Am J Cancer Res 2024;14(11):5400-5410
www.ajcr.us /ISSN:2156-6976/ajcr0160875

Original Article

Machine learning models using multiparametric MRI for preoperative risk stratification in endometrial cancer

Vu Pham Thao Vy[1,2], Jerry Chin-Wei Chien[3,4], Wiwan Irama[5], Hao-Yang Wu[3], Tzu-I Wu[6,7,8], Wei-Yu Chen[9,10], Chia-Hao Liang[11], Truong Nguyen Khanh Hung[12], Wilson T Lao[3,4], Wing P Chan[3,4]

[1]International Ph.D. Program in Medicine, College of Medicine, Taipei Medical University, Taipei 110, Taiwan; [2]Department of Radiology, Thai Nguyen National Hospital, Thai Nguyen 24000, Vietnam; [3]Department of Radiology, Wan Fang Hospital, Taipei Medical University, Taipei 116, Taiwan; [4]Department of Radiology, School of Medicine, College of Medicine, Taipei Medical University, Taipei 110, Taiwan; [5]Department of Radiology, Keelung Hospital, Ministry of Health and Welfare, Keelung 201, Taiwan; [6]Department of Obstetrics and Gynecology, Wan Fang Hospital, Taipei Medical University, Taipei 116, Taiwan; [7]Cancer Center, Wan Fang Hospital, Taipei Medical University, Taipei 116, Taiwan; [8]Department of Obstetrics and Gynecology, School of Medicine, College of Medicine, Taipei Medical University, Taipei 110, Taiwan; [9]Department of Pathology, Wan Fang Hospital, Taipei Medical University, Taipei 116, Taiwan; [10]Department of Pathology, School of Medicine, College of Medicine, Taipei Medical University, Taipei 110, Taiwan; [11]Ph.D. Program of Interdisciplinary Medicine, National Yang Ming Chiao Tung University, Taipei 112, Taiwan; [12]Department of Orthopedic and Trauma, Cho Ray Hospital, Ho Chi Minh 70000, Vietnam

Received September 30, 2024; Accepted November 10, 2024; Epub November 15, 2024; Published November 30, 2024

Abstract: This study evaluated the efficacy of machine learning and radiomics of preoperative multiparameter MRIs in predicting low- vs high-risk histopathologic features and early vs advanced FIGO stage (IA vs IB or higher) in endometrial cancer. This retrospective study of patients with endometrial cancer histologically confirmed from 2008 through 2023 excluded those with: (a) previous treatment for endometrial carcinoma, (b) incomplete MRI examinations or low-quality MR images, (c) incomplete pathology reports, (d) non-visualized tumors on MRI, or (e) distant metastases. In total, 110 radiomic features were extracted using commercial PACS built-in software following segmentation after sagittal T2-weighted imaging (T2WI), contrast enhanced T1-weighted imaging (CE-T1WI), and diffusion weighted imaging (DWI). The radiomic features from each imaging sequence were utilized for initial modeling. A combined model, which included features retained from all 3 sequences, was then established. The area under the receiver operating characteristic curve (AUC) determined the efficacy of each model. For 5 specific histopathologic features, the combined model achieved AUCs of 0.87 (95% CI, 0.85-0.90), 0.90 (95% CI, 0.88-0.92), 0.88 (95% CI, 0.87-0.90), 0.88 (95% CI, 0.86-0.92), and 0.87 (95% CI, 0.86-0.90). This model incorporated 38 radiomic features: 12 from T2WI, 17 from CE-T1WI, and 9 from DWI. In conclusion, an MRI radiomics-based model can differentiate between early- and advanced-stage endometrial cancer and between low- and high-risk histologic markers, giving it the potential to predict high risk and stratify preoperative risk in those with endometrial cancer. The findings may aid personalized preoperative assessments to guide clinical decision-making in endometrial cancer.

Keywords: Endometrial cancer, machine learning, magnetic resonance imaging (MRI), radiomics, uterus

Introduction

Endometrial cancer (EC) is a type of cancer that arises from the inner epithelial lining of the uterus. More than 417000 people were diagnosed with EC in 2020 [1], and it is estimated that more than 67000 people in the US will be diagnosed with cancer of the uterus in 2024 [2]. Although EC occurs most frequently in postmenopausal women, each year, a greater proportion of younger women is diagnosed with EC in Asia [3, 4]. Therefore, it is important to classify patients at an early age and manage them appropriately. The International Federation of Gynecology and Obstetrics (FIGO) staging method divides all ECs into 1 of 4 pathologic stages [5]. People with stage I comprise 80% of all ECs, and the 5-year survival rate exceeds 95%, the highest of all gynecological cancers [6]. Comparatively, those with stage III or stage IV

ECs have 5-year survival rates of 50-65% [7] or 15%-19% [8, 9], respectively. Stage I-II ECs are considered to carry low or intermediate risks and have low recurrence rates and favorable overall survival rates [10]. Early diagnosis is associated with a better prognosis and a greater 5-year survival rate; therefore, an accurate initial diagnosis and timely treatment play important roles in EC management. Importantly, because the treatment plan and follow-up therapy depend on the stage of EC, detection during the early stages can decrease the need for extensive surgical intervention or adjuvant therapies, consequently reducing medical costs, morbidity, and mortality.

Magnetic resonance imaging (MRI) plays roles in diagnosing EC, detecting advanced disease, strategizing radiation ports, tracking response to treatment, and evaluating post-treatment surveillance to detect recurrence. In conventional MRI, ECs are viewed using T1- and T2-weighted imaging (T2WI) as well as contrast-enhanced (CE) imaging to detect malignant lesions [11]. Both T2WI and CE imaging along with diffusion-weighted imaging (DWI) are essential for evaluating myometrial invasion of EC. Because DWI enhances tumor detection and characterization as well as visualization of small implants in peritoneal carcinomatosis, it is an essential MRI sequence. In clinical practice, the DWI protocol should incorporate at least 1, but preferably 2, planes: an axial oblique plane along the uterus and a sagittal plane. Obtaining both T2WI and DWI on the same plane allows image fusion and enhances anatomical correlation. To avoid pitfalls, it is important to evaluate DWI images with the corresponding apparent dispersion coefficient (ADC) maps and anatomical images. Several studies explored the role of DWI with ADC for evaluating tumor grade, but results proved ambiguous [12-15]. In EC, incorporating CE T1WI with T2WI improved diagnostic accuracy up to 92% [16, 17].

The stage of EC as determined by imaging depends on visual assessments by trained radiologists, and variability will be found between observers and institutional protocols [18]. The imaging tumor profile includes a large number of quantitative imaging characteristics that can be used in multidimensional models for defining diseases and predicting clinical and biological results, opening the way to a promising approach for tailoring treatment strategies. Several recent studies have also examined MRI-based tumor characteristics in EC and linked them to aggressive phenotypes [18-20]. Radiomics is a cost-effective and non-invasive method used to characterize tissue intensity, shape, and texture by quantifying the imaging phenotype of the region of interest (ROI). The process involves several fundamental steps, such as image acquisition and preprocessing, ROI annotation, feature extraction and selection, and model construction and prediction, with the primary objective of connecting extensive extracted image information with clinical and biological data. Radiomics serves not only as a clinical decision making tool but also as a research tool for uncovering new molecular disease pathways [21, 22]. Hence, we hypothesized that using MRI-based radiomics can enhance differentiation between the various types of EC.

We performed this study to assess the diagnostic accuracy of 3-dimensional radiomics-based XGBoost models using MRI to differentiate low-risk from high-risk histopathologic features and early FIGO stage (IA) from advanced FIGO stage (IB or higher) EC. We extracted radiomic features to establish 4 radiomic models: 3 single-sequence models, each based on 1 imaging technique, and a combined model that used radiomic features from all 3 sequences. The models were compared in their abilities to predict histopathologic features in EC.

Materials and methods

This retrospective study was approved by the Joint Institutional Review Board of the Taipei Medical University (Approval No. N 202312087). The need for informed consent was waived by the Joint Institutional Review Board of the Taipei Medical University because of the retrospective nature of the study. All methods were performed in accordance with standardized guidelines and regulations.

Study participants

We identified potential study participants using the cancer registry database at our hospital, seeking those who received diagnoses of histologically confirmed EC between January, 2008 and December, 2023 after undergoing total hysterectomies with bilateral salpingo-oophorectomies subsequent to 1.5-T MRI. We exclud-

ed those who had (a) previous treatments for EC, (b) incomplete MRI protocols or low-quality MR images, such as motion artifacts, (c) incomplete pathology reports, (d) non-visualized MRI tumors to obtain the ground truth, or (e) distant metastases. The final sample was randomly divided into 2 groups, a training set and a test set.

MR imaging

A 3-T GE scanner (Discovery MR750; GE Healthcare, Waukesha, WI) or a 1.5-T Siemens scanner (MAGNETOM Avanto 1.5T, Siemens Healthcare, Erlangen) with phased-array abdominal coils was used to obtain MRI sequences. All patients were required to breathe freely while in the supine position during data acquisition. The following sequences were obtained: axial T1WI, axial T2WI, coronal and sagittal T2-weighted imaging with fat saturation, DWI with b values of 0 and 880 s/mm^2 and an ADC map, and CE-T1WI with fat saturation. All images were performed with the following settings: matrix, 512 × 512; FOV, 207 × 207 mm^2; and slice thickness, 3.0 mm.

Original MRI staging reports were also recorded for comparison. These reports included image reading by 2 gynecologic radiologists who followed the imaging criteria for staging EC [23].

Surgical histopathologic analyses

Every patient underwent a comprehensive surgical procedure including total hysterectomy, bilateral salpingo-oophorectomy, and bilateral pelvic lymphadenectomy sampling. The resected uterus was meticulously sliced every 3-4 mm to ensure thorough histopathologic examination, and the slices were stained for analysis with hematoxylin and eosin. Gynecologic pathologists examined the specimens to determine tumor subtype, histopathologic grade, the presence of deep myometrial invasion (MI), adenomyosis status, and lymphovascular space invasion (LVSI). Tumor grade was either low grade (including FIGO grades 1 and 2 endometrioid carcinoma) or high grade (FIGO grade 3 endometrioid carcinoma or non-endometrioid histologic findings for the purposes of radiomic modeling). To assess the risk of prognosis, grade was further divided into early stage (IA) or advanced stage (IB or higher) according to the European Society for Medical Oncology guidelines [24].

Radiomic feature extraction and analysis

Tumor ROIs were segmented on sagittal T2WI, CE-T1WI, and DWI because this is well-accepted for identifying the region of EC. Using QUIBIM, a PACS built-in software, 2 senior radiologists manually delineated the ROIs, and lesions were identified and manually segmented on the images from each MRI sequence. The ROIs were defined as areas of abnormal hypointensity compared to the normal endometrium. All contours were reviewed by a radiologist with more than 20 years of experience in gynecological MRI. Contour modifications, if necessary, were manually performed by a senior radiologist, and consensus between 2 readers was reached for annotation. All radiologists were blinded to the clinical and histopathologic results. These annotated images were deemed to act as ground truth for further model training and validation datasets. Using QUIBIM, 110 radiomic features were extracted after segmenting pelvic MR images, and the radiomic results were stored in a database. The extracted features were divided into eight groups: (i) first-order statistics, (ii) shape-based 3-dimensional features, (iii) shape-based 2-dimensional features, (iv) grey-level co-occurrence matrices, (v) grey-level run length matrices, (vi) grey-level size zone matrices, (vii) neighboring grey tone difference matrices, and (viii) grey-level dependence matrices. For each set of preprocessing parameters, radiomic features were extracted; the set demonstrating the greatest feature stability under ROI variations was chosen. Radiomic features preselected for their reproducibility, lack of correlation, and discriminatory power were the only ones utilized in constructing the XGBoost model.

Feature stability was assessed using a 2-way random effect, single rater, absolute agreement model to obtain the intraclass correlation coefficient (ICC) based on data extracted from the annotations of 2 readers. Features were deemed stable if the lower bound of the ICC 95% confidence interval was ≥ 0.75. A MinMax scaler (range, 0-1) was fitted to the training data and then used to transform the dataset. Subsequently, features with low variance (< 0.01) and those that were highly intercorrelated (Pearson correlation coefficient ≥ 0.80) were removed. Using the remaining features, an initial XGBoost classifier was trained to assess the predictive power of each feature. The most

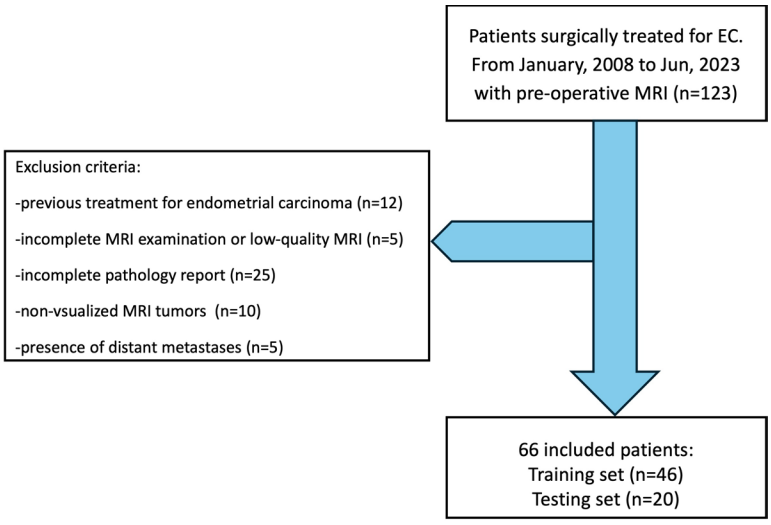

Figure 1. Patient selection flowchart.

predictive features were selected using SelectKBest, and a final XGBoost classifier, trained using these features, was applied to an independent test set. Models were trained on the training set for histopathologic feature classification, and diagnostic performance was then evaluated on both the training and test sets, using bootstrapping to generate confidence intervals.

Statistical analyses of the diagnostic performances of radiomics versus radiologists

The radiomic features from each image sequence (CE-T1W, T2W, and DWI) were utilized to establish the initial models. A combined model, which incorporated the features from all 3 sequences, was also established, then dimension reduction was applied.

Based on the retained features, the final XGBoost classifier was trained on the training set, then applied to the test set. Its diagnostic performance and that of radiologist staging were compared to the ground truth as determined by histology. The estimated value of the extracted volume was reported in the form of an area under the curve (AUC). The DeLong test was used to compare the AUCs of various receiver operating characteristic curves, thus determining diagnostic efficacy. The correlations between each radiomic characteristic and histopathological results were investigated, finding either the Spearman rank correlation coefficient or the points-biserial correlation coefficient, which were plotted on a correlation heat map. The Mann-Whitney test (or Kruskal-Wallis test, as appropriate) was used to nonparametrically compare clinical variables with the radiomic characteristics in each set.

Results

From January, 2008 to December, 2023, a total of 123 patients underwent surgical treatment for EC. The application of exclusion criteria excluded 57 patients, resulting in a study set of 66 patients (**Figure 1**). Of these, 46 were assigned to the training set (median age, 56 years; range, 34-81 years), and 20 were assigned to the test set (median age, 56 years; range, 35-80 years). In total, 31 diagnoses were confirmed as early FIGO stage (IA), and 35 were confirmed as advanced FIGO stage (IB or higher). All clinical and pathologic information is presented in **Table 1**.

Using preoperative MRI, radiologists correctly assessed FIGO stage in 53 ECs (see Tables S1 and S2). Of the 33 ECs found to be stage IA, 28 were correctly classified; however, 3 were actually stage IB, 1 was actually stage II, and 1 was actually stage III. The sensitivity and specificity of MRI in stage IA were 90.3% and 87.8%, respectively. Compared with postoperative pathology, the diagnostic accuracy of preoperative MRI was 0.80 (53 of 66 patients).

The ROI of each tumor was manually drawn independently (**Figure 2**) by 2 radiologists. A total of 110 radiomic features were extracted across all 3 sequences based on QUIBIM after segmentation. The constituent ratios of features yielding acceptable ICCs (≥ 0.75) were 87.3% (96 of 110) on CE-T1WI, 81.2% (90 of 110) on T2WI, and 78.2% (86 of 110) on DWI. Features that remained consistent despite changes in image preprocessing parameters and variations in contour delineation (with ICC ≥ 0.75) and which were also not highly correlated with each other or had low variance were retained for subsequent analysis (**Figure 3**).

Table 1. Demographic characteristics and histopathologic findings in both the training and test sets of patients

Characteristics	Training set (N=46)	Test set (n=20)	*P* Value
Age (mean ± SD, year)	56.13	56.6	
Overall FIGO stage			0.27
IA	22	9	
IB-IV	24	11	
Deep myometrial invasion			0.48
Absent	9	4	
Present	37	16	
Lymphovascular space invasion			0.18
Absent	29	12	
Present	17	8	
Histopathologic grade			0.54
Low (grade 1 or 2)	41	17	
High (grade 3 or non endometrioid)	6	3	
Pelvis lymph node metastasis			0.51
Absent	38	17	
Present	8	3	
Adenomyosis			0.16
Absent	31	13	
Present	15	7	

From CE-T1WI, QUIBIM extracted 110 radiomic features, 14 of which were excluded from further analysis because of poor reproducibility (ICC < 0.75). The remaining 96 features were retained for the next step.

Using the variance and pairwise correlation requirements, another 28 features were excluded, resulting in 68 features retained in total. Using SelecKBest, the top 25 features were selected for training the XGBoost model. Similar processes were used on the T2WI and DWI sequences. Finally, a combined model was established by merging the retained radiomic features from all 3 sequences, and this model was compared to each single-sequence model.

The AUCs for classifying deep MI, FIGO stage, grade, adenomyosis status, and LVSI were 0.88 (95% confidence interval [CI], 0.83-0.89), 0.74 (95% CI, 0.71-0.76), 0.73 (95% CI, 0.70-0.76), 0.80 (95% CI, 0.77-0.82), and 0.79 (95% CI, 0.77-0.83), respectively (**Figure 4**). The accuracy metrics of all models are shown in **Table 2**. The DeLong test showed no statistically significant differences in AUCs among the 3 models (all P > 0.05).

The combined model demonstrated AUCs of 0.87 (95% CI, 0.85-0.90), 0.90 (95% CI, 0.88-0.92), 0.88 (95% CI, 0.87-0.90), 0.88 (95% CI, 0.86-0.92), and 0.87 (95% CI, 0.86-0.90) for the same 5 histopathologic features, respectively. This model incorporated 38 radiomic features (17 from CE-T1WI, 12 from T2WI, and 9 from DWI). The 10 features with the greatest importance scores based on the combined model are shown in Table S3. The combined model performed better than each individual model (all P < 0.05).

Discussion

In this study, radiomics-based machine learning models were developed to noninvasively identify high-risk histopathologic features of ECs using multiparametric MRI. We found that a combined model based on XGBoost algorithms performed well when AUC was the basis for assessment (deep MI, 0.87; FIGO stage, 0.90; grade, 0.88; adenomyosis status, 0.88; and LVSI, 0.87). Comparatively, the radiomics analysis was superior to our human specialist in MRI staging EC (0.79).

Several recent studies of EC have investigated the relationships between radiomic features in MRI and histopathologic results. Some of these studies used a single MRI sequence for radiomic feature extraction, while others extracted features from 2 or 3 planes of up to 2 sequences. In this study, we focused only on the sagittal plane as viewed using CE-T1WI, T2WI, and DWI.

A major challenge in treatment planning and prognostication for EC is the preoperative assessment of risk factors such as deep MI, LVSI, adenomyosis status, and nodal metastasis, all of which are critical for tailoring surgery and subsequent therapy. Standard MRI has demonstrated high sensitivity (81%-90%) and specificity (82%-91%) in evaluating deep MI [25-27]. A systematic review reported that the use of MRI for assessing LVSI in EC is diagnostically accurate (AUC, 0.82; sensitivity, 73%; and

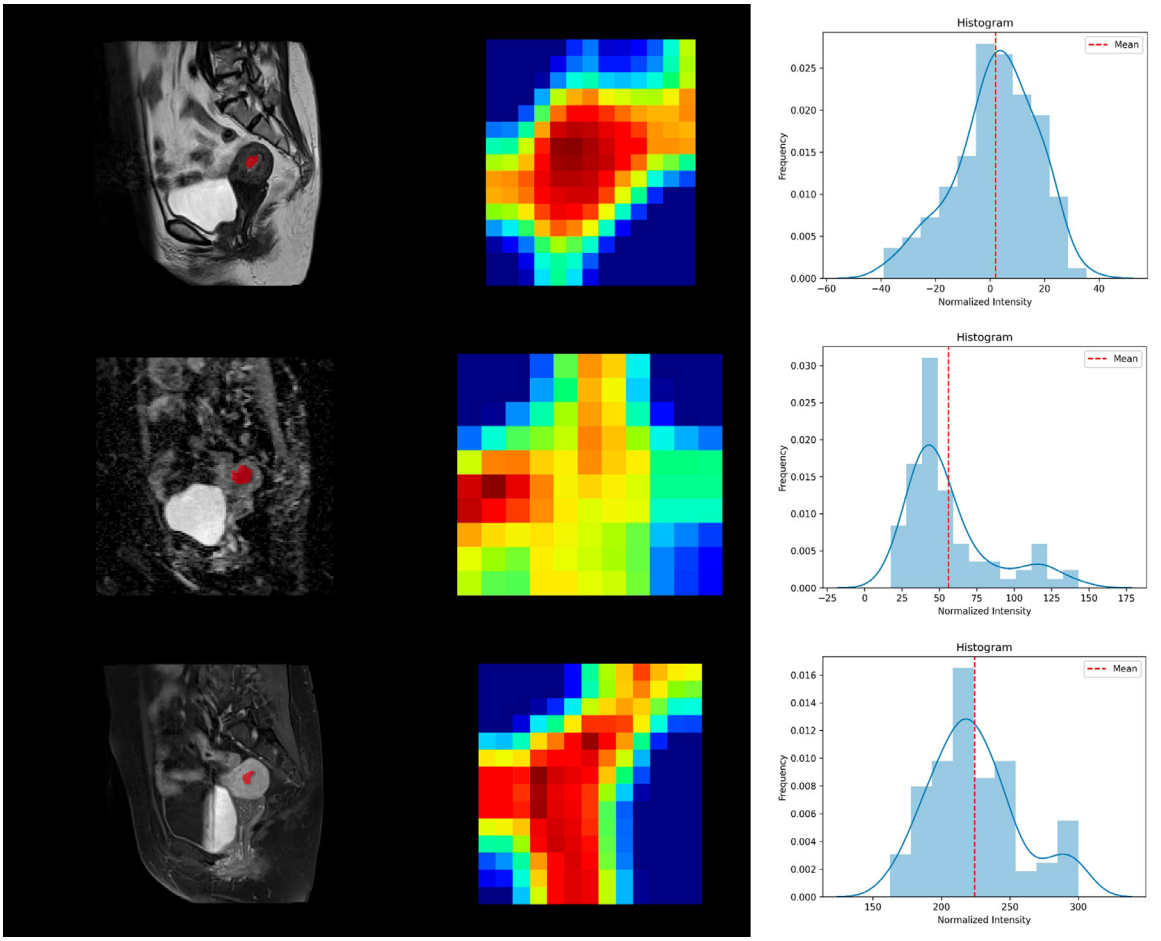

Figure 2. Tumor segmentation, feature extraction, and histogram analysis using QUIBIM. The region of interest was segmented on the sagittal plane in each imaging modality: T2WI, DWI, and CE-T1WI.

specificity, 77%) [28]. However, the accuracy of standard qualitative MRI evaluation is highly dependent on reader experience, leading to considerable interobserver variability. For example, Chen et al. [29] reported an AUC of 0.847 when identifying deep MI, whereas Dong et al. [30] reported an AUC of 0.792. Comparatively, our combined model yielded an AUC of 0.87. However, when including cases with superficial MI, which has additional risk factors such as lymph node metastases, risk stratification can be confounded. According to the European Society for Medical Oncology guidelines, FIGO stage IA lesions are truly considered low risk [5]. Therefore, distinguishing FIGO stage IA lesions from higher stage lesions can be more clinically relevant than radiomics studies that consider only the presence or absence of deep MI without assessing FIGO stage [31, 32] because this can potentially confound results. When classifying FIGO stage, we used a radiomics signature that was more rigorous than that proposed by Yan et al. [33] and similar to that proposed by Yang et al. [34]. Notably, our models demonstrated superior performance in detecting LVSI compared to the model proposed by Luo et al. [35], which combined features extracted from T2WI, DWI, and ADC maps.

MRI interpretations may vary between trained radiologists due to differences in image perception, experience levels and subtle nuances in feature identification. These interpretation discrepancies may introduce biases, potentially affecting the selection of features, model performance, and overall reproducibility of the results. Recognizing these limitations and discussing possible ways of reducing interpretation variability, such as standardized reporting

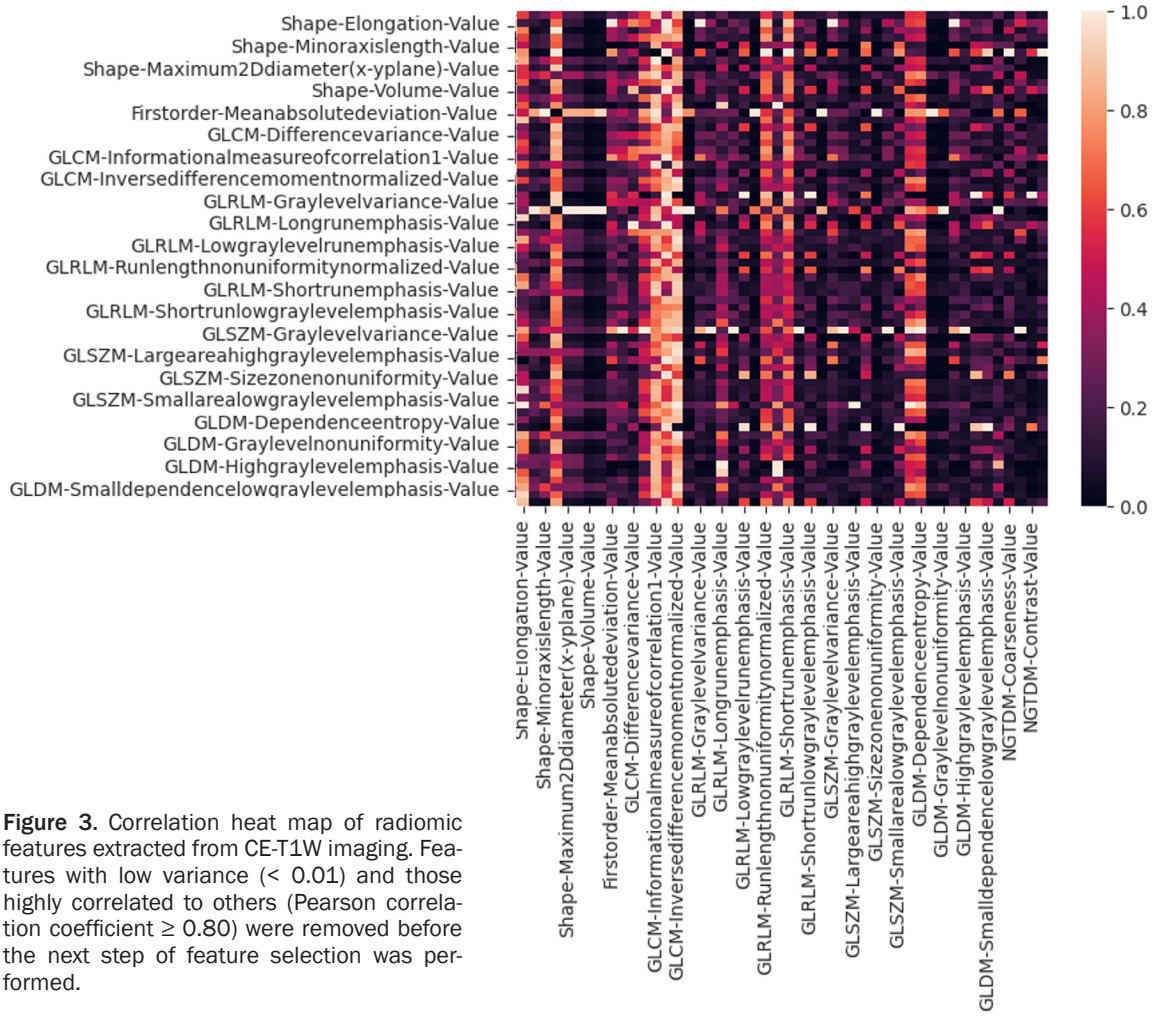

Figure 3. Correlation heat map of radiomic features extracted from CE-T1W imaging. Features with low variance (< 0.01) and those highly correlated to others (Pearson correlation coefficient ≥ 0.80) were removed before the next step of feature selection was performed.

protocols or consensus readings, will strengthen the validity and reproducibility of the study. In our study, features were considered stable only if the lower bound of the 95% confidence interval for the ICC was ≥ 0.75. The AUC values for each single-sequence model were greater than 0.7, indicating that single-sequence models have some predictive ability. Furthermore, the combined model demonstrated significantly greater predictive performance compared to each of the 3 single-sequence models (all $P < 0.05$). In the future, perhaps T2WI and DWI can be combined without the benefit of CE-T1WI, thus reducing the use of contrast, but without losing diagnostic quality. It should be noted that positron emission tomography/computed tomography has not been reliable in predicting high-risk EC because it tends to have high false-negative rates [36]. The radiomics models developed in this study could support the preoperative classification of EC risk factors, consistent with preliminary studies that used MRI-based texture analysis on small sample sizes [37].

This study has some limitations. First, its retrospective design suggests a potential risk of selection bias, emphasizing the importance of confirming our results through prospective validation. Second, the sample was relatively small and was collected in a single center, also potentially introducing selection bias. The limited sample size in model training may affect its ability to generalize effectively. Small data sets are susceptible to overfitting, which can result in inconsistent performance in real-world scenarios. Future studies should be aimed at expanding sample size, especially by collecting data from multiple centers, in order to improve the relevance of the model in different clinical settings. Finally, our sample was not balanced in the distribution of histopathologic features,

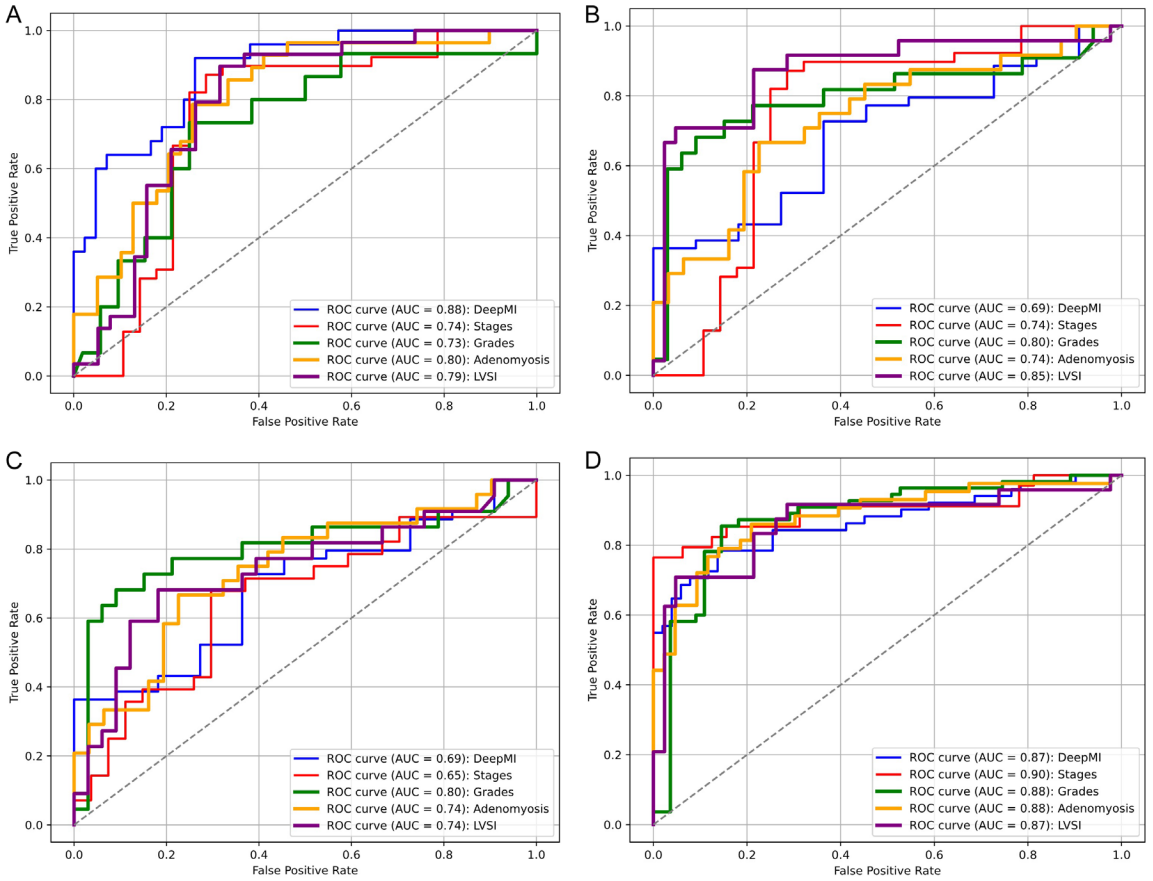

Figure 4. The ROC curves for classifying deep myometrial invasion, International Federation of Gynecology and Obstetrics stage, high-grade tumors, adenomyosis status, and lymphovascular space invasion for each model: (A) The sagittal CE-T1WI model, (B) The sagittal T2WI model, (C) The sagittal DWI model, and (D) The combined model. *ROC* receiver operating characteristic, *AUC* area under the ROC curve.

Table 2. Performance of the 4 models in classifying 5 histopathologic features

Histopathologic features	CE T1W model		T2W model		DWI model		Combined model	
	AUC	95% CI	AUC	95% CI	AUC	95% CI	AUC	95% CI
Deep myometrial invasion	0.88	0.83-0.89	0.69	0.65-0.71	0.69	0.66-0.71	0.87	0.85-0.90
FIGO stages	0.74	0.71-0.76	0.74	0.70-0.77	0.65	0.63-0.68	0.90	0.88-0.92
Grades	0.73	0.70-0.76	0.80	0.76-0.82	0.80	0.77-0.81	0.88	0.87-0.90
Adenomyosis status	0.80	0.77-0.82	0.74	0.72-0.78	0.74	0.70-0.79	0.88	0.86-0.92
Lymphovascular space invasion	0.79	0.77-0.83	0.85	0.83-0.86	0.74	0.72-0.78	0.87	0.86-0.90

The combined model achieved the best AUC for most features.

specifically histopathologic grade. For machine learning strategies to be effective in building reliable models based on small sample sizes, data distribution should be similar between the training and test sets. This imbalance might compromise the model's reliability and generalizability. Future studies should consider strategies to address this issue, such as data augmentation techniques, SMOTE, to improve the representation of minority classes, or the use of weight loss functions and balanced sampling during model training to reduce biases. In addition, targeted data collection efforts to obtain a more balanced sample across histopathological levels would strengthen the clinical applicability and predictive accuracy of the model.

Our model demonstrated strong diagnostic performance in identifying high-risk cases and holds promise for preoperative risk stratifica-

tion using MRI in patients with endometrial cancer. Potential real-world challenges, such as the training needs of clinical staff, data compatibility with existing electronic health records (EHRs) and regulatory considerations, will be addressed by including examples or case scenarios showing how their model could improve patient outcomes, such as personalizing surgical planning and following-up management.

Conclusions

In this study, an MRI radiomics model was developed by integrating radiomic features from 3 common imaging sequences, CE-T1WI, T2WI, and DWI. While further evidence is necessary before radiomics can be employed for clinical decision-making, our model exhibited promising diagnostic performance in predicting high risk and showed the potential for use in preoperative risk stratification using MRI in those with EC.

Acknowledgements

This work was supported by grant from Wan Fang Hospital, Taipei Medical University, Taiwan (113-wf-eva-13).

Disclosure of conflict of interest

C.H. Liang is an employee of Herricane Medical Care Company in Taiwan. No funding was received from the Herricane Medical Care Company for this research or publication activities.

Address correspondence to: Dr. Wing P Chan, Department of Radiology, Wan Fang Hospital, Taipei Medical University, No. 111 Hsing Long Road, Section 3, Taipei 116, Taiwan. Tel: +886-2 29307930 Ext. 1300; Fax: +886-2-29316809; E-mail: wingchan@tmu.edu.tw

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021; 71: 209-249.

[2] Siegel RL, Giaquinto AN and Jemal A. Cancer statistics, 2024. CA Cancer J Clin 2024; 74: 12-49.

[3] Koppikar S, Oaknin A, Babu KG, Lorusso D, Gupta S, Wu LY, Rajabto W, Harano K, Hong SH, Malik RA, Strebel H, Aggarwal IM, Lai CH, Dejthevaporn T, Tangjitgamol S, Cheng WF, Chay WY, Benavides D, Hashim NM, Moon YW, Yunokawa M, Anggraeni TD, Wei W, Curigliano G, Maheshwari A, Mahantshetty U, Sheshadri S, Peters S, Yoshino T and Pentheroudakis G. Pan-Asian adapted ESMO Clinical Practice Guidelines for the diagnosis, treatment and follow-up of patients with endometrial cancer. ESMO Open 2023; 8: 100774.

[4] Mohammed S, Polymeros K, Wickham-Joseph R, Luqman I, Charadva C, Morris T, Collins A, Barber S, Khunti K and Moss EL. Comparing characteristics of endometrial cancer in women of South Asian and White Ethnicity in England. Cancers (Basel) 2021; 13: 6123.

[5] Berek JS, Matias-Guiu X, Creutzberg C, Fotopoulou C, Gaffney D, Kehoe S, Lindemann K, Mutch D and Concin N; Endometrial Cancer Staging Subcommittee, FIGO Women's Cancer Committee. FIGO staging of endometrial cancer: 2023. Int J Gynaecol Obstet 2023; 162: 383-394.

[6] Trojano G, Olivieri C, Tinelli R, Damiani GR, Pellegrino A and Cicinelli E. Conservative treatment in early stage endometrial cancer: a review. Acta Biomed 2019; 90: 405-410.

[7] Bansal S, Buck AM, Herzog TJ, Deutsch I, Burke WM and Wright JD. Stage IIIA endometrial carcinoma: outcome and predictors of survival. Obstet Gynecol 2009; 114: 100-105.

[8] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2018. CA Cancer J Clin 2018; 68: 7-30.

[9] Haight PJ, Riedinger CJ, Backes FJ, O'Malley DM and Cosgrove CM. The right time for change: a report on the heterogeneity of IVB endometrial cancer and improved risk-stratification provided by new 2023 FIGO staging criteria. Gynecol Oncol 2023; 175: 32-40.

[10] Sasada S, Yunokawa M, Takehara Y, Ishikawa M, Ikeda S, Kato T and Tamura K. Baseline risk of recurrence in stage I-II endometrial carcinoma. J Gynecol Oncol 2018; 29: e9.

[11] Takahashi S, Murakami T, Narumi Y, Kurachi H, Tsuda K, Kim T, Enomoto T, Tomoda K, Miyake A, Murata Y and Nakamura H. Preoperative staging of endometrial carcinoma: diagnostic effect of T2-weighted fast spin-echo MR imaging. Radiology 1998; 206: 539-547.

[12] Kishimoto K, Tajima S, Maeda I, Takagi M, Ueno T, Suzuki N and Nakajima Y. Endometrial cancer: correlation of apparent diffusion coefficient (ADC) with tumor cellularity and tumor grade. Acta Radiol 2016; 57: 1021-1028.

[13] Woo S, Cho JY, Kim SY and Kim SH. Histogram analysis of apparent diffusion coefficient map of diffusion-weighted MRI in endometrial cancer: a preliminary correlation study with histological grade. Acta Radiol 2014; 55: 1270-1277.

[14] Bharwani N, Miquel ME, Sahdev A, Narayanan P, Malietzis G, Reznek RH and Rockall AG. Diffusion-weighted imaging in the assessment of tumour grade in endometrial cancer. Br J Radiol 2011; 84: 997-1004.
[15] Takahashi M, Kozawa E, Tanisaka M, Hasegawa K, Yasuda M and Sakai F. Utility of histogram analysis of apparent diffusion coefficient maps obtained using 3.0T MRI for distinguishing uterine carcinosarcoma from endometrial carcinoma. J Magn Reson Imaging 2016; 43: 1301-1307.
[16] Kinkel K, Kaji Y, Yu KK, Segal MR, Lu Y, Powell CB and Hricak H. Radiologic staging in patients with endometrial cancer: a meta-analysis. Radiology 1999; 212: 711-718.
[17] Sala E, Crawford R, Senior E, Shaw A, Simcock B, Vrotsou K, Palmer C, Rajan P, Joubert I and Lomas D. Added value of dynamic contrast-enhanced magnetic resonance imaging in predicting advanced stage disease in patients with endometrial carcinoma. Int J Gynecol Cancer 2009; 19: 141-146.
[18] Lefebvre TL, Ueno Y, Dohan A, Chatterjee A, Vallières M, Winter-Reinhold E, Saif S, Levesque IR, Zeng XZ, Forghani R, Seuntjens J, Soyer P, Savadjiev P and Reinhold C. Development and validation of multiparametric MRI-based radiomics models for preoperative risk stratification of endometrial cancer. Radiology 2022; 305: 375-386.
[19] Lin Z, Wang T, Li Q, Bi Q, Wang Y, Luo Y, Feng F, Xiao M, Gu Y, Qiang J and Li H. Development and validation of MRI-based radiomics model to predict recurrence risk in patients with endometrial cancer: a multicenter study. Eur Radiol 2023; 33: 5814-5824.
[20] Tan Q, Wang Q, Jin S, Zhou F and Zou X. Network Evolution Model-based prediction of tumor mutation burden from radiomic-clinical features in endometrial cancers. BMC Cancer 2023; 23: 712.
[21] Tomaszewski MR and Gillies RJ. The biological meaning of radiomic features. Radiology 2021; 298: 505-516.
[22] Hyun SH, Ahn MS, Koh YW and Lee SJ. A machine-learning approach using PET-based radiomics to predict the histological subtypes of lung cancer. Clin Nucl Med 2019; 44: 956-960.
[23] Edge SB and American Joint Committee on Cancer ACS. AJCC cancer staging handbook: from the AJCC cancer staging manual. Springer; 2010.
[24] Oaknin A, Bosse TJ, Creutzberg CL, Giornelli G, Harter P, Joly F, Lorusso D, Marth C, Makker V, Mirza MR, Ledermann JA and Colombo N; ESMO Guidelines Committee. Electronic address: clinicalguidelines@esmo.org. Endometrial cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. Ann Oncol 2022; 33: 860-877.
[25] Liyanage A, Cardoza S, Kasabia D and Moore H. Accuracy of MRI in predicting deep myometrial invasion in endometrial cancer and the influence of leiomyoma, adenomyosis and the microcystic elongated and fragmented tumour pattern. J Med Imaging Radiat Oncol 2024; 68: 235-242.
[26] Yasmeen T, Nasir S and Naqvi H. Diagnostic accuracy of pre-operative magnetic resonance imaging in assessing prognostic pathological parameters in endometrial carcinoma. Liaquat Natl Tumor Board J 2020; 2: 20-25.
[27] Yıldız G, Turan K and Yıldız P. The efficiency of MRI technique in determining the depth of myometrial invasion in endometrial cancer cases. Eastern Journal of Medicine 2021; 26.
[28] Meng X, Yang D, Deng Y, Xu H, Jin H and Yang Z. Diagnostic accuracy of MRI for assessing lymphovascular space invasion in endometrial carcinoma: a meta-analysis. Acta Radiol 2024; 65: 133-144.
[29] Chen X, Wang Y, Shen M, Yang B, Zhou Q, Yi Y, Liu W, Zhang G, Yang G and Zhang H. Deep learning for the determination of myometrial invasion depth and automatic lesion identification in endometrial cancer MR imaging: a preliminary study in a single institution. Eur Radiol 2020; 30: 4985-4994.
[30] Dong HC, Dong HK, Yu MH, Lin YH and Chang CC. Using deep learning with convolutional neural network approach to identify the invasion depth of endometrial cancer in myometrium using MR images: a pilot study. Int J Environ Res Public Health 2020; 17: 5993.
[31] Rodríguez-Ortega A, Alegre A, Lago V, Carot-Sierra JM, Ten-Esteve A, Montoliu G, Domingo S, Alberich-Bayarri Á and Martí-Bonmatí L. Machine learning-based integration of prognostic magnetic resonance imaging biomarkers for myometrial invasion stratification in endometrial cancer. J Magn Reson Imaging 2021; 54: 987-995.
[32] Stanzione A, Cuocolo R, Del Grosso R, Nardiello A, Romeo V, Travaglino A, Raffone A, Bifulco G, Zullo F, Insabato L, Maurea S and Mainenti PP. Deep myometrial infiltration of endometrial cancer on MRI: a radiomics-powered machine learning pilot study. Acad Radiol 2021; 28: 737-744.
[33] Yan BC, Li Y, Ma FH, Feng F, Sun MH, Lin GW, Zhang GF and Qiang JW. Preoperative assessment for high-risk endometrial cancer by developing an MRI- and clinical-based radiomics nomogram: a multicenter study. J Magn Reson Imaging 2020; 52: 1872-1882.

[34] Yang J, Cao Y, Zhou F, Li C, Lv J and Li P. Combined deep-learning MRI-based radiomic models for preoperative risk classification of endometrial endometrioid adenocarcinoma. Front Oncol 2023; 13: 1231497.
[35] Luo Y, Mei D, Gong J, Zuo M and Guo X. Multiparametric MRI-based radiomics nomogram for predicting lymphovascular space invasion in endometrial carcinoma. J Magn Reson Imaging 2020; 52: 1257-1262.
[36] Stewart KI, Chasen B, Erwin W, Fleming N, Westin SN, Dioun S, Frumovitz M, Ramirez PT, Lu KH, Wong F, Aloia TA and Soliman PT. Preoperative PET/CT does not accurately detect extrauterine disease in patients with newly diagnosed high-risk endometrial cancer: a prospective study. Cancer 2019; 125: 3347-3353.
[37] Ueno Y, Forghani B, Forghani R, Dohan A, Zeng XZ, Chamming's F, Arseneau J, Fu L, Gilbert L, Gallix B and Reinhold C. Endometrial carcinoma: MR imaging-based texture model for preoperative risk stratification-a preliminary analysis. Radiology 2017; 284: 748-757.

Table S1. Comparison of preoperative MRI staging by radiologists and pathological staging

MRI staging	Pathological Staging					Total
	Ia	Ib	II	III	IV	
Ia	28	3	1	1	0	33
Ib	1	15	0	1	0	17
II	1	1	4	2	0	8
III	1	0	0	3	1	5
IV	0	0	0	0	3	3
Total	31	19	5	7	4	66

Table S2. Preoperative staging of MRI diagnostic test evaluation index

MRI staging	Sensitivity	Specificity
Ia	90.30%	87.80%
Ib	78.90%	94.10%
II	80.00%	93.90%
III	62.50%	96.70%
IV	75.00%	95.00%

Table S3. The 10 most important features based on feature importance score in the combined model

Radiomic feature	MRI sequence	Feature importance score
Firstorder-Kurtosis-Value	CE-T1W	0.97826
GLSZM-Zoneentropy-Value	CE-T1W	0.95652
GLSZM-Graylevelvariance-Value	T2W	0.95652
GLRLM-Highgraylevelrunemphasis-Value	DWI	0.94392
GLRLM-Graylevelvariance-Value	CE-T1W	0.94333
GLSZM-Zoneentropy-Value	T2W	0.93342
Shape-Majoraxislength-Value	T2W	0.92344
GLRLM-Graylevelnonuniformity-Value	DWI	0.91282
GLCM-Inversedifferencemomentnormalized-Value	CE-T1W	0.91182
GLRLM-Graylevelnonuniformity-Value	CE-T1W	0.90281