European Radiology (2021) 31:411–422
https://doi.org/10.1007/s00330-020-07099-8

IMAGING INFORMATICS AND ARTIFICIAL INTELLIGENCE

Radiologists with MRI-based radiomics aids to predict the pelvic lymph node metastasis in endometrial cancer: a multicenter study

Bi Cong Yan[1] · Ying Li[1] · Feng Hua Ma[2] · Guo Fu Zhang[2] · Feng Feng[3] · Ming Hua Sun[4] · Guang Wu Lin[5] · Jin Wei Qiang[1]

Received: 22 January 2020 / Revised: 31 May 2020 / Accepted: 21 July 2020 / Published online: 4 August 2020

Abstract

Objective To construct a MRI radiomics model and help radiologists to improve the assessments of pelvic lymph node metastasis (PLNM) in endometrial cancer (EC) preoperatively.

Methods During January 2014 and May 2019, 622 EC patients (age 56.6 ± 8.8 years; range 27–85 years) from five different centers (A to E) were divided into training set, validation set 1 (351 cases from center A), and validation set 2 (271 cases from centers B–E). The radiomics features were extracted basing on T2WI, DWI, ADC, and CE-T1WI images, and most related radiomics features were selected using the random forest classifier to build a radiomics model. The ROC curve was used to evaluate the performance of training set and validation sets, radiologists based on MRI findings alone, and with the aid of the radiomics model. The clinical decisive curve (CDC), net reclassification index (NRI), and total integrated discrimination index (IDI) were used to assess the clinical benefit of using the radiomics model.

Results The AUC values were 0.935 for the training set, 0.909 and 0.885 for validation sets 1 and 2, 0.623 and 0.643 for the radiologists 1 and 2 alone, and 0.814 and 0.842 for the radiomics-aided radiologists 1 and 2, respectively. The AUC, CDC, NRI, and IDI showed higher diagnostic performance and clinical net benefits for the radiomics-aided radiologists than for the radiologists alone.

Conclusions The MRI-based radiomics model could be used to assess the status of pelvic lymph node and help radiologists improve their performance in predicting PLNM in EC.

Key Points

- *A total of 358 radiomics features were extracted. The 37 most important features were selected using the random forest classifier.*
- *The reclassification measures of discrimination confirmed that the radiomics-aided radiologists performed better than the radiologists alone, with an NRI of 1.26 and an IDI of 0.21 for radiologist 1 and an NRI of 1.37 and an IDI of 0.24 for radiologist 2.*

Keywords Magnetic resonance imaging · Radiomics · Lymphatic metastasis · Lymph node · Endometrial neoplasms

Bi Cong Yan and Ying Li contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-020-07099-8) contains supplementary material, which is available to authorized users.

✉ Jin Wei Qiang
dr.jinweiqiang@163.com

[1] Department of Radiology, Jinshan Hospital, Fudan University, 1508 Longhang Road, Shanghai 201508, China

[2] Departments of Radiology, Obstetrics & Gynecology Hospital, Fudan University, 128 ShenYang Road, Shanghai 200090, China

[3] Department of Radiology, Cancer Hospital of Nantong University, 30 North Tong Yang Road, 536 Chang Le Road, Nantong 226361, Jiangsu, China

[4] Department of Radiology, Shanghai First Maternity and Infant Hospital, Tongji University School of Medicine, Shanghai 201204, China

[5] Department of Radiology, Huadong Hospital of Fudan University, Fudan University, 221 West Yan'an Road, Shanghai 200040, China

Abbreviations

CDC	Clinical decisive curve
CI	Confidence interval
EC	Endometrial cancer
ER	Estrogen receptor
IDI	Integrated discrimination index
LNM	Lymph node metastasis
NRI	Net reclassification index
PLNM	Pelvic lymph node metastasis
PR	Progesterone receptor
SMOTE	Synthetic minority oversampling technique

Introduction

Lymph node metastasis (LNM) is an important factor affecting endometrial cancer (EC) prognosis. Multiple studies have suggested that lymphadenectomy can help surgical staging and modulate or eliminate the need for adjuvant therapy [1, 2]. However, controversy exists regarding whether lymphadenectomy should be performed in early-stage EC. Studies show that systematic lymphadenectomy contributes to a higher incidence of complications and produces little evidence for adjuvant therapy in early-stage EC [3, 4]. Thus, for optimizing surgical methods and clinical outcomes, dedicated efforts should be redirected in identifying EC patients who have preoperative pelvic LNM (PLNM) and need lymphadenectomy, thereby minimizing overtreatment for EC patients without PLNM.

Magnetic resonance imaging (MRI) is a noninvasive technique with high resolution for visualizing soft tissue. However, a meta-analysis indicated that MRI has low sensitivity in diagnosing PLNM [5]. A recent study used positron emission tomography/computed tomography (PET/CT) to preoperatively assess the lymph node status of EC but resulted in a high false positive rate [6]. Radiomics, a method of high-throughput quantitative information extraction from medical images such as MR images, may offer valuable information for EC with objective and reproducible modes [7–10]. Radiomics has been confirmed to be a useful tool in several cancers [11–14]. One study showed a moderate performance in preoperatively diagnosing LNM based on 2D tumor MRI texture features of EC [15]. Based on a limited sample in a single center, another study showed that PET radiomics is a valuable tool for detecting LNM in EC [16]. Furthermore, immunohistochemical profiles, such as estrogen receptor (ER), progesterone receptor (PR), P53, and Ki-67 levels, are correlated with LNM and are good predictors of lymph node status and prognosis in patients with EC [17, 18]. However, no study has shown the correlation of these immunohistochemical indexes with the radiomics image biomarkers.

We assumed that MRI-based radiomics could be a useful tool in preoperatively diagnosing PLNM. The purpose of this study was to explore whether MRI-based radiomics could improve the diagnostic performance of radiologists in the assessment of PLNM in EC patients, in multiple centers and with a large sample size. Our secondary aim was to investigate the correlation of radiomics features of PLNM with immunohistochemical indexes.

Materials and methods

Study participants

This retrospective study was performed after approval by the institutional review boards of all participating centers, and informed consent was waived. During January 2014 and May 2019, 664 consecutive EC patients with preoperative MRI were reviewed at the centers A to E. The inclusion criteria were as follows: (1) histopathologically proven EC based on total hysterectomy and bilateral salpingo-oophorectomy (THBSO) and lymphadenectomy; (2) MRI sequences including axial T2-weighted imaging (T2WI), diffusion-weighted imaging (DWI), apparent diffusion coefficient (ADC), and contrast-enhanced T1-weighted imaging (CE-T1WI); (3) an interval time between MRI examination and surgery less than 35 days; and (4) complete immunohistochemical information (including ER, PR, P53, and Ki-67). The exclusion criteria were as follows: (1) lack of any of the abovementioned MRI sequences ($n = 5$); (2) imaging with obvious motion artifacts or sequences that could not match well with each other ($n = 1$); (3) tumor too small to be visible ($n = 4$); (4) an interval time between MRI examination and surgery more than 35 days ($n = 0$); (5) insufficient pathological or surgical information (without lymphadenectomy) ($n = 25$); (6) patients with para-aortic LNM without PLNM ($n = 5$); and (7) chemotherapy/radiotherapy performed prior to the surgery ($n = 2$). Finally, 622 EC patients (age 56.6 ± 8.8 years; range 27–85 year) were included in this multicenter study. Figure 1 shows the workflow of this study.

MR imaging

Magnetic resonance imaging was performed using 1.5/3.0-T scanners with phased-array abdominal coils. The patients laid in a supine position and breathed freely during the acquisition. The following sequences were obtained: axial T1-weighted imaging (T1WI), T2WI with and without fat saturation (FS), DWI with a *b* value of 800 or 1000 s/mm^2, and CE-T1WI with FS, which was performed immediately after the intravenous administration of gadopentetate dimeglumine at a dose of 0.2 mmol/kg of body weight and a rate of 2 to 3 ml/s. The scanning details are shown in Table 1.

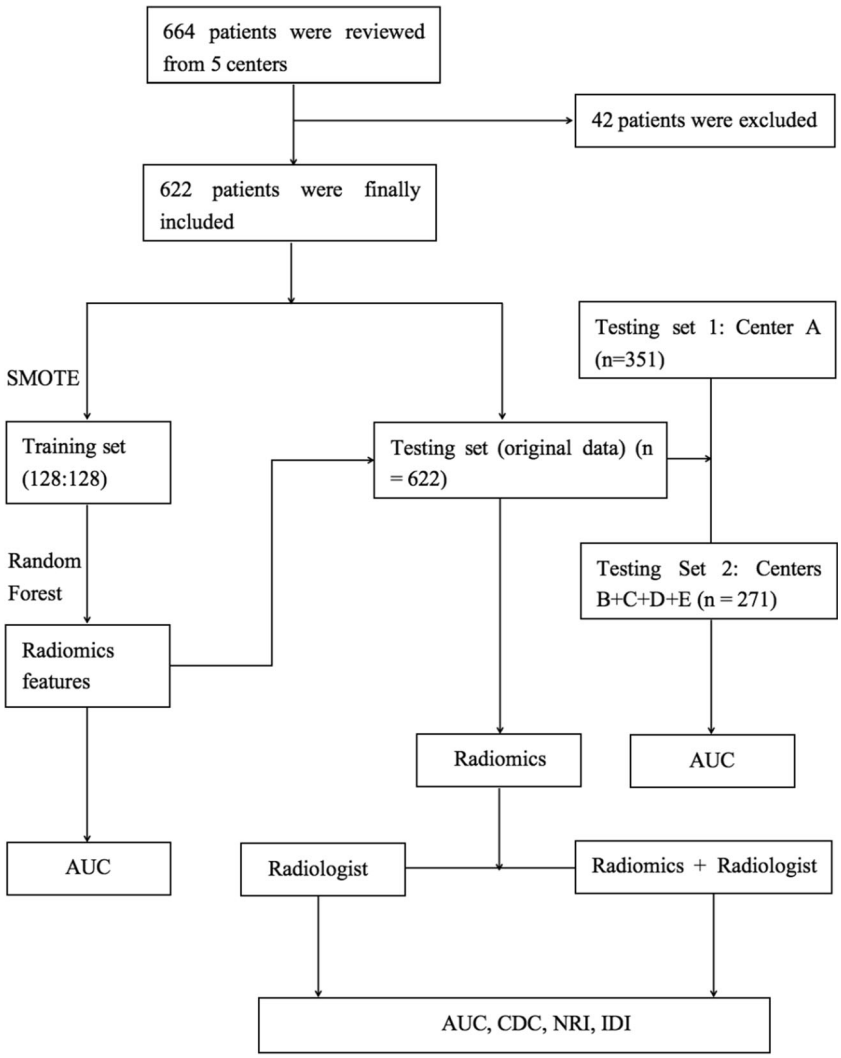

Fig. 1 The workflow of this study. CDC, clinical decisive curve; IDI, integrated discrimination index; NRI, net reclassification index; SMOTE, synthetic minority oversampling technique

Image segmentation

Using MitkWorkbench (http://mitk.org/wiki/The_Medical_Imaging_Interaction_Toolkit_(MITK)), the multisequence images from axial T2WI, DWI, ADC mapping, and CE-T1WI were matched. The region of interest (ROI) for each tumor was manually drawn along the margin of each tumor slice based on axial T2WI by referring to the DWI, ADC, and CE-T1WI at the delayed phase and carefully avoiding nearby normal myometrium or endometrium. The volumetric region of interest (VOI) of each tumor was segmented. All ROI drawing was performed by an experienced radiologist (reader 1) blinded to the patients' histopathology. With 1-month intervals, 50 patients were randomly chosen, and the ROI drawing of tumors was repeated by two radiologists (reader 1 and reader 2). Inter- and intra-observer agreements of each extracted feature were determined by calculating the intraclass correlation coefficient (ICC).

Image preprocessing and radiomics feature extraction

Radiomics feature extraction based on the VOIs from T2WI, DWI, ADC, and CE-T1WI at the delayed phase was performed using Pyradiomics software (https://pypi.org/project/pyradiomics/) following the IBSI recommendation (http://arxiv.org/abs//1612.07003). The VOIs from DWI, ADC, and CE-T1WI were aligned to those from T2WI using in-house software based on the Insight Segmentation and Registration Toolkit (ITK, version 4.7.2; https://itk.org/). The VOIs were resampled with obtained isotropic voxels ($3 \times 3 \times 3$ mm). Normalization was performed by subtracting the mean value from each voxel and then dividing by the standard deviation and scaling the values to within a 0–600 range. To ensure better comparability of MRI gray values, a fixed bin width of 1 was used to compute textural features.

Table 1 MRI examination's parameters

	Center A	Center B	Center C	Center D	Center E
MRI	1.5-T Avanto Siemens	3.0-T Trio Siemens	3.0-T Verio Siemens	1.5-T OPTIMA 360 General Electric	3.0-T Verio Siemens
T1WI	TSE, TR/TE = 761/10, matrix = 512 × 512, thickness = 4 mm, FOV = 360 × 280 mm; thickness = 4 mm, FOV = 360 × 280 mm	TSE, TR/TE = 700/11, matrix = 512 × 512, thickness = 5 mm, FOV = 380 × 280 mm	TSE, TR/TE = 340/10, matrix = 512 × 512, thickness = 4 mm, FOV = 340 × 280 mm	FSE, TR/TE = 788/15, matrix = 512 × 512, thickness = 4 mm, FOV = 350 × 270 mm	VIBE, TR/TE = 3.4/1.3, matrix = 320 × 320, thickness = 3 mm, FOV = 380 × 280 mm
T2WI	SE, TR/TE = 4000/98 or 8000/83, matrix = 512 × 512	SE, TR/TE = 3300/88, matrix = 512 × 512, thickness = 3 mm, FOV = 380 × 280 mm	TSE, TR/TE = 8000/83, matrix = 512 × 512, thickness = 4 mm, FOV = 340 × 270 mm	FSE, TR/TE = 2705/60, matrix = 512 × 512, thickness = 4 mm, FOV = 380 × 270 mm	TSE, TR/TE = 2770/64, matrix = 320 × 320, thickness = 4 mm, FOV = 380 × 280 mm
DWI	EPI, TR/TE = 4000/100, *b* = 800 (or 1000), matrix = 256 × 256, thickness = 5 mm, FOV = 360 × 80 mm	EPI, TR/TE = 2500/100, *b* = 800 (or 1000), matrix = 256 × 256, thickness = 4 mm, FOV = 380 × 280 mm	EPI, TR/TE = 4000/100, *b* = 1000, matrix = 256 × 256, thickness = 5 mm, FOV = 340 × 280 mm	EPI, TR/TE = 3708/76, *b* = 800, matrix = 256 × 256, thickness = 5 mm, FOV = 380 × 270 mm	EPI, TR/TE = 7100/79, *b* = 1000, matrix = 256 × 256, thickness = 4 mm, FOV = 380 × 280 mm
CE-T1WI	FLASH, TR/TE = 196/2.9, matrix = 512 × 512, thickness = 4 mm, FOV = 360 × 280 mm	TSE, TR/TE = 700/11, matrix = 512 × 512, thickness = 5 mm, FOV = 380 × 280 mm	TSE, TR/TE = 196/2.9, matrix = 512 × 512, thickness = 4 mm, FOV = 340 × 280 mm	LAVA, TR/TE = 3.7/1.7, matrix = 512 × 512, thickness = 4 mm, FOV = 350 × 270 mm	VIBE, TR/TE = 3.4/1.3, matrix = 320 × 320, thickness = 3 mm, FOV = 380 × 280 mm

CE, contrast enhanced; *DWI* diffusion-weighted imaging; *EPI*, echo planar imaging; *FLASH*, fast low-angle shot sequence; *FSE*, fast spin echo; *LAVA*, liver acquisition with volume acceleration; *MRI*, magnetic resonance imaging; *SE*, spin echo; *T1WI*, T1-weighted imaging; *T2WI*, T2-weighted imaging; *TE*, echo time; *TR*, repetition time; *TSE*, turbo spin echo; *VIBE*, volumetric interpolated breath-hold examination

Redundant and low-reproducible feature elimination and imbalance adjustment

To eliminate the effect of different MRI scanning protocols and improve the classification efficiency of the diagnostic models, a compensation method named "Combat" was used to realign feature distributions computed from different MRI equipment and protocols [19]. Due to the potential uncertainty introduced by manual ROI delineation, some features may have low reproducibility (ICC < 0.75) and would need to be removed to stabilize our model. To identify redundant features, Pearson correlation matrixes were built using pairwise feature correlations. The mean absolute correlation of each feature was calculated. If two features have a high correlation ($r > 0.9$), the feature with the largest mean absolute correlation was removed.

A minimum redundancy maximum relevance (MRMR) method was applied to identify the importance of the remaining features, and the top 50 important features were preserved. Because positive/negative samples were not uniformly distributed, the synthetic minority oversampling technique (SMOTE) method was used to oversample the minority class (positive PLNM) and undersample the majority class (negative PLNM) to balance the data set to improve the classification performance of the machine learning model [20, 21].

Building of radiomics model and validation

After eliminating redundant features, the remainders were processed by the random forest models (number of trees = 500). The variables involved in the random forest model leading to the smallest out-of-bag (OOB) error were selected. The process of model building using random forest is shown in Fig. 2.

The entire original data set was divided into validation set 1: patients from center A, and validation set 2: patients from centers B to E for radiomics model validation. Receiver operating characteristic (ROC) curve and calibration curve were used to assess the diagnostic performance and the goodness of fit of the radiomics model for the training and validation data sets.

Clinical application of the radiomics model

To investigate the clinical application of the radiomics model, two radiologists (radiologist 1 with 5 years and radiologist 2 with 10 years of experience in pelvic MRI)

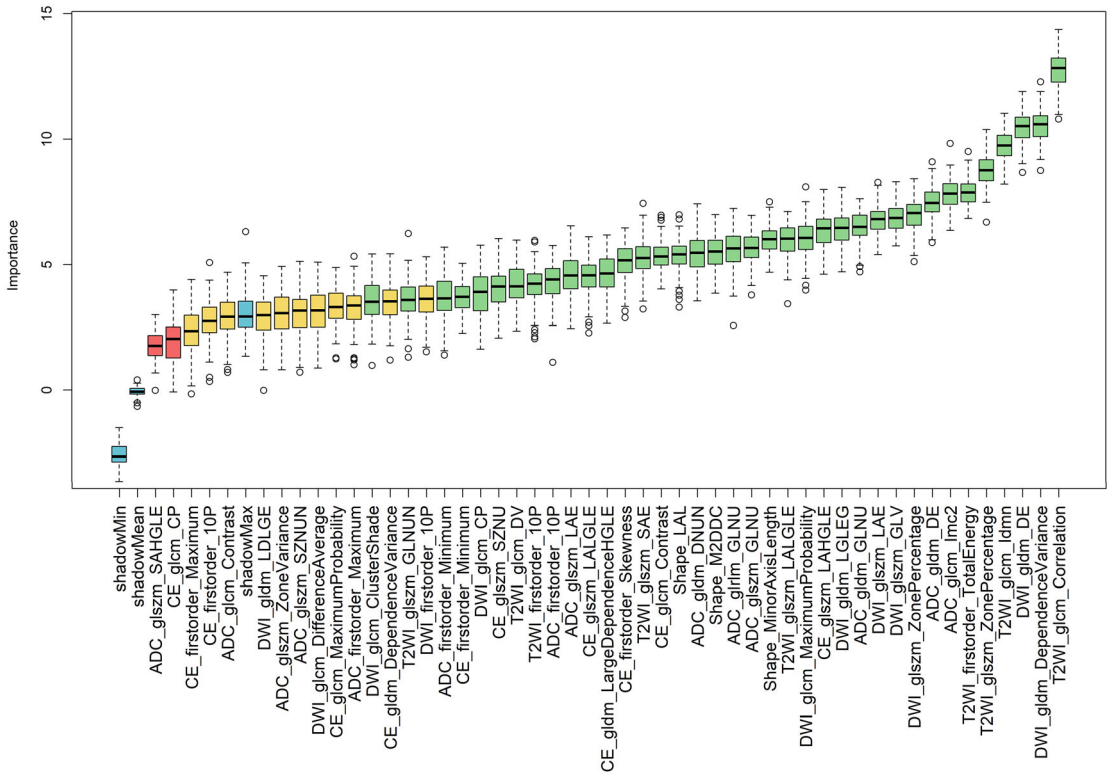

Fig. 2 The top 50 radiomics features (green boxes) associated with PLNM of EC are identified using the training set from the random forest model. The importance of the radiomics features is compared with the importance of shadow attributes (importance threshold, blue boxes) created by shuffling the original attributes. The features that have significantly worse importance than the shadow features (blue boxes) are consecutively dropped (red boxes). On the other hand, attributes that are significantly better than the shadow attributes are included (green boxes). The tentative radiomics features are plotted in yellow boxes. EC, endometrial cancer; PLNM, pelvic lymph node metastasis

who were blinded to histopathological data and the radiomics results reviewed the entire MRI series of each case to identify PLNM basing on the positive criteria of pelvic lymph node (short axis diameter > 8 mm, or with any one of following definite morphological abnormalities: irregular contour, hyperintensity on DWI, or central necrosis and non-homogeneous enhancement on CE-T1WI images) [22]. After a period of 30 days, all cases were repeatedly reviewed by the same radiologists, who were blinded to histopathological data, and given the prediction result of the radiomics model. If a lymph node was > 8 mm and was negative predicted by radiomics, the repeat review found none of LN morphological abnormalities, then a negative PLNM would be reported. If a lymph node was ≤ 8 mm and was positive predicted by radiomics, the repeat review found inconclusive morphological abnormalities, then a positive PLNM would be reported; otherwise, a negative PLNM would be maintained. The performances of the radiomics model and the radiologists with and without the aid of the radiomics model in assessing PLNM were evaluated by ROC curve and compared using the clinical decision curve (CDC), net reclassification index (NRI), and total integrated discrimination index (IDI).

Correlations of radiomics features with immunohistochemical indexes

Finally, correlations between the selected radiomics features and immunohistochemical indexes (ER, PR, P53, and Ki-67) were calculated.

Statistical analysis

All statistical analyses were performed using R software (version 3.6.1; http://www.r-project.org). Student's *t* test was used to compare quantitative variables, and the Mann-Whitney *U* test, chi-squared test, or Fisher's exact test was used to compare qualitative variables. The Pearson or Spearman correlations were calculated to explore associations between radiomics features and immunohistochemical indexes. ROC curve was used to evaluate the diagnostic performance. The DeLong test was used to assess the goodness of fit of the radiomics model and of the radiologists without and with the radiomics model. The inter-rater reliability of the results obtained from the two radiologists was also calculated by Cohen's kappa test. The "ComBatHarmonization" package, "rms" package, "pROC" package, "dca.R" package, and

"PredictABEL" package were used for analyses. A p value < 0.05 was considered statistically significant.

Results

Study participants

A group of 664 patients was reviewed, and 42 patients were excluded. Finally, 622 patients (351 from center A: validation set 1 and a joint group of 271 patients from centers B–E: validation set 2) were included. No significant differences in baseline clinical characteristics between patients with positive and negative PLNM were shown. The patients' clinical characteristics are presented in Table 2. Patients were staged according to the 2014 FIGO classification [23].

Data processing

A total of 358 radiomics features were initially extracted. Out of the 358 features, 4 features (first-order statistics-total energy from T2WI, DWI, the ADC mapping, and CE-T1WI) available in Pyradiomics were not defined in the IBSI. All the remaining features were calculated according to IBSI's definitions. The details are shown in Supplementary Table 1.

Features with either inter-observer or intra-observer ICC < 0.75 were removed, leaving 235/358 features (65.6%). Features with Pearson correlation coefficients > 0.9 were removed, leaving 114 features. After MRMR, 50 features were preserved. After the SMOTE method adjusted the sample imbalance, a ratio of 1:1 (128 positive PLNM vs. 128 negative PLNM) was achieved. After random forests screened, we finally included 37 radiomics features to build the radiomics model. The selected features for constructing the diagnosis model and distribution diagram of importance are shown in Fig. 2 and Supplementary Table 2.

Diagnostic performance

The random forest results showed that the glcm_Correlation feature from T2WI made the greatest contribution to the diagnosis for differentiating positive from negative PLNM. The AUCs of training set, validation set 1, and validation set 2 before and after "ComBat" compensation are shown in Fig. 3a–c, which indicated that after eliminating the protocol effect, a higher PLNM prediction performance was achieved. The calibration curve demonstrated a goodness of fit for the radiomics model in the three data sets (Fig. 3d, e).

By referring radiomics results, lymph node status was changed from positive to negative (n = 23 [3.7%] for radiologist 1, and 11 [1.8%] for radiologist 2) and from negative to positive (n = 34 [5.5%] for radiologist 1, and 51 [8.2%] for radiologist 2). Twelve positive PLNM, which were predicted by radiomics and confirmed by histopathology, could not be found by radiologists and still diagnosed as negative. The AUCs for predicting PLNM with ComBat were 0.935 (95% CI: 0.90.06–0.964) for the training set, 0.909 (95% CI: 0.866–0.951) for validation set 1, 0.885 (95% CI: 0.815–0.955) for validation set 2, 0.623 (95% CI: 0.564–0.683) for the radiologist 1 alone, 0.643 (95% CI: 0.584–0.703) for the radiologist 2 alone, 0.814 (95% CI: 0.756–0.871) for the radiomics-aided radiologist 1, and 0.842 (95% CI: 0.798–0.896) for the radiomics-aided radiologist 2 (Fig. 4). The accuracy was 88.3% for the training set, 80.3% for validation set 1, 88.6% for validation set 2, 84.6% and 88.1% for the radiologists 1 and 2 alone, and 90.2% and 90.4% for the radiomics-aided radiologists 1 and 2. The kappa value of PLNM prediction between two radiologists was 0.81 for radiologists alone and 0.92 for the radiomics-aided radiologists.

Clinical application

All data sets (n = 622) were used for CDC analyses for the radiomics model and the radiologists without and with the aid of the radiomics model for predicting PLNM in EC, which is shown in Fig. 5. The net benefit of the radiologists working with aid of the radiomics model was higher than that of the radiologists alone or the radiomics model alone in the risk (positive PLNM) threshold probabilities' range of 0.1–0.5. The reclassification measures of discrimination confirmed that the radiomics-aided radiologists performed better than the radiologists alone, with an NRI of 1.26 (95% CI: 1.03–1.48) and an IDI of 0.21 (95% CI: 0.16–0.26) for radiologist 1 and an NRI of 1.37 (95% CI: 1.16–1.58) and an IDI of 0.24 (95% CI: 0.19–0.29) for radiologist 2 (Fig. 6).

Correlations between radiomics features and immunohistochemical indexes

As shown in Supplementary Figure 1 and Supplementary Table 3, the co-occurrence matrix plots of PLNM and the immunohistochemical indexes of EC suggest that radiomics features are correlated with ER, PR, P53, and Ki-67.

Discussion

This retrospective multicenter study revealed that the radiomics features extracted from multiparametric MRI could preoperatively add useful information in the assessment of the pelvic lymph node status of EC. Radiomics could aid radiologists to improve their performance in predicting PLNM of EC, especially in helping radiologists to rule out false positive PLNM (> 8 mm) and to rule out false negative PLNM (≤ 8 mm). Furthermore, the radiomics features were correlated

Table 2 Characteristics of included endometrial cancer patients

	Validation set 1	Validation set 2	*p*
Hospital	Center A	Centers B to E	
Number	351	271	
Age	57.13 ± 9.11	55.39 ± 8.51	0.068
Non-menopause /menopause	125/226	95/155	0.879
PLNM			
Negative PLNM	314	244	0.918
Positive PLNM	37	27	
Myometrial invasion			
≤ 50%	260	209	0.434
> 50%	91	62	
Tumor grade			
Low (G1+G2)	286	215	0.570
High (G3)	29	32	
CSI			
Negative CSI	300	237	0.551
Positive CSI	51	34	
LVSI			
Negative LVSI	265	225	0.029
Positive LVSI	86	46	
FIGO staging			
I			
IA	222	180	0.285
IB	49	30	
II	38	20	
III			
IIIA	5	9	
IIIB	0	1	
IIIC1	21	20	
IIIC2	11	6	
IV			
IVA	1	3	
IVB	4	2	
Histological subtypes			
Endometrioid adenocarcinoma	315	247	0.439
Carcinosarcoma	5	6	
Clear cell carcinoma	4	3	
Mixed adenocarcinoma	9	3	
Serous adenocarcinoma	18	11	
Others	0	1	
Immunohistochemistry (mean ± SD)			
ER	0.63 ± 0.32	0.85 ± 6.05	0.554
PR	0.54 ± 0.35	0.44 ± 0.31	0.000
Ki-67	0.31 ± 0.23	0.33 ± 0.24	0.318
P53	0.74 ± 0.39	0.25 ± 0.24	< 0.001
Rad scores	0.10 ± 0.17	0.10 ± 0.13	0.766
Radiologist diagnosis PLNM			
Negative PLNM	318	250	0.483
Positive PLNM	33	21	

CSI, cervical stromal invasion; *DMI*, deep myometrial invasion; *EC*, endometrial carcinoma; *EI*, extrauterine invasion; *ER*, estrogen receptor; *FIGO*, International Federation of Obstetrics and Gynecology; *LA*, lymphadenectomy; *LNM*, lymph node metastasis; *LVSI*, lymphovascular space invasion; *PLNM*, pelvic lymph node metastasis; *PR*, progesterone receptor

with immunohistochemical indexes ER, PR, P53, and Ki-67 of EC.

For EC patients without LNM, a limiting surgery that includes hysterectomy and oophorectomy is optimal without reducing the survival rate, whereas for patients with LNM, a comprehensive staging surgery including lymphadenectomy is mandatory. Therefore, preoperatively assessing lymph node status is crucial to guide the management of patients with EC. Unfortunately, due to the frustrating performance of conventional MRI and PET-CT in detecting LNM [24–26], many early-stage EC patients undergo unnecessary lymphadenectomy, which results in surgical complications, such as increased operation time and bleeding, extending the postoperative days in the hospital and increasing the incidence of lymphedema.

Recent studies suggested that MRI-based texture feature analysis could be helpful for differentiating the presence of LNM in

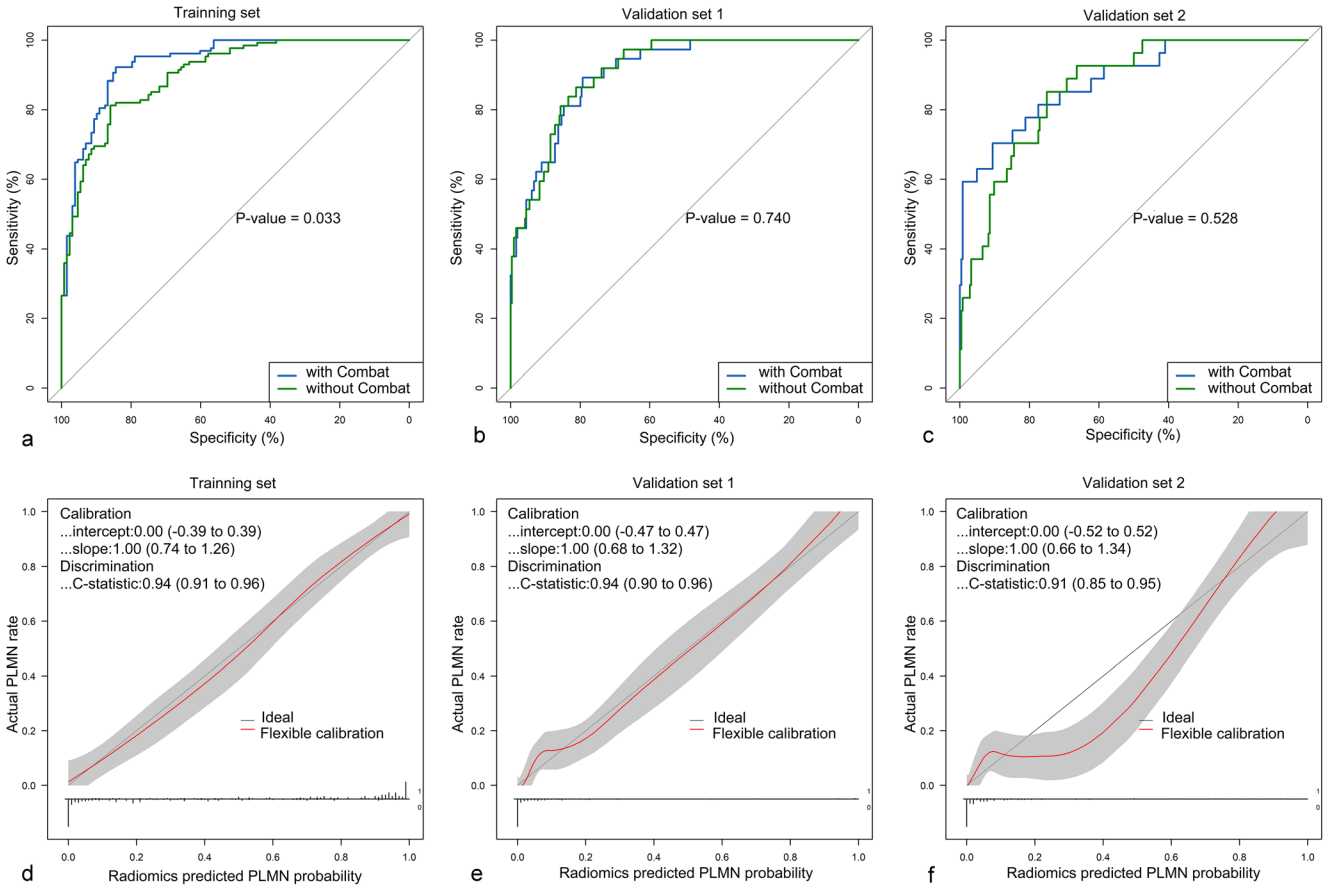

Fig. 3 Areas under the ROC curve of radiomics in diagnosis PLMN with and without ComBat in training set (**a**), validation set 1 (**b**), and validation set 2 (**c**). The calibration curves show good fits for the radiomics model in the training set (**d**), validation set 1 (**e**), and validation set 2 (**f**). The AUC for predicting PLNM without ComBat was 90.1% (95% CI: 86.5–93.7%) for the training set, 91.3% (95% CI: 87.5–95.2%) for validation set 1, and 86.4% (95% CI: 80.2–92.5%) for validation set 2. EC, endometrial cancer; PLNM, pelvic lymph node metastasis

EC patients, with sensitivity, specificity, and accuracy of 0.68, 0.73, and 0.72, respectively [27]. A latest study by Xu et al suggested that the radiomics combined with the clinical parameters (CA125 and lymph node size) achieved an excellent predictive accuracy (especially in normal size lymph nodes) for the LNM in EC [22]. In our study, whole-volume

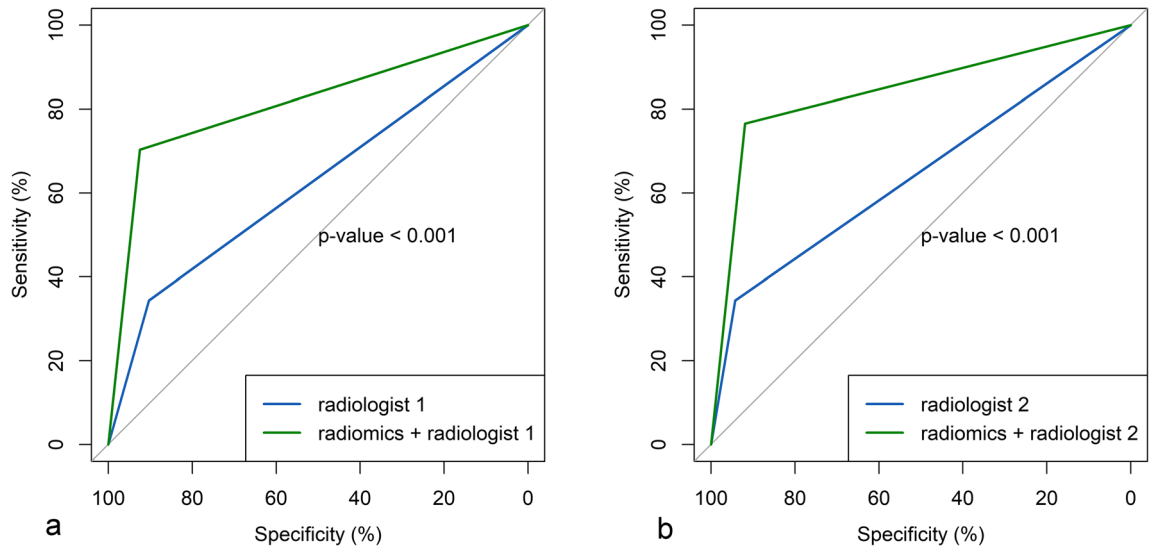

Fig. 4 Areas under the ROC curve of radiologist 1 (**a**) and radiologist 2 (**b**) with and without radiomics aids in diagnosis of PLMN in EC. EC, endometrial cancer; PLNM, pelvic lymph node metastasis

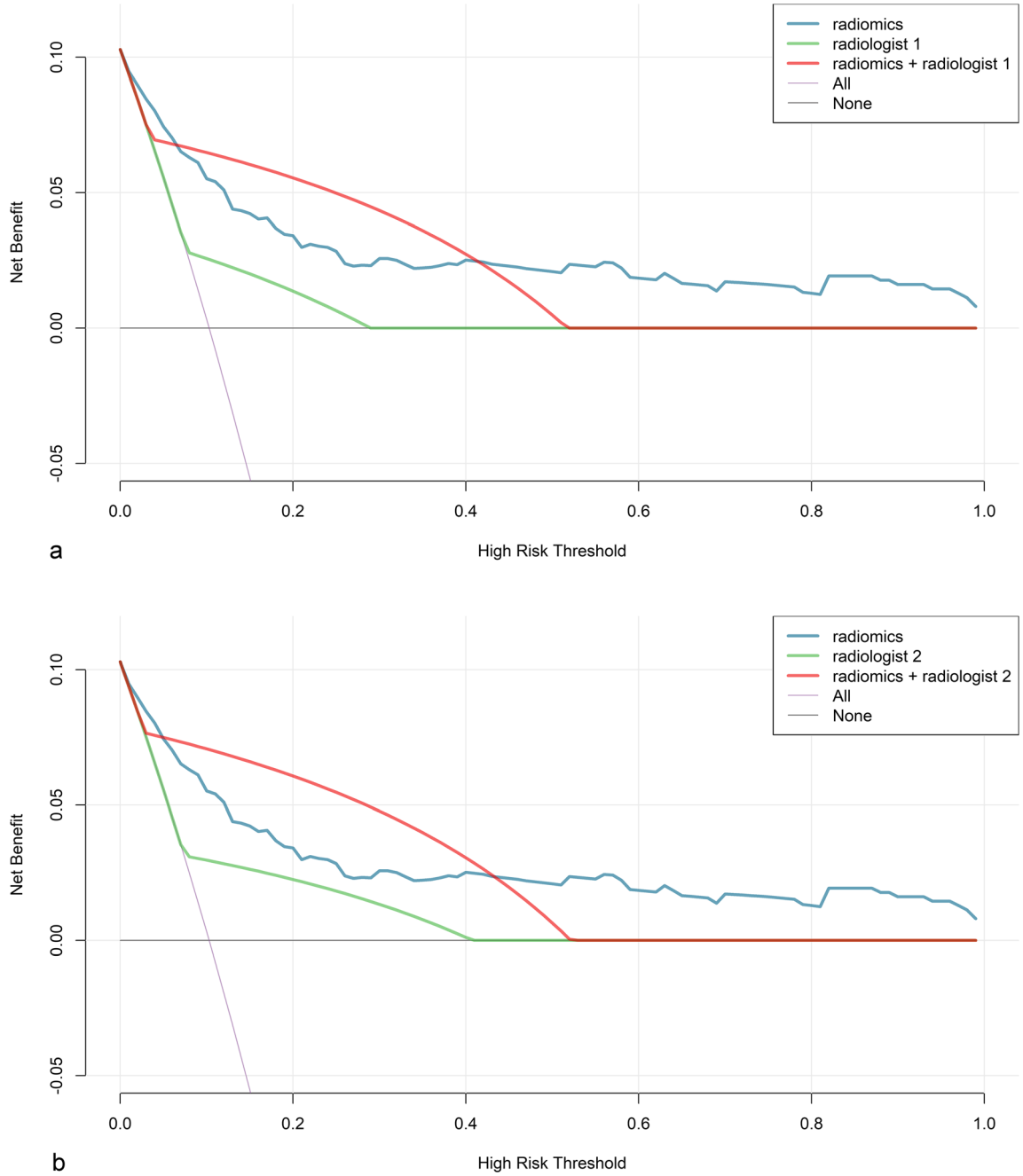

Fig. 5 Clinical decision curve analyses for the radiologist 1 (**a**) and radiologist 2 (**b**) with and without radiomics aid for predicting PLNM in EC. The *y*-axis measures the net benefit, calculated by subtracting the proportion of all patients who were false positive from the proportion who were true positive, weighted by the relative harm of forgoing treatment compared with the negative consequences of an unnecessary treatment. EC, endometrial cancer; PLNM, pelvic lymph node metastasis

multiparametric MRI radiomics features were extracted based on a relatively large sample size and multicenter data, which may have contributed to a good diagnostic performance by unraveling more comprehensive information about tumor heterogeneity [7, 28]. Our results showed that this MRI radiomics model had a good diagnostic performance for the assessment of pelvic lymph node status in patients with EC. The selected features for PLNM such as Shape_MinorAxisLength were consistent with those of a previous study [29]. In addition, owing to the low positive PLNM rate of EC patients, the SMOTE method was used to balance the data set to improve the classification performance of a machine learning model. Further validation of the model was performed using two validation sets and resulted in AUCs of 0.909 and 0.885 without significant fluctuation. These results indicated that this computer-based data analysis method could be a helpful tool to assess the presence of PLNM in EC patients.

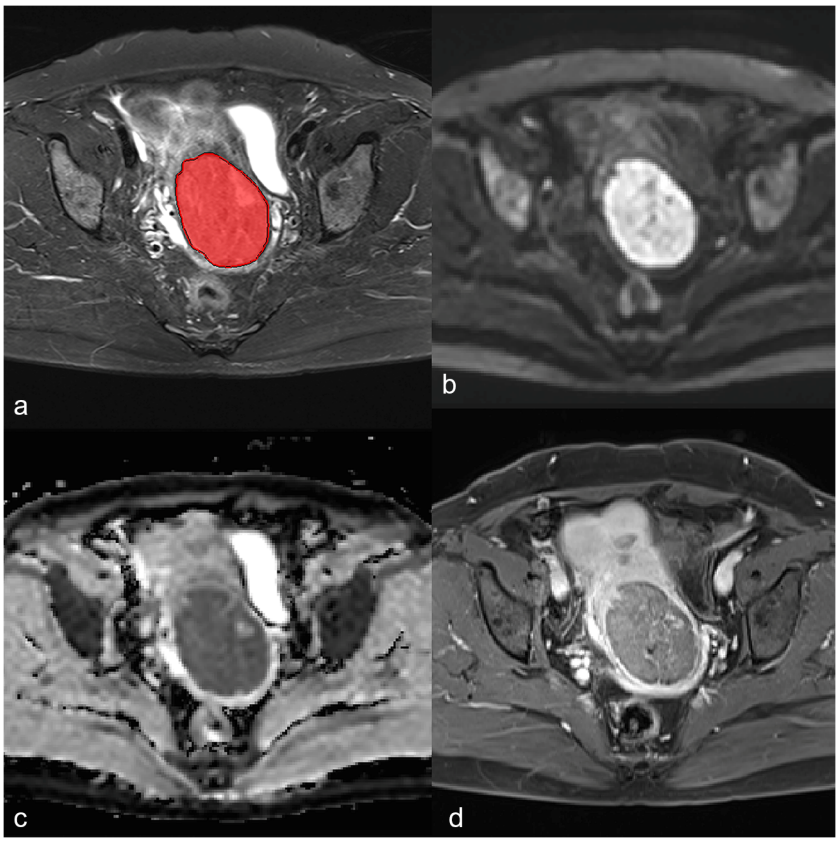

Fig. 6 MR images in a 54-year-old woman with endometrial cancer. **a** Axial T2-weighted imaging (T2WI) is marked with a region of interest. **b** Axial apparent diffusion coefficient imaging. **c** Axial diffusion-weighted imaging (DWI) ($b = 1000$ s/mm^2). **d** Axial contrast-enhanced T1-weighted imaging (CE-T1WI)

CDC analysis was applied to evaluate the net benefit of the radiomics model in aiding the radiologist in predicting PLNM of EC. The net benefit was calculated by subtracting the proportion of all patients who were false positive from the proportion who were true positive, weighted by the relative harm of forgoing treatment compared with the negative consequences of an unnecessary treatment. The results showed that the radiologists could have the higher net benefit with the aid of the radiomics model in the risk (positive PLNM) range of 0.1–0.5. Furthermore, given the known limitations of measures of association and ROC curves [30], we used the reclassification framework to provide an outcome prediction analysis of clinical decision-making. Our reclassification results showed that the clinical benefits were significantly improved, with IDIs of 0.21 and 0.24, which indicated that out of 100 patients, more than 20 patients benefited from the radiomics-aided radiologists' prediction compared with the radiologists alone. Moreover, we found that under certain conditions, when radiologists were informed of the radiomics prediction of PLNM, they could not find potential metastatic lymph nodes confirmed by histopathology. The reason for this disadvantage could be attributed to the presence of small size metastatic lymph nodes, partially obscured by the uterus, to the partial volume effect or not clearly visible due to the relatively limited spatial resolution of MRI. The use of radiomics models for the analysis of EC could be an additional noninvasive method to improve MRI preoperative staging of EC, with a good capability to predict lymph node status before surgery.

Many studies have indicated that some immunohistochemical indexes, such as ER, PR, P53, and Ki-67, are correlated with LNM [17, 18]. Our study explored the correlation of immunohistochemical indexes with the radiomics features that contributed to the assessment of PLNM. The expression of ER and PR can provide information for endocrinotherapy. The expression of Ki-67, an extensively investigated marker of cell proliferation, reflects the proportion of malignant cells and is associated with tumor progression and metastasis, as well as prognosis. P53 can be used for redefining the POLE mutation in EC [31]. Considering the correlation of the radiomics features with the immunohistochemical indexes of EC, we should be able to obtain comprehensive information on the tumors in EC patients.

Our study had some limitations. First, we excluded five patients who had para-aortic LNM without PLNM because not all of the included patients underwent abdominal MRI scans according to daily clinical routines. Second, the CE-T1WI scanning time and *b* value of the DWI were not uniform, since this was a retrospective study. Third, we did not draw the ROI based on each visible pelvic lymph node because it was time consuming and difficult to match every lymph node between images and postoperative resection

samples. Fourth, we did not include the high-order wavelet features because during our analysis process, the wavelet features were not stable and lacked reasonable clinical interpretation [32].

Conclusion

MRI-based radiomics analysis could be used to predict the presence of PLNM in EC. The radiomics model could aid radiologists in improving their performance for assessing PLNM in EC. The radiomics features had correlations with the immunohistochemical indexes of EC.

Funding information This study has received funding from the National Natural Science Foundation of China (No. 81971579).

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Jin Wei Qiang.

Conflict of interest The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry One of the authors has significant statistical expertise.

Informed consent Written informed consent was not required for this study because of the retrospective nature of the study.

Ethical approval Local institutional Review Board approval was obtained.

Methodology
- retrospective
- diagnostic study
- multicenter study

References

1. Straughn JMJ, Huh WK, Kelly FJ et al (2002) Conservative management of stage I endometrial carcinoma after surgical staging. Gynecol Oncol 84:194–200
2. Cragun JM, Havrilesky LJ, Calingaert B et al (2005) Retrospective analysis of selective lymphadenectomy in apparent early-stage endometrial cancer. J Clin Oncol 23:3668–3675
3. Kitchener H, Swart AM, Qian Q, Amos C, Parmar MK (2009) Efficacy of systematic pelvic lymphadenectomy in endometrial cancer (MRC ASTEC trial): a randomised study. Lancet 373: 125–136
4. Creutzberg CL, van Putten WL, Koper PC et al (2000) Surgery and postoperative radiotherapy versus surgery alone for patients with stage-1 endometrial carcinoma: multicentre randomised trial. PORTEC Study Group. Post Operative Radiation Therapy in Endometrial Carcinoma. Lancet 355:1404–1411
5. Bi Q, Chen Y, Wu K et al (2020) The diagnostic value of MRI for preoperative staging in patients with endometrial cancer: a meta-analysis. Acad Radiol 27:960–968
6. Stewart KI, Chasen B, Erwin W et al (2019) Preoperative PET/CT does not accurately detect extrauterine disease in patients with newly diagnosed high-risk endometrial cancer: a prospective study. Cancer 125:3347–3353
7. Gillies RJ, Kinahan PE, Hricak H (2016) Radiomics: images are more than pictures, they are data. Radiology 278:563–577
8. Kumar V, Gu Y, Basu S et al (2012) Radiomics: the process and the challenges. Magn Reson Imaging 30:1234–1248
9. Lambin P, Rios-Velazquez E, Leijenaar R et al (2012) Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 48:441–446
10. Rizzo S, Botta F, Raimondi S et al (2018) Radiomics: the facts and the challenges of image analysis. Eur Radiol Exp 2:36
11. Ji GW, Zhang YD, Zhang H et al (2019) Biliary tract cancer at CT: a radiomics-based model to predict lymph node metastasis and survival outcomes. Radiology 290:90–98
12. Wu S, Zheng J, Li Y et al (2018) Development and validation of an MRI-based radiomics signature for the preoperative prediction of lymph node metastasis in bladder cancer. EBioMedicine 34:76–84
13. Wibmer A, Hricak H, Gondo T et al (2015) Haralick texture analysis of prostate MRI: utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores. Eur Radiol 25:2840–2850
14. Gu D, Hu Y, Ding H et al (2019) CT radiomics may predict the grade of pancreatic neuroendocrine tumors: a multicenter study. Eur Radiol 29:6880–6890
15. Ueno Y, Forghani B, Forghani R et al (2017) Endometrial carcinoma: MR imaging-based texture model for preoperative risk stratification-a preliminary analysis. Radiology 284:748–757
16. De Bernardi E, Buda A, Guerra L et al (2018) Radiomics of the primary tumour as a tool to improve (18)F-FDG-PET sensitivity in detecting nodal metastases in endometrial cancer. EJNMMI Res 8: 86
17. Yu C, Jiang X, Li B, Gan L, Huang J (2015) Expression of ER, PR, C-erbB-2 and Ki-67 in endometrial carcinoma and their relationships with the clinicopathological features. Asian Pac J Cancer Prev 16:6789–6794
18. Gülseren V, Kocaer M, Özdemir İA, Çakır İ, Sancı M, Güngördük K (2020) Do estrogen, progesterone, P53 and Ki67 receptor ratios determined from curettage materials in endometrioid-type endometrial carcinoma predict lymph node metastasis? Curr Probl Cancer 44:100498
19. Orlhac F, Frouin F, Nioche C, Ayache N, Buvat I (2019) Validation of a method to compensate multicenter effects affecting CT radiomics. Radiology 291:53–59
20. Seo JH, Kim YH (2018) Machine-learning approach to optimize SMOTE ratio in class imbalance dataset for intrusion detection. Comput Intell Neurosci 2018:9704672
21. Feng Z, Rong P, Cao P et al (2018) Machine learning-based quantitative texture analysis of CT images of small renal masses: differentiation of angiomyolipoma without visible fat from renal cell carcinoma. Eur Radiol 28:1625–1633
22. Xu X, Li H, Wang S et al (2019) Multiplanar MRI-based predictive model for preoperative assessment of lymph node metastasis in endometrial cancer. Front Oncol 9:1007
23. FIGO Committee on Gynecologic Oncology (2014) FIGO staging for carcinoma of the vulva, cervix, and corpus uteri. Int J Gynaecol Obstet 125:97–98
24. Hu J, Zhang K, Yan Y, Zang Y, Wang Y, Xue F (2019) Diagnostic accuracy of preoperative (18)F-FDG PET or PET/CT in detecting pelvic and para-aortic lymph node metastasis in patients with endometrial cancer: a systematic review and meta-analysis. Arch Gynecol Obstet 300:519–529

25. Bian L, Wang M, Gong J et al (2019) Comparison of integrated PET/MRI with PET/CT in evaluation of endometrial cancer: a retrospective analysis of 81 cases. PeerJ 7:e7081
26. Duncan KA, Drinkwater KJ, Frost C, Remedios D, Barter S (2012) Staging cancer of the uterus: a national audit of MRI accuracy. Clin Radiol 67:523–530
27. Ytre-Hauge S, Dybvik JA, Lundervold A et al (2018) Preoperative tumor texture analysis on MRI predicts high-risk disease and reduced survival in endometrial cancer. J Magn Reson Imaging 48: 1637–1647
28. Sala E, Mema E, Himoto Y et al (2017) Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. Clin Radiol 72:3–10
29. Lavaud P, Fedida B, Canlorbe G, Bendifallah S, Darai E, Thomassin-Naggara I (2018) Preoperative MR imaging for ESMO-ESGO-ESTRO classification of endometrial cancer. Diagn Interv Imaging 99:387–396
30. Cook NR (2007) Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation 115:928–935
31. Talhouk A, McConechy MK, Leung S et al (2015) A clinically applicable molecular-based classification for endometrial cancers. Br J Cancer 113:299–310
32. Fiset S, Welch ML, Weiss J et al (2019) Repeatability and reproducibility of MRI-based radiomic features in cervical cancer. Radiother Oncol 135:107–114