Rodriguez Ortega, A.; Alegre, A.; Lago, V.; Carot Sierra, JM.; Ten-Esteve, A.; Montoliu, G.; Domingo, S.... (2021). Machine Learning-Based Integration of Prognostic Magnetic Resonance Imaging Biomarkers for Myometrial Invasion Stratification in Endometrial Cancer. Journal of Magnetic Resonance Imaging. 54(3):987-995. https://doi.org/10.1002/jmri.27625

TITLE:

Machine learning-based integration of prognostic MR imaging biomarkers for myometrial invasion stratification in endometrial cancer

ABSTRACT

Background

Estimation of the depth of myometrial invasion (MI) in endometrial cancer is pivotal in the preoperatively staging. Magnetic Resonance (MR) reports suffer from human subjectivity. Multiparametric MR imaging radiomics and parameters may improve the diagnostic accuracy.

Purpose

To discriminate between patients with MI≥50% using a machine learning-based model combining texture features and descriptors from preoperatively MR images.

Study Type

Retrospective.

Population

143 women with endometrial cancer were included. The series was split into training (n=107, 46 with MI≥50%) and test (n=36, 16 with MI≥50%) cohorts.

Field Strength/Sequences

Fast spin echo T2-weighted (T2W), diffusion-weighted (DW) and T1-weighted gradient echo dynamic contrast-enhanced (DCE) sequences were obtained at 1.5 or 3T magnets.

Assessment

Tumors were manually segmented slice-by-slice. Texture metrics were calculated from T2W and ADC map images. Also, the apparent diffusion coefficient (ADC), wash-in slope, wash-out slope, initial area under the curve at 60 seconds and at 90 seconds, initial slope, time to peak and peak amplitude maps from DCE sequences were obtained as parameters. MR diagnostic models using single-sequence features and a combination of features and parameters from the three sequences were built to estimate MI using Adaboost methods. The pathological depth of MI was used as gold standard.

Statistical Test

Area under the receiver operating characteristic curve (AUROC), sensitivity, specificity, accuracy, positive predictive value, negative predictive value, precision and recall were computed to assess the Adaboost models performance.

Results

The diagnostic model based on the features and parameters combination showed the best performance to depict patient with MI≥50% in the test cohort (accuracy=86.1% and AUROC=87.1%). The rest of diagnostic models showed a worse accuracy (accuracy=41.67-63.89% and AUROC=41.43-63.13%).

Data Conclusion

The model combining the texture features from T2W and ADC map images with the semi-quantitative parameters from DW and DCE series allow the preoperative estimation of myometrial invasion.

KEYWORDS

Endometrial cancer; Magnetic Resonance; Radiomics; Diffusion, Perfusion.

INTRODUCTION

Endometrial carcinoma is the most common gynecologic malignancy in developed countries (1). Five-year survival rates vary between 96% for an early stage I or II, and 20% for stage IV disease (2). International Federation of Gynecology and Obstetrics (FIGO) proposes to perform hysterectomy and bilateral salpingo-oophorectomy, as well as pelvic and para-aortic lymphadenectomy, for a proper staging depending on tumor characteristics (3). However, endometrial cancer is frequently present in patients with obesity, hypertension or diabetes, where surgery might be complex and risky (4). Preoperative staging determines the risk group, assessing the risk-benefit ratio of surgery in order to prevent overtreatment and aids in selection of and planning for the optimal treatment and surgery. Therefore, the role of imaging as a non-invasive, accurate diagnostic tool is crucial in determining preoperatively the best treatment and prognosis estimations of patients with endometrial carcinoma (5).

Myometrial invasion (MI) in endometrial cancer shows a high correlation with prevalence of metastasis and survival of patients (6). As MI can only be measured when the uterus is removed (7), an initial preoperative estimation of the MI factor may be useful in treatment planning and follow-up strategies (8, 9).

MI has been evaluated through several imaging modalities, including MR, computed tomography (CT), and ultrasound (US) (10, 11). However, MR has a much higher diagnostic accuracy compared to CT and US (12, 13). MR images usually used in the evaluation of endometrial cancer include high-resolution T2-weighted (T2W), diffusion-weighted (DW), and dynamic contrast-enhanced (DCE) sequences. The T2W and DCE acquisitions have been shown to be useful in tumor grading and neovascularity assessment (14) as well as to estimate the depth of MI (15). DW imaging improves the

preoperative staging, providing information on tumor cellularity, myometrial penetration, and lymphovascular invasion (16).

However, the diagnostic accuracy of MR imaging related to radiologist reading subjectivity in current clinical setting.

Imaging biomarkers allow for resolving this limitations by providing objective information through the quantification of structural and dynamic features related to relevant clinical endpoints of the disease (17). Specifically, texture analysis applies mathematical models on digital images to evaluate the different levels of pixel intensities as well as their distribution and interrelation, objectively characterizing imaging patterns that are invisible to radiologists. In endometrial cancer, some texture analysis solutions have been suggested to be useful in this setting (18) as well as the use of well-established dynamic parameters such as apparent diffusion coefficient (ADC), obtained from DW images, and DCE metrics (19). The combination of T2W, DW and DCE derived metrics might also improve the MI estimation(7, 18, 20–22) . Therefore, we hypothesized that the inclusion of texture features from different MR images will increase the accuracy of the diagnostic models based on machine-learning for MI estimation.

The aim of this study was to construct and assess the ability of a machine learning-based model combining texture-based features from T2W and ADC maps and statical descriptors of the ADC maps and semi-quantitative maps from DCE images to estimate myometrial invasion in endometrial cancers prior to surgery.

METHODS AND MATERIALS

Study Population

The hospital Institutional Ethical Review Board approved our observational single-center study, and the need to obtain informed consent was waived. Consecutive patients with histologically verified endometrial cancer having a pelvic MR exam prior surgery and a minimum of 2-year-follow-up in our center were selected for the study. No neoadjuvant therapy (neither radiotherapy nor chemotherapy) was performed before surgery. All MR exams included T2W, DW and DCE images, having been subjectively check for quality and absence of relevant respiratory artifacts, which could cause misdiagnosis in the pelvic region. This initial series included 150 cases over a period of 6 years (from January 2013 to December 2018). The exclusion criteria were: (a) inadequate histopathological reports (n=4); (b) MR images with a wrong acquisition planning (n=2); and (c) tumor not visible on MR images (n=1). The final study population consisted of 143 patients (64.7±10.7 years, mean ± standard deviation (SD)). The patient selection process is shown in Figure 1.

Surgical Characteristics

All 143 patients were evaluated by a multidisciplinary tumor board. Staging was mainly based on the preoperative MR and US imaging findings. All patients underwent total hysterectomy and bilateral adnexectomy. Post-surgery pathological tumor staging was used as reference standard to evaluate the performance of the models. The final histopathological results showed that 62 (43%) of the patients had MI 50% or greater while in 81 (57%) patients the MI was below 50% (Table 1).

Surgical results also showed lymphovascular space invasion (LSVI) in 16 (11%) patients. Regarding tumor grade, 26 (18%) were high-grade (G3), 20 (14%) intermediate grade (G2), and 97 (68%) low-grade tumors. Finally, 104 (73%) patients were classified

according to FIGO recommendations as stage 1 (IA and IB), 20 (14%) were stage 2 (II), 18 (7%) were stage 3 (III, IIIA, IIIB and IIIC), and 1 (1%) were stage 4 (IV).

MR Imaging

MR exams were preoperatively performed at either 1.5 T or 3T (Signa HDxt, GE Healthcare, Milwaukee, WI) with an eight-channel receive surface coil. The sequences acquired included T2W, DW, and DCE images (Table 2). The MR acquisition plane orientation was transverse oblique (perpendicular to the main axis of the uterus) for the T2W sequence and transverse for the DW and DCE sequences. The DW gradients were averaged in three orthogonal directions. DCE images were acquired with one dynamic prior to and three dynamics after the injection of a gadolinium-based contrast agent (Multihance®, Bracco, Milan, Italy), for a total of four dynamics with temporal resolution of 80 seconds between dynamics. A bolus injection of 0.1 mmol/kg contrast agent followed by 40mL of saline flush was administered at 3 ml/s using a power injector (Optistar® Elite, Guerbet Headquarter, Paris, France).

Image Analysis

The image analysis pipeline is depicted in Figure 2. All MR images were exported in Digital Imaging and Communication on Medicine (DICOM) format from the Picture Archiving and Communication System (PACS) and anonymized for the study. DICOM files were converted to Neuroimaging Informatics Technology Initiative (NIFTI) files (dicom2nii software). Two radiologists (AA & LMB) with 4 and 30 years of experience manually drew, slice by slice, the region of interest (ROI) delimiting the endometrial cancer on the T2W images by using ITK-SNAP 3.6.0 (23). Each radiologist segmented half of cases and checked the segmentation performed by the other radiologist.

Discrepancies were resolved by consensus. The radiologists were blinded to the pathological results, including tumor grade, myometrial penetration, lymphovascular invasion and FIGO staging.

The preprocessing steps included the application of filters and registration to improve and harmonize image quality across MR scanners. For the DW images, the Smallest Univalue Segment Assimilating Nucleus (SUSAN) filtering method was used to reduce noise while preserving the internal image structure (24). For the DCE images, a Gaussian filter with a window of 5x5 was applied. The filters applied to the T2W images included a denoising filter based on the non-local means algorithms (25) and a filter to correct for the non-uniformity of the low frequency intensity during MR acquisition based on the N4 bias field correction algorithm (26). DWI and DCE series were co-registered to the lowest b-value (b=0 s/mm^2) and to the first dynamic (prior to bolus injection), respectively, to reduce spatial misalignments (intra-sequence registration). After that, DW and DCE series were also co-registered to the T2W series (inter-sequence registration) to guarantee the spatial coherence within a common reference space. Rigid and non-rigid registration steps were performed for the intra- and inter-sequence registration, respectively, using Elastix toolbox v.4.8. (27, 28).

All images were processed and analyzed using an in-house program using MATLAB (2015a, The MathWorks). DW images (b=0 and 800 s/mm^2) were used to calculate the apparent diffusion coefficient (ADC) maps for each voxel with a mono-exponential signal decay. The DCE images were analyzed through time-intensity curve (TIC) methods (29). A semi-quantitative analysis was used to computed the contrast-enhancement curve shape descriptors to provide pathophysiology information of the tumor (30). The extracted maps were the wash-in slope (WIS), wash-out slope (WOS), initial area under the curve at 60 (IAUC60) and 90 (IAUC90) seconds, initial slope (IS), time to peak (TTP), and peak

amplitude (Peak) (29). Figure 3 shows two representative examples of the ADC and IAUC60 maps for two patients. The segmented ROIs in the T2W-space were also used to extract statistical descriptors, including the mean, median, standard deviation (SD) and percentiles (25%, 75% and 90%) from the ADC and the different DCE maps.

Feature Selection

Texture-based features were calculated on the ROI for both the T2W and ADC map images using first-order statistics (n=19), shape-based 3D (n=16) and 2D (n=10), the Gray Level Co-Occurrence Matrix (GLCM, n=24), Grey Level Run Length Matrix (GLRLM, n=16) and Gray Level Size Zone Matrix (GLSZM, n=16) Neighbouring Gray Tone Difference Matrix (NGTDM, n=5) and Gray Level Dependence Matrix (GLDM, n=14). One-hundred and twenty texture features were calculated using the Pyradiomics software (31). All texture features meaning can be consulted on the Pyradiomics software webpage (https://pyradiomics.readthedocs.io/en/latest/features.html).

Two steps were applied for feature selection. First, the ANOVA F-value was calculated for each feature, using MI as the dependent variable. All features were organized from highest to lowest according to their calculated F-values. Features that showed non-significant differences (p-value ≥0.05) were removed. The Pearson coefficient correlation was then calculated from remaining features and, when two or more features showed a correlation greater than 0.8, the feature with the highest F-value was selected, dropping the others. Finally, the magnetic field strength, B_0, was included after the feature selection process in order to evaluate B_0 bias on the diagnosis model.

Model Training

The Adaboost machine learning method was applied to develop the diagnostic models to stratify between two different MI stages (patients with MI ≥ 50% and patients with MI < 50%) (18).

Four Adaboost models were built using inputs obtained from: (I) the T2W texture features, (II) ADC maps texture features, (III) statistical descriptors from all semi-quantitative maps, and (IV) a combination from the above. All Adaboost models were trained by a 10-fold cross-validation method to choose the optimal model hyperparameters. The hyperparameters to be optimized were maximum number of estimators, the learning rate and the used algorithm. During this training phase, many models were built by varying these hyperparameters for the MI. The best performed model with the highest area under the receiver operating characteristic curve (AUROC) was selected. Gini importance, or Mean Decrease in Impurity calculation, was used to know how much each feature influences the best performing model.

For each Adaboost model, the patient cohort was randomly split into training, 107 patients (75%), and a test cohort, 36 patients (25%), with the same proportion of each depth of myometrial invasion. All features of the training cohort were standardized using the robust scale normalization method. The same standardization was also applied to the test cohort.

All models and statistical analyses were performed using Python 3.6.10 (Python Software Foundation, https://www.python.org/). The Adaboost diagnostic models were built using the "sklearn" (32) and "panda" (33) packages. For each model, the sensitivity, specificity, accuracy, positive predictive value (PPV), negative predictive value (NPV), AUROC, precision and recall were computed for the test cohort.

RESULTS

The diagnostic models' ability to stratify the myometrial invasion is summarized in Table 3. The model with the best performance used all texture features obtained from T2W and ADC mapping images and the statistical descriptors obtained from ADC map images and the semi-quantitative map images (accuracy: 86%, recall: 83% and AUROC: 87%). The model with the DCE descriptors alone had the worst performance (accuracy: 42%, recall: 40% and AUROC: 41%).

The optimal hyperparameters for the combined Adaboost model were as follows: maximum number of estimators = 100, learning rate = 1.0, and algorithm='SAMME.R'. To train the combined model, 32 extracted features were selected. Table 4 shows the selected inputs and their Gini importance in the combined texture and descriptors model.

DISCUSSION

Our results showed that the combination of texture features, derived from T2W and ADC maps (measured from DW images), and statistical descriptors from ADC and semi-quantitative maps (measure from DCE images) has the highest diagnostic accuracy for discriminating patients with MI higher than 50%. This approach is in agreement with a study by Ueno et al. (18), who built a prediction random forest model with texture parameters from T2W, DW and DCE images to estimate MI in 137 women. However, our model (accuracy = 86% and AUROC = 87%) slightly outperformed theirs (accuracy = 81% and AUROC = 84%) for depth of MI. Our improved results might be due to the use of a classification model more robust to overfitting (34). Adaboost was selected because of its low generalization error its strength towards overfitting to find the diagnostic model to stratify the different invasion stages. In addition, the use of our feature selection method may also have influenced in these differences. Our feature

selection methods allowed to work with those features that had the greatest relevance in the diagnosis of the depth of MI, removing from further analysis those that had zero variance or a high correlation with other features. Despite these differences, our results support that the combination of radiomics features (T2W and ADC images) and statical descriptors (from ADC and semi-quantitative images) improves the accuracy of myometrial invasion prediction (21). In this sense, the diagnostic accuracy of our model was better than the reported by expert radiologists (78- 83%)(18).

It is worth noting the importance of shape features for the Adaboost diagnostic model. This could be expected since previous studies have shown that both tumor size and volume, measured from MR images, are linked with the depth of MI (15, 21, 35, 36). Another relevant result is that the influence of MR field strength on the diagnostic model was negligible. This is important because it demonstrates that our diagnostic model can be used with standard-of-care MR images, regardless of magnetic field strength. This latter would have to be tested in future studies in which the Adaboost model is validated with MR images acquired from different manufactures.

Limitations

The use of Real-World Data introduces challenges to models regarding input data quality and stability. We were able to minimize variability as in our hospital the same MR machines and quite similar MR protocols were used during the 6 years period of the study. Our models should be validated in larger cohort of patients from different hospitals and different MR scanners and protocols to evaluate the reproducibility of the results to establish an accurate preoperative initial prognostic estimation of deep MI. If proven reproducible, this will provide stability and strength to our diagnostic model to be transfer into clinical practice as a decision support tool.

Our diagnostic models were based on Adaboost methods because the number of biomarkers was large, and these models can work with any dimensionality (34). Other ensemble learning methods can be tested in order to study the methods best fitting the data. Finally, other tumor endpoints such as FIGO stages or histologic grade were not studied because samples were unbalanced and, therefore, accuracies could not be properly calculated.

Conclusion

An Adaboost diagnostic model based on textural features and statistical descriptors from standard-of-care multiparametric MR exam allows preoperative differentiation between 50% depth of MI, regardless of the magnetic field strength. Before implementation in a real clinical setting, this machine learning model should be validated with patients from different hospitals and MR scanners. The use of diagnostic models based on feature and descriptors extraction from MR images has a high potential to provide more accurate tools for the presurgical diagnosis of myometrial invasion.

REFERENCES

1. Ghandili S, Bardenhagen J, Izbicki JR: Pancreatic Metastasis from Endometrial Carcinoma. *J Gastrointest Surg* 2018; 23:377–378.

2. Siegel RL, Miller KD, Jemal A: Cancer statistics, 2015. *CA Cancer J Clin* 2015; 65:5–29.

3. Pecorelli S: Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium. *Int J Gynecol Obstet* 2009; 105:103–104.

4. Colombo N, Creutzberg C, Amant F, et al.: ESMO-ESGO-ESTRO Consensus Conference on Endometrial Cancer: diagnosis, treatment and follow-up. *Ann Oncol* 2015; 27:16–41.

5. Morice P, Leary A, Creutzberg C, Abu-Rustum N, Darai E: Endometrial cancer. *Lancet* 2016; 387:1094–1108.

6. Lin G, Ng K-K, Chang C-J, et al.: Myometrial Invasion in Endometrial Cancer: Diagnostic Accuracy of Diffusion-weighted 3.0-T MR Imaging—Initial Experience. *Radiology* 2009; 250:784–792.

7. 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.

8. Dogan D, Inan N, Sarisoy HT, et al.: Preoperative evaluation of myometrial invasion in endometrial carcinoma: diagnostic performance of 3T MRI. *Abdom Imaging* 2012; 38:388–396.

9. Frei KA, Kinkel K, Bonél HM, Lu Y, Zaloudek C, Hricak H: Prediction of Deep

Myometrial Invasion in Patients with Endometrial Cancer: Clinical Utility of Contrast-enhanced MR Imaging-A Meta-analysis and Bayesian Analysis. *Radiology* 2000; 216:444–449.

10. Ytre-Hauge S, Salvesen O, Krakstad C, Trovik J, Haldorsen IS: Tumour texture features from preoperative CT predict high-risk disease in endometrial cancer. *Clin Radiol* 2021; 76:79.e13-79.e20.

11. Cubo-Abert M, Díaz-Feijoo B, Bradbury M, et al.: Diagnostic performance of transvaginal ultrasound and magnetic resonance imaging for the preoperative evaluation of low-grade endometrioid endometrial carcinoma: a single-center prospective comparative study. *Ultrasound Obstet Gynecol* 2021.

12. Otero-García MM, Mesa-Álvarez A, Nikolic O, et al.: Role of MRI in staging and follow-up of endometrial and cervical cancer: pitfalls and mimickers. *Insights Imaging* 2019; 10.

13. Nougaret S, Lakhman Y, Vargas HA, et al.: From Staging to Prognostication. *Magn Reson Imaging Clin N Am* 2017; 25:611–633.

14. Kinkel K, Forstner R, Danza FM, et al.: Staging of endometrial cancer with MRI: Guidelines of the European Society of Urogenital Imaging. *Eur Radiol* 2009; 19:1565–1574.

15. 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; 276:797–808.

16. Hori M, Kim T, Onishi H, et al.: Endometrial cancer: preoperative staging using three-dimensional T2-weighted turbo spin-echo and diffusion-weighted MR imaging at

3.0 T: a prospective comparative study. *Eur Radiol* 2013; 23:2296–2305.

17. Martí-Bonmatí L, Alberich-Bayarri A: *Imaging Biomarkers: Development and Clinical Integration*. Springer; 2016.

18. Ueno Y, Forghani B, Forghani R, et al.: Endometrial Carcinoma: MR Imaging–based Texture Model for Preoperative Risk Stratification—A Preliminary Analysis. *Radiology* 2017; 284:748–757.

19. Fasmer KE, Bjørnerud A, Ytre-Hauge S, et al.: Preoperative quantitative dynamic contrast-enhanced MRI and diffusion-weighted imaging predict aggressive disease in endometrial cancer. *Acta radiol* 2017; 59:1010–1017.

20. Bhosale P, Ma J, Iyer R, et al.: Feasibility of a reduced field-of-view diffusion-weighted (rFOV) sequence in assessment of myometrial invasion in patients with clinical FIGO stage I endometrial cancer. *J Magn Reson Imaging* 2015; 43:316–324.

21. Nougaret S, Horta M, Sala E, et al.: Endometrial Cancer MRI staging: Updated Guidelines of the European Society of Urogenital Radiology. *Eur Radiol* 2018; 29:792–805.

22. Beddy P, Moyle P, Kataoka M, et al.: Evaluation of depth of myometrial invasion and overall staging in endometrial cancer: Comparison of diffusion-weighted and dynamic contrast-enhanced MR imaging. *Radiology* 2012; 262:530–537.

23. Yushkevich PA, Piven J, Hazlett HC, et al.: User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* 2006; 31:1116–1128.

24. Smith SM, Brady JM: SUSAN: a new approach to low level image processing. *Int J Comput Vis* 1997; 23:45–78.

25. Buades A, Coll B, Morel JM: A non-local algorithm for image denoising. In *Proc - 2005 IEEE Comput Soc Conf Comput Vis Pattern Recognition, CVPR 2005*. *Volume II*. IEEE Computer Society; 2005:60–65.

26. Tustison NJ, Avants BB, Cook PA, et al.: N4ITK: Improved N3 bias correction. *IEEE Trans Med Imaging* 2010; 29:1310–1320.

27. Klein S, Staring M, Murphy K, Viergever MA, Pluim J: elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Trans Med Imaging* 2010; 29:196–205.

28. Shamonin DP, Bron EE, Lelieveldt BPF, Smits M, Klein S, Staring M: Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease. *Front Neuroinform* 2014; 7:50.

29. Assili S, Kazerooni AF, Aghaghazvini L, Rad HRS, Islamian JP: Dynamic contrast magnetic resonance imaging (DCE-MRI) and diffusion weighted MR imaging (DWI) for differentiation between benign and malignant salivary gland tumors. *J Biomed Phys Eng* 2015; 5:157.

30. Petrillo A, Fusco R, Petrillo M, et al.: Standardized Index of Shape (SIS): a quantitative DCE-MRI parameter to discriminate responders by non-responders after neoadjuvant therapy in LARC. *Eur Radiol* 2015; 25:1935–1945.

31. van Griethuysen JJM, Fedorov A, Parmar C, et al.: Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Res* 2017; 77:e104–e107.

32. Pedregosa F, Michel V, Grisel O, et al.: *Scikit-Learn: Machine Learning in Python*. *Volume 12*; 2011.

33. Mckinney W: *Data Structures for Statistical Computing in Python*. 2010.

34. Rätsch G, Onoda T, Müller KR: Soft margins for AdaBoost. *Mach Learn* 2001; 42:287–320.

35. Mainenti PP, Pizzuti LM, Segreto S, et al.: Diffusion volume (DV) measurement in endometrial and cervical cancer: A new MRI parameter in the evaluation of the tumor grading and the risk classification. *Eur J Radiol* 2016.

36. Yan B, Liang X, Zhao T, Niu C, Ding C, Liu W: Preoperative prediction of deep myometrial invasion and tumor grade for stage I endometrioid adenocarcinoma: a simple method of measurement on DWI. *Eur Radiol* 2019; 29:838–848.

TABLES

Table 1. Patient and tumor characteristics.

Variable	**Data (n=143)**
Age at diagnosis (years), mean ± SD	64.7 ± 10.7
FIGO Stage, n (%)	
IA	62 (43.4)
IB	42 (29.4)
II	20 (14.0)
III	1 (0.7)
IIIA	4 (2.8)
IIIB	3 (2.1)
IIIC	10 (7.0)
IV	1 (0.7)
Depth of Myometrial Invasion, n (%)	
<50%	81 (56.6)
≥50%	62 (43.4)
Histologic grade, n (%)	
G1	97 (67.8)
G2	20 (14.0)
G3	26 (18.2)
LVSI, n (%)	
Yes	16 (11.2)
No	127 (88.8)

Data are reported as number of patient and percentage of total in parentheses, unless otherwise noted. FIGO: International Federation of Gynecology and Obstetrics, SD: standard deviation, LVSI: lymphovascular space invasion.

Table 2. MR imaging sequences and parameters

Field Strength	Scan Type	Sequence	Acquisition Plane	TR/TE (ms)	Reconstruction Matrix	NEX	FOV (cm^2)	Slice Thickness (mm)	Slice Gap (mm)	b-values (s/mm^2)
3 Tesla	T2W	FSE	Axial,	1800/94.7	512x512	2	18×18	3	0.3	
3 Tesla	DW	SE-EPI	Axial	6275/95.4	256x256	4	38x38	5.5	1	0 & 800
3 Tesla	DCE	3D T1W GE	Axial	3.78/1.8	256x256	1	28x28	2	-1	
1.5 Tesla	T2W	FSE	Axial,	3320/109.1	512x512	2	20x20	3.5	0.5	
1.5 Tesla	DW	SE-EPI	Axial	11500/90.0	256x256	4	36x36	5	1.5	0 & 800
1.5 Tesla	DCE	3D T1W GE	Axial	3.3/1.5	512x512	1	28x28	2	-1	

T2W: T2-weighted, DW: diffusion-weighted, DCE: dynamic contrast-enhanced, TR: repetition time, TE: echo time, NEX: number of excitations, FOV: field of view, FSE: fast spin echo, SE: spin echo, EPI: echo planar imaging, T1W: T1-weighted, GE: gradient echo

Table 3. Diagnostic performance in percent of Adaboost methods in the different myometrial invasion models.

MR Imaging features dataset	Sensitivity	Specificity	Accuracy	PPV	NPV	AUROC	Precision	Recall
T2W Texture	50.0	68.75	58.33	52.38	66.66	59.38	52.38	68.75
ADC Texture	70.0	56.25	63.89	60.0	63.89	63.13	60.0	56.25
DCE descriptors	42.86	40.0	41.67	33.33	50.0	41.43	33.33	40.0
T2W Texture + ADC Texture + DCE descriptors + ADC descriptors	80.95	93.33	86.11	77.78	94.44	87.14	77.78	93.33

MR: magnetic resonance, T2W: T2-weighted, ADC: apparent diffusion coefficient, DCE: dynamic contrast-enhanced, PPV: positive predictive value, NPV: negative predictive value, AUROC: area under the receiver operating characteristic curve.

Table 4. Features selected for best Adaboost model of myometrial invasion. The table shows the feature name, the sequence from which the feature has been extracted and the Gini importance. A greater Gini value indicates a greater importance of the feature in the diagnostic model.

Feature/Parameter	Texture Class	Image type	Importance
Elongation	Shape 2D	T2W	0.09
Minimum	First Order	T2W	0.08
Flatness (ADC)	Shape 3D	DWI	0.08
Coarseness (ADC)	NGTDM	DWI	0.07
Small Dependence Low Gray Level Emphasis	GLDM	T2W	0.07
Large Area High Gray Level Emphasis	GLSZM	T2W	0.06
WIS_p90		DCE	0.06
Busyness (ADC)	NGTDM	DWI	0.05
Zone Percentage	GLSZM	T2W	0.05
Cluster Shade (ADC)	GLCM	DWI	0.04
WIS_Kurtosis		DCE	0.04
Strength	NGTDM	T2W	0.04
Least Axis Length (ADC)	Shape 3D	DWI	0.03
Sphericity (ADC)	Shape 2D	DWI	0.03
Kurtosis		DCE	0.03
Inverse Difference Normalized (ADC)	GLCM	DWI	0.03
Surface Area to Volume ratio (ADC)	Shape 3D	DWI	0.03
Skewness	First Order	T2W	0.02
Inverse Difference Moment Normalized	GLCM	T2W	0.02
WIS_Std		DCE	0.01
Large Dependence High Gray Level Emphasis	GLDM	T2W	0.01
Total Energy (ADC)	First Order	DWI	0.01
Informational Measure of Correlation2	GLCM	T2W	0.01
Zone Entropy	GLSZM	T2W	0.01
Correlation (ADC)	GLCM	DWI	0.01
Skewness (ADC)	First Order	DWI	0.01
Kurtosis	First Order	T2W	0.01

Least Axis Length	Shape 3D	T2W	0
Small Area Low Gray Level Emphasis	GLSZM	T2W	0
Surface Area to Volume ratio	Shape 3D	T2W	0
Cluster Shade	GLCM	T2W	0
Flatness	Shape 3D	T2W	0
B0		--	0

FIGURES LEGENDS

Figure 1. Flowchart of the patient selection process.

Figure 2. Schematic overview of the MR image analysis pipeline.
T2W: T2-weighted, DWI: diffusion-weighted imaging, DCE: dynamic contrast-enhanced, P0-P3: from the initial dynamic acquisition to last dynamic acquisition, ADC: apparent diffusion coefficient, IAUC60: initial area under the curve at 60 seconds, IAUC90: initial area under the curve at 90 seconds, WIS: wash-in slope, WOS: wash-out slope, IS: initial slope, TTP: time to peak, PEAK: peak amplitude.

Figure 3. Axial oblique T2-weighted image (left column) and parametric maps from diffusion-weighted (DW) (Apparent diffusion coefficient (ADC) maps [mm2/s] – middle column) and dynamic contrast-enhanced (DCE) (initial area under the curve at 60 seconds (IAUC60) maps – right column). A) A 51-year-old female with histopathologically proven grade 1 endometrial cancer, International Federation of Gynecology and Obstetrics (FIGO) stage IA, without myometrial and lymphovascular space invasion. B) A 63-years-old female with histopathologically proven grade 1 endometrial cancer, FIGO stage III, without lymphovascular space invasion and with deep myometrial.

Figure 1.

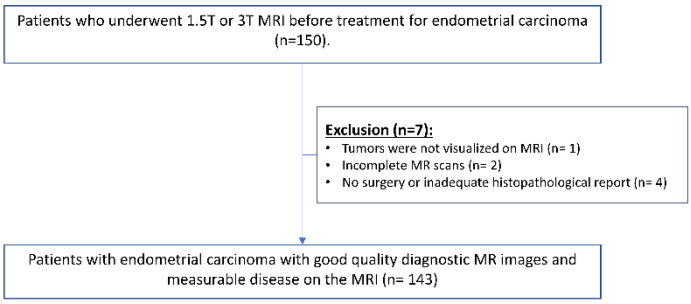
Patients who underwent 1.5T or 3T MRI before treatment for endometrial carcinoma (n=150).
Exclusion (n=7):
Tumors were not visualized on MRI (n= 1)
Incomplete MR scans (n= 2)
No surgery or inadequate histopathological report (n= 4)
Patients with endometrial carcinoma with good quality diagnostic MR images and measurable disease on the MRI (n= 143)

Figure 2.

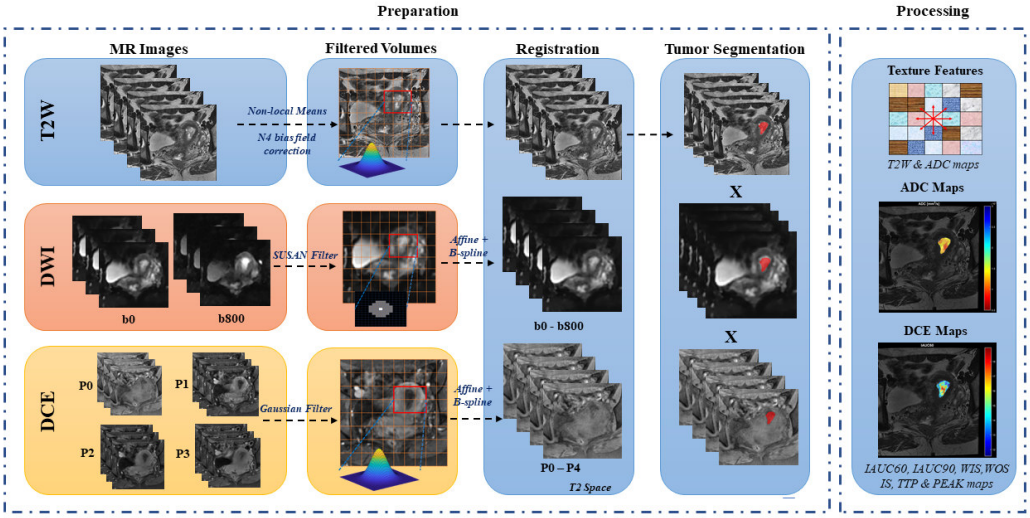
Preparation
Processing
MR Images
Filtered Volumes
Registration
Tumor Segmentation
T2W
Non-local Means
N4 bias field correction
DWI
SUSAN Filter
Affine + B-spline
b0
b800
b0 - b800
DCE
P0
P1
P2
P3
Gaussian Filter
Affine + B-spline
P0 – P4
T2 Space
X
X
Texture Features
T2W & ADC maps
ADC Maps
DCE Maps
IAUC60, IAUC90, WIS, WOS
IS, TTP & PEAK maps

Figure 3.