The Influence of Different Ultrasonic Machines on Radiomics Models in Prediction Lymph Node Metastasis for Patients with Cervical Cancer

Jinling Yi, MS[1],*, Xiyao Lei, MS[1],*, Lei Zhang, BS[1], Qiao Zheng, MS[1], Juebin Jin, MS[2], Congying Xie, PhD[1,3], Xiance Jin, PhD[1,4] 🔴 , and Yao Ai, MS[1] 🔴

Abstract

Objective To investigate the effects of different ultrasonic machines on the performance of radiomics models using ultrasound (US) images in the prediction of lymph node metastasis (LNM) for patients with cervical cancer (CC) preoperatively. **Methods** A total of 536 CC patients with confirmed histological characteristics and lymph node status after radical hysterectomy and pelvic lymphadenectomy were enrolled. Radiomics features were extracted and selected with US images acquired with ATL HDI5000, Voluson E8, MyLab classC, ACUSON S2000, and HI VISION Preirus to build radiomics models for LNM prediction using support vector machine (SVM) and logistic regression, respectively. **Results** There were 148 patients (training vs validation: 102:46) scanned in machine HDI5000, 75 patients (53:22) in machine Voluson E8, 100 patients (69:31) in machine MyLab classC, 110 patients (76:34) in machine ACUSON S2000, and 103 patients (73:30) in machine HI VISION Preirus, respectively. Few radiomics features were reproducible among different machines. The area under the curves (AUCs) ranged from 0.75 to 0.86, 0.73 to 0.86 in the training cohorts, and from 0.71 to 0.82, 0.70 to 0.80 in the validation cohorts for SVM and logistic regression models, respectively. The highest difference in AUCs for different machines reaches 17.8% and 15.5% in the training and validation cohorts, respectively. **Conclusions** The performance of radiomics model is dependent on the type of scanner. The problem of scanner dependency on radiomics features should be considered, and their effects should be minimized in future studies for US images.

---

[1] Radiotherapy Center, The 1st Affiliated Hospital of Wenzhou Medical University, Wenzhou, China
[2] Department of Medical Engineering, The 1st Affiliated Hospital of Wenzhou Medical University, Wenzhou, China
[3] Department of Radiation and Medical Oncology, The 2nd Affiliated Hospital of Wenzhou Medical University, Wenzhou, China
[4] School of Basic Medical Science, Wenzhou Medical University, Wenzhou, China

*Authors contribute equally.

Corresponding Authors:
Xiance Jin, Radiotherapy Center, The 1st Affiliated Hospital of Wenzhou Medical University, Wenzhou, China.
Email: jinxc1979@hotmail.com

Yao Ai, Radiotherapy Center, The 1st Affiliated Hospital of Wenzhou Medical University, Wenzhou, China.
Email: aiyaowy@wmu.edu.cn

Introduction

Over the past decade, with the development of imaging technologies and efficient learning strategies, radiomics has emerged and is widely accepted in precision medicine by extracting mineable high-dimensional features from multimodal medical images, for example, computed tomography (CT), positron emission tomography (PET), MRI, or ultrasound (US).[1,2] Studies demonstrated that radiomics features could characterize the internal structure of tissue—such as tumor heterogeneity to enhance the diagnostic, prognostic, and therapeutic outcomes for a broad range of diseases.[3,4] Studies also indicated that with advanced analysis of images using artificial intelligence, radiomics could correlate imaging phenotypes with genomic and proteomic signatures,[5,6] and subsequently complement the information of tissue sampling and circulating biomarkers to reform clinical decision-making.[7,8]

However, one major shortcoming of radiomics is that features could be influenced by many factors, such as the type of scanner, imaging settings, reconstruction parameters, delineation of the tumor, etc,[9–11] which results in a low reproducibility of radiomics features and makes it difficult to compare and interpret radiomics studies for clinical application.[9,10] Standardization in imaging acquisition and analysis had been proposed to improve feature stability, but these standardization approaches are not always feasible as most radiomics studies are retrospective in nature.[3] Awareness of the shortcomings of the radiomics features and studies is of critical significance for the feasibility and applicability of these radiomics studies.

Thanks to its advantages of nonionizing radiation, portability, accessibility, and cost-effectiveness, US has been one of the most used imaging modalities for screening and diagnosis in obstetrics and gynecology.[12] Radiomics studies indicated that US radiomics features were highly associated with breast biologic characteristics,[13] gestational age,[14] neonatal respiratory morbidity,[15] etc. Studies also demonstrated that US-based radiomics was able to predict lymph node metastasis (LNM) for patients with papillary thyroid carcinoma and cervical cancer (CC).[16,17] However, US images also suffer from some limitations, such as low imaging quality caused by noise and artifacts, high inter- and intra-observe variability across different scanners and institutes, and highly operator or diagnostician experience dependent, which will hinder the clinical application of US-based radiomics models.[18] The purpose of this study is to investigate the influence of different US scanners on the reproducibility of radiomics features and the accuracy of LNM prediction for patients with CC.

Materials and Methods

Patients

By searching electronic medical records, a total of 1723 CC patients underwent radical hysterectomy and pelvic lymphadenectomy in our hospital between January 2014 and November 2018 were retrospectively reviewed. Inclusion criteria: (a) patients should have standard ultrasonography within 2 weeks before hysterectomy and (b) patients had confirmed histological characteristics and lymph node status after the operation. Exclusion criteria: (a) patients with incomplete or incorrect clinical data, (b) patients who had other malignant tumors, (c) patients who were treated by chemotherapy or radiotherapy before the operation, and (d) patients with unclear or missing images. A total of 536 patients were enrolled, including 148 with HDI5000, 75 cases of Voluson E8, 100 cases of MyLab classC, 110 cases of ACUSON S2000, and 103 cases of HI VISION Preirus. This retrospective study was conducted following the Declaration of Helsinki and approved by the Ethics Committee in Clinical Research (ECCR) of Wenzhou Medical University First Affiliated Hospital (ECCR No. 2019059). The written informed consent was waived by the ECCR due to the retrospective nature of this study with confirmation of following the Declaration of Helsinki with patient data confidentiality. Besides, all of the patient details in this study have been anonymized. The reporting of this study conforms to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guideline.[19]

US Imaging and Machines

Transvaginal US images were acquired when patients lie in a lithotomy position with an empty bladder. In this study, the following 5-color Doppler ultrasonic machines were used to acquire the US images: ATL HDI 5000（Philips）using the transducer C8-4v at 4-8 MHz; Voluson-E8 (GE Healthcare) using the transducer RIC5-9-D at 5-9 MHz; Mylab classC (Esaote) using the transducer EC1123 at 3-9 MHz; ACUSON S2000 (Siemens) using the transducer MC9-4 at 1.5-6.0 MHz and HI VISION Preirus (Hitachi Ltd) using the transducer EUP-U531 at 4-8 MHz. All images were stored in PNG format and archived in the hospital's DICOM system.

Segmentation and Feature Extraction

Usually, each patient will have 10 to 20 standard US images collected. Regions of interest were delineated manually by a radiologist with 10 years of experience in gynecological US diagnosis using the LIFEx package (http://www.lifexsoft.org),[20] and confirmed by another senior radiologist with 15 years of experience in gynecological US diagnosis. A typical transvaginal US image and contoured target volume are shown in Figure 1. Image preprocessing was performed before feature extraction. A total of 449 radiomics features were extracted from the contoured volumes after intensity normalization using Pyradiomics package (PyRadiomics, https://www.python.org), which includes shape features, first-order histogram statistics, gray-level co-occurrence matrix (GLCM), neighborhood gray-level different matrix, gray-level run length matrix (GLRLM), gray level size zone matrix (GLSZM), and wavelet filtered first-order, GLCM, gray-level different matrix (GLDM), GLRLM, and GLSZM, respectively.[21,22]

Feature Selection and Model Building

Optimal features correlated with LNM were further selected with a 2-step method in the training cohort. First, the Mann-Whitney U test was used to select the radiomics features which had a potential correlation with LNM with a significant

p-value <.05. Second, the least absolute shrinkage and selection operator was applied to select the optimal features with 5-fold cross-validation to adjust the elastic network parameters to select the best combination of feature parameters to avoid overfitting. The dependent variables in the study are dichotomous and we use the parameters family="binomial", α=1, nlambda=100, type.measure="auc."</div><div class="p" xmltag="p">To avoid overfitting the data and achieve stable and superior performance, radiomics models for different machines were built with 2 machine learning classifiers: support vector machine (SVM) and logical regression, which are currently used most commonly in radiomics research studies. The DeLong test was adopted to compare the discrimination performances of differential radiomics models. A total of 10 models were built with radiomics features from 5 different ultrasonic machines. The receiver operating characteristic (ROC) curves with area under the curves (AUCs) were applied to evaluate the maximum diagnostic effect of these 10 models for the prediction of LNM for patients with CC in both the training and validation cohorts.
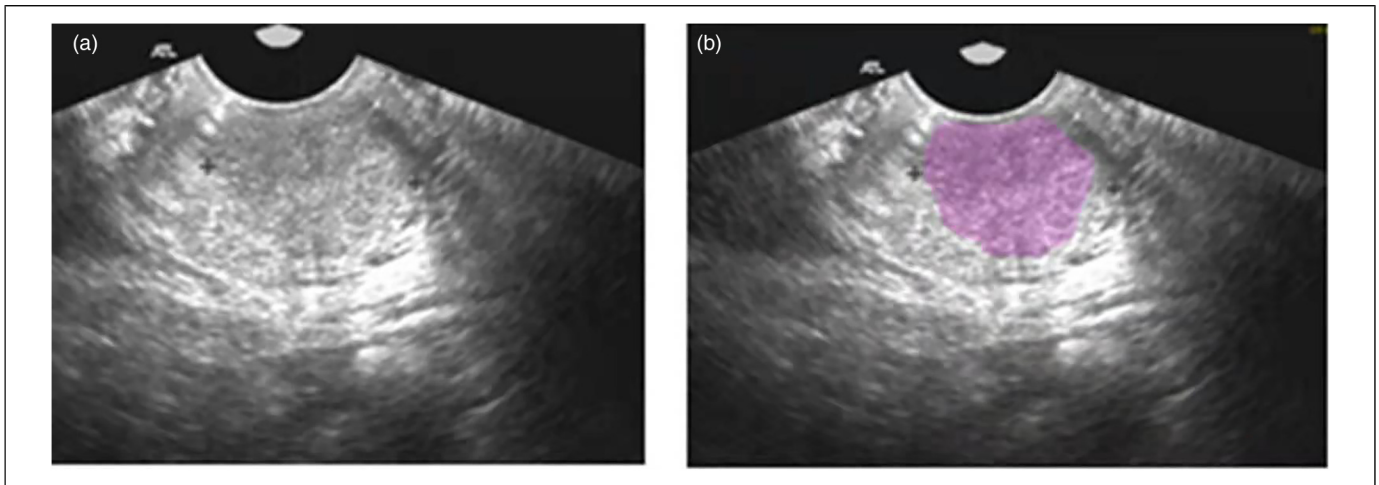

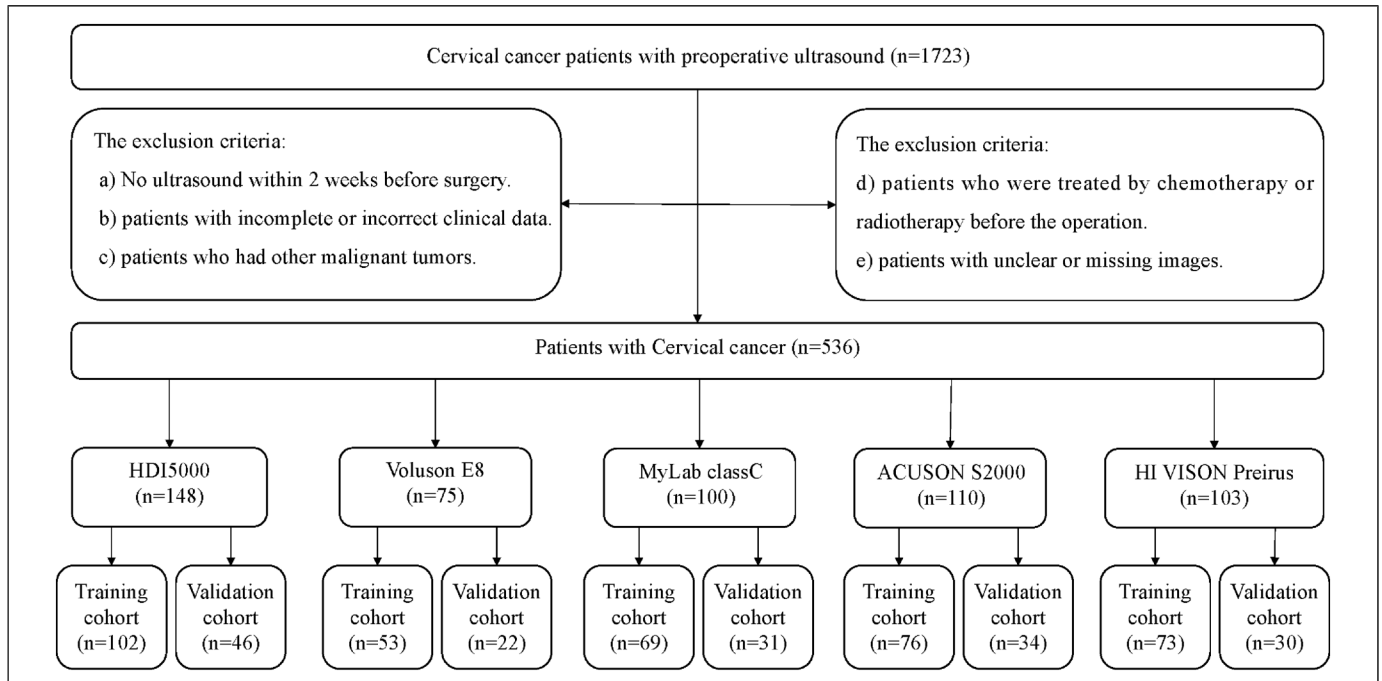**Figure 1.** A typical transvaginal ultrasound image with contoured target volume.

Figure 2. The flow chart of patient selection

Table 1. The Characteristics of Enrolled Patients in the Training and Validation Data Sets.

Characteristics	Training sets			Validation sets		
	LNM+	LNM−		LNM+	LNM−	
Total patients	(n = 73)	(n = 294)	p	(n = 42)	(n = 127)	p
Age			.1			.35
Average	52.23	54.55		51.3	53.21	
Median	51	55		51.5	52	
Range	27-79	27-79		30-76	34-78	
SD	12.14	10.31		10.86	10.44	
Histological type			.51			.97
Squamous cell carcinoma	62	268		37	112	
Adenocarcinoma	5	19		4	8	
Adenosquamous cell carcinoma	5	11		0	4	
Other/not reported	1	6		1	3	
Tumor stage			.37			.3
I	32	146		15	57	
II	41	148		27	20	
HDI5000	LNM+	LNM−	p	LNM+	LNM−	p
	(n = 24)	(n = 78)		(n = 5)	(n = 41)	
Age			.29			.97
Average	48.13	51.22		53.2	53.34	
Median	48.5	50		55	53	
Range	30-65	33-75		43-60	35-78	
SD	10.25	10.56		6.57	11.06	
Histological type			.004			.41
Squamous cell carcinoma	18	74		4	38	
Adenocarcinoma	4	4		1	0	
Adenosquamous cell carcinoma	2	0		0	2	
Other/not reported	0	0		0	1	
Tumor stage			.13			.88
I	8	40		2	15	
II	16	38		3	26	
HI VISION Preirus	LNM+	LNM−	p	LNM+	LNM−	p
	(n = 21)	(n = 52)		(n = 4)	(n = 26)	
Age			.07			.26
Average	50.1	55.42		49.5	55.15	
Median	52	56		49	53.5	
Range	29-72	32-79		40-60	42-71	
SD	12.46	9.84		8.81	7.99	
Histological type			.1			.57
Squamous cell carcinoma	18	49		4	24	
Adenocarcinoma	1	2		0	0	
Adenosquamous cell carcinoma	1	0		0	1	
Other/not reported	1	1		0	1	
Tumor stage			.98			.006
I	8	20		0	19	
II	13	32		4	7	
MyLab classC	LNM+	LNM−	p	LNM	LNM−	p
	(n = 18)	(n = 51)		(n = 5)	(n = 26)	
Age			.87			.79
Average	55.44	55.49		57.4	58.85	
Median	55	57		58	59	
Range	43-70	27-78		46-70	45-78	
SD	9.38	10.7		10.85	8.54	
Histological type			.2			.66
Squamous cell carcinoma	17	42		5	25	
Adenocarcinoma	1	5		0	1	

(continued)

Table 1. (continued)

Characteristics	Training sets			Validation sets		
	LNM+	LNM−		LNM+	LNM−	
Total patients	($n = 73$)	($n = 294$)	p	($n = 42$)	($n = 127$)	p
Adenosquamous cell carcinoma	0	3		0	0	
Other/not reported	0	1		0	0	
Tumor stage			.81			.53
I	8	21		1	9	
II	10	30		4	17	
ACUSON S2000	LNM+	LNM−	p	LNM+	LNM−	p
	($n = 16$)	($n = 60$)		($n = 5$)	($n = 29$)	
Age			.83			.043
Average	53.81	55.07		62.8	52.24	
Median	52.5	54.5		60	51	
Range	27-75	30-76		51-79	36-69	
SD	14.09	11.02		10.23	8.67	
Histological type			.18			.48
Squamous cell carcinoma	15	48		4	27	
Adenocarcinoma	1	6		0	2	
Adenosquamous cell carcinoma	0	4		1	0	
Other/not reported	0	2		0	0	
Tumor stage			.3			.36
I	7	35		2	18	
II	9	25		3	11	
Voluson E8	LNM+	LNM−	p	LNM+	LNM−	p
	($n = 12$)	($n = 41$)		($n = 5$)	($n = 17$)	
Age			.19			.81
Average	49.25	54.08		49.2	53.06	
Median	44.5	2		45	52	
Range	33-76	33-74		44-66	34-75	
SD	14.22	10.11		9.47	12.46	
Histological type			.86			.87
Squamous cell carcinoma	10	32		4	13	
Adenocarcinoma	0	4		0	1	
Adenosquamous cell carcinoma	0	3		1	2	
Other/not reported	2	2		0	1	
Tumor stage			.06			.79
I	9	18		2	8	
II	3	23		3	9	

Abbreviations: LNM, lymph node metastasis; −, negative; +, positive; SD, standard deviation.

Notes: (1) *p*-value is calculated from the univariate association test between subgroups; (2) Fisher's exact test and chi-square test were used for categorized variables; (3) 2-sample *t*-test was used for continuous variables.

Statistical Analysis

Statistical analyses were performed with the R analysis platform (version 4.0.4, MathSoft, http://www. Rproject.org) and SPSS 19.0. The selection of key features and logistic regression model building was performed with the "glmnet" software package. SVM model building with the confusion matrix was performed using the "caret" package in the "e1071" software package. For all tests, a $p<.05$ was considered statistically significant.

Result

A total of 536 CC patients with confirmed histological characteristics and lymph node status after radical hysterectomy and pelvic lymphadenectomy were enrolled in this study. Patients were randomly divided into training and validation cohorts for each US imaging machine at a ratio close to 7:3. There were 148 patients (102:46) scanned in machine HDI5000, 75 patients (53:22) in machine Voluson E8, 100 patients (69:31) in machine MyLab classC, 110 patients (76:34) in machine ACUSON S2000, and 103 patients (73:30) in machine HI VISION Preirus, respectively. The flowchart is shown in Figure 2. Detailed characteristics of enrolled total patients and patients scanned with different machines are presented in Table 1.

Figure 3 shows the selection of optimal radiomics for LNM prediction for different machines. Based on the elastic-net method through tuning the parameter (λ) with a 5-fold
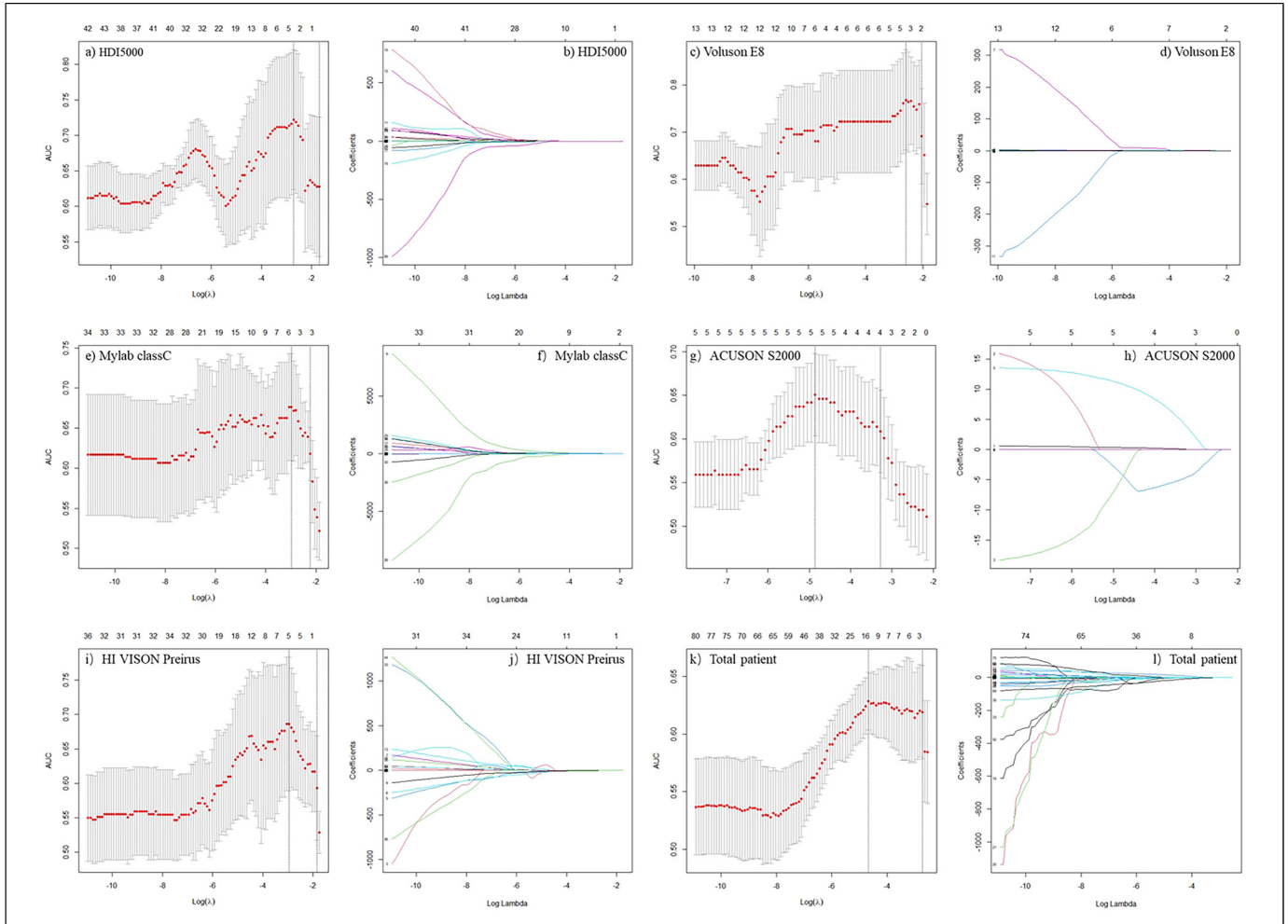
Figure 3. The selection of optimal radiomics for lymph node metastasis prediction based on the elastic-net method through tuning the parameter (λ) with a 5-fold cross-validation via maximum area under the curves: (a, b) for machine HDI 5000; (c, d) for machine Voluson E8; (e, f) for machine MyLab classC; (g, h) for machine ACUSON S2000; (i, j) for machine HI VISION Preirus; and (k, l) for all the machines.

cross-validation via maximum AUCs in the training cohorts, radiomics features with nonzero coefficients were selected. Eventually, 4 radiomics features with potential correlation with LNM were selected from machine HDI 5000, 3 features from machine Voluson E8, 6 features from machine MyLab classC, 5 features from machine ACUSON S2000, and 5 features from machine HI VISION Preirus, respectively. Details of the selected radiomics features are presented in Table 2.

The performance of LNM prediction models with radiomics features from different machines is presented in Figure 4. The AUCs ranged from 0.75 to 0.86 and 0.73 to 0.86 in the training cohorts and from 0.71 to 0.82 and 0.70 to 0.80 in the validation cohorts for SVM and logistic regression models, respectively. Detailed performance of these models for different machines and total patients are shown in Table 3. DeLong test found that the performances of differential radiomics models were not significantly different (both, $p > .05$), as shown in Tables 4 and 5.

Discussion

LNM is one of the most important risk factors for recurrence and survival for patients with CC.[23,24] Undiagnosed or inaccurately assessed LNM is a major cause of suboptimal treatment. For instance, 90% of routinely resected lymph modes are not metastatic, and unnecessary pelvic lymphadenectomy leads to complications, such as prolonged surgery, blood loss, infection, nerve or vascular injury, lymphocyst formation, etc.[25,26] Therefore, an accurate noninvasive technique for lymph node assessment is vital in the clinical management of patients with cervical. A noninvasive LNM prediction method with radiomics has been intensively investigated for CC preoperatively with MRI,[27,28] ^{18}F-fluorodeoxyglucose PET/CT,[29,30] as well as US images.[17] In this study, the effects of US machines on the reproducibility of radiomics features and radiomics models for the prediction of LNM for patients with CC were investigated.

A noninvasive LNM prediction method using radiomics is of critical clinical benefit in the management of CC. Jin et al[17]

Table 2. List of Selected Radiomics Features with Potential Correlation with Lymph Node Metastasis for Patients with Cervical Cancer from Different Ultrasound Machines.

Models	Feature	*P*-value
HDI 5000		
	wavelet.HH_firstorder_Energy	.012
	wavelet.HL_glrlm_RunLengthNonUniformity	.000
	wavelet.LH_glszm_ZoneVariance	.005
	wavelet.HL_glszm_LargeAreaLowGrayLevelEmphasis	.014
Voluson E8		
	original_glrlm_ShortRunHighGrayLevelEmphasis	.008
	original_gldm_SmallDependenceHighGrayLevelEmphasis	.015
	original_glszm_SmallAreaHighGrayLevelEmphasis	.021
MyLab classC		
	wavelet.HL_firstorder_Minimum	.043
	wavelet.LH_glcm_Imc1	.034
	wavelet.LH_glcm_InverseVariance	.009
	wavelet.HH_gldm_DependenceVariance	.004
	wavelet.LH_glszm_SizeZoneNonUniformityNormalized	.022
	wavelet.LH_glszm_SmallAreaHighGrayLevelEmphasis	.006
ACUSON S2000		
	wavelet.HL_glrlm_RunVariance	.047
	original_glszm_SizeZoneNonUniformityNormalized	.022
	original_glszm_SmallAreaEmphasis	.026
	wavelet.HH_glszm_GrayLevelNonUniformityNormalized	.032
	wavelet.HH_glszm_LargeAreaLowGrayLevelEmphasis	.047
HI VISION Preirus		
	original_glcm_MaximumProbability	.035
	wavelet.HH_glcm_Imc2	.038
	original_glszm_SmallAreaHighGrayLevelEmphasis	.038
	wavelet.HL_glszm_LargeAreaHighGrayLevelEmphasis	.009
	wavelet.LL_glszm_ZoneEntropy	.003

reported an AUC of 0.79 and 0.77 in the training and validation cohorts with a US images-based radiomics model from one Philips machine (IU22) in predicting LNM for early-stage CC preoperatively. Our study demonstrated similar AUCs around 0.70 to 0.82 for radiomics models with US images from different machines. However, the highest difference in AUCs for different machines reaches 17.8% and 15.5% in the training and validation cohorts, respectively, as shown in Figure 4.

As shown in Table 2, the selected optimal features from different US machines were variable. Few radiomics features were reproducible among different machines. Only features of wavelet.HL_glszm_LargeAreaLowGrayLevelEmphasis repeated in 3 machines (HDI 5000, ACUSON S2000, and HI VISION Preirus) and wavelet.LH_glszm_SmallAreaHigh GrayLevelEmphasis repeated in 3 machines (Voluson E8, MyLab classC, and HI VISION Preirus). This indicated that radiomic features from US images are highly scanner-dependent. Previously, Yasaka et al[31] demonstrated that radiomics features from CT images were also scanner-dependent based on a phantom study due to differences in imaging acquisition parameters and the scanner design. Mackin et al[32] indicated that the variation of radiomic features among different CT scanners was similar to their variation in CT images for 20 nonsmall cell lung cancer patients. Li et al[33] assessed the reproducibility of radiomics features from different US scanners, acquisition parameters, segmentation location, and extraction platforms, and also found that the wavelet features showed the best reproducibility among different scanners. The problem of scanner dependency on radiomics features should be considered, and their effects should be minimized in future studies for US images.

Table 3 also demonstrated that radiomics models with combined features from different machines performed inferior to models with features from 1 machine only. This is also a dilemma for current radiomics studies. A large cohort of patients from multiple centers is needed in order to generate convincing results to transfer radiomics as a prognostic tool for potential clinical application.[34,35] However, data from multiple centers is unavoidable with variability in the scanner, models, acquisition protocols, reconstruction settings, etc, which further affects the reproducibility of radiomics features and models.[9,36] As is similarly shown in this study, combined radiomics features from multiple US machines are inferior to the performance of radiomics models. By using the same image acquisition parameters, reconstruction technique, image segmentation, and method of feature extraction, the radiomics
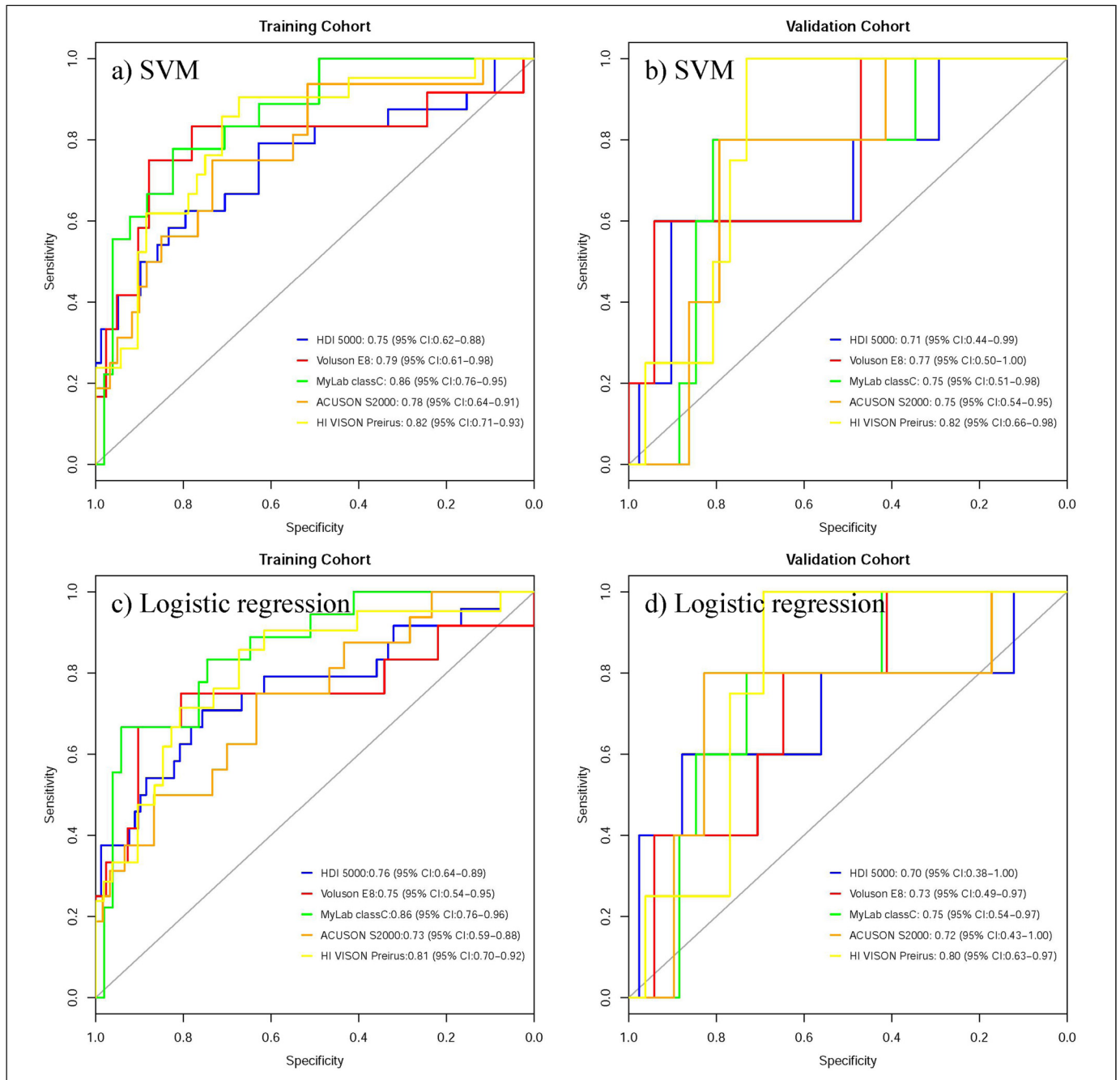
Figure 4. The performance of radiomics models with ultrasound images from different ultrasonic machines (a, b) using support vector machine and (c, d) using logistic regression with training and validation cohorts, respectively.

features repeatability problem caused by different scanners would be solved. Although harmonizing images and features had been aggressively investigated, there is much more work needed to realize the potential clinical value of radiomics in the future.[37,38] Application of deep learning to harmonize US images from different scanners would be a solution for future radiomics studies.

There were several limitations in our study. First, lack of power analysis for estimation of sample size. Second, radiomics feature reliability assessed by the intraclass correlation coefficient was not considered in this study. Finally, the amount of dataset in several scanners is inadequate. Commonly, a larger amount of dataset will improve the confidence and performance of our model.

Conclusion

The effects of different US machines on the performance of radiomics models based on US images in the prediction of

Table 3. The Performance of SVM and Logistic Regression Models with Radiomics Features from Different Machines for the Training and Validation Cohorts.

Models	Machines	Training cohort				Validation cohort			
		AUC (95% CI)	Accuracy	Sensitivity	Specificity	AUC (95% CI)	Accuracy	Sensitivity	Specificity
Logistic regression	Total patients	0.66(0.59-0.73)	0.80	0.77	0.54	0.61(0.50-0.72)	0.76	0.45	0.81
	HDI 5000	0.76(0.64-0.89)	0.84	0.71	0.76	0.70(0.38-1.00)	0.83	0.60	0.88
	Voluson E8	0.75(0.54-0.95)	0.83	0.67	0.90	0.73(0.49-0.97)	0.64	0.80	0.65
	MyLab classC	0.86(0.76-0.96)	0.84	0.67	0.94	0.75(0.54-0.97)	0.61	0.80	0.73
	ACUSON S2000	0.73(0.59-0.88)	0.82	0.75	0.63	0.72(0.43-1.00)	0.83	0.80	0.83
	HI VISION Preirus	0.81(0.70-0.92)	0.77	0.86	0.67	0.80(0.63-0.97)	0.73	1.00	0.69
SVM	Total patients	0.65(0.58-0.71)	0.81	0.93	0.29	0.60(0.50-0.70)	0.75	0.64	0.60
	HDI5000	0.75(0.62-0.88)	0.81	0.63	0.80	0.71(0.44-0.99)	0.89	0.60	0.90
	Voluson E8	0.79(0.61-0.98)	0.79	0.75	0.88	0.77(0.50-1.00)	0.68	0.60	0.94
	MyLab classC	0.86(0.76-0.95)	0.81	0.78	0.82	0.75(0.51-0.98)	0.65	0.80	0.81
	ACUSON S2000	0.78(0.64-0.91)	0.79	0.75	0.73	0.75(0.54-0.95)	0.85	0.80	0.79
	HI VISION Preirus	0.82(0.71-0.93)	0.77	0.91	0.67	0.82(0.66-0.98)	0.77	1.00	0.73

Abbreviations: SVM, support vector machine; AUC, area under the curve; CI, confidence interval.

Table 4. The DeLong Test of Support Vector Machine Models from Different Machines for Validation Cohorts.

	HDI 5000	Voluson E8	MyLab classC	ACUSON S2000	HI VISION Preirus
HDI 5000	—	0.7889	0.8546	0.8537	0.5220
Voluson E8	—	—	0.9182	0.9080	0.7405
MyLab classC	—	—	—	0.9934	0.6243
ACUSON S2000	—	—	—	—	0.5892
HI VISION Preirus	—	—	—	—	—

Table 5. The DeLong Test of Logistic Regression Models from Different Machines for Validation Cohorts.

	HDI 5000	Voluson E8	MyLab classC	ACUSON S2000	HI VISION Preirus
HDI 5000	—	0.8973	0.7965	0.9229	0.6137
Voluson E8	—	—	0.8826	0.9784	0.6545
MyLab classC	—	—	—	0.8719	0.7516
ACUSON S2000	—	—	—	—	0.6701
HI VISION Preirus	—	—	—	—	—

LNM status for CC preoperatively were investigated. The optimal features for prediction models were scanner dependent. The maximum AUC differences for models with images from the different scanners were 17.8% and 15.5% in the training and validation cohorts, respectively.

Ethics Statement

The ECCR of the authors' hospital approved this retrospective study (ECCR no. 2019059). The written informed consent was waived by the ECCR due to the retrospective nature of this study with confirmation of following the Declaration of Helsinki with patient data confidentiality.

ORCID iDs

Xiance Jin https://orcid.org/0000-0002-4117-5953
Yao Ai https://orcid.org/0000-0001-9992-0250

References

1. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology*. 2016;278(2):563-577. doi:10.1148/radiol.2015151169

2. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun*. 2014;5:4644. doi:10.1038/ncomms5006

3. Lambin P, Leijenaar R, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14(12):749-762. doi:10.1038/nrclinonc.2017.141

4. Incoronato M, Aiello M, Infante T, et al. Radiogenomic analysis of oncological data: a technical survey. *Int J Mol Sci*. 2017;18(4):805. doi:10.3390/ijms18040805

5. Afshar P, Mohammadi A, Plataniotis KN, et al. From handcrafted to deep-learning-based cancer radiomics: challenges and opportunities. *IEEE Signal Process Mag*. 2019; 36(4):132-160. doi: 10.1109/MSP.2019.2900993

6. Hosny A, Parmar C, Quackenbush J, et al. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018;18(8):500-510. doi:10.1038/s41568-018-0016-5

7. Zhang B, Tian J, Dong D, et al. Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma. *Clin Cancer Res*. 2017;23(15):4259-4269. doi:10.1158/1078-0432.CCR-16-2910

8. Leijenaar RT, Carvalho S, Hoebers FJ, et al. External validation of a prognostic CT-based radiomic signature in oropharyngeal squamous cell carcinoma. *Acta Oncol*. 2015;54(9):1423-1429. doi:10.3109/0284186X.2015.1061214

9. Yan J, Chu-Shern JL, Loi HY, et al. Impact of image reconstruction settings on texture features in ^{18}F-FDG PET. *J Nucl Med*. 2015;56(11):1667-1673. doi:10.2967/jnumed.115.156927

10. Van Velden FH, Kramer GM, Frings V, et al. Repeatability of radiomic features in non-small-cell lung cancer [^{18}F]FDG-PET/CT studies: impact of reconstruction and delineation. *Mol Imaging Biol*. 2016;18(5):788-795. doi:10.1007/s11307-016-0940-2

11. Balagurunathan Y, Gu Y, Wang H, et al. Reproducibility and prognosis of quantitative features extracted from CT images. *Transl Oncol*. 2014;7(1):72-87. doi:10.1593/tlo.13844

12. Abinader R, Warsof SL. Benefits and pitfalls of ultrasound in obstetrics and gynecology. *Obstet Gynecol Clin North Am*. 2019;46(2):367-378. doi:10.1016/j.ogc.2019.01.011

13. Guo Y, Hu Y, Qiao M, et al. Radiomics analysis on ultrasound for prediction of biologic behavior in breast invasive ductal carcinoma. *Clin Breast Cancer*. 2018;18(3):e335-e344. doi:10.1016/j.clbc.2017.08.002

14. Cobo T, Bonet-Carne E, Martínez-Terrón M, et al. Feasibility and reproducibility of fetal lung texture analysis by automatic quantitative ultrasound analysis and correlation with gestational age. *Fetal Diagn Ther*. 2012;31(4):230-236. doi:10.1159/000335349

15. Bonet-Carne E, Palacio M, Cobo T, et al. Quantitative ultrasound texture analysis of fetal lungs to predict neonatal respiratory morbidity. *Ultrasound Obstet Gynecol*. 2015;45(4):427-433. doi:10.1002/uog.13441

16. Liu T, Zhou S, Yu J, et al. Prediction of lymph node metastasis in patients with papillary thyroid carcinoma: a radiomics method based on preoperative ultrasound images. *Technol Cancer Res Treat*. 2019;18:1533033819831713. doi:10.1177/1533033819831713

17. Jin X, Ai Y, Zhang J, et al. Noninvasive prediction of lymph node status for patients with early-stage cervical cancer based on radiomics features from ultrasound images. *Eur Radiol*. 2020;30(7):4117-4124. doi:10.1007/s00330-020-06692-1

18. Salomon LJ, Winer N, Bernard JP, et al. A score-based method for quality control of fetal images at routine second-trimester ultrasound examination. *Prenat Diagn*. 2008;28(9):822-827. doi:10.1002/pd.2016

19. von Elm E, Altman DG, Egger M, et al. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med*. 2007;147:573-577. doi: 10.7326/0003-4819-147-8-200710160-00010

20. Nioche C, Orlhac F, Boughdad S, et al. LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res*. 2018;78(16):4786-4789. doi:10.1158/0008-5472.CAN-18-0125

21. Garg M, Dhiman G. A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants. *Neural Comput Appl*. 2021;33:1311-1328. doi:10.1007/s00521-020-05017-z

22. Durgamahanthi V, Anita Christaline J, Shirly Edward A, GLCM and GLRLM based texture analysis: application to brain cancer diagnosis using histopathology images. In: Dash SS, Das S, Panigrahi BK eds., *Intelligent computing and applications. Advances in Intelligent Systems and Computing*, Vol. 1172. Singapore: Springer, 2021; pp. 691–706. doi:10.1007/978-981-15-5566-4_61

23. Biewenga P, van der Velden J, Mol BW, et al. Prognostic model for survival in patients with early stage cervical cancer. *Cancer*. 2011;117(4):768-776. doi:10.1002/cncr.25658

24. Bourgioti C, Chatoupis K, Moulopoulos LA. Current imaging strategies for the evaluation of uterine cervical cancer. *World J Radiol*. 2016;8(4):342-354. doi:10.4329/wjr.v8.i4.342

25. Lv K, Guo HM, Lu YJ, et al. Role of 18F-FDG PET/CT in detecting pelvic lymph-node metastases in patients with early-stage uterine cervical cancer: comparison with MRI findings. *Nucl Med Commun*. 2014;35(12):1204-1211. doi:10.1097/MNM.0000000000000198

26. Wang HY, Sun JM, Tang J. Sentinel lymph nodes detection in patients with cervical cancer undergoing radical hysterectomy. *Zhonghua fu chan ke za zhi*. 2014;39(1):7-9. doi: 10.1111/j.1525-1438.2006.00402.x

27. Wu Q, Shi D, Dou S, et al. Radiomics analysis of multiparametric MRI evaluates the pathological features of cervical squamous cell carcinoma. *J Magn Reson Imaging*. 2019;49(4):1141-1148. doi:10.1002/jmri.26301

28. Wu Q, Shi D, Dou S, et al. Radiomic signature as a predictive factor for lymph node metastasis in early-stage cervical cancer. *J Magn Reson Imaging*. 2019;49(1):304-310. doi:10.1002/jmri.26209

29. Li K, Sun H, Lu Z, et al. Value of [18F]FDG PET radiomic features and VEGF expression in predicting pelvic lymphatic metastasis and their potential relationship in early-stage cervical squamous cell carcinoma. *Eur J Radiol*. 2018;106:160-166. doi:10.1016/j.ejrad.2018.07.024

30. Shen WC, Chen SW, Liang JA, et al. [18]Fluorodeoxyglucose positron emission tomography for the textural features of cervical cancer associated with lymph node metastasis and histological type. *Eur J Nucl Med Mol Imaging*. 2017;44(10):1721-1731. doi:10.1007/s00259-017-3697-1

31. Yasaka K, Akai H, Mackin D, et al. Precision of quantitative computed tomography texture analysis using image filtering: a phantom study for scanner variability. *Medicine (Baltimore)*. 2017;96(21):e6993. doi:10.1097/MD.0000000000006993

32. Mackin D, Fave X, Zhang L, et al. Measuring computed tomography scanner variability of radiomics features. *Invest Radiol*. 2015;50(11):757-765. doi:10.1097/RLI.0000000000000180

33. Li MD, Cheng MQ, Chen LD, et al. Reproducibility of radiomics features from ultrasound images: influence of image acquisition and processing. *Eur Radiol*. 2022:1-9. doi: 10.1007/s00330-022-08662-1

34. Sun C, Tian X, Liu Z, et al. Radiomic analysis for pretreatment prediction of response to neoadjuvant chemotherapy in locally advanced cervical cancer: a multicentre study. *EBioMedicine*. 2019;46:160-169. doi:10.1016/j.ebiom.2019.07.049

35. Lucia F, Visvikis D, Vallières M, 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(4):864-877. doi:10.1007/s00259-018-4231-9

36. Peerlings J, Woodruff HC, Winfield JM, et al. Stability of radiomics features in apparent diffusion coefficient maps from a multicentre test-retest trial. *Sci Rep*. 2019;9(1):4800. doi:10.1038/s41598-019-41344-5

37. Luo R, Wang J, Zhong H, et al. OC-0160: radiomics features harmonization for CT and CBCT in rectal cancer. *Radiother Oncol*. 2017;123:S81-S82. doi:10.1016/S0167-8140(17)30603-5

38. Li Y, Han G, Wu X, et al. Normalization of multicenter CT radiomics by a generative adversarial network method. *Phys Med Biol*. 2021;66(5):055030. doi:10.1088/1361-6560/ab8319