

# Development and Validation of Multiparametric MRI-based Radiomics Models for Preoperative Risk Stratification of Endometrial Cancer

Citation for published version (APA):

Lefebvre, T. L., Ueno, Y., Dohan, A., Chatterjee, A., Vallières, M., Winter-Reinhold, E., Saif, S., Levesque, I. R., Zeng, X. Z., Forghani, R., Seuntjens, J., Soyer, P., Savadjiev, P., & Reinhold, C. (2022). Development and Validation of Multiparametric MRI-based Radiomics Models for Preoperative Risk Stratification of Endometrial Cancer. *Radiology*, 305(2), 375-386. Article 212873. <https://doi.org/10.1148/radiol.212873>

## Document status and date:

Published: 01/11/2022

## DOI:

[10.1148/radiol.212873](https://doi.org/10.1148/radiol.212873)

## Document Version:

Publisher's PDF, also known as Version of record

## Document license:

Taverne

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# Development and Validation of Multiparametric MRI–based Radiomics Models for Preoperative Risk Stratification of Endometrial Cancer

Thierry L. Lefebvre, MSc • Yoshiko Ueno, MD, PhD • Anthony Dohan, MD, PhD • Avishek Chatterjee, PhD • Martin Vallières, PhD • Eric Winter-Reinhold • Sameh Saïf, MD • Ives R. Levesque, PhD • Xing Ziggy Zeng, MD • Reza Forghani, MD, PhD • Jan Seuntjens, PhD • Philippe Soyer, MD, PhD • Peter Savadjiev, PhD\* • Caroline Reinhold, MD, MSc\*

From the Medical Physics Unit, Department of Oncology (T.L.L., A.C., M.V., I.R.L., J.S.), Department of Diagnostic Radiology (Y.U., S.S., R.F., P. Savadjiev, C.R.), and School of Computer Science (P. Savadjiev), McGill University, Montreal General Hospital Site, 1650 Cedar Ave, Montreal, QC, Canada H3G 1A4; Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, England (T.L.L.); Department of Radiology, Kobe University Graduate School of Medicine, Kobe, Japan (Y.U.); Department of Radiology, Cochin Hospital, AP-HP, Centre, Paris, France (A.D., P. Soyer); Université de Paris, Faculté de Médecine, Paris, France (A.D., P. Soyer); Department of Precision Medicine, GROW–School for Oncology and Developmental Biology, Maastricht University, Maastricht, the Netherlands (A.C.); Department of Computer Science, Université de Sherbrooke, Sherbrooke, Quebec, Canada (M.V.); Augmented Intelligence & Precision Health Laboratory (E.W.R., R.F., C.R.), Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada; Department of Obstetrics and Gynecology, McGill University Health Centre, Montreal, Quebec, Canada (X.Z.Z.); and Montreal Imaging Experts, Montreal, Quebec, Canada (R.F., C.R.). Received November 16, 2021; revision requested January 11, 2022; revision received April 19; accepted May 16. Address correspondence to C.R. (email: [caroline.reinhold@mcgill.ca](mailto:caroline.reinhold@mcgill.ca)).

This work was jointly funded by the Fonds de recherche du Québec–Santé et the Fondation de l'association des radiologistes du Québec. T.L.L. acknowledges financial support from the NSERC Alexander Graham Bell Canada Graduate Scholarship. M.V. acknowledges funding from the Canada CIFAR AI Chairs Program.

\* P. Savadjiev and C.R. are co–senior authors.

Conflicts of interest are listed at the end of this article.

See also the editorial by Kido and Nishio in this issue.

Radiology 2022; 000:1–12 • <https://doi.org/10.1148/radiol.212873> • Content codes: **OB** **GU** **MR**

**Background:** Stratifying high-risk histopathologic features in endometrial carcinoma is important for treatment planning. Radiomics analysis at preoperative MRI holds potential to identify high-risk phenotypes.

**Purpose:** To evaluate the performance of multiparametric MRI three-dimensional radiomics-based machine learning models for differentiating low- from high-risk histopathologic markers—deep myometrial invasion (MI), lymphovascular space invasion (LVSI), and high-grade status—and advanced-stage endometrial carcinoma.

**Materials and Methods:** This dual-center retrospective study included women with histologically proven endometrial carcinoma who underwent 1.5-T MRI before hysterectomy between January 2011 and July 2015. Exclusion criteria were tumor diameter less than 1 cm, missing MRI sequences or histopathology reports, neoadjuvant therapy, and malignant neoplasms other than endometrial carcinoma. Three-dimensional radiomics features were extracted after tumor segmentation at MRI (T2-weighted, diffusion-weighted, and dynamic contrast-enhanced MRI). Predictive features were selected in the training set with use of random forest (RF) models for each end point, and trained RF models were applied to the external test set. Five board-certified radiologists conducted MRI-based staging and deep MI assessment in the training set. Areas under the receiver operating characteristic curve (AUCs) were reported with balanced accuracies, and radiologists' readings were compared with radiomics with use of McNemar tests.

**Results:** In total, 157 women were included: 94 at the first institution (training set; mean age, 66 years  $\pm$  11 [SD]) and 63 at the second institution (test set; 67 years  $\pm$  12). RF models dichotomizing deep MI, LVSI, high grade, and International Federation of Gynecology and Obstetrics (FIGO) stage led to AUCs of 0.81 (95% CI: 0.68, 0.88), 0.80 (95% CI: 0.67, 0.93), 0.74 (95% CI: 0.61, 0.86), and 0.84 (95% CI: 0.72, 0.92), respectively, in the test set. In the training set, radiomics provided increased performance compared with radiologists' readings for identifying deep MI (balanced accuracy, 86% vs 79%;  $P = .03$ ), while no evidence of a difference was observed in performance for advanced FIGO stage (80% vs 78%;  $P = .27$ ).

**Conclusion:** Three-dimensional radiomics can stratify patients by using preoperative MRI according to high-risk histopathologic end points in endometrial carcinoma and provide nonsignificantly different or higher performance than radiologists in identifying advanced stage and deep myometrial invasion, respectively.

© RSNA, 2022

Online supplemental material is available for this article.

In developed countries, endometrial carcinoma is the most common gynecologic malignant neoplasm, with increasing incidence and associated mortality over the past decade (1). Important prognostic factors include the International Federation of Gynecology and Obstetrics (FIGO) stage, histopathologic tumor grade, and lymphovascular space invasion (LVSI) (2–4). Meta-analyses have shown the diagnostic value of MRI for predicting the depth of

myometrial invasion (MI) (5,6). As such, MRI has become essential in the preoperative staging of patients with endometrial carcinoma (7,8).

However, imaging staging of endometrial carcinoma relies on visual assessment by trained radiologists, which is subject to interobserver variability and does not readily translate from one institution to another. For example, it has been shown that high MRI staging performance

## Abbreviations

ADC = apparent diffusion coefficient, AUC = area under the ROC curve, DCE-MRI = dynamic contrast-enhanced MRI, FIGO = International Federation of Gynecology and Obstetrics, LVSI = lymphovascular space invasion, MI = myometrial invasion, RF = random forest, ROC = receiver operating characteristic, 3D = three-dimensional

## Summary

Three-dimensional radiomics-based machine learning models can stratify women with endometrial carcinoma at preoperative multiparametric MRI based on high-risk histopathologic features and can identify advanced-stage cancer (International Federation of Gynecology and Obstetrics system) reproducibly across sets acquired at different centers.

## Key Results

- In this dual-center study of 157 women with endometrial cancer, multiparametric MRI-based three-dimensional radiomics models differentiated deep myometrial invasion (MI), lymphovascular space invasion, tumor grade, and International Federation of Gynecology and Obstetrics (FIGO) stage, with areas under the receiver operating characteristic curve of 0.81, 0.80, 0.74, and 0.84, respectively, in the test set ( $n = 63$ ).
- Radiomics provided higher or nonsignificantly different performance than radiologists for identifying deep MI (balanced accuracy, 86% vs 79%, respectively;  $P = .03$ ) and advanced FIGO stage (80% vs 78%;  $P = .27$ ) in the training set ( $n = 94$ ).

reported from single-center studies was not replicated in a tertiary referral center that received MRI examinations acquired with heterogeneous protocols from a large number of external referring hospitals (9). In addition, although routinely performed for initial screening and risk stratification, preoperative biopsies tend to underestimate tumor grade, particularly for larger tumors (10,11). A recent meta-analysis showed only a 67% agreement between endometrial sampling and the final posthysterectomy assignment of tumor grade (12). Thus, given the high sampling variability of biopsy and the importance of capturing intratumoral heterogeneity (13), noninvasive quantitative whole-tumor assessment at standard-of-care MRI before surgery—so-called virtual biopsy—holds potential for improving the risk stratification of women with endometrial carcinoma.

Radiomics quantifies the image-based phenotypes of entire lesions and predicts patient outcomes according to mathematically defined image descriptors (14). Several studies have investigated the role of three-dimensional (3D) radiomics derived from standard-of-care MRI for staging of endometrial carcinoma by using histopathologic findings as the standard of reference (15–23). Most of these were single-center studies without an external test set, limiting the generalizability of the results (15–22). To ensure clinical translatability of such models, an external test set is essential, but, to our knowledge, only one recent study included an independent test set (23).

The aim of this study was to evaluate the diagnostic performance of 3D radiomics-based machine learning models using multiparametric MRI to differentiate low- from high-risk histopathologic markers (deep MI, LVSI, and high-grade status) and early-FIGO stage (IA) from advanced-FIGO stage (IB or higher) endometrial carcinoma. The latter classification task is performed to assess the potential of the MRI-based radiomics signature for risk stratification, since the European Society for

Medical Oncology considers FIGO IA the only low-risk stage (24). The methodologic and design contributions relative to the other recent radiomics studies referred to earlier (15–22) can be summarized as follows: first, latest recommendations from the Image Biomarker Standardization Initiative (or IBSI) on the development of robust and reproducible radiomics models were applied (25); second, a full 3D analysis was implemented at multiparametric MRI; and finally, proposed machine learning models were validated on an independent external test set and compared with preoperative readings from radiologists.

## Materials and Methods

### Patients

Institutional review board approval and waiver of informed consent were obtained at both institutions, McGill University Health Centre, Montreal, Canada (institution 1), and Hôpital Lariboisière, Assistance Publique-Hôpitaux de Paris, Paris, France (institution 2), for this dual-center retrospective study. Clinical and imaging data obtained in women who underwent MRI before surgery between January 2011 and July 2015 at the first institution made up the training set and were used to train machine learning models to identify the presence of deep MI, LVSI, advanced FIGO stage, and high-grade tumors. The diagnostic performance and reproducibility of the predictive model were tested on images in an independent group of women recruited at the second institution, referred to as the test set.

A total of 165 consecutive women at the first institution and 99 women at the second institution with histologically proven endometrial carcinoma were deemed eligible after undergoing 1.5-T MRI before total hysterectomy with bilateral salpingo-oophorectomy. Consecutive women were identified using the cancer registry database of each hospital. Exclusion criteria were (a) a maximum tumor diameter less than 1 cm; (b) an incomplete MRI examination; (c) an incomplete pathology report; (d) neoadjuvant therapy; and (e) malignant neoplasms other than endometrial carcinoma. The final sample comprised 94 women in the training set and 63 women in the test set. Eighty-six of the 94 women recruited at the first institution (training set) overlap with the patient sample of a preliminary single-center two-dimensional radiomics study on the classification of deep MI, LVSI, and high-grade lesions (26). In the current study, 3D radiomics analyses following IBSI standards were conducted, including an external test set from an independent institution, and with expanded aims including the classification of advanced-stage lesions.

### MRI Protocol

All pelvic MRI examinations were performed on 1.5-T systems (institution 1: Signa Excite, GE Healthcare; institution 2: Magnetom AvantoVB, Siemens Healthineers) using a vendor-provided pelvic phased-array surface coil. Examinations included standard sequences (T2-weighted imaging, diffusion-weighted imaging [ $b = 0$  and  $1000 \text{ sec/mm}^2$ ], and dynamic contrast-enhanced MRI [DCE-MRI]). DCE-MRI was performed after administration of 0.1 mmol per kilogram of body weight of gadolinium chelate (Gadovist, Bayer). Images were acquired during multiple phases following administration of intravenous

gadolinium chelate in either sagittal or axial oblique planes (before and at 25, 60, and 120 seconds after contrast material administration in the sagittal plane and 240 seconds after contrast material administration in the axial oblique plane). MRI acquisition protocols are summarized in Table E1 (online).

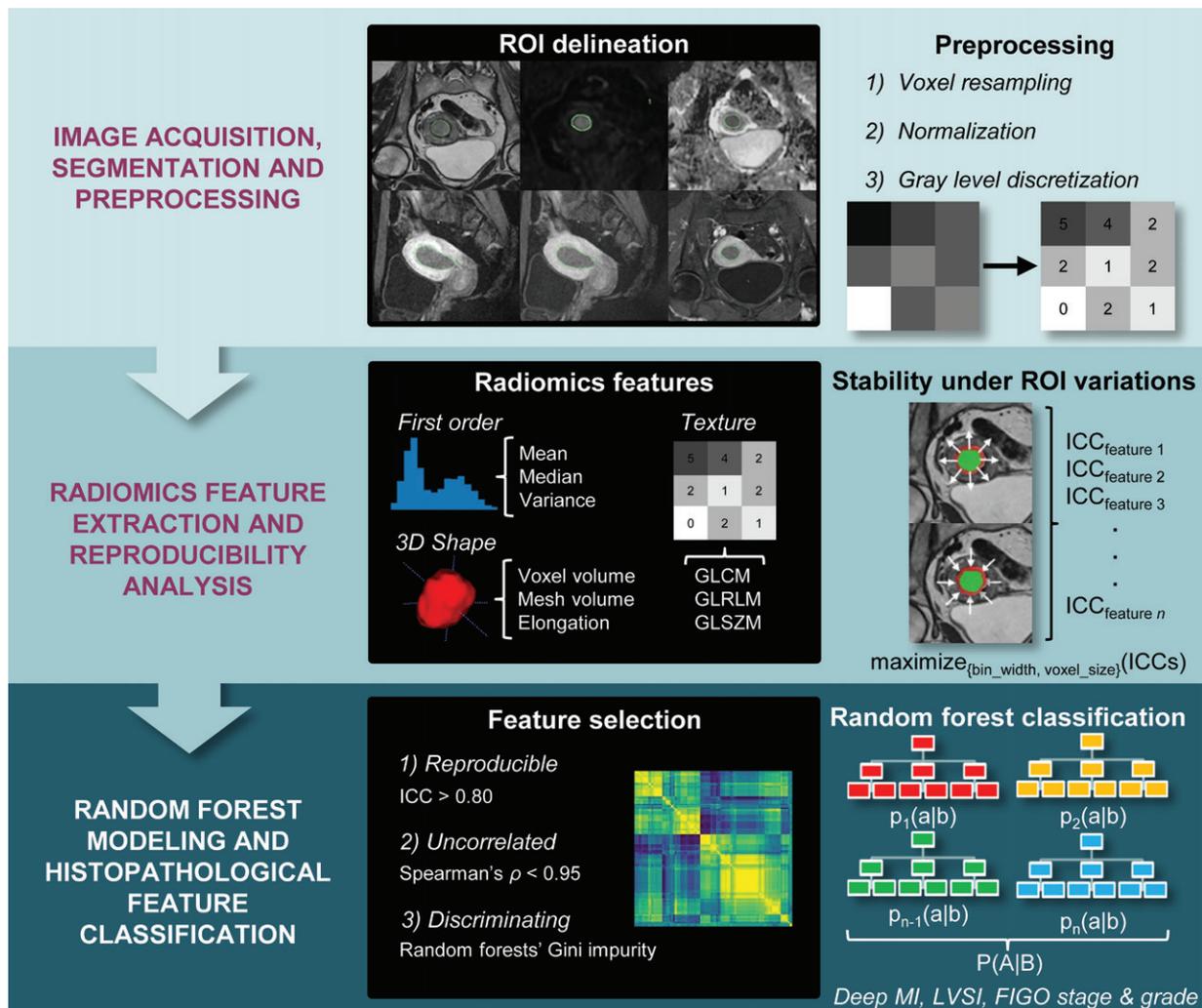
### Surgical Histopathologic Analyses

All women recruited at both institutions underwent total hysterectomy, bilateral salpingo-oophorectomy, and bilateral pelvic lymph node sampling. To ensure full histopathologic sampling of the resected specimen, the uterus was sliced contiguously every 3–4 mm, and slices were stained with hematoxylin and eosin. Histopathologic analysis was conducted by gynecologic pathologists as part of clinical standard of care, and histopathologic reports were obtained to determine tumor subtype, histopathologic grade, the presence of deep MI ( $\geq 50\%$ ) and LVSI. For the purpose of radiomics modeling, tumor grade was dichotomized

as follows: (a) low grade: FIGO grades 1 and 2 endometrioid carcinoma and (b) high grade: FIGO grade 3 endometrioid carcinoma or nonendometrioid histologic findings. Furthermore, FIGO stage was dichotomized into early (IA) and advanced (IB or higher) for risk stratification in accordance with European Society for Medical Oncology guidelines (24).

### Lesion Segmentation

Contours defining the 3D tumor region of interest for each patient were manually drawn section by section on image volumes from the six different sequences, in either the sagittal or axial oblique plane: T2-weighted images, diffusion-weighted images ( $b = 1000 \text{ sec/mm}^2$ ), apparent diffusion coefficient (ADC) map (generated from the combination of  $b = 0 \text{ sec/mm}^2$ , 1000  $\text{sec/mm}^2$ ), and the second, third, and delayed phases of DCE-MRI. The region of interest of each tumor was manually drawn independently (Fig 1) by one radiologist at each institution



**Figure 1:** Diagram shows radiomics pipeline used for the classification of histopathologic features of endometrial carcinoma at multiparametric MRI. Lesions are manually segmented on images acquired with each MRI sequence. Radiomics features are extracted for different sets of preprocessing parameters, and the set leading to the highest feature stability under region of interest (ROI) variations is selected. Only preselected reproducible, uncorrelated, and discriminating radiomics features are included in random forest model building. Models are trained on the training set for histopathologic feature classification, and diagnostic performance is reported for both the training and test sets with bootstrapped CIs. 3D = three-dimensional, FIGO = International Federation of Gynecology and Obstetrics, GLCM = gray-level co-occurrence matrix, GLRLM = gray-level run length matrix, GLSZM = gray-level size zone matrix, ICC = intraclass correlation coefficient, LVSI = lymphovascular space invasion, MI = myometrial invasion.

(Y.U. and A.D., both with 6 years of experience in pelvic MRI). All contours were reviewed by an experienced fellowship-trained subspecialty radiologist with more than 20 years of experience (C.R.). Contour modifications were done manually by the senior radiologist on sets from both institutions if deemed necessary, but were not repeated independently. The three radiologists were aware of the diagnosis of endometrial carcinoma but blinded to clinical and histopathologic outcomes.

### Radiomics Analysis

The developed radiomics analysis pipeline is depicted in Figure 1. Briefly, radiomics features were computed using the IBSI-compliant Pyradiomics 3.0 Python package (27). Feature reproducibility was then assessed using a pipeline complying with IBSI recommendations. Features that were robust to changes in image preprocessing parameters and variations in contour delineation (intraclass correlation coefficient with two-way random effects model >0.80) (28) and also not highly correlated with each other (Spearman  $\rho$  <0.95) were retained for subsequent analysis. Based on those retained features, an initial random forest (RF) classifier was trained to determine the predictive power of individual features. The top five features were retained, and a final RF classifier based on these five features was trained and then applied to the independent test set. To account for class imbalance, the majority class was downsampled in bootstrapped folds during the training with use of the imbalanced-learn package (29).

### Radiologists' Readings

In the training set, five board-certified radiologists (with at least 5 years of experience in gynecologic MRI interpretation) prospectively performed FIGO staging at multiparametric MRI in accordance with recent guidelines during routine clinical assessment (30,31). In addition, each woman's complete imaging file including the pelvic MRI examinations, histologic results, and physical examination was reviewed by an interdisciplinary gynecologic tumor board panel before surgery for quality assurance, as is standard of care.

### Statistical Analysis

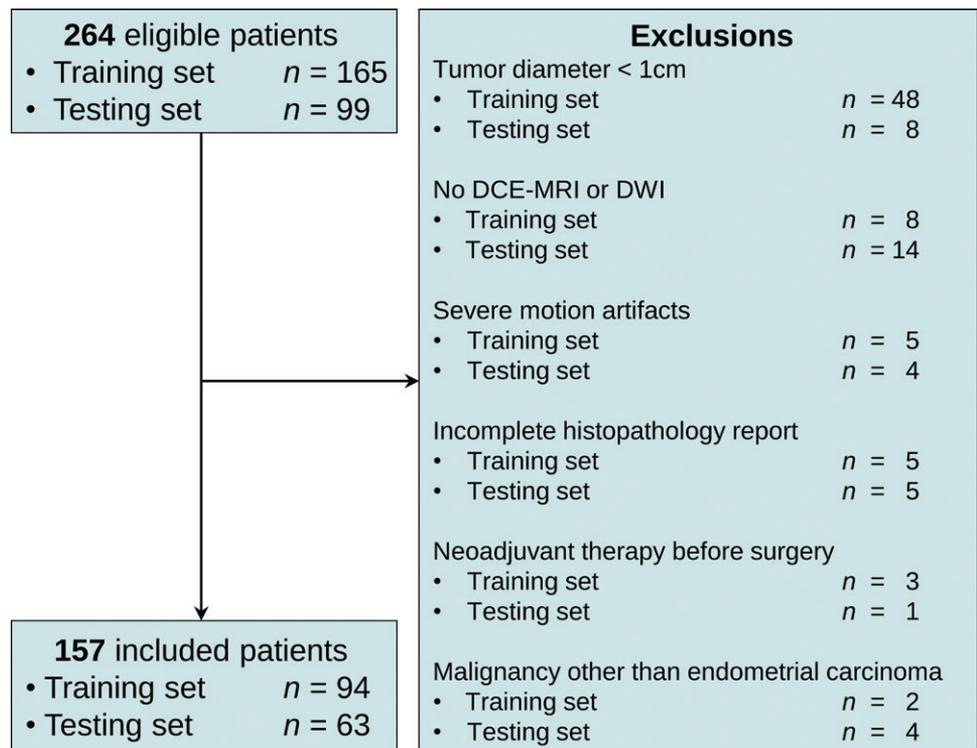
Diagnostic performance metrics obtained on both sets include receiver operating characteristic (ROC) curves with associated areas under the ROC curve (AUCs), sensitivity, specificity, balanced accuracy, positive predictive value, and negative predictive value

for the threshold maximizing the Youden index. Balanced accuracies (ie, [sensitivity + specificity]/2) were reported to mitigate the effect of imbalanced class distribution in the training and test sets in the RF models. McNemar tests were performed to compare the binary predictions of radiomics models and radiologists' readings (32).

ROC analysis for classifying tumors in dichotomized groups for each histopathologic outcome was performed individually for each single radiomics feature. Based on AUCs obtained for the whole training set, the top three most discriminative features were also identified. As a reference, MRI volumetry was also computed on T2-weighted images, and the predictive value of the extracted volumes, expressed as the AUC, was reported.

Radiomics feature correlation with surgical and histopathologic outcomes of interest and other clinical variables (age, FIGO stage, cervical stromal invasion, MI, grade, LVSI, lymph node metastasis, and adenomyosis) was explored, and Spearman rank or point-biserial correlation coefficients were reported in a correlation heatmap. Nonparametric comparisons of clinical variables and radiomics features in each set and each disease group, respectively, were performed with a Mann-Whitney *U* test for binary groups and a Kruskal-Wallis test for multiple groups.

All RF modeling and statistical analyses were conducted in Python 3.7.4 and using the Scikit-learn 0.24.2 machine learning package (33). Code used for radiomics analyses can be found online in the repository for the article (<https://www.github.com/thierleft/mri-3Dradiomics-endometrialcancer>). Additional methodologic details regarding the developed radiomics pipeline are provided in Appendix E1 (online), with a particular focus on reproducibility.



**Figure 2:** Flowchart of patient selection at the two recruiting institutions (training and test sets). DCE-MRI = dynamic contrast-enhanced MRI, DWI = diffusion-weighted imaging.

## Results

### Patient Characteristics and Surgical Histopathologic Findings

Of the 165 and 99 women eligible at institutions 1 and 2, respectively, 48 (29%) and eight (8.0%) had a maximum tumor diameter less than 1 cm; eight (4.8%) and 14 (14%) had an incomplete MRI examination; five (3.0%) and four (4.0%) had severe motion artifacts; five (3.0%) and five (5.1%) had an incomplete histopathology report; three (1.8%) and one (1.0%) received neoadjuvant therapy; and two (1.2%) and four (4.0%) had malignant neoplasms other than endometrial carcinoma (Fig 2). The final sample comprised 94 women in the training set (institution 1; mean age, 66 years  $\pm$  11 [SD]) and 63 women in the test set (institution 2; mean age, 67 years  $\pm$  12). Characteristics of women recruited at both institutions are presented in Table 1.

Final postoperative histopathologic assessment was used to diagnose advanced FIGO stage in 52 of 94 women (55%) and 44 of 63 women (70%) in the training and test sets, respectively; deep MI in 43 of 94 (46%) and 36 of 63 (57%); LVSI in 49 of 94 (52%) and 19 of 63 (30%); and high-grade tumor in 33 of 94 (35%) and 15 of 63 (24%). The histopathologic subtype was endometrioid carcinoma in 80 of 94 women (85%) and 60 of 63 women (95%) in the training and test sets, respectively; serous carcinoma in 11 of 94 (12%) and two of 63 (3.2%); clear cell carcinoma in two of 94 (2.1%) and one of 63 (1.6%); and mixed or undifferentiated carcinoma in one of 94 (1.1%) and 0 of 63. Histopathologic findings did not differ between training and test sets (range,  $P = .05$ –.80), except for the distribution of women with LVSI (training set, 49 of 94 [52%] vs test set, 19 of 63 [30%];  $P = .007$ ).

### Radiomics Analysis

The initial reproducibility analysis revealed that the preprocessing parameters that led to the highest overall intraclass correlation coefficient across all MRI series were resampled isotropic voxels of 1 mm<sup>3</sup> and normalized absolute discretized images with a fixed bin size of 25. With this set of preprocessing steps, 56 radiomics features were excluded from further analysis because their associated intraclass correlation coefficients were below the predefined 0.80 reproducibility threshold (Appendix E1 [online]). Thus, there were 580 features left from the 636 (91%) originally extracted features (ie, 106 features per MRI series) after this first feature selection step. Next, highly correlated multicollinear features were removed (Spearman  $\rho > 0.95$ ), resulting in the retention of only 361 of the 580 selected reproducible features (57%

of the 636 originally extracted features). Before RF modeling was conducted, the top three most discriminative radiomics features based on ROC analysis, as ranked by the resulting AUC for each classification task, were identified and are reported in Table 2, in addition to tumor volumes computed on T2-weighted images for reference. The volumes extracted from segmentations performed on each MRI series did not differ from each other within women ( $P > .99$ ).

The sets of five radiomics features selected by the RF feature selection and included in each final model are shown in Table 3 along with the MRI sequence from which they were derived and their importance during the RF feature selection process. These selected features were individually among the most discriminating in single-feature ROC analysis performed on the full training set (Table 2), confirming that the RF selection process was able to detect important features. There was no strong correlation (Spearman rank or point-biserial correlation coefficient  $< |0.5|$ ) between these selected radiomics features and clinical variables in the training set (Fig 3A). The five selected features for dichotomizing advanced FIGO stage lesions and the presence of deep MI, LVSI, and high-grade tumors were

**Table 1: Surgical Histopathologic Findings in Women Recruited at Each Enrolling Institution (Training and Test Sets)**

Finding	Training Set, Institution 1 ( $n = 94$ )	Test Set, Institution 2 ( $n = 63$ )	$P$ Value
Age (y)*	66 $\pm$ 11 [43–90]	67 $\pm$ 12 [44–88]	.55
Overall FIGO stage			.80
IA	42 (45)	19 (30)	
IB	21 (22)	17 (27)	
II	3 (3.2)	13 (21)	
IIIA	6 (6.4)	5 (7.9)	
IIIC1	13 (14)	3 (4.8)	
IIIC2	8 (8.5)	2 (3.2)	
IVB	1 (1.1)	4 (6.3)	
Deep myometrial invasion			.13
Absent	51 (54)	27 (43)	
Present	43 (46)	36 (57)	
Lymphovascular space invasion			.007
Absent	45 (48)	44 (70)	
Present	49 (52)	19 (30)	
Histopathologic subtype			.16
Endometrioid	80 (85)	60 (95)	
Serous	11 (12)	2 (3.2)	
Clear cell	2 (2.1)	1 (1.6)	
Mixed or undifferentiated	1 (1.1)	0 (0)	
Histopathologic grade			.22
Low (grade 1 or 2)	61 (65)	48 (76)	
High (grade 3 or nonendometrioid)	33 (35)	15 (24)	
Pelvic lymph node metastasis			.05
Absent	71 (76)	57 (90)	
Present	23 (24)	6 (10)	

Note.—Unless otherwise specified, data are numbers of women, with percentages in parentheses. FIGO = International Federation of Gynecology and Obstetrics.

\* Data are means  $\pm$  SDs, with ranges in brackets.

**Table 2: Most Discriminative Single Radiomics Feature Based on Receiver Operating Characteristic Curve Analysis for Each Histopathologic End Point of Endometrial Cancer and Comparison with Tumor Volume for MRI Volumetry Reference**

Histopathologic Feature and Radiomics Feature*	AUC (Training Set, Institution 1)	AUC (Test Set, Institution 2)
<b>Advanced FIGO stage</b>		
1. Zone entropy, GLSZM, delayed-phase DCE-MRI (ZoneEntropy <sub>GLSZM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</sub> )	0.81 (0.74, 0.88)	0.75 (0.64, 0.86)
2. Large dependence high gray-level emphasis, GLDM, second-phase DCE-MRI (LargeDependenceHighGrayLevelEmphasis <sub>GLDM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</sub> )	0.78 (0.69, 0.85)	0.77 (0.66, 0.88)
3. Least axis length, T2WI (LeastAxisLength <sub>MORPH, T2WI, CSI:3D:1 mm</sub> )	0.77 (0.69, 0.84)	0.77 (0.66, 0.87)
MRI volumetry reference: Volume, T2WI (VoxelVolume <sub>MORPH, T2WI, CSI:3D:1 mm</sub> )	0.74 (0.66, 0.82)	0.77 (0.67, 0.87)
<b>Deep myometrial invasion</b>		
1. Least axis length, delayed-phase DCE-MRI (LeastAxisLength <sub>MORPH, DCE-MRI, CSI:3D:1 mm</sub> )	0.83 (0.73, 0.91)	0.75 (0.60, 0.86)
2. Dependence entropy, GLDM, delayed-phase DCE-MRI (DependenceEntropy <sub>GLDM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</sub> )	0.81 (0.71, 0.88)	0.70 (0.56, 0.83)
3. Zone entropy, GLSZM, delayed-phase DCE-MRI (ZoneEntropy <sub>GLSZM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</sub> )	0.80 (0.70, 0.88)	0.66 (0.51, 0.79)
MRI volumetry reference: Volume, T2WI (VoxelVolume <sub>MORPH, T2WI, CSI:3D:1 mm</sub> )	0.73 (0.64, 0.82)	0.72 (0.62, 0.83)
<b>Lymphovascular space invasion</b>		
1. Large dependence high-gray level emphasis, GLDM, third-phase DCE-MRI (LargeDependenceHighGrayLevelEmphasis <sub>GLDM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</sub> )	0.75 (0.64, 0.84)	0.76 (0.61, 0.87)
2. Inverse difference moment normalized, GLCM, ADC (InverseDifferenceMomentNormalized <sub>GLCM, ADC, FBS:25, RS:3σ, CSI:3D:1 mm</sub> )	0.75 (0.64, 0.83)	0.76 (0.60, 0.87)
3. Minor axis length, T2WI (MinorAxisLength <sub>MORPH, T2WI, CSI:3D:1 mm</sub> )	0.74 (0.62, 0.83)	0.68 (0.51, 0.80)
MRI volumetry reference: Volume, T2WI (VoxelVolume <sub>MORPH, T2WI, CSI:3D:1 mm</sub> )	0.71 (0.63, 0.80)	0.71 (0.60, 0.83)
<b>High grade</b>		
1. Inverse difference moment normalized, GLCM, DWI (InverseDifferenceMomentNormalized <sub>GLCM, DWI, FBS:25, RS:3σ, CSI:3D:1 mm</sub> )	0.71 (0.60, 0.80)	0.70 (0.55, 0.83)
2. Inverse difference normalized, GLCM, DWI (InverseDifferenceMomentNormalized <sub>GLCM, ADC, FBS:25, RS:3σ, CSI:3D:1 mm</sub> )	0.71 (0.58, 0.80)	0.69 (0.55, 0.82)
3. Long run high gray level emphasis, GLRLM, third-phase DCE-MRI (LongRunHighGrayLevelEmphasis <sub>GLRLM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</sub> )	0.70 (0.57, 0.80)	0.69 (0.52, 0.82)
MRI volumetry reference: Volume, T2WI (VoxelVolume <sub>MORPH, T2WI, CSI:3D:1 mm</sub> )	0.59 (0.47, 0.70)	0.68 (0.56, 0.81)

Note.—Data in parentheses are bootstrapped 95% CIs. 3D = three-dimensional, ADC = apparent diffusion coefficient, AUC = area under the receiver operating characteristic curve, CSI = cubic spline interpolation, DCE-MRI = dynamic contrast-enhanced MRI, DWI = diffusion-weighted imaging, FBS = fixed bin size, FIGO = International Federation of Gynecology and Obstetrics, GLCM = gray-level co-occurrence matrix, GLDM = gray-level dependence matrix, GLRLM = gray-level run length matrix, GLSZM = gray-level size zone matrix, MORPH = morphology, RS = resegmentation, T2WI = T2-weighted imaging.

\* Image Biomarker Standardization Initiative nomenclature is included in parentheses to provide a comprehensive description of feature extraction performed for each feature.

all significantly different between the groups with and without the histopathologic outcome of interest (range,  $P < .001$  to  $P = .03$ ) (Fig 3B). Figure 4 shows an example of two correctly classified lesions belonging to different histopathologic categories.

### Diagnostic Performance of the Radiomics-based RF Models

In Table 4, the diagnostic performance of RF modeling is reported for each of the four end points, with ROC curves in Figure 5. For the four outcomes, the developed RF models resulted in higher diagnostic performance than simple MRI volumetry using tumor volumes extracted on T2-weighted images ( $P < .05$ ) (Table 2).

### Diagnostic Performance of Radiologists

The diagnostic performance of the radiologists' readings for FIGO staging and assessment of deep MI is detailed in Table 4. In the sample recruited at the first institution, the diagnostic performance of the radiomics models was not significantly different from the radiologists' readings for identifying advanced FIGO stage (balanced accuracy, 80% vs 78%, respectively;  $P = .27$ ) and was superior for identifying deep MI (balanced accuracy, 86% vs 79%;  $P = .03$ ).

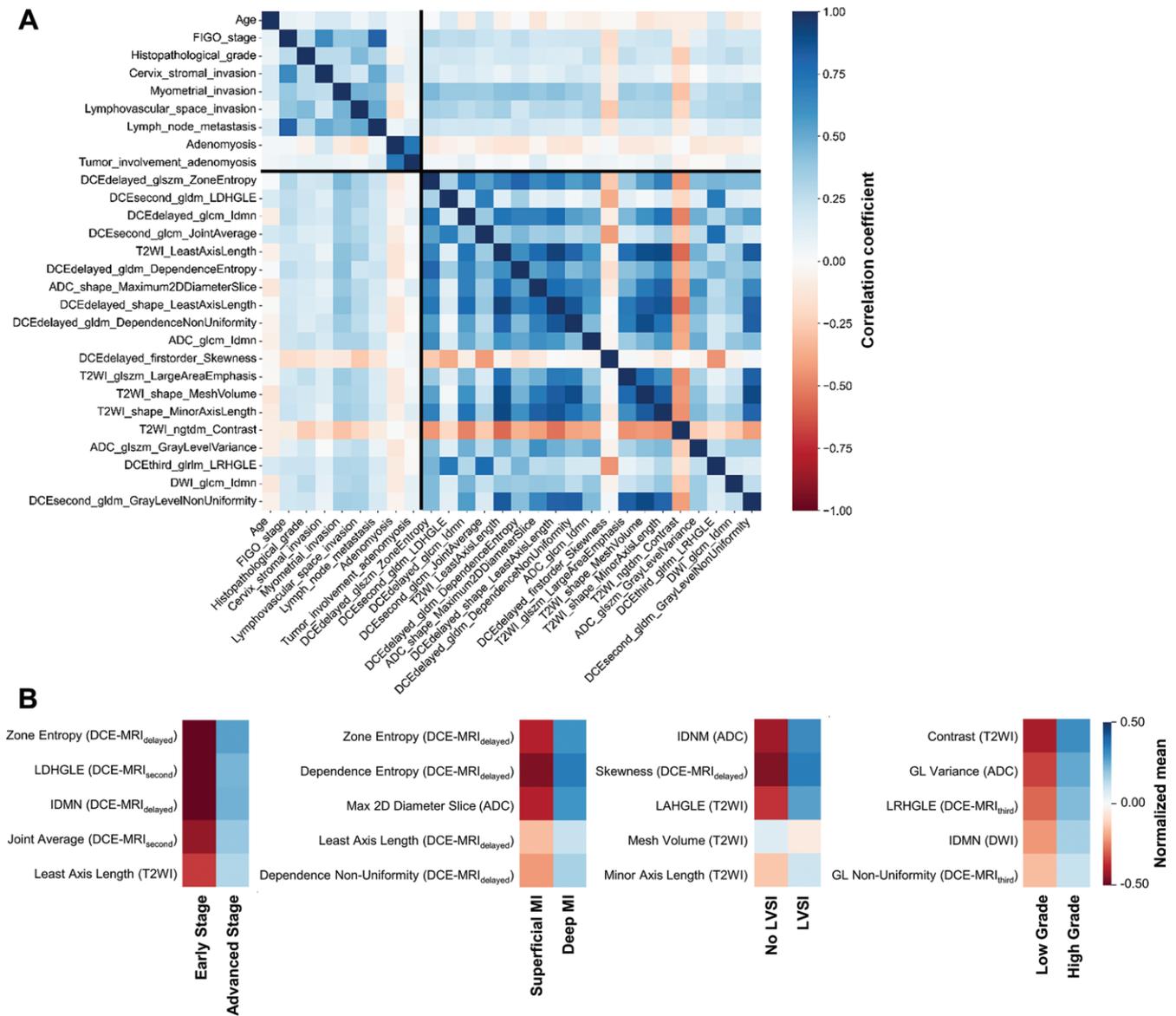
### Discussion

In this study, radiomics-based diagnostic machine learning models were developed for the noninvasive identification of

**Table 3: Description of the Selected Radiomics Features in the Random Forest Models and Their Importance for Classifying Each Histopathologic End Point of Interest**

Histopathologic Feature and Radiomics Feature	Category	MRI Sequence	IBSI Nomenclature	Mann-Whitney <i>U</i> Test <i>P</i> Value	Feature Importance
<b>Advanced FIGO stage</b>					
Zone entropy	GLSZM, texture	Delayed-phase DCE-MRI	<i>ZoneEntropy</i> <small>GLSZM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</small>	<.001	0.043
Large dependence high gray-level emphasis	GLDM, texture	Second-phase DCE-MRI	<i>LargeDependenceHighGrayLevelEmphasis</i> <small>GLDM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</small>	.002	0.020
Inverse difference moment normalized	GLCM, texture	Delayed-phase DCE-MRI	<i>InverseDifferenceMomentNormalized</i> <small>GLCM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</small>	.002	0.014
Joint average	GLCM, texture	Second-phase DCE-MRI	<i>JointAverage</i> <small>GLCM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</small>	.007	0.011
Least axis length	Shape	T2WI	<i>LeastAxisLength</i> <sub>MORPH, T2WI, CSI:3D:1 mm</sub>	.03	0.010
<b>Deep myometrial invasion</b>					
Zone entropy	GLSZM, texture	Delayed-phase DCE-MRI	<i>ZoneEntropy</i> <small>GLSZM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</small>	<.001	0.040
Dependence entropy	GLDM, texture	Delayed-phase DCE-MRI	<i>DependenceEntropy</i> <small>GLDM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</small>	<.001	0.023
Maximum two-dimensional diameter section	Shape	ADC	<i>Maximum2DdiameterSlice</i> <small>MORPH, ADC, CSI:3D:1 mm</small>	<.001	0.017
Least axis length	Shape	Delayed-phase DCE-MRI	<i>LeastAxisLength</i> <sub>MORPH, DCE-MRI, CSI:3D:1 mm</sub>	<.001	0.014
Dependence nonuniformity	GLDM, texture	Delayed-phase DCE-MRI	<i>DependenceNonUniformity</i> <small>GLDM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</small>	<.001	0.011
<b>Lymphovascular space invasion</b>					
Inverse difference moment normalized	GLCM, texture	ADC	<i>InverseDifferenceMomentNormalized</i> <small>GLCM, ADC, FBS:25, RS:3σ, CSI:3D:1 mm</small>	<.001	0.024
Skewness	First order statistics	Delayed-phase DCE-MRI	<i>Skewness</i> <sub>IH, DCE-MRI, RS:3σ, CSI:3D:1 mm</sub>	.002	0.020
Large area high gray-level emphasis	GLSZM, texture	T2WI	<i>LargeAreaHighGrayLevelEmphasis</i> <small>GLSZM, T2WI, FBS:25, RS:3σ, CSI:3D:1 mm</small>	<.001	0.014
Mesh volume	Shape	T2WI	<i>MeshVolume</i> <sub>MORPH, T2WI, CSI:3D:1 mm</sub>	<.001	0.011
Minor axis length	Shape	T2WI	<i>MinorAxisLength</i> <sub>MORPH, T2WI, CSI:3D:1 mm</sub>	<.001	0.011
<b>High grade</b>					
Contrast	NGTDM, texture	T2WI	<i>Contrast</i> <sub>NGTDM, T2WI, FBS:25, RS:3σ, CSI:3D:1 mm</sub>	.003	0.019
Gray-level variance	GLSZM, texture	ADC	<i>GrayLevelVariance</i> <small>GLSZM, ADC, FBS:25, RS:3σ, CSI:3D:1 mm</small>	.03	0.015
Long run high gray-level emphasis	GLRLM, texture	Third-phase DCE-MRI	<i>LongRunHighGrayLevelEmphasis</i> <small>GLRLM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</small>	.01	0.014
Inverse difference moment normalized	GLCM, texture	DWI	<i>InverseDifferenceMomentNormalized</i> <small>GLCM, DWI, FBS:25, RS:3σ, CSI:3D:1 mm</small>	.001	0.012
Gray-level nonuniformity	GLDM, texture	Second-phase DCE-MRI	<i>GrayLevelNonUniformity</i> <small>GLDM, DCE-MRI, FBS:25, RS:3σ, CSI:3D:1 mm</small>	.02	0.010

Note.—Feature importance in the random forest is assessed as the mean decrease in Gini impurity across the forest. *P* values for the comparison of radiomics features in dichotomized groups for each end point are reported. Image Biomarker Standardization Initiative (IBSI) nomenclature is included to provide a comprehensive description of feature extraction performed for each feature. 3D = three-dimensional, ADC = apparent diffusion coefficient, CSI = cubic spline interpolation, DCE-MRI = dynamic contrast-enhanced MRI, DWI = diffusion-weighted imaging, FBS = fixed bin size, FIGO = International Federation of Gynecology and Obstetrics, GLCM = gray-level co-occurrence matrix, GLDM = gray-level dependence matrix, GLRLM = gray-level run length matrix, GLSZM = gray-level size zone matrix, IH = intensity histogram, MORPH = morphology, NGTDM = neighboring gray-tone difference matrix, RS = ressegmentation, T2WI = T2-weighted imaging.

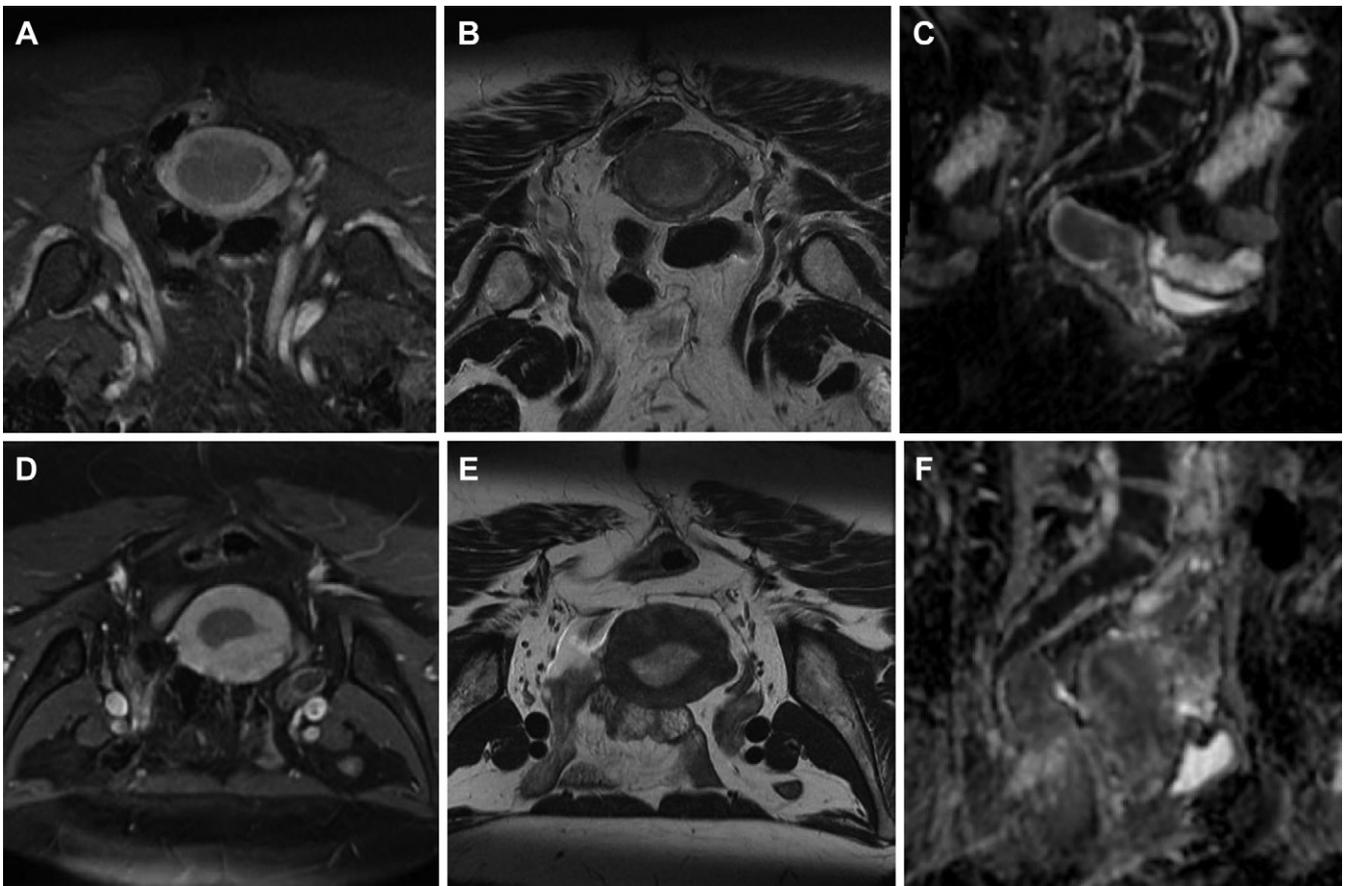


**Figure 3:** (A) Correlation heatmap of clinical variables and top five radiomics features selected with random forest modeling for each outcome. (B) Normalized mean of radiomics features in dichotomized disease groups with or without advanced International Federation of Gynecology and Obstetrics (FIGO) stage, deep myometrial invasion (MI), lymphovascular space invasion (LVSI), and high-grade endometrial carcinoma. Black lines overlaid on the heatmap represent the split between clinical variables and radiomics features. Spearman rank correlation coefficients are reported except for dichotomous variables, where point-biserial correlation coefficients are reported. ADC = apparent diffusion coefficient, DCE = dynamic contrast enhancement, DWI = diffusion-weighted imaging, GL = gray level, glcm = GL co-occurrence matrix, gldm = GL dependence matrix, glrlm = GL run length matrix, glszm = GL size zone matrix, IDMN = inverse difference moment normalized, LAHGLE = large area high GL emphasis, LDHGLE = large dependence high GL emphasis, LRHGLE = long run high GL emphasis, ngtdm = neighboring gray-tone difference matrix, 2D = two-dimensional, T2WI = T2-weighted imaging.

high-risk histopathologic features of endometrial carcinoma at multiparametric MRI. Recently, several groups of authors have studied associations between radiomics signatures derived from clinical MRI examinations and histopathologic outcomes in women with endometrial carcinoma (15–23,26,34). Some used a single MRI series to extract radiomics features (16,17,20) or extracted radiomics features on two-dimensional sections only (26,34). In our study, we focused on standardized validation and reproducibility of computational methods. We observed a consistent and clinically acceptable performance between the training and test sets (test set areas under the receiver operating characteristic

curve [AUCs], 0.84 for advanced International Federation of Gynecology and Obstetrics [FIGO] stage; 0.81 for deep myometrial invasion [MI]; 0.80 for lymphovascular space invasion [LVSI]; and 0.74 for high-grade status), despite factors such as contouring variability and image acquisition on MRI scanners from different vendors, which are likely to increase variability between data coming from two independent institutions. This points to the robustness of the methodology proposed in our study.

In contrast with our study, many prior studies did not include an external independent test set (15–22). For identifying advanced-FIGO stage endometrial carcinoma, the



**Figure 4:** Endometrial carcinoma in two women at preoperative multiparametric MRI. **(A)** Axial oblique T2-weighted image, **(B)** axial oblique delayed-phase dynamic contrast-enhanced MRI (DCE-MRI) scan 240 seconds after gadolinium chelate injection, and **(C)** sagittal apparent diffusion coefficient (ADC) map in a 75-year-old woman with International Federation of Gynecology and Obstetrics (FIGO) stage IIIA endometrial carcinoma with deep myometrial invasion (MI) and lymphovascular space invasion (LVSI) (upper row); and **(D)** axial oblique T2-weighted image, **(E)** axial oblique delayed-phase DCE-MRI scan 240 seconds after gadolinium chelate injection, and **(F)** sagittal ADC map in a 53-year-old woman with FIGO stage IA endometrial carcinoma without deep MI or LVSI (lower row). Lesions were correctly classified for FIGO stage (radiomics risk, 0.89 in the first woman and 0.16 in the second), deep MI status (0.94 and 0.11), LVSI status (0.88 and 0.20), and grade status (0.76 and 0.22). The axial oblique plane was perpendicular to the endometrial cavity, resulting in a short-axis view.

extracted radiomics signature had diagnostic performance (AUC = 0.84) similar to that reported by Yan et al (23) in their validation group (AUC = 0.85). Many previous studies have performed radiomics analyses for the identification of deep MI, which constitutes a predictive risk factor. However, the inclusion of lesions with superficial MI but with other risk factors, such as lymph node metastases, confounds the risk stratification of women with endometrial carcinoma, as only FIGO stage IA lesions are truly low risk according to European Society for Medical Oncology guidelines (24). Differentiating lesions with FIGO stage IA from lesions with higher stage may be of greater clinical relevance (23) compared with radiomics studies considering only the presence or absence of deep MI without explicitly accounting for FIGO stage, potentially confounding the results (15,17,34).

For detecting LVSI, our radiomics signature was similar to that proposed by Luo et al (19), which was built from two shape-based features (one of which was extracted from T2-weighted imaging) and two texture-based features extracted from DCE-MRI and ADC maps, akin to our model. This

model similarity suggests consistency and reproducibility of radiomics-based signatures of LVSI.

Prior studies have also shown that features extracted from ADC maps could differentiate low-grade from high-grade endometrial carcinomas (16,35). In our sample, one feature (from a texture feature category) extracted from ADC maps was included in the final radiomics model for this outcome. Interestingly, a recent study by Yamada et al (16) also included ADC-based texture features from the same category in the final RF model, suggesting a consistent radiomics-based signature of high-grade tumors. As ADC maps characterize the density of tumor cellularity by assessing the impedance of water molecule diffusion, the radiomics feature selected by our model reflects the discriminative value of ADC maps in detecting the more heterogeneous cellularity seen in high-grade endometrial lesions. In the study by Yamada et al (16), the RF model achieved a higher AUC than in our study for predicting high grade (0.89 vs 0.74), albeit using a single-institution set. In contradistinction, Fasmer et al (18) reported similar AUC for identifying nonendometrioid (0.73) and grade 3 (0.63) endometrial carcinoma with their radiomics signatures extracted from delayed-phase DCE-MRI.

**Table 4: Diagnostic Performance of the Best-Performing Random Forest Models for Each Histopathologic Feature**

Histopathologic Feature and Data Set	AUC	Sensitivity (%)	Specificity (%)	Balanced Accuracy (%)	PPV (%)	NPV (%)	<i>P</i> Value
<b>Advanced</b>							.27
<b>FIGO stage</b>							
Training, institution 1	0.85 (0.77, 0.92)	87 (45/52) [76, 96]	74 (31/42) [62, 88]	80 [72, 88]	80 (45/56) [68, 89]	82 (31/38) [68, 92]	
Test, institution 2	0.84 (0.72, 0.92)	84 (37/44) [75, 94]	79 (15/19) [55, 92]	82 [71, 91]	90 (37/41) [75, 97]	68 (15/22) [47, 88]	
Radiologists' readings	...	75 (39/52) [61, 86]	81 (34/42) [65, 91]	78 [68, 85]	83 (39/47) [69, 92]	72 (34/47) [57, 84]	
<b>Deep myometrial invasion</b>							.03
Training, institution 1	0.86 (0.75, 0.93)	84 (36/43) [71, 93]	88 (45/51) [77, 95]	86 [78, 92]	84 (36/43) [71, 93]	88 (45/52) [75, 94]	
Test, institution 2	0.81 (0.68, 0.88)	86 (31/36) [68, 94]	74 (20/27) [60, 93]	80 [70, 91]	82 (31/38) [67, 91]	80 (20/25) [60, 90]	
Radiologists' readings	...	81 (35/43) [65, 91]	77 (39/51) [63, 87]	79 [68, 85]	74 (35/47) [59, 85]	83 (39/47) [69, 92]	
<b>Lymphovascular space invasion</b>							...
Training, institution 1	0.86 (0.76, 0.92)	84 (41/49) [71, 91]	84 (38/45) [72, 92]	84 [76, 90]	84 (41/49) [72, 93]	84 (38/45) [71, 93]	
Test, institution 2	0.80 (0.67, 0.93)	84 (16/19) [65, 100]	75 (33/44) [60, 85]	80 [63, 90]	61 (16/26) [46, 75]	89 (33/37) [76, 98]	
<b>High grade</b>							...
Training, institution 1	0.74 (0.63, 0.83)	91 (30/33) [76, 98]	61 (37/61) [52, 72]	76 [57, 86]	55 (30/54) [41, 69]	92 (37/40) [80, 98]	
Test, institution 2	0.74 (0.61, 0.86)	67 (10/15) [46, 84]	83 (40/48) [70, 92]	75 [65, 86]	56 (10/18) [41, 68]	89 (40/45) [75, 95]	

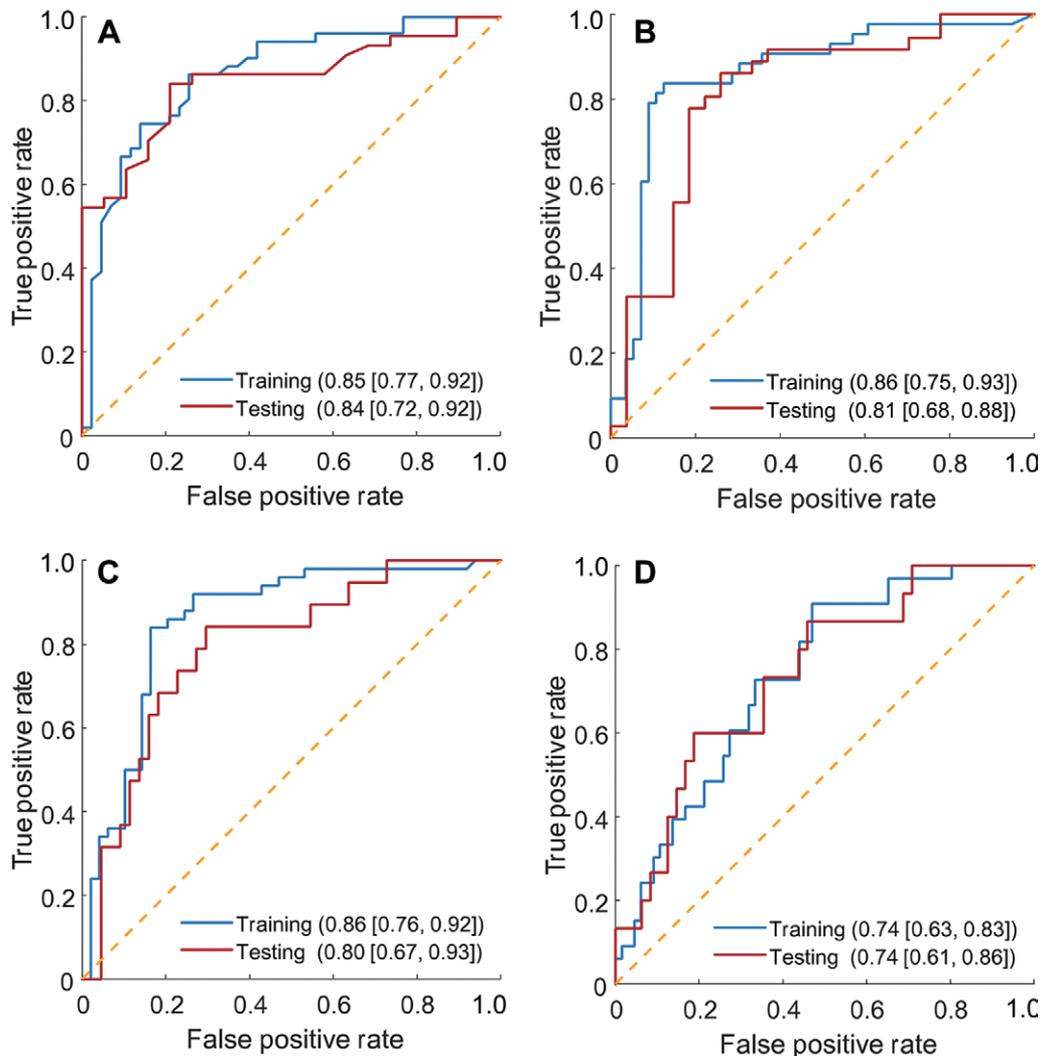
Note.—Radiologists' readings of the institution 1 data set are included for paired comparisons with radiomics-based random forest modeling for identifying advanced International Federation of Gynecology and Obstetrics (FIGO) stage and deep myometrial invasion. Data in parentheses are numerators and denominators, with 95% CIs in brackets. *P* values denote the statistical significance of the difference between radiologists and radiomics performance in the set acquired at institution 1, which was calculated using a McNemar test. AUC = area under the receiver operating characteristic curve, NPV = negative predictive value, PPV = positive predictive value.

The reported performance of radiomics modeling for differentiating early from advanced FIGO stage was not significantly different from that of the clinical readings of the five radiologists in the training set (*P* = .27), as evidenced by the largely overlapping confidence intervals, in particular for balanced accuracy. The dedicated radiomics model for identifying deep MI outperformed radiologists in preoperative patient stratification in the training set (*P* = .03). Considering the current workload of radiologists, having preoperative MRI-derived radiomics risk stratification with similar to higher diagnostic performance when compared with readings of experienced radiologists may provide more junior radiologists with enhanced decision-making capabilities while alleviating the pressure on radiology departments.

Our study had limitations. First, regions of interest were drawn by two different radiologists for the training and test sets. Second, our sample was relatively small. Third, the comparison between radiomics and radiologist readings was performed on the training set, as radiologist readings were not available for the test set. This is a limitation because the performance of radiomics on the training set is generally optimistic. However, our results from an independent test set do not

show evidence of overfitting; thus, the optimistic effect of comparison based on the training data can be considered limited. Finally, our sample had an imbalance in the distribution of histopathologic features, specifically histopathologic grade. For machine learning strategies to be effective in small sample sizes, the distribution of data for model building should ideally be similar between the training and test sets.

In conclusion, our MRI three-dimensional radiomics-based machine learning models provided consistent clinically acceptable performance for differentiating early- from advanced-International Federation of Gynecology and Obstetrics (FIGO) stage endometrial carcinoma and low- from high-risk histologic markers (tumor grade, myometrial invasion [MI], and lymphovascular space invasion) as evaluated in two independent sets from different institutions at preoperative multiparametric MRI. Our classifiers had performance similar to or higher than radiologists in identifying advanced FIGO stage and deep MI, respectively. Accurate noninvasive preoperative identification of women with low-risk endometrial carcinoma may decrease surgical time by obviating the need for sentinel lymph node dissection and the potential attendant lymphadenectomy-



**Figure 5:** Receiver operating characteristic (ROC) curves in the training and the test sets for classifying (A) advanced International Federation of Gynecology and Obstetrics stage, (B) deep myometrial invasion, (C) lymphovascular space invasion, and (D) high-grade tumors with use of radiomics-based random forest modeling. Areas under the ROC curves are reported in parentheses in the legend, with 95% CIs in brackets. The dashed yellow lines represent the ROC curve of a random classifier.

related complications. Although more evidence is needed before radiomics is used for clinical decision-making, our study suggests that the proposed technique may provide a comprehensive tool for preoperative risk stratification at MRI in women with endometrial cancer.

**Author contributions:** Guarantors of integrity of entire study, T.L.L., C.R.; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, T.L.L., A.D., R.F., J.S., P. Savadjiev, C.R.; clinical studies, T.L.L., Y.U., A.D., E.W.R., S.S., X.Z.Z., P. Soyer, C.R.; experimental studies, T.L.L., A.D., I.R.L., P. Savadjiev; statistical analysis, T.L.L., Y.U., A.D., A.C., M.V., P. Savadjiev, C.R.; and manuscript editing, T.L.L., Y.U., A.D., A.C., M.V., E.W.R., I.R.L., X.Z.Z., R.F., J.S., P. Soyer, P. Savadjiev, C.R.

**Disclosures of conflicts of interest:** T.L.L. No relevant relationships. Y.U. No relevant relationships. A.D. No relevant relationships. A.C. No relevant relationships. M.V. No relevant relationships. E.W.R. No relevant relationships. S.S. No relevant relationships. I.R.L. No relevant relationships. X.Z.Z. No relevant relationships. R.F. No relevant relationships. J.S. No relevant relationships.

P. Soyer No relevant relationships. P. Savadjiev No relevant relationships. C.R. No relevant relationships.

## References

- Parkin DM, Whelan SL, Ferlay J, Storm H. Cancer Incidence in Five Continents: Vol I to VIII. IARC CancerBase No. 7. Lyon, France: IARC Press, 2005.
- Uharek P, Mlyncek M, Ravinger J, Matejka M. Prognostic factors in women 45 years of age or younger with endometrial cancer. *Int J Gynecol Cancer* 2008;18(2):324–328.
- Sorbe B. Predictive and prognostic factors in definition of risk groups in endometrial carcinoma. *ISRN Obstet Gynecol* 2012;2012:325790.
- Ruz-Caracuel I, Ramón-Patino JL, López-Janeiro Á, et al. Myoinvasive pattern as a prognostic marker in low-grade, early-stage endometrioid endometrial carcinoma. *Cancers (Basel)* 2019;11(12):E1845.
- Andreano A, Rechichi G, Rebora P, Sironi S, Valsecchi MG, Galimberti S. MR diffusion imaging for preoperative staging of myometrial invasion in patients with endometrial cancer: a systematic review and meta-analysis. *Eur Radiol* 2014;24(6):1327–1338.
- Bi Q, Chen Y, Wu K, et al. The diagnostic value of MRI for preoperative staging in patients with endometrial cancer: a meta-analysis. *Acad Radiol* 2020;27(7):960–968.
- Sala E, Rockall AG, Freeman SJ, Mitchell DG, Reinhold C. The added role of MR imaging in treatment stratification of patients with gynec-

- colonic malignancies: what the radiologist needs to know. *Radiology* 2013;266(3):717–740.
8. Beddy P, O'Neill AC, Yamamoto AK, Addley HC, Reinhold C, Sala E. FIGO staging system for endometrial cancer: added benefits of MR imaging. *RadioGraphics* 2012;32(1):241–254.
  9. Soneji ND, Bharwani N, Ferri A, Stewart V, Rockall A. Pre-operative MRI staging of endometrial cancer in a multicentre cancer network: can we match single centre study results? *Eur Radiol* 2018;28(11):4725–4734.
  10. Helpman L, Kupets R, Covens A, et al. Assessment of endometrial sampling as a predictor of final surgical pathology in endometrial cancer. *Br J Cancer* 2014;110(3):609–615.
  11. Leitao MM Jr, Kehoe S, Barakat RR, et al. Comparison of D&C and office endometrial biopsy accuracy in patients with FIGO grade 1 endometrial adenocarcinoma. *Gynecol Oncol* 2009;113(1):105–108.
  12. Visser NCM, Reijnen C, Massuger LFAG, Nagtegaal ID, Bulten J, Pijnenborg JMA. Accuracy of endometrial sampling in endometrial carcinoma: a systematic review and meta-analysis. *Obstet Gynecol* 2017;130(4):803–813.
  13. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012;366(10):883–892.
  14. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014;5(1):4006.
  15. Rodríguez-Ortega A, Alegre A, Lago V, et al. Machine learning-based integration of prognostic magnetic resonance imaging biomarkers for myometrial invasion stratification in endometrial cancer. *J Magn Reson Imaging* 2021;54(3):987–995.
  16. Yamada I, Miyasaka N, Kobayashi D, et al. Endometrial carcinoma: texture analysis of apparent diffusion coefficient maps and its correlation with histopathologic findings and prognosis. *Radiol Imaging Cancer* 2019;1(2):e190054.
  17. Stanzione A, Cuocolo R, Del Grosso R, et al. Deep myometrial infiltration of endometrial cancer on MRI: a radiomics-powered machine learning pilot study. *Acad Radiol* 2021;28(5):737–744.
  18. Fasmer KE, Hodneland E, Dybvik JA, et al. Whole-volume tumor MRI radiomics for prognostic modeling in endometrial cancer. *J Magn Reson Imaging* 2021;53(3):928–937.
  19. Luo Y, Mei D, Gong J, Zuo M, Guo X. Multiparametric MRI-based radiomics nomogram for predicting lymphovascular space invasion in endometrial carcinoma. *J Magn Reson Imaging* 2020;52(4):1257–1262.
  20. Xu X, Li H, Wang S, et al. Multiplanar MRI-based predictive model for preoperative assessment of lymph node metastasis in endometrial cancer. *Front Oncol* 2019;9:1007.
  21. Long L, Sun J, Jiang L, et al. MRI-based traditional radiomics and computer-vision nomogram for predicting lymphovascular space invasion in endometrial carcinoma. *Diagn Interv Imaging* 2021;102(7-8):455–462.
  22. Bereby-Kahane M, Dautry R, Matzner-Lober E, et al. Prediction of tumor grade and lymphovascular space invasion in endometrial adenocarcinoma with MR imaging-based radiomic analysis. *Diagn Interv Imaging* 2020;101(6):401–411.
  23. Yan BC, Li Y, Ma FH, et al. 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(6):1872–1882.
  24. Colombo N, Preti E, Landoni F, et al; ESMO Guidelines Working Group. Endometrial cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2013;24(Suppl 6):vi33–vi38.
  25. Zwanenburg A, Vallières M, Abdalah MA, et al. The image biomarker standardization initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* 2020;295(2):328–338.
  26. Ueno Y, Forghani B, Forghani R, et al. Endometrial carcinoma: MR imaging-based texture model for preoperative risk stratification—a preliminary analysis. *Radiology* 2017;284(3):748–757.
  27. van Griethuysen JJM, Fedorov A, Parmar C, et al. Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 2017;77(21):e104–e107.
  28. Benchoufi M, Matzner-Lober E, Molinari N, Jannot AS, Soyer P. Interobserver agreement issues in radiology. *Diagn Interv Imaging* 2020;101(10):639–641.
  29. Lemaître G, Nogueira F, Aridas CK. Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning. *J Mach Learn Res* 2017;18(17):1–5. <https://jmlr.org/papers/v18/16-365.html>.
  30. Amant F, Mirza MR, Koskas M, Creutzberg CL. Cancer of the corpus uteri. *Int J Gynaecol Obstet* 2018;143(Suppl 2):37–50.
  31. Freeman SJ, Aly AM, Kataoka MY, Addley HC, Reinhold C, Sala E. The revised FIGO staging system for uterine malignancies: implications for MR imaging. *RadioGraphics* 2012;32(6):1805–1827.
  32. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika* 1947;12(2):153–157.
  33. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in Python. *J Mach Learn Res* 2011;12(85):2825–2830. <https://www.jmlr.org/papers/v12/pedregosa11a.html>.
  34. Ytre-Hauge S, Dybvik JA, Lundervold A, et al. Preoperative tumor texture analysis on MRI predicts high-risk disease and reduced survival in endometrial cancer. *J Magn Reson Imaging* 2018;48(6):1637–1647.
  35. Nougaret S, Reinhold C, Alsharif SS, et al. Endometrial cancer: combined MR volumetry and diffusion-weighted imaging for assessment of myometrial and lymphovascular invasion and tumor grade. *Radiology* 2015;276(3):797–808.