Radiomics Analysis on T2-MR Image to Predict Lymphovascular Space Invasion in Cervical Cancer

Shuo Wang[1, 4†], Xi Chen[3†], Zhenyu Liu[1†], Qingxia Wu[2], Yongbei Zhu[1], Meiyun Wang*[2], Jie Tian*[1, 4]

1 CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China

2 Department of Radiology, Henan Provincial People's Hospital, Henan, 450003, China

3 School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China

4 University of Chinese Academy of Sciences, Beijing, 100080, China

ABSTRACT

Lymphovascular space invasion (LVSI) is an important determinant for selecting treatment plan in cervical cancer (CC). For CC patients without LVSI, conization is recommended; otherwise, if LVSI is observed, hysterectomy and pelvic lymph node dissection are required. Despite the importance, current identification of LVSI can only be obtained by pathological examination through invasive biopsy or after surgery. In this study, we provided a non-invasive and preoperative method to identify LVSI by radiomics analysis on T2-magnetic resonance image (MRI), aiming at assisting personalized treatment planning. We enrolled 120 CC patients with T2 image and clinical information, and allocated them into a training set (n = 80) and a testing set (n= 40) according to the diagnostic time. Afterwards, 839 image features were extracted to reflect the intensity, shape, and high-dimensional texture information of CC. Among the 839 radiomic features, 3 features were identified to be discriminative by Least absolute shrinkage and selection operator (Lasso)-Logistic regression. Finally, we built a support vector machine (SVM) to predict LVSI status by the 3 radiomic features. In the independent testing set, the radiomics model achieved area under the receiver operating characteristic curve (AUC) of 0.7356, classification accuracy of 0.7287. The radiomics signature showed significant difference between non-LVSI and LVSI patients ($p<0.05$). Furthermore, we compared the radiomics model with clinical model that uses clinical information, and the radiomics model showed significant improvement than clinical factors (AUC=0.5967 in the validation cohort for clinical model).

Keywords: T2 weighted magnetic resonance image, radiomics, lymphovascular space invasion, cervical cancer, machine learning.

---

† Contribute equally: Shuo Wang, Xi Chen and Zhenyu Liu.

*Co-corresponding authors: Jie Tian (tian@ieee.org;), Meiyun Wang (marian9999@163.com).

1. INTRODUCTION

Cervical cancer is the second leading cause of cancer death in women [1]. Lymphovasucular space invasion (LVSI) has been proved to be an important independent prognostic factor [2] and an important

determinant for treatment planning in cervical cancer (CC). The NCCN (National Comprehensive Cancer Network, NCCN) guidelines for CC consider LVSI as an important determinant in selecting appropriate treatment plan. Early cervical cancer with or without LVSI has completely different therapeutic plan. For cervical cancer patients with FIGO IA1 stage, conization is recommended for tumor without LVSI, however, patients with LVSI need radical hysterectomy and pelvic lymph node dissection, even with stage of IA1. In addition, according to SEDLIS criteria, pathologically proved LVSI positive is the main indication for external pelvic radiotherapy after radical hysterectomy [3]. Despite the importance of detecting LVSI status, current diagnosis of LVSI can only be obtained by pathological examination through invasive biopsy or after surgery.

Benefiting from the ability to noninvasively visualize a cancer's appearance on a macroscopic level [4], medical imaging demonstrated strong value in tumor analysis such as tumor volume measurement [5, 6] and predicting genetic mutation status and prognostic risk [7-9]. Magnetic resonance image (MRI), as a routinely used medical imaging modality, contains many mineable features associated with outcomes of cancer [10] such as in reflecting the treatment response of CC [11]. Quantitative analysis methods using medical images for clinical outcome prediction is defined as "radiomics" analysis [4]. Instead of analyzing medical images through radiologists' experience, radiomics analysis provides unified and quantized measurement of image information to avoid inter-observer variance. In addition, radiomics methods can mine thousands of image features from visual-level characteristics to high-dimensional non-linear features that cannot be described visually. Most importantly, through the automatic feature selection and model building procedures, radiomics model selects and combines the most important image features that are associated with specific clinical outcomes.

Since very few methods can be used to non-invasively predict LVSI status before operation, we provide a non-invasive and preoperative method to identify LVSI status by radiomics analysis on T2-weightd MRI, aiming at assisting personalized treatment planning.

2. METHODS

2.1. Patients and data acquisition

With the institutional review board approval, we collected 120 patients from Henan Provincial People's Hospital between October 2015 to December 2017. The patients were allocated into a training set and a testing set according to the diagnostic time at a ratio of 2:1. Therefore, the first 80 patients comprised the training set while the remaining 40 patients comprised the testing set.

T2-weightd MRI (T2-MRI) and clinical characteristics including age, menstrual status, the stage of International Federation of Gynecology and Obstetrics (FIGO stage) and tumor size were collected for all the patients. The LVSI status of patients were confirmed by histological examination. Parameters for the T2-MRI were: TR = 2800 ms, TE = 70 ms, FOV = 340 * 340 mm, Matrix = 256*256, Thickness = 4.0 mm, and Gap = 1.0 mm.

Table1 listed all the clinical characteristics of patients from the training and testing sets. This table indicates that all the clinical characteristics do not have significant difference between the training set and the testing set. Here, we used independent samples t test to assess the difference of age between the two sets; and used chi-square test to evaluate the difference of menstrual status, FIGO stage, tumor size and LVSI status between the training set and the testing set.

Table.1. Clinical characteristics of the patients

Characteristics	Training Set	Validation Set	*P*-value

Age (mean, range, years)	49.20 (29 , 67)	50.45 (32 , 75)	0.214
Menstrual status			0.602
Premenopausal	44	24	
Postmenopausal	36	16	
FIGO Stage			0.161
IB	37	25	
IIA	32	13	
IIB	11	2	
Tumor size			0.354
≤4cm	64	29	
＞4cm	16	11	
LVSI status			0.108
LVSI	25	7	
Non-LVSI	55	33	

Note: All the clinical characteristics do not have significant difference between the training set and the testing set.

2.2. Radiomics analysis

We used a radiomics analysis method on T2-MR image to predict LVSI status non-invasively, which includes four steps: 1) tumor segmentation, 2) feature extraction, 3) feature reduction and 4) model building. The following figure 1 illustrated the radiomics analysis process for LVSI status prediction.

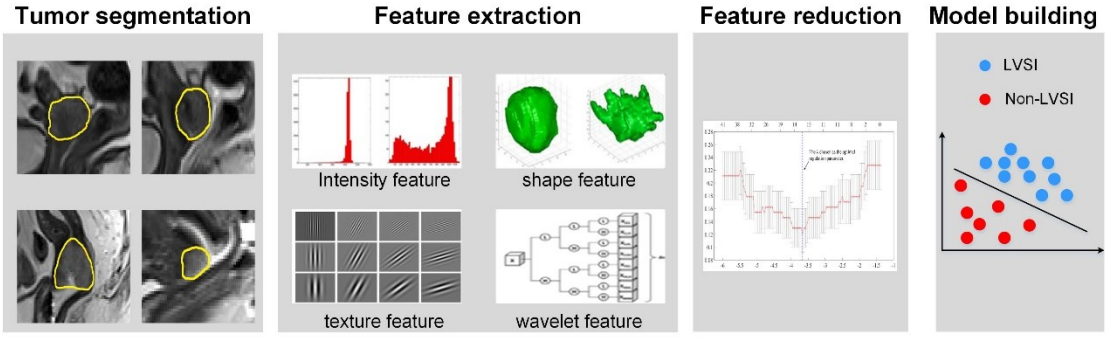

Fig.1. The radiomics analysis process for LVSI prediction in cervical cancer.

Since we predict the LVSI status according to tumor information, tumor segmentation is necessary for further feature extraction. We manually segmented tumor boundary by ITK-SNAP toolkit (www.itksnap.org) on each axial slice of the T2-MRI for each patient. The following analysis was performed only inside the tumor area (defined as region of interest, ROI). Within the tumor ROIs, we extracted 839 image features including intensity features, shape features, texture features and wavelet features by PyRadiomics toolkit [12]. The intensity features, shape features and texture features were extracted in the original T2-MRI; and the wavelet features were extracted from the wavelet-decomposed images on three directions of the original T2-MRI. Afterwards, we used logistic regression with L1 regularization for feature reduction using the following loss function $\mathrm{L}(w)$.

$$\mathrm{L}(w) = \frac{1}{N}\sum_{n=1}^{N}[y_n log p_n + (1-y_n) log(1-y_n)] + \lambda|w|$$

In this equation, p_n is the output of the logistic regression, and y_n is the LVSI status (1 for LVSI, 0 for non-LVSI). $|w|$ is the L1 regularization term. Benefiting from the L1-regularization, the logistic

regression selects features by shrinking the coefficients of trivial features towards zero. The regulation parameter λ determines the amount of reserved features: a larger λ would reserve less features. In this study, the regulation parameter λ (log-scale) was chosen with grid-mesh method, where the minimal λ was set as 0.1; and the number of λ was set as 100. The λ achieving the minimal misclassification error in the training set with five-fold cross-validation was chosen as the optimal regulation parameters.

Based on the selected features, we used a support vector machine (SVM) model to build the associations between the images features and LVSI status. The SVM model used radial basis function as kernel function and predicts a probability of the tumor being LVSI. To quantitatively evaluate the performance of the radiomics model, we trained the model in the training set and validated its performance in the independent testing set. The radiomics model was assessed by the area under the Receiver Operating Characteristic (ROC) curve (AUC), classification accuracy, sensitivity and specificity.

3. RESULTS

After feature reduction, 3 discriminative features were selected from the 839 features. These three features are all derived from wavelet-decomposed T2-DWI images, including: 1) the energy of the wavelet-transformed image in HHL direction, 2) the high gray level emphasis of the wavelet-transformed image in HLH direction, and 3) the 90 percentile of the intensity of wavelet-transformed image in LLL direction.

Based on the selected features, the radiomics model achieved AUC=0.8094, accuracy=0.7108 in the training set, and AUC=0.7356, accuracy=0.7287 in the independent testing set. Table 1 listed the detailed results of the radiomics model. Moreover, figure 2a plotted the ROC curves of the radiomics model in the training and testing sets. In addition, we illustrated the distribution of the radiomics-predicted score in figure 2b, which indicates that the radiomics score has significant difference between LVSI patients and non-LVSI patients ($p < 0.05$). More importantly, we compared the radiomics model with clinical model that utilizes clinical characteristics in table 1. The performance of the clinical model is only AUC=0.6326 in the training set and AUC=0.5967 in the testing set. Therefore, the radiomics model demonstrated much improvement than clinical characteristics, and provides a non-invasive method for LVSI status prediction.

Table 2 Performance of the radiomics model

Evaluation Metrics	Training Set	Testing Set
AUC (95% CI)	0.8094 (0.7761, 0.8427)	0.7356 (0.6790, 0.7920)
Accuracy (95% CI)	0.7108 (0.6740, 0.7475)	0.7287 (0.6780, 0.7793)
Sensitivity (95% CI)	0.7493 (0.7007, 0.7978)	0.8336 (0.7696, 0.8975)
Specificity (95% CI)	0.6633 (0.6067, 0.7198)	0.6431 (0.5706, 0.7154)

Note: AUC is area under the receiver operating characteristic curve; CI is confidential interval.

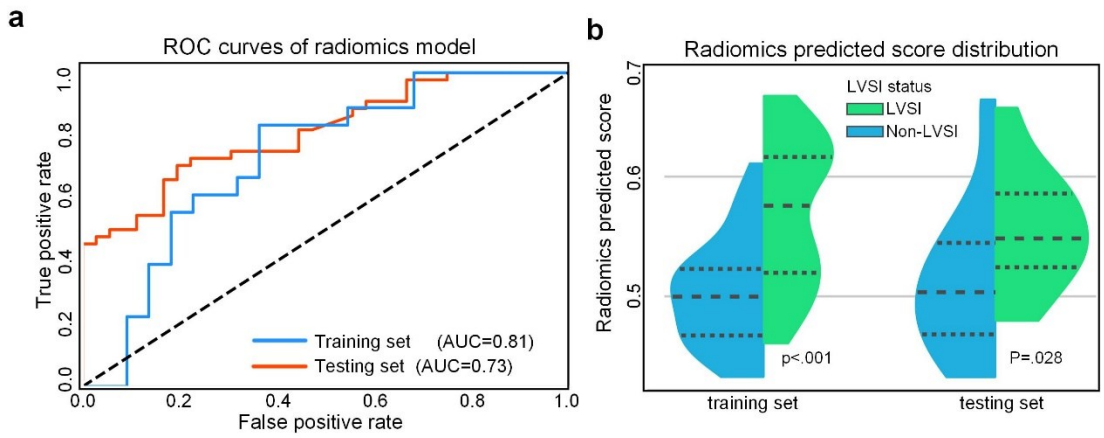

Fig.2. a) The ROC curves of the radiomics model: b) Distribution of the radiomics predicted score.

4. CONCLUSIONS

In this study, we proposed a radiomics analysis method on T2-MR image to predict LVSI status in patients with cervical cancer. The identification of LVSI status is important for cervical cancer since it affects the following treatment planning. However, the LVSI status can only be identified by invasive histological examination currently. Therefore, we provided a non-invasive model to predict the LVSI status by T2-MR image that is routinely used for cervical cancer diagnosis. Furthermore, the radiomics model showed much improvement than using clinical characteristics for prediction (AUC=0.7356 for radiomics model and AUC=0.5967 for clinical model in the testing set). Therefore, the radiomics analysis of T2-MR image data can provide a non-invasive method for LVSI status prediction in cervical cancer, which can further assist treatment planning.

REFERENCES

[1] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA: A Cancer Journal for Clinicians,* vol. 68, pp. 7-30, 2018.

[2] D. Herr, J. Konig, V, K. Koretz, R. Kreienberg, and C. Kurzeder, "Prognostic impact of satellite-lymphovascular space involvement in early-stage cervical cancer," *Annals of Surgical Oncology,* vol. 16, pp. 128-132, 2009.

[3] A. Sedlis, ., B. N. Bundy, M. Z. Rotman, S. S. Lentz, L. I. Muderspach, and R. J. Zaino, "A randomized trial of pelvic radiation therapy versus no further therapy in selected patients with stage IB carcinoma of the cervix after radical hysterectomy and pelvic lymphadenectomy: A Gynecologic Oncology Group Study," *Gynecologic Oncology,* vol. 73, p. 177, 1999.

[4] P. Lambin, R. T. Leijenaar, T. M. Deist, J. Peerlings, E. E. de Jong, J. van Timmeren*, et al.*, "Radiomics: the bridge between medical imaging and personalized medicine," *Nature Reviews Clinical Oncology,* vol. 14, p. 749, 2017.

[5] S. Wang, M. Zhou, Z. Liu, Z. Liu, D. Gu, Y. Zang*, et al.*, "Central focused convolutional neural networks: Developing a data-driven model for lung nodule segmentation," *Medical Image Analysis,* vol. 40, pp. 172-183, 2017.

[6] S. Wang, M. Zhou, O. Gevaert, Z. Tang, D. Dong, Z. Liu*, et al.*, "A multi-view deep convolutional neural networks for lung nodule segmentation," in *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*, 2017, pp. 1752-1755.

[7] S. Wang, Z. Liu, Y. Rong, B. Zhou, Y. Bai, W. Wei*, et al.*, "Deep learning provides a new computed tomography-based prognostic biomarker for recurrence prediction in high-grade serous ovarian cancer," *Radiotherapy and Oncology,* 2018.

[8] S. Wang, J. Shi, Z. Ye, D. Dong, D. Yu, M. Zhou*, et al.*, "Predicting EGFR Mutation Status in Lung Adenocarcinoma on CT Image Using Deep Learning," *European Respiratory Journal,* p. 1800986, 2019.

[9] S. Wang, Z. Liu, X. Chen, Y. Zhu, H. Zhou, Z. Tang*, et al.*, "Unsupervised Deep Learning Features for Lung Cancer Overall Survival Analysis," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018, pp. 2583-2586.

[10] Z. Liu, X.-Y. Zhang, Y.-J. Shi, L. Wang, H.-T. Zhu, Z.-C. Tang*, et al.*, "Radiomics analysis for evaluation of pathological complete response to neoadjuvant chemoradiotherapy in locally advanced rectal cancer," *Clinical Cancer Research,* p. clincanres. 1038.2017, 2017.

[11] S. R. Bowen, Y. Wtc, D. S. Hippe, W. Wu, S. C. Partridge, S. Elias*, et al.*, "Tumor radiomic heterogeneity: Multiparametric functional imaging to characterize variability and predict response following cervical cancer radiation therapy," *Journal of Magnetic Resonance Imaging,* vol. 47, 2017.

[12] J. J. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan*, et al.*, "Computational radiomics system to decode the radiographic phenotype," *Cancer research,* vol. 77, pp. e104-e107, 2017.