the number of selected features where a larger λ retains more features. To obtain an optimal feature number and avoid over-fitting, we used 5-fold cross-validation in the training cohort to choose the optimal λ . The λ value that maximized the C-index in the training cohort was selected as the optimal regularization parameter, and the feature number was therefore determined automatically by the λ .

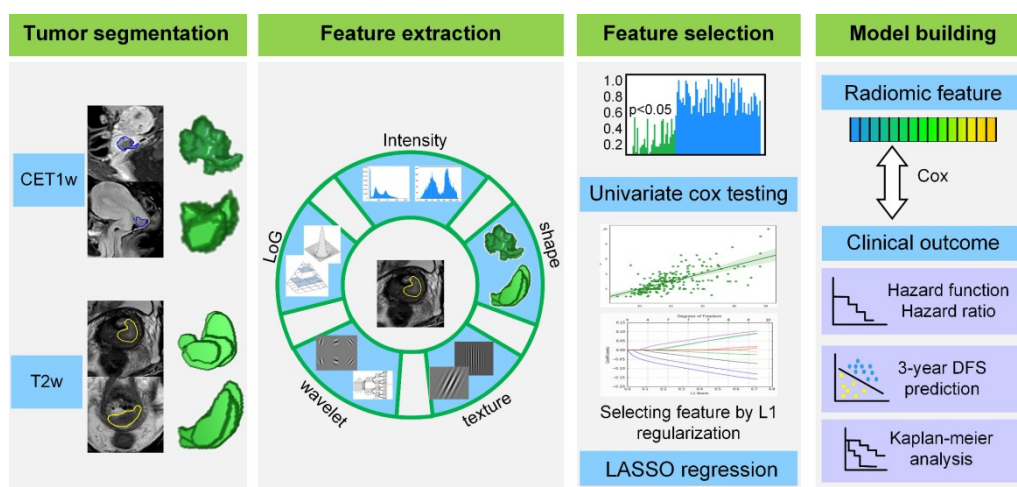


Figure 1. Radiomics framework of predicting the DFS of patients with cervical cancer. DFS: disease-free survival.

Creation of clinicopathological model for DFS prediction

We included 7 candidate clinicopathological findings which showed potential prognostic values to build the clinicopathological model, including age, histological type, differentiation grade, HPV, SCCA level, LNM and LVI status. Among the clinicopathological findings, we firstly used the univariate Cox proportional hazards model to select the significant prognostic factors in the training cohort. Then, significant variables in the univariate Cox model ($P < 0.05$) were included in the multivariate Cox model to build a clinicopathological model for DFS prediction. Detailed calculating process of the Cox model and the Rad-score were provided in Supplementary Methods 3. The performance of the model was evaluated by C-index. The C-index indicates the concordance between the model-predicted DFS and the actual DFS on all the patients, where a C-index around 0.5 means poor predictive value and a C-index around 0.7 indicates good predictive value.

Development and validation of radiomic score and combined model on DFS prediction

In the training cohort, we used a multivariate Cox proportional hazard model to predict a radiomic score (Rad-score) indicating the relative disease progression hazard for each patient. For a given patient, the Cox model used the radiomic feature of this patient to generate a Rad-score larger than zero. A small Rad-score means that the disease progression risk is relatively low, and the DFS for this patient is consequently long; a large Rad-score means high-risk of disease progression, and a relatively short DFS. As described in Supplementary Methods 3, the Cox model used an exponential combination of the

selected radiomic feature to generate the Rad-score.

To further evaluate whether the clinicopathological findings can improve the performance of the Rad-score, we used the Rad-score and the significant clinicopathological features to build a combined multivariable Cox model (combined model) for DFS prediction. The prognostic performance of the Rad-score and the combined model were also evaluated by C-index.

3-year DFS probability prediction of various models

Considering the Cox model is capable of predicting the DFS probability at a given time point, we also used the Cox model to estimate 3-year DFS. Receiver-operating characteristic (ROC) analyses were performed to estimate the prognostic performance of the three models in predicting 3-year DFS.

Statistical Analysis

All the statistical analyses in this study were performed with SPSS 21 and python 2.7. The t test or Mann-Whitney U test of independent samples were conducted to assess the significance of age, neutrophils, lymphocyte, platelet, Histological type, HPV and SCCA level between the training cohort and the validation cohort. The Chi-squared test was conducted to evaluate the significance of the categorical variables such as FIGO stage, differentiation grade, LNM and LVI between the training and validation cohorts. The LASSO-based feature selection and Cox proportional hazards model building were implemented using "scikit-learn" and "lifelines" package. Kaplan-Meier curve was analyzed using the "rms" package and Log-rank test. A two-tailed P -value less than 0.05 was considered statistically significant.

Results

Patient characteristics

A total of 248 patients were included from two cohorts. The patient characteristics were presented in Table 1. The mean age of patients was 47.77 ± 9.89 years. The median follow-up time was 30 months (range, 6-96 months). The results demonstrated that there was no significant difference between the training cohort and the validation cohort ($P = 0.151-0.990$).

DFS prediction performance of the radiomic score

A total of 18 radiomic features were selected for radiomic score building (Supplementary Table S1, Figure S1). In the training cohort, the radiomic score showed good performance on DFS prediction (C-index, 0.786; 95% CI: 0.753–0.820). In the validation cohort, the performance of the radiomic score was further confirmed (C-index, 0.753; 95% CI: 0.696–0.805). The hazard ratio (HR) for radiomic score was 2.259 (95% CI: 2.124–2.394) in the training cohort.

Kaplan-Meier analysis of radiomic score

According to the Rad-score, we further divided patients into high-risk and low-risk groups, and performed Kaplan-Meier analysis to validate the prognostic value of the Rad-score. We used the mean hazard score of the training cohort as a cut-off value to divide patients into high-risk and low-risk groups. As shown in Figure 2, higher Rad-scores were significantly associated with worse DFS in the training cohort and the validation cohort (both with $P < 0.001$). Figure 4 showed two representative patients with distinctly different DFS time (14 months vs 64 months). Although they had almost the same clinicopathological features, their Rad-scores (2.046 vs 0.237) were significantly different.

Performance and validation of the combined model on DFS prediction

Only two clinical features (LNM and LVI) were selected to create a clinicopathological model. This model achieved a poor performance in DFS estimation, with a C-index of 0.711 (95% CI: 0.671-0.753) in the training cohort and 0.632 (95% CI: 0.567-0.700) in the validation cohort. The combined model incorporating the radiomic score and the two clinicopathological features showed improvement in the training cohort (C-index, 0.813; 95% CI: 0.780-0.845), but showed no improvement in the validation cohort (C-index, 0.714; 95% CI: 0.642-0.784) when compared with radiomic score.

Table 1. Comparison of clinical characteristics of patients between the training and validation cohorts

Characteristics	Training cohort (n=166)	Validation cohort (n=82)	p
Age (years, mean \pm SD)	47.69 \pm 9.46	47.93 \pm 10.77	0.864
Neutrophils (10 ⁹ /L, mean \pm SD)	4.77 \pm 2.79	6.23 \pm 9.38	0.151
Lymphocyte (10 ⁹ /L, mean \pm SD)	2.06 \pm 4.13	1.84 \pm 1.05	0.545
Platelet (10 ⁹ /L, mean \pm SD)	264.5 \pm 82.54	270.7 \pm 87.26	0.542
SCC (ng/mL)	4.84 \pm 7.51	5.48 \pm 8.09	0.237
HPV	535.53 \pm 764.31	601.87 \pm 847.68	0.236
Histological type (%)			0.669
Squamous cell carcinoma	141 (84.94%)	68 (82.93%)	
Adenocarcinoma	22 (13.25%)	11 (13.41%)	
Adenosquamous carcinoma	1 (0.60%)	2 (2.44%)	
Small cell carcinoma	2 (1.20%)	1 (1.22%)	
FIGO stage (%)			0.712
IB	110 (66.27%)	57 (69.51%)	
IIA	56 (33.73%)	25 (30.49%)	
Differentiation (%)			0.400
Low grade	100 (60.24%)	46 (56.10%)	
Middle grade	64 (38.55%)	33 (40.24%)	
High grade	2 (1.20%)	3 (3.66%)	
Lymph node metastasis (%)			>0.990
Non-metastasis	131 (78.92%)	64 (78.05%)	
Metastasis	35 (21.08%)	18 (21.95%)	
Lymphovascular space invasion (LVI, %)			0.452
Non-LVI	132 (79.52%)	61 (74.39%)	
LVI	34 (20.48%)	21 (25.61%)	
Mean DFS time (months, mean \pm std)	34.97 \pm 20.44	35.79 \pm 21.38	0.747

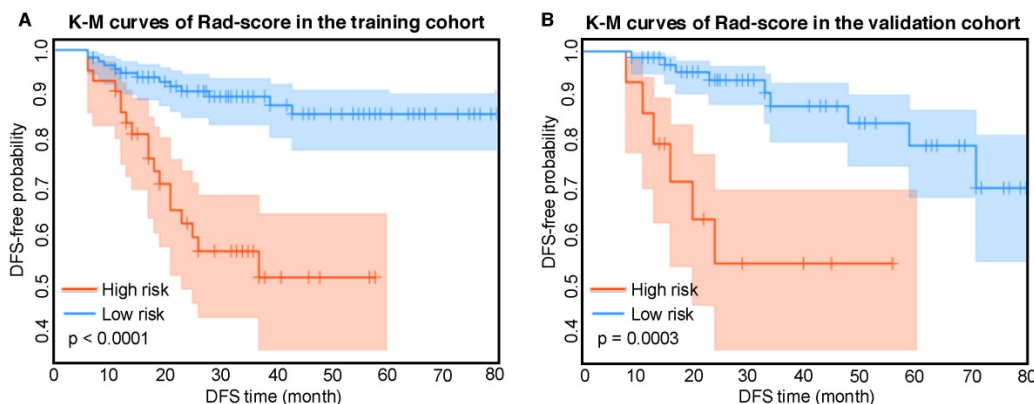


Figure 2. Kaplan-Meier analysis of the Rad-score. A) Kaplan-Meier curves of the Rad-score in the training cohort. Vertical lines indicate censored data. Shadows represent 95% CI. B) Kaplan-Meier curves of the Rad-score in the validation cohort. DFS: disease-free survival. CI: confidence interval.

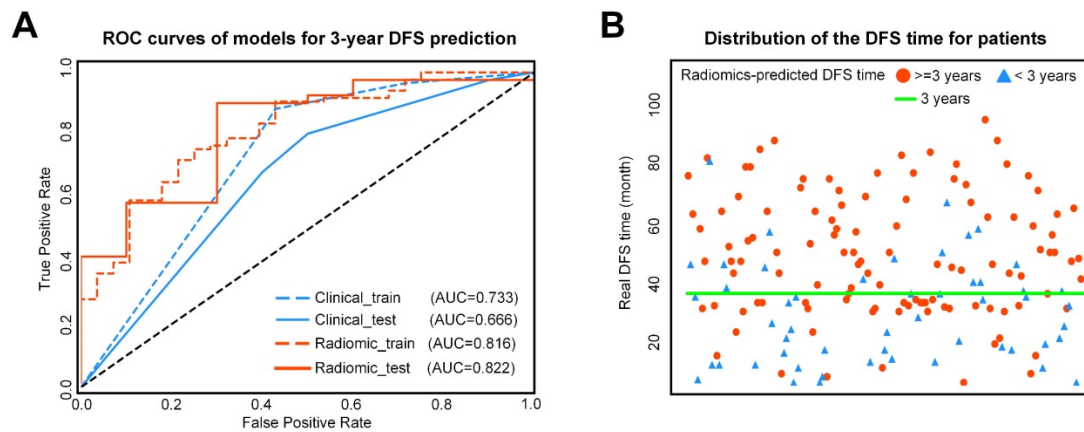


Figure 3. A) ROC curves of the two models for 3-year DFS probability prediction. B) Distribution of the DFS time for patients. The red circle represents patients who are predicted to have DFS time longer than 3 years by the Rad-score, and the blue triangle represents patients who are predicted to have DFS time less than 3 years by the Rad-score. Most patients who are predicted to have DFS timelonger than 3 years distribute above the patients who are predicted to have DFS time less than 3 years. ROC: receiver-operating characteristic; DFS: disease-free survival.

Table 2. Model performance on predicting DFS and 3-year DFS probability

Models	Cohorts	C-Index (95% CI)	AUC (95% CI)	ACC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Clinical model	training	0.711 (0.671, 0.753)	0.733 (0.689, 0.773)	0.754 (0.723, 0.785)	0.802 (0.769, 0.836)	0.607 (0.530, 0.678)
	validation	0.632 (0.567, 0.700)	0.666 (0.595, 0.742)	0.745 (0.700, 0.796)	0.805 (0.760, 0.857)	0.500 (0.372, 0.635)
Radiomic score	training	0.786 (0.753, 0.820)	0.816 (0.779, 0.854)	0.746 (0.713, 0.780)	0.756 (0.718, 0.794)	0.714 (0.649, 0.784)
	validation	0.753 (0.696, 0.805)	0.822 (0.765, 0.882)	0.765 (0.718, 0.813)	0.780 (0.729, 0.833)	0.700 (0.583, 0.817)
Combined model	training	0.813 (0.780, 0.845)	0.849 (0.816, 0.880)	0.754 (0.722, 0.786)	0.756 (0.719, 0.795)	0.750 (0.679, 0.811)
	validation	0.714 (0.642, 0.784)	0.759 (0.678, 0.843)	0.765 (0.717, 0.817)	0.780 (0.730, 0.835)	0.700 (0.581, 0.822)

Note: CI represents confidence interval. C-Index represents Harrell's concordance index, which measures the performance of the DFS prediction. AUC represents area under the receiver operating characteristic curve, and ACC is accuracy. AUC and ACC evaluate the performance of the 3-year DFS prediction.

3-year DFS probability prediction of various models

For 3-year DFS probability prediction, the clinicopathological model achieved an area under the receiver operating characteristic curve (AUC) of 0.666 (95% CI: 0.595-0.742), sensitivity of 0.805 (95% CI: 0.760-0.857), specificity of 0.500 (95% CI: 0.372, 0.635), and accuracy of 0.745 (95% CI: 0.700-0.796) (Table 2, Figure 3). The radiomic score yielded an AUC of 0.822 (95% CI: 0.765-0.882), sensitivity of 0.780 (95% CI: 0.729-0.833), specificity of 0.700 (95% CI: 0.583-0.817), and accuracy of 0.765 (95% CI: 0.718-0.813) (Table 2, Figure 3). The Rad-score showed significant difference between patients with DFS time >3 years and <3 years ($P < 0.0001$ in the training cohort; $P = 0.0010$ in the validation cohort). In Supplementary Table S2, we also provided the performance of the Rad-score on predicting DFS at multiple time points.

The combined model yielded an AUC of 0.759 (95% CI: 0.678-0.843), sensitivity of 0.780 (95% CI: 0.730-0.835), specificity of 0.700 (95% CI: 0.581-0.822), and accuracy of 0.765 (95% CI: 0.717-0.817) (Table 2, Figure 3). Therefore, the combined radiomic model showed no performance improvement in 3-year DFS estimation when compared with the Rad-score.

Discussion

In our study, we evaluated the prognostic value of MR-derived radiomic features on patients with IB-IIA cervical cancer treated by radical hysterectomy. The results showed that LASSO-Cox based radiomic score had favorable predictive performance in DFS estimation. Our study would help to determine whether more intensive observation and aggressive treatment regimens should be administered in patients with worse DFS, with the aim of assisting clinical treatment and healthcare decisions.

Radiomics provided underlying diagnostic, therapeutic, and prognostic information by noninvasively extracting useful imaging features from medical images [22]. Some previous studies applied PET radiomics to predict survival of cervical cancer, which was more accurate than conventional clinical factors [23-25]. However, these studies had smaller sample size and did not focus on early-stage cervical cancer treated by surgery. Their findings were of limited clinical relevance because the treatment patients received was a strong indicator for survival. However, Brooks FJ et al. showed that radiomic features based on PET could not provide additional information for small tumor lesions due to the limited spatial resolution of PET imaging with the large voxel

size sampling [26]. The value of MRI radiomic model has been proven in predicting the LNM and LVI status in patients with cervical cancer preoperatively [27, 28]. Giving this background, radiomic features based on MRI was available to predict the survival of early-stage cervical cancer patients.

The radiomic score that combined the T2w and CET1w images yielded a C-index of 0.753 (95% CI: 0.696-0.805) on DFS prediction and an AUC of 0.822 (95% CI: 0.765-0.882) on 3-year DFS prediction in the validation cohort, which were higher than either clinical model or the combined model. This indicated that the radiomic score may already contain information in clinical factors, and can mine more prognostic information than clinical factors by observing the whole tumor scope and extracting high-dimensional features (e.g., wavelet and LoG features). Thus, it could be used as a surrogate biomarker to improve the prognostic ability pretreatment. Meanwhile, the Rad-score could stratify patients into high-risk and low-risk groups. Patients with higher Rad-scores have the worse DFS, which suggested that some low-risk patients would have received unnecessary radical hysterectomy treatment; whereas for patients with high risk of recurrence and metastasis, a systemic adjuvant treatment would be more beneficial. Our findings would initiate a pivotal step to enable surgeons to tailor treatment basing on the specific clinical and radiomic features for high-risk and low-risk patients with early-stage cervical cancer.

Of the 1299 radiomic features, 18 were identified to be predictive for DFS, which includes 10 features derived from CET1w and 8 features extracted from T2w images. This may indicate that CET1w images

probably contains more prognostic information than T2w images. Importantly, shape flatness was included in the 10 CET1w- derived features. This feature characterized the shape of tumor: small flatness value indicates an irregular tumor shape. In the radiomic score, a tumor with small flatness generates a poor prognostic outcome, which is in line with a prior study that tumor sphericity is a poor prognostic marker for breast cancer [11]. In addition, eight wavelet features and one LoG feature were selected from the CET1w images. The wavelet features reflected tumor information in eight space domains, and the LoG features reflected tumor information from three frequency domains. This result demonstrated that the raw CET1w images may include limited prognostic features; however, through a wavelet and LoG transform of the raw CET1w images, much prognostic information can be mined. This further reflected the advantages of radiomics method since it is good at mine high-dimensional information that is difficult to be sensed manually. For example, the “SumEntropy” in a wavelet subspace and the “Skewness” in a LoG subspace were selected, indicating that in a high-dimensional wavelet and LoG spaces, the tumor heterogeneity described by entropy and the intensity of tumor have prognostic value. Similarly, one intensity feature, two texture features, one wavelet feature and four LoG features were selected from the T2w images.

In a previous study, pretreatment HPV genotype was reported as a prognostic biomarker in cervical cancer [29, 30]. So far, the correlation between the HPV level and survival time in cervical cancer patients is seldom studied and remains unknown. The

result of our data showed that there is no significant correlation between them. The potential reason for the unconformity result is that the HPV test involves all of the risk HPV genotypes, with no specific categories, whereas some high-risk HPV genotypes are not associated with the prognosis of cervical cancer [31]. Our study demonstrated that LVI and LNM were identified as significant independent prognostic factors for DFS, which is consistent with previous studies [32-34]. The mean DFS is 24.75 months for patients with LVI, and 38.56 months for patients without LVI ($P < 0.001$). Similarly, the mean DFS is 29.91 months for patients with LNM, and 37.50 months for patients without LNM ($P = 0.009$). In this study, the clinical stage

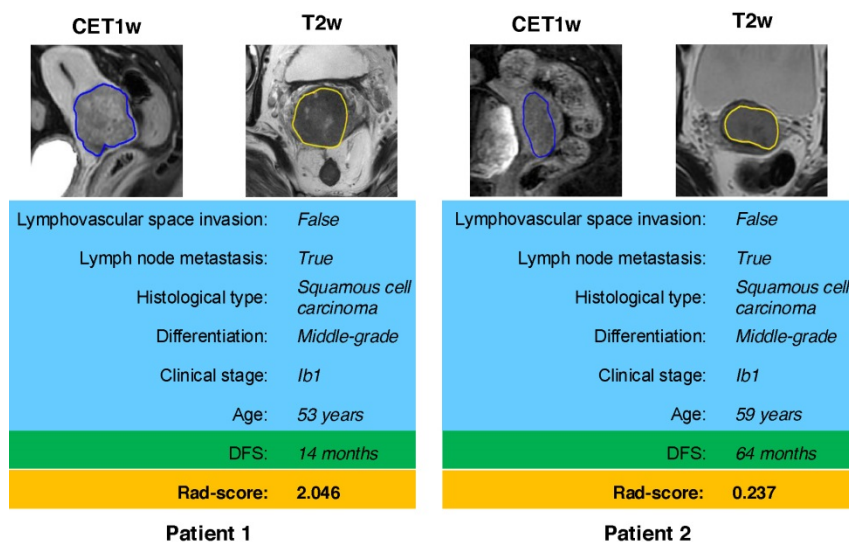


Figure 4. MR images of two patients with significantly different DFS time. Although patient 1 and patient 2 have similar clinicopathological characteristics, their Rad-score are different and discriminative. DFS: disease-free survival.

(IB-IIA) was not associated with DFS; the possible reason is that the survival of patients in early-stage is relatively good.

Despite the favorable results of the radiomic score, our study also has some limitations. First, although the number of patients in this study was large as compared with previous studies, larger and prospective datasets would be needed to optimize the performance of the model in the future. In addition, since the DFS of patients in the two institutions was different; therefore we combined the data from two institutions when dividing the training and validation cohorts. In future work, we hope to perform external validation in a different center. Second, because of the multicenter scanner settings, the MR images may be affected by different scanners and protocols. Therefore, we combined images from multiple scanners and tried to eliminate the device-effect by image standardization and robust 2D feature extraction. However, a more comprehensive method to balance the scanner-variance worth future exploring. Third, we mined 18 prognostic features from MR image and compared their performance with clinical factors. However, the association between the radiomic feature and biological level events was not explained. In the future, we will explain the radiomic features at biological level by combining gene profile. In addition, deep learning as an emerging method in medical image analysis may provide valuable features that are complementary to radiomic features[35]. Due to limited training samples, directly using deep learning in this study can probably cause overfitting; however, combining deep learning features with the selected radiomic features worth future exploring. For example, we can use transfer learning to extract deep learning features without the need for large training data, and then combine the deep learning features with the selected radiomic features. This strategy can combine the advantages of both deep learning and hand-crafted radiomic features.

In conclusion, this study provides a noninvasive and pretreatment prognostic biomarker for the DFS of cervical cancer based on MRI. Moreover, for each cervical cancer patient, the radiomic score can predict the hazard risk of the patient being disease-free and stratify the patient into high-risk and low-risk groups. The study may offer some important insights into precise treatment, providing valuable guidance for clinical physicians regarding the treatment strategies including radical hysterectomy or chemoradiation on early-stage cervical cancer patients.

Abbreviations

MRI: Magnetic Resonance Imaging; DFS:

disease-free survival; LASSO: Least absolute shrinkage and selection operator; LNM: lymph node metastasis; LVI: lymph-vascular invasion; CT: computed tomography; PET/CT: positron emission tomography-computed tomography; T2w: T2-weighted; CET1w: contrast-enhanced T1-weighted; DICOM: Digital Imaging and Communications in Medicine; PACS: Picture Archiving and Communication System; ROI: Regions-of-interest; LoG: Laplacian of Gaussian; HPV: human papillomavirus; SCCA: squamous cell carcinoma antigen; ROC: Receiver-operating characteristic; AUC: Area under the receiver operating characteristic curve.

Supplementary Material

Supplementary figures and tables.

<http://www.thno.org/v10p2284s1.pdf>

Acknowledgements

We acknowledge financial support from the National Natural Science Foundation of China (81571664, 81871323, 81801665, 81772012, 81922040); the National Natural Science Foundation of Guangdong Province (2018B030311024); the Scientific Research General Project of Guangzhou Science Technology and Innovation Commission (201707010328); the China Postdoctoral Science Foundation (2016M600145); the Beijing Natural Science Foundation under Grant No. 7182109; the National Key Research and Development Plan of China under Grant Nos. 2017YFA0205200; and the Youth Innovation Promotion Association CAS (grant number 2019136).

Authors' contribution

Conception and design: Shuixing Zhang, Jie Tian, Zhenyu Liu, Lizhi Liu, Yingying Ding. Collection and assembly of data: Jin Fang, Yan Jin, Fei Wang, Qiuying Chen, Liting Chen, Yueyue Li, Zhuozhi Chen, Minmin Li. Data analysis and interpretation: Jie Tian, Zhenyu Liu, Shuo Wang, Jin Fang, Bin Zhang. Manuscript writing: Jin Fang, Bin Zhang, Shuo Wang, Shuixing Zhang, Jie Tian, Zhenyu Liu, Lizhi Liu. Final approval of manuscript: All authors.

Competing Interests

The authors have declared that no competing interest exists.

References

1. Banerjee S. Bevacizumab in cervical cancer: a step forward for survival. *Lancet*. 2017; 390: 1626-8.
2. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2018; 68: 394-424.

3. Quinn MA, Benedet JL, Odicino F, Maisonneuve P, Beller U, Creasman WT, et al. Carcinoma of the cervix uteri. FIGO 26th Annual Report on the Results of Treatment in Gynecological Cancer. *Int J Gynaecol Obstet.* 2006; 95 (Suppl 1): S43-103.
4. Gregg S, Scaffa C. Surgical management of early cervical cancer: the shape of future studies. *Curr Oncol Rep.* 2012; 14: 527-34.
5. Moore KN, Java JJ, Slaughter KN, Rose PG, Lanciano R, DiSilvestro PA, et al. Is age a prognostic biomarker for survival among women with locally advanced cervical cancer treated with chemoradiation? An NRG Oncology/Gynecologic Oncology Group ancillary data analysis. *Gynecol Oncol.* 2016; 143: 294-301.
6. Creasman WT, Kohler MF. Is lymph vascular space involvement an independent prognostic factor in early cervical cancer? *Gynecol Oncol.* 2004; 92: 525-9.
7. Skret-Magierlo JE, Soja PJ, Skret A, Kruczek A, Kaznowska E, Wicherek L. Perineural space invasion in cervical cancer (FIGO IB1-IIb) accompanied by high-risk factors for recurrence. *J Cancer Res Ther.* 2014; 10: 957-61.
8. Choe J, Lee SM, Do KH, Lee JB, Lee SM, Lee JG, et al. Prognostic value of radiomic analysis of iodine overlay maps from dual-energy computed tomography in patients with resectable lung cancer. *Eur Radiol.* 2019; 29: 915-23.
9. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer.* 2012; 48: 441-6.
10. Zhang B, Tian J, Dong D, Gu D, Dong Y, Zhang L, et al. Radiomics Features of Multiparametric MRI as Novel Prognostic Factors in Advanced Nasopharyngeal Carcinoma. *Clin Cancer Res.* 2017; 23: 4259-69.
11. Park H, Lim Y, Ko ES, Cho HH, Lee JE, Han BK, et al. Radiomics Signature on Magnetic Resonance Imaging: Association with Disease-Free Survival in Patients with Invasive Breast Cancer. *Clin Cancer Res.* 2018; 24: 4705-14.
12. Sun C, Tian X, Liu Z, Li W, Li P, Chen J, et al. Radiomic Analysis for Pretreatment Prediction of Response to Neoadjuvant Chemotherapy in Locally Advanced Cervical Cancer: A Multicentre Study. *Ebiomedicine.* 2019; 46:160-9.
13. Liu Z, Wang S, Dong D, Wei J, Fang C, Zhou X, et al. The Applications of Radiomics in Precision Diagnosis and Treatment of Oncology: Opportunities and Challenges. *Theranostics.* 2019; 9: 1303-22.
14. Liu Z, Li Z, Qu J, Zhang R, Zhou X, Li L, et al. Radiomics of multi-parametric MRI for pretreatment prediction of pathological complete response to neoadjuvant chemotherapy in breast cancer: a multicenter study. *Clin Cancer Res.* 2019; 25: 3538-47.
15. Jiang Y, Yuan Q, Lv W, Xi S, Huang W, Sun Z, et al. Radiomic signature of 18F fluorodeoxyglucose PET/CT for prediction of gastric cancer survival and chemotherapeutic benefits. *Theranostics.* 2018;8(21):5915-28.
16. Kircher M, Ihne S, Brumberg J, Morbach C, Knop S, Kortüm KM, et al. Detection of cardiac amyloidosis with (18)F-Florbetaben-PET/CT in comparison to echocardiography, cardiac MRI and DPD-scintigraphy. *Eur J Nucl Med Mol Imaging.* 2019; 46: 1407-16.
17. Rauch GM, Kaur H, Choi H, Ernst RD, Klopp AH, Boonsirikamchai P, et al. Optimization of MR imaging for pretreatment evaluation of patients with endometrial and cervical cancer. *Radiographics.* 2014; 34:1082-98.
18. Park JJ, Kim CK, Park SY, Park BK. Parametrial invasion in cervical cancer: fused T2-weighted imaging and high-b-value diffusion-weighted imaging with background body signal suppression at 3 T. *Radiology.* 2015; 274:734-41.
19. Van Griethuysen JJM, Fedorov A, Parmar C, Hosny A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res.* 2017; 77: e104-7.
20. Shen C, Liu Z, Guan M, Song J, Lian Y, Wang S, et al. 2D and 3D CT Radiomics Features Prognostic Performance Comparison in Non-Small Cell Lung Cancer. *Transl Oncol.* 2017; 10: 886-94.
21. Liu Z, Zhang XY, Shi YJ, Wang L, Zhu HT, Tang Z, et al. Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer. *Clin Cancer Res.* 2017; 23: 7253-62.
22. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology.* 2016; 278: 563-77.
23. 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.
24. Chung HH, Kang SY, Ha S, Kim JW, Park NH, Song YS, et al. Prognostic value of preoperative intratumoral FDG uptake heterogeneity in early stage uterine cervical cancer. *J Gynecol Oncol.* 2016; 27: e15.
25. Ho KC, Fang YH, Chung HW, Yen TC, Ho TY, Chou HH, et al. A preliminary investigation into textural features of intratumoral metabolic heterogeneity in (18)F-FDG PET for overall survival prognosis in patients with bulky cervical cancer treated with definitive concurrent chemoradiotherapy. *Am J Nucl Med Mol Imaging.* 2016; 6: 166-75.
26. Brooks FJ, Grigsby PW. The effect of small tumor volumes on studies of intratumoral heterogeneity of tracer uptake. *J Nucl Med.* 2014; 55: 37-42.
27. Li Z, Li H, Wang S, Dong D, Yin F, Chen A, et al. MR-Based Radiomics Nomogram of Cervical Cancer in Prediction of the Lymph-Vascular Space Invasion preoperatively. *J Magn Reson Imaging.* 2018; 49: 1420-6.
28. Wu Q, Wang S, Chen X, Wang Y, Li D, Liu Z, et al. Radiomics analysis of magnetic resonance imaging improves diagnostic performance of lymph node metastasis in patients with cervical cancer. *Radiother Oncol.* 2019; 138: 141-8.
29. Cuschieri K1, Brewster DH, Graham C, Nicoll S, Williams AR, Murray GI, et al. Influence of HPV type on prognosis in patients diagnosed with invasive cervical cancer. *Int J Cancer.* 2014; 135: 2721-6.
30. Chong GO, Lee YH, Han HS, Lee HJ, Park JY, Hong DG, et al. Prognostic value of pre-treatment human papilloma virus DNA status in cervical cancer. *Gynecol Oncol.* 2018; 148: 97-102.
31. Hang D, Jia M, Ma H, Zhou J, Feng X, Lyu Z, et al. Independent prognostic role of human papillomavirus genotype in cervical cancer. *BMC Infect Dis.* 2017; 17: 391.
32. Nanthamongkolkul K, Hanprasertpong J. Longer waiting times for early stage cervical cancer patients undergoing radical hysterectomy are associated with diminished long-term overall survival. *J Gynecol Oncol.* 2015; 26: 262-9.
33. Ho CM, Chien TY, Huang SH, Wu CJ, Shih BY, Chang SC. Multivariate analysis of the prognostic factors and outcomes in early cervical cancer patients undergoing radical hysterectomy. *Gynecol Oncol.* 2004; 93: 458-64.
34. Raspagliesi F, Ditto A, Quattrone P, Solima E, Fontanelli R, Dousias V, et al. Prognostic factors in microinvasive cervical squamous cell cancer: long-term results. *Int J Gynecol Cancer.* 2005; 15: 88-93.
35. Li H, Xu C, Xin B, Zheng C, Zhao Y, Hao K, et al. 18F-FDG PET/CT Radiomic Analysis with Machine Learning for Identifying Bone Marrow Involvement in the Patients with Suspected Relapsed Acute Leukemia. *Theranostics.* 2019;9(16):4730-9.