Radiology

ORIGINAL RESEARCH • GENITOURINARY IMAGING

Diagnostic Algorithm to Differentiate Benign Atypical Leiomyomas from Malignant Uterine Sarcomas with Diffusion-weighted MRI

Cendos Abdel Wahab, MD • Anne-Sophie Jannot, MD • Pietro A. Bonaffini, MD • Camille Bourillon, MD • Caroline Cornou, MD • Marie-Aude Lefrère-Belda, MD • Anne-Sophie Bats, MD, PhD • Isabelle Thomassin-Naggara, MD, PhD • Alexandre Bellucci, MD • Caroline Reinhold, MD, PhD • Laure S. Fournier, MD, PhD

From the Departments of Radiology (C.A.W., C.B., A.B., L.S.F.), Medical Informatics and Public Health (A.S.J.), Gynecologic and Breast Oncologic Surgery (C.C., A.S.B.), and Pathology (M.A.L.B.), AP-HP, Hôpital Européen Georges Pompidou, 20 Rue Leblanc, Université de Paris, F-75015 Paris, France; Department of Radiology McGill University Health Centre, Montreal, Canada (P.A.B., C.R.); Department of Radiology, AP-HP, Hôpital Tenon, Sorbonne Université, Paris, France (I.T.N.); and Université de Paris, PARCC, INSERM, France (A.B., L.F.). Received August 22, 2019; revision requested November 4; revision received June 29, 2020; accepted July 28. **Address correspondence to** L.F. (e-mail: *laure.fournier@aphp.fr*).

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

See also the editorial by Méndez in this issue.

Radiology 2020; 297:361–371 • https://doi.org/10.1148/radiol.2020191658 • Content codes: GU MR

Background: Improving the differentiation of uterine sarcomas from atypical leiomyomas remains a clinical challenge and is needed to avoid inappropriate surgery.

Purpose: To develop a diagnostic algorithm including diffusion-weighted MRI criteria to differentiate malignant uterine sarcomas from benign atypical leiomyomas.

Materials and Methods: This case-control retrospective study identified women with an atypical uterine mass at MRI between January 2000 and April 2017, with surgery or MRI follow-up after 1 year or longer. A diagnostic algorithm including T2-weighted MRI and diffusion-weighted imaging (DWI) signal and apparent diffusion coefficient (ADC) values was developed to predict for sarcoma. The training set consisted of 51 sarcomas and 105 leiomyomas. Two external validation sets were used to evaluate interreader reproducibility (16 sarcomas; 26 leiomyomas) and impact of reader experience (29 sarcomas; 30 leiomyomas). Wilson confidence intervals (CIs) were calculated for sensitivity and specificity.

Results: Evaluated were 156 women (median age, 50 years; interquartile range, 44–63 years). Predictive MRI criteria for malignancy were enlarged lymph nodes or peritoneal implants, high DWI signal greater than that in endometrium, and ADC less than or equal to 0.905×10^{-3} mm^2/sec. Conversely, a global or focal area of low T2 signal intensity and a low or an intermediate DWI signal less than that in endometrium or lymph nodes allowed readers to confidently diagnose as benign a uterine mass demonstrating one or more of these signs ($P < .001$) in 100% cases in all three data sets. The sensitivities and specificities of the algorithm for diagnosis of malignancy were 98% (50 of 51 masses; 95% CI: 90%, 100%) and 94% (99 of 105 masses; 95% CI: 88%, 98%) in the training set; 88% (14 of 16 masses; 95% CI: 64%, 97%) and 100% (26 of 26 masses; 95% CI: 87%, 100%) in the validation set; and 83% (24 of 29 masses; 95% CI: 65%, 92%) and 97% (29 of 30 masses; 95% CI: 83%, 99%) for the less experienced reader, respectively.

Conclusion: A diagnostic algorithm with predictive features including lymphadenopathy, high diffusion-weighted imaging signal with reference to endometrium, and low apparent diffusion coefficient enabled differentiation of malignant sarcomas from atypical leiomyomas, and it may assist inexperienced readers.

Online supplemental material is available for this article.

Uterine leiomyoma is the most frequent uterine tumor in women, with a prevalence reaching 70%–80% among women aged 50 years and older (1). Conversely, uterine sarcoma is a rare malignant uterine tumor of mesenchymal origin, representing 3%–7% of primary malignant uterine tumors with an incidence of 0.7 per 100 000 women (2). It has a poor prognosis, with a 5-year survival rate of 17%–55%, even when it is discovered at an early stage.

Current practice defines typical leiomyomas as masses with low signal intensity on both T2-weighted images and diffusion-weighted images. However, up to 65% of leiomyomas manifest with degenerative change (3) and therefore will not have this typical appearance on MRI scans. The differential diagnosis with uterine sarcoma may be difficult, and 0.01%–0.5% of resected tumors diagnosed preoperatively as leiomyomas are confirmed as leiomyosarcomas at pathology (4,5).

In postmenopausal women, nonconservative surgical treatment such as hysterectomy can be offered for leiomyomas, but uterine preservation is often preferred in premenopausal women. However, it is imperative to exclude the diagnosis of sarcoma before any conservative treatment is implemented (6). Conservative myomectomy in cases of sarcoma can result in dissemination with morcellation. Conversely, inappropriate hysterectomies should be avoided for atypical leiomyomas. It is therefore necessary to have a diagnostic algorithm that optimizes both sensitivity and specificity.

Abbreviations

ADC = apparent diffusion coefficient, CI = confidence interval, DWI = diffusion-weighted imaging

Summary

A diagnostic algorithm including diffusion-weighted MRI criteria may help even inexperienced readers distinguish uterine sarcoma from atypical leiomyoma.

Key Results

- An MRI diagnostic algorithm that included enlarged lymph nodes, peritoneal implants, high diffusion-weighted MRI signal greater than that in endometrium, and apparent diffusion coefficient less than or equal to 0.905×10^{-3} mm²/sec enabled identification of leiomyosarcoma and correct classification in 40 of 42 (95%) patients in a validation set and 53 of 59 (95%) of examinations interpreted by an inexperienced reader.
- A global or focal area of low T2 myometrial signal, a low or intermediate diffusion-weighted MRI signal less than that in endometrium, or absence of enlarged lymph nodes allowed a diagnosis of a uterine mass in two independent validation sets.

The purpose of our study was to develop an MRI-based diagnostic algorithm to differentiate malignant uterine sarcomas from atypical benign leiomyomas and to test this algorithm in two independent external validation data sets.

Materials and Methods

We performed a retrospective, noninterventional case-control study. Institutional review board approval was obtained, and the need for written informed consent was waived because patients were informed of the reuse of their data for research purposes in conformity with European and national legislation.

Patients

We extracted pathology and MRI reports in consecutive women from the clinical data warehouse (7) of the Hôpital Européen Georges Pompidou, Paris, France (a tertiary center) between January 2000 and April 2017. Patients were selected according to the following inclusion criteria: *(a)* age 18 years or older; *(b)* at least one uterine mass, with atypical features for leiomyoma on MRI scans (ie, not demonstrating low signal intensity on T2-weighted and diffusion-weighted images (8) and localization that was not obviously endometrial); and *(c)* surgical procedure with pathologic confirmation, or MRI follow-up for 1 year or longer in the absence of operation. Exclusion criteria were as follows: *(a)* absence of MRI scans in picture archiving and communication system; *(b)* absence of a uterine mass (eg, endometrial thickening), or histologic type other than sarcoma or leiomyoma; and *(c)* leiomyomas without follow-up or with follow-up less than 1 year.

Each patient with sarcoma was matched to two control patients (ie, patients without sarcoma) from the same management year. Figure 1 displays the study flowchart.

MRI Acquisition

Our retrospective study spanned 17 years, and acquisition protocols varied during this period, including the *b* value of diffusion-weighted imaging (DWI). A standard acquisition protocol is detailed in Table E1 (online) and was compliant with European Society of Urogenital Radiology recommendations (9). Acquisition consisted of T2-weighted images in three planes (repetition time seconds/echo time msec, 4.5–7.8/89.2–93.5); axial gradient-echo T1-weighted sequence, reconstructed as in-phase, out-of-phase, water and fat images (Dixon) or gradient-echo sequences with and without fat suppression (repetition time msec/echo time msec, 320/12.7); axial DWI (repetition time seconds, 9.7; echo time, minimum; b = 500, 800, or 1000 sec/mm²); dynamic contrast-enhanced three-dimensional T1-weighted (perfusion) sequence (repetition time seconds/echo time msec, 5.1/1.4) with a temporal resolution of 6 seconds, 50 acquisition phases, for a total acquisition time of 5:20 minutes; and delayed axial T1-weighted Dixon images after the administration of contrast material. A bolus injection of 0.2 mL/kg of gadoteric acid (Dotarem; Guerbet, Villepinte, France) at an injection rate of 2 mL/sec followed by a 25-mL saline flush given through a power injector (Optistar LE; Mallinckrodt, Liebel-Flarsheim Company, Cincinnati, Ohio) was performed after the acquisition of three phases of the perfusion sequence.

Data Analysis

The clinical, pathologic, and imaging parameters listed in Table 1 were reported and managed with Research Electronic Data Capture (REDCap; the REDCap Consortium) (10). Clinical data were collected from the electronic medical records. Pathologic data collected included size and histologic type; the examination was performed by the same pathologist during the study period, who specialized in gynecologic oncology (M.A.L.B., with >20 years of experience) and served as our reference standard.

Diagnoses were based on the Stanford criteria, supplemented by *World Health Organization Classification of Tumours of the Female Reproductive Organs* (11). Although they are not strictly malignant, tumors of uncertain malignant potential such as endometrial stromal nodule and smooth muscle tumor of uncertain malignant potential were included in the malignant category because diagnosis before operation should be suggested to ensure appropriate surgical management.

Imaging data were reinterpreted by a single reader (C.A.W., a junior radiologist with 1 year of experience in pelvic imaging) who was blinded only to pathology. In cases of multiple masses, only the largest mass that met the inclusion criteria was analyzed. Figure E1 (online) illustrates each imaging sign.

Four MRI features were assessed, as follows: *(a)* presence of enlarged lymph nodes or peritoneal implants, *(b)* presence of focal region or global low T2 signal intensity, *(c)* visual analysis of diffusion-weighted images relative to the myometrium and endometrium/lymph nodes, and *(d)* the apparent diffusion coefficient (ADC) value. Enlarged lymph nodes were defined as lymph nodes with small axis 10 mm or greater. If the tumor was heterogeneous, the highest signal intensity of the solid component was reported. An intermediate T2 signal was defined as higher than that of gluteal muscle and lower than that of water. Low T2 signal concerned solid components and not what had previously been reported as "dark T2 areas" (12), corresponding to

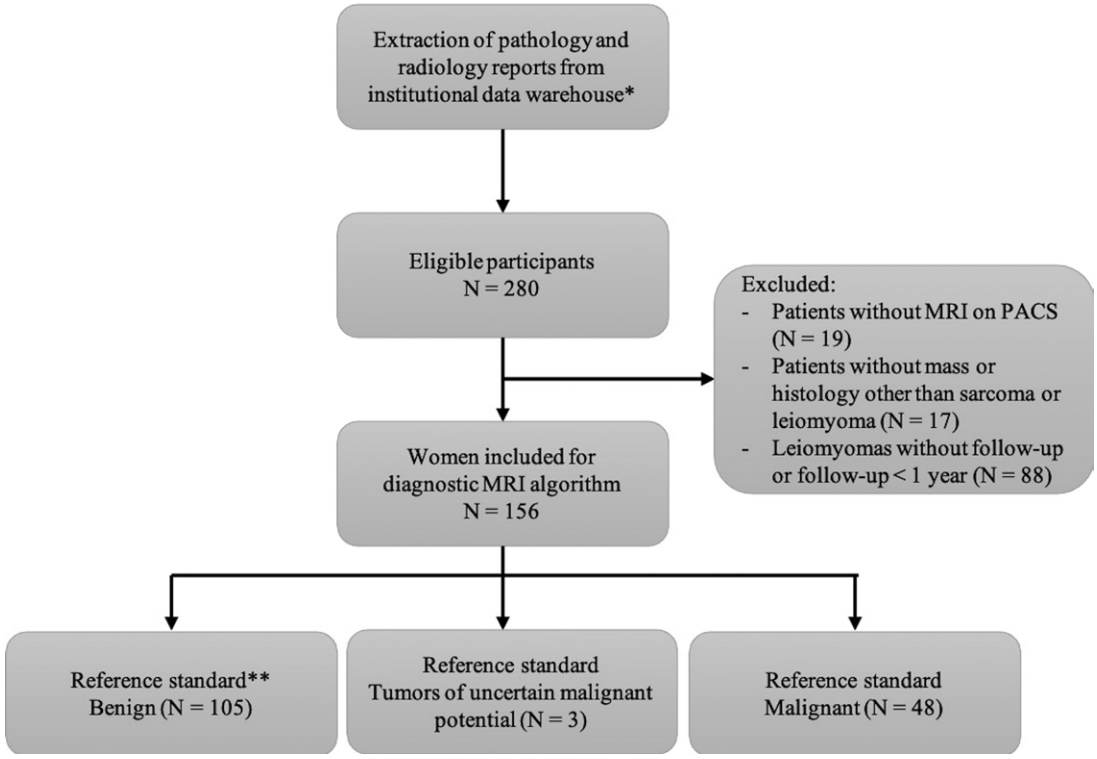

Figure 1: Study flowchart according to Standards for Reporting Diagnostic Accuracy Studies. *Search terms for data warehouse were as follows: For sarcomas: *pathology report* and *hyst-* or *uterus* or *uterin-* and *sarcom-* or *STUMP* or *nodul-* and *endometrial stroma*; for leiomyomas: *pathology report* and *hyst-* and *myom-* and *myxoid* or *oedema*; or *REMNO* (code for MRI in electronic records) and *pelv-* and *leiomyoma* or *fibroid* and *atypical* or *cellular* or *oedema*. Thus, all cases of uterine sarcoma or tumor of uncertain malignant potential according to histologic type were included in our study without reference to their MRI appearance. Atypical leiomyomas were identified for inclusion based on either pathology report or MRI report. **Reference standard was pathologic evaluation for all sarcomas or tumors of uncertain malignant potential and for 88 of 105 benign leiomyomas, and follow-up MRI after more than 1 year for 17 of 105 benign leiomyomas. PACS = picture archiving and communication system.

hemorrhagic portions (high T1 signal and absence of enhancement on contrast-enhanced sequences).

DWI signal intensity was analyzed visually in three categories on the basis of the highest signal intensity within the mass (Fig 2). A low signal intensity was defined as lower or equal to that of the myometrium, an intermediate signal intensity as higher than that in the myometrium but lower than or equal to that in the endometrium and/or lymph nodes, and a high signal intensity as higher than that in the endometrium and/or lymph nodes.

ADC maps and perfusion analysis were performed on a workstation (ADW; GE Healthcare, Chicago, Ill). ADC values of each tumor were obtained by placing a circular region of interest (24 mm^2 by default) on the ADC map in the area with the lowest ADC, avoiding hemorrhagic, necrotic, or cystic portions within the lesion by referring to T1-weighted, T2-weighted, and contrast-enhanced images. Several regions of interest were placed on the maps, and the lowest value of mean ADC was recorded.

For perfusion analysis, two regions of interest (24 mm^2) were drawn in the external myometrium and in the earliest enhancing part of solid tissue. We classified the enhancement of solid tissue by using the time-intensity curve classification developed for adnexal masses (13).

External Validation Set 1 to Evaluate Generalizability and Interreader Variability

This independent validation set was obtained from the Hôpital Tenon (Paris, France) and included the cohort from a previously published study (14) that had examined characteristics of leiomyomas versus sarcomas but had not applied the same algorithm as the one we developed. Women who had a diagnosis of mesenchymal tumor of the uterus were identified in the institution's pathology database. Only women who had undergone MRI of the pelvis before surgery and had a solitary mass were selected. By applying the same inclusion criteria for MRI signal as in the training set, we included only atypical masses, resulting in the selection of 42 of the 51 women cited in our study. Two readers (C.A.W. and L.S.F., with 1 year and 15 years of experience in female pelvic imaging, respectively) independently read the MRI scans and applied the algorithm to define certainly benign, probably benign, and highly suspicious masses.

External Validation Set 2

External validation set 2 is an independent validation set obtained from McGill University Health Center (Montreal, Canada) and it included 59 women. The reader had moderate experience in female pelvic imaging (P.A.B., 5 years of experi-

Table 1: Clinical, Pathologic, and Imaging Data Collected for Training Set from Patient Electronic Medical Records

Type of Data Collected
Clinical data
Age
Menopausal status
Hormone therapy at time of diagnosis or history of breast cancer with hormone therapy
History of pelvic radiation therapy
Pelvic pain
Abnormal vaginal bleeding
Palpation of a pelvic mass
Histologic classification
Malignant tumors
Leiomyosarcoma
Endometrial stromal sarcoma
Adenosarcoma
Undifferentiated endometrial sarcoma
Carcinosarcoma (or malignant mixed müllerian tumor (MMMT)
Liposarcoma
Tumors of uncertain malignant potential
Smooth muscle tumor of uncertain malignant potential
Endometrial stromal nodule
Benign tumors
Edematous leiomyoma
Cellular leiomyoma
Degenerative hyaline leiomyoma
Degenerative cystic leiomyoma
Myxoid leiomyoma
Hemorrhagic leiomyoma
Lipoleiomyoma
Infarcted leiomyoma
Leiomyoma with no histologic degeneration
Imaging data
Morphologic criteria
No. of uterine masses
Size at initial and follow-up MRI (when performed)
Tumor margin (irregular or smooth)
Homogeneity on T2-weighted images
Presence of focal portion with low T2 signal (ie, solid tissue enhancing after administration of contrast material but equal to or lower than signal in gluteal muscle at T2-weighted imaging)
Presence of intratumoral hemorrhage (ie, high T1 signal and no enhancement)
Presence of cystic alterations (ie, high T2, low T1 signal and no enhancement)
Delayed enhancement
Presence of enlarged lymph nodes (small axis ≥10 mm)
Presence of peritoneal implants
Presence of pelvic fluid regardless of quantity
Functional criteria
Visual analysis of diffusion-weighted imaging signal intensity (low, intermediate, high)
Apparent diffusion coefficient of the most restricted solid portion of the mass
Perfusion curve type relative to outer myometrium

Note.—There were 156 patients from whom data were collected.

ence) and no experience using the algorithm. The pathology database was searched for the keyword *leiomyosarcoma* in the reports, and the radiology database was searched for keyword *leiomyosarcoma* or *cannot exclude leiomyosarcoma,* the latter to identify atypical leiomyomas on MRI scans during the same study period.

A first read was performed by the reader (P.A.B.) after an explanation of the meaning of each imaging sign included in the

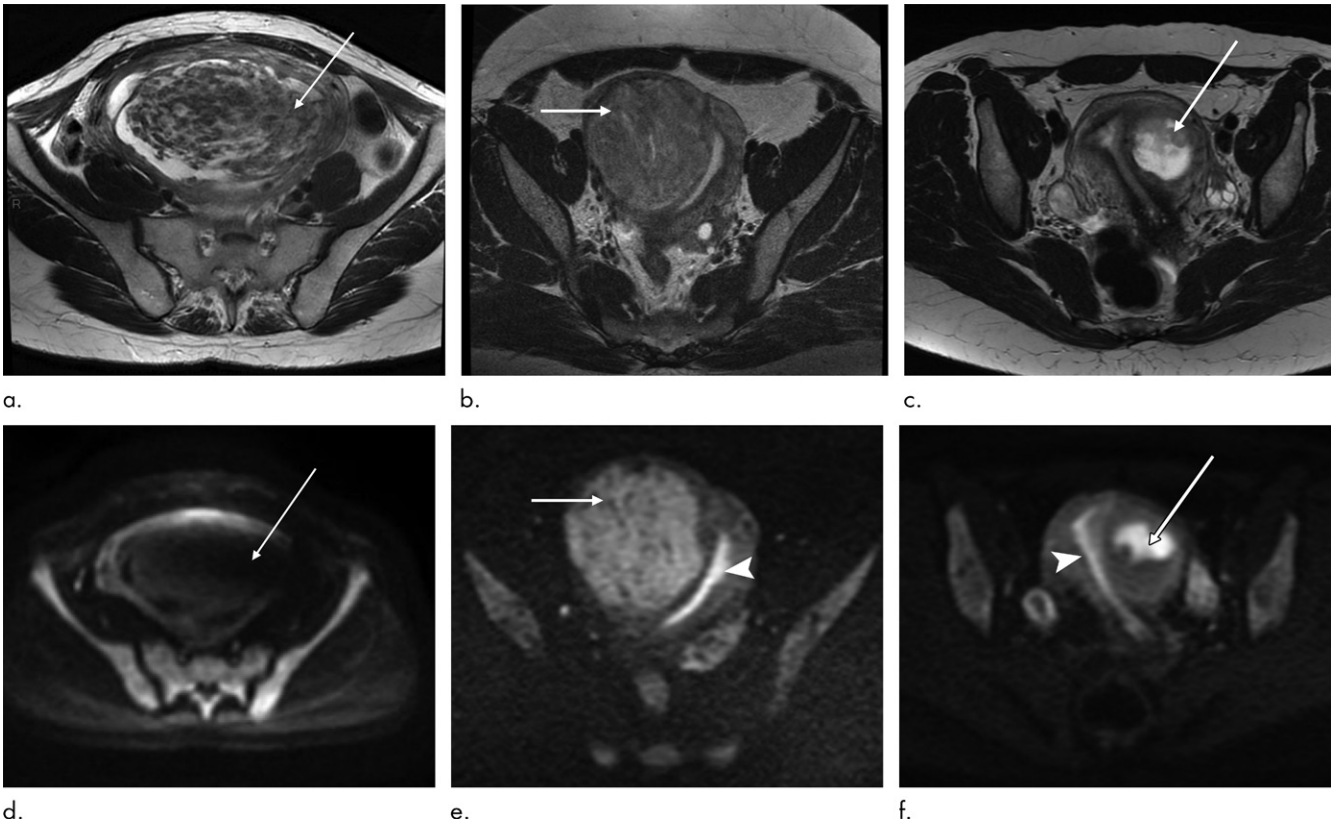

Figure 2: Examples of visual categories of diffusion-weighted imaging signal intensity (b = 1000 sec/mm^2) in three masses with corresponding T2 sequences. **(a, d)** Images depict atypical leiomyoma (arrow) with, respectively, heterogeneous signal intensity on T2-weighted images and low signal intensity on diffusion-weighted images. **(b, e)** Images represent atypical leiomyoma (arrow) with, respectively, intermediate signal intensity on T2-weighted images and intermediate diffusion-weighted imaging signal intensity higher than that in myometrium but lower than in endometrium (arrowhead). **(c, f)** Images show endometrial stromal sarcoma (arrow) with, respectively, intermediate signal intensity on T2-weighted images and high diffusion-weighted imaging signal intensity higher than that in endometrium (arrowhead).

algorithm (focal or global low T2 signal; low, intermediate, or high DWI signal; and how to measure ADC) by using slides with examples. A consensus read (P.A.B., C.A.W., L.S.F.) was then performed to identify whether false-negative and false-positive findings were because of lack of experience in interpreting the imaging signs or errors in the algorithm.

Statistical Analysis

To select relevant variables for the diagnostic algorithm of malignancy (including sarcomas and tumors of uncertain malignant potential), we performed a Wilcoxon signed ranked test for quantitative variables and a Fisher test for qualitative variables. A recursive partitioning model was used with statistically significant imaging binary features to build a diagnostic algorithm to classify tumors as benign or malignant by using the rpart package (15). Continuous features were dichotomized beforehand by choosing the cut-off value to maximize the Youden index in univariate models. Sensitivity and specificity of the diagnostic algorithm were estimated along with their Wilson confidence interval (CI) in the training sample and the two independent validation sets. We considered missing data as the most pejorative (ie, we classified them by default in the worst-case category: presence of intratumoral hemorrhage or necrosis was classified as present when missing, DWI signal was classified as higher than in endometrium and lymph nodes, and ADC was classified below 0.905×10^{-3} mm^2/sec). Interreader variability was evaluated by using the Cohen κ coefficient.

Results

Patients

Tables 2 and 3 show, respectively, the demographics and pathologic findings in the 156 women included in our study (median age, 51 years; interquartile range, 45–66 years).

Image Analysis

Perfusion and DWI sequences were performed in 136 of 156 (87%) and 137 of 156 (88%) women, respectively, and were both available in 123 of 156 (79%) women (96 leiomyomas and 27 sarcomas). For DWI, the b values used were 0 and 800 or 1000 sec/mm^2 (96 leiomyomas and 30 sarcomas) or 0 and 500 sec/mm^2 (four leiomyomas and seven sarcomas). Delayed contrast enhancement was acquired for 152 of 156 (97%) women (104 leiomyomas and 48 sarcomas).

Data Analysis

There was a difference in the median pathologic size between sarcomas and leiomyomas (80 mm [interquartile range, 50–120 mm] vs 65 mm [interquartile range, 45–83 mm], respectively;

Table 2: Patient Characteristics in Three Data Sets

	Training Set		External Validation Set 1		External Validation Set 2	
Characteristic	Leiomyoma (n = 105)	Sarcoma or Tumor of Uncertain Malignant Potential (n = 51)	Leiomyoma n = 26	Sarcoma or Tumor of Uncertain Malignant Potential (n = 16)	Leiomyoma (n = 30)	Sarcoma or Tumor of Uncertain Malignant Potential (n = 29)
Median age (y)	48 [42;53]	64 [53–72]*	46 [22–69]	55 [40–81]*	56 [31–79]	56 [36–86] NS
Menopausal status	25 (24)	43 (84)*	3 (12)	10 (63)*	12 (40)	27 (93)*
History of hormone therapy†	8/105 (8)	5/49 (10) NS	NA	NA	NA	NA
History of pelvic radiation therapy	0/105 (0)	1/49 (2) NS	NA	NA	NA	NA
Pelvic pain	44/105 (42)	15/50 (30) NS	NA	NA	NA	NA
Abnormal vaginal bleeding	70/105 (67)	37/50 (74) NS	NA	NA	NA	NA
Palpation of a pelvic mass	27/102 (26)	14/47 (30) NS	NA	NA	NA	NA

Note.—Unless otherwise indicated, data are numbers of patients in whom information was available; data in parentheses are percentages and data in brackets are interquartile range. The denominators used to calculate the percentages vary according to the data available. NA = not available, NS = not statistically significant between leiomyomas and sarcomas or tumors of uncertain malignant potential.

* Indicates statistically significantly different values between leiomyomas and sarcomas or tumors of uncertain malignant potential with P value less than .001.

† Hormone therapy given at time of diagnosis, or history of breast cancer with hormone therapy.

P = .04). Women with sarcomas were older than women with leiomyomas (median age, 64 years [interquartile range, 42–53 years] vs 48 years [interquartile range, 53–72 years], respectively; $P < .001$), with large overlap in range (39–91 years for women with sarcomas and 29–83 years for women with leiomyomas). Menopause was more frequent in women with sarcomas (84% [43 of 51 women] vs 24% [25 of 105 women]; $P < .001$).

History and clinical symptoms were not different in women with sarcomas and women with leiomyomas: abnormal vaginal bleeding, 74% (37 of 50 women) versus 67% (70 of 105 women; P = .45); pelvic pain, 30% (15 of 50 women) versus 42% (44 of 105 women; P = .16); palpation of pelvic mass, 30% (14 of 47 women) versus 26% (27 of 102 women; P = .69), respectively.

Detailed results of MRI criteria to discriminate malignant sarcomas from benign leiomyomas are in Table E2 (online). Fifty-one percent (26 of 51) of sarcomas manifested as a solitary uterine mass versus 27% (28 of 105) of the atypical leiomyomas. Features that could exclude sarcoma in our sample were presence of an area of low T2 signal intensity regardless of DWI signal, a low DWI signal regardless of T2 signal intensity, an intermediate DWI signal higher than in myometrium but lower than in endometrium or lymph nodes, and a type 1 perfusion curve (present in only 4% [four of 100] of leiomyomas).

All sarcomas had at least a portion with a high DWI signal higher than that of endometrium and/or lymph nodes. The most discriminating ADC cutoff value for differentiating sarcomas from leiomyomas was 0.905×10^{-3} mm^2/sec by using the Youden index. When excluding tumors with b values other than 1000 sec/mm^2 (122 of 137 tumors), Youden index maximization yielded the same cutoff value (0.905×10^{-3} mm^2/sec).

Figure 3 summarizes the diagnostic algorithm obtained by using recursive partitioning. This algorithm allowed identification of certainly benign lesions (no sarcomas with the appropriate imaging signs), probably benign lesions (one of 51 sarcomas and 10 of 105 myomas), and highly suspicious lesions (50 of 51 sarcomas and one of 105 myomas). Three more lesions were considered inconclusive (all leiomyomas) because no DWI examination was available, yielding a sensitivity of 98% (50 of 51 lesions; 95% CI: 90%, 100%) and specificity of 96% (101 of 105 lesions; 95% CI: 91%, 99%).

The misclassified uterine masses were one false-negative finding (one endometrial stromal nodule with ADC $> 0.905 \times 10^{-3}$ mm^2/sec; Fig 4) and four false-positive findings (three leiomyomas for which no DWI examination was available and one cellular leiomyoma with ADC $\leq 0.905 \times 10^{-3}$ mm^2/sec). If we analyzed only women with DWI acquisitions, which we consider mandatory in order to correctly characterize uterine masses, there was one false-positive and one false-negative finding, with a sensitivity of 97% (35 of 36 masses; 95% CI: 84%, 100%) and specificity of 99% (99 of 100 masses; 95% CI: 94%, 100%).

External Validation Set 1

Tables 2 and 3 show details of the 42 women in this set. The diagnostic algorithm correctly classified 40 of 42 cases for the first reader, and 40 of 42 cases for the second reader (Table 4). For reader 1, the misclassified uterine masses were two false-negative findings, classified in the probably benign category (one low-grade endometrial stromal sarcoma and one smooth muscle tumor of uncertain malignant potential; ADC $> 0.905 \times 10^{-3}$ mm^2/sec); that is, a sensitivity of 88% (14 of 16 masses; 95% CI: 64%, 97%) and specificity of 100% (26 of 26 masses; 95% CI: 87%, 100%). For reader 2, there was one false-negative finding (the smooth muscle tumor of uncertain malignant potential), but also one false-positive finding, a benign leiomyoma in which a portion with restricted diffusion (ADC, 0.740×10^{-3} mm^2/sec) was interpreted as highly suspicious; that is, a sensitivity of 94% (15 of 16 masses; 95% CI: 72%, 99%) and specificity of 96% (25 of 26 masses; 95% CI:

Table 3: Details of Training and Validation Sets with Pathologic Findings for Masses in Which Surgery Was Performed

Characteristic	Training Set	Validation Set 1	Validation Set 2
Leiomyomas	105*	26†	30‡
Edematous	6 (6)	1 (4)	1 (3)
Cellular	7 (7)	1 (4)	2 (7)
Degenerated hyaline	42 (40)	…	7 (23)
Degenerated cystic	4 (4)	7 (27)§	…
Myxoid	3 (3)	…	1 (3)
Lipoleiomyoma	3 (3)	1 (4)	1 (3)
Hemorrhagic	1 (1)	…	…
Necrobiotic	17 (17)	…	1 (3)
Hydropic	…	…	1 (3)
Mitotically active	…	…	1 (3)
"Bizarre"	…	1 (4)	…
No signs of degeneration at pathologic evaluation	5 (5)	15 (57)	5 (17)
Follow-up at MRI only	17 (16)	...	10 (33)
Tumors of uncertain malignant potential	3	3	2
STUMP	2 (67)	3 (100)	2 (100)
ESN	1 (33)	…	…
Sarcomas	48	13	27
Leiomyosarcomas	17 (35)	2 (15)	6 (22)
Liposarcomas	2 (4)	…	…
Rhabdomyosarcoma	…	1 (8)	…
Endometrial stromal sarcoma	7 (15)	1 (8)	2 (7)
Adenosarcoma	1 (2)	…	…
Carcinosarcomas	19 (40)	2 (15)	18 (67)
Undifferentiated endometrial sarcomas	2 (4)	7 (54)	1 (4)
Total	156	42	59

Note.—Data are numbers of masses; data in parentheses are percentages. ESN = endometrial stromal nodule, STUMP = smooth muscle tumor of uncertain malignant potential.

* Surgery was performed in 88 of 105 leiomyomas (84%); in 17 of 105 (16%), MRI follow-up was used as standard of reference.

† Surgery was performed in 26 of 26 (100%).

‡ Surgery was performed in 20 of 30 (67%); in 10 of 30 (33%), MRI follow-up was used as standard of reference.

§ Described as degenerated on pathology report with no detail regarding type.

81%, 99%). There were three cases of discrepancies between readers; that is, a 93% overall agreement and a κ coefficient of 0.86 (95% CI: 0.70, 1.00).

External Validation Set 2

Table 3 shows details of the 59 women in this set. After the first read by the inexperienced reader, the algorithm correctly classified 53 of 59 cases (sensitivity 83%; 24 of 29 cases; 95% CI: 65%, 92%; and specificity 97%; 29 of 30 cases; 95% CI: 83%, 99%) (Table 4). The misclassified uterine masses were one false-positive finding (one leiomyoma with ADC $< 0.905 \times 10^{-3}$ mm^2/sec classified as highly suspicious) and five false-negative findings. The consensus read resulted in reclassification of three false-negative findings, corresponding to initial errors of interpretation of the algorithm. In two cases, the inexperienced reader had considered the DWI signal to be lower than that in the endometrium, whereas there were areas with a signal intensity higher than in endometrium and with ADC less than 0.905×10^{-3} mm^2/sec, resulting in reclassification of the mass as highly suspicious. In one case, a low T2 area was considered to be benign according to the algorithm. On the consensus read, it appeared that this area was hemorrhage (Fig 5) and solid portions showed a signal at DWI that was higher than that in endometrium and an ADC less than 0.905×10^{-3} mm^2/sec; this case was reclassified as highly suspicious.

Following consensus reading, the diagnostic algorithm correctly classified 56 of 59 patients (Table 4) (sensitivity, 93% [27 of 29 patients; 95% CI: 78%, 98%]; and specificity, 97% [29 of 30 patients; 95% CI: 83%, 98%]). The uterine masses incorrectly classified by the algorithm were one false-positive

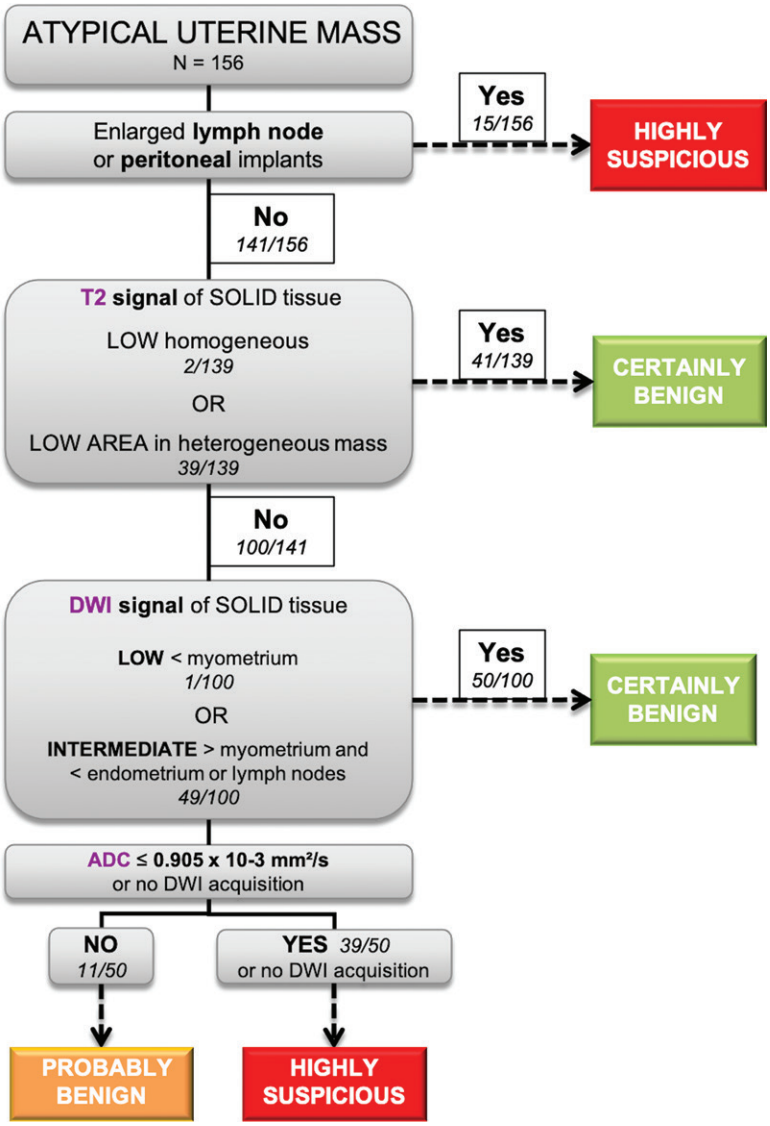

Figure 3: Proposed diagnostic algorithm for reading MRI scans to differentiate benign atypical leiomyomas from malignant sarcomas. Number of masses for each sign is reported in italics as number of masses manifesting sign versus total number of masses. ADC = apparent diffusion coefficient, DWI = diffusion-weighted imaging.

finding (a leiomyoma classified as highly suspicious [ADC $<$ 0.905×10^{-3} mm²/sec) and two false-negative findings (one carcinosarcoma classified as certainly benign with a low T2 portion and one smooth muscle tumor of uncertain malignant potential classified as probably benign, with intermediate DWI signal higher than in the myometrium but lower than in the endometrium).

Discussion

By using a recursive partitioning technique, we developed an easy-to-use diagnostic algorithm to help radiologists distinguish atypical leiomyoma from sarcoma in routine clinical practice, without the need of a calculator.

The four MRI criteria included in this algorithm were presence of enlarged lymph nodes or peritoneal implants; presence of focal region or global low T2 signal intensity; visual analysis of DWI signal intensity as low, intermediate, or high in reference to myometrium and endometrium and/or lymph nodes; and an ADC value of 0.905×10^{-3} mm²/sec or less.

Uterine masses were classified as certainly benign if they showed a global or focal area of low T2 signal and/or a low or an intermediate DWI signal (lower than that in the endometrium or lymph nodes). If they did not meet these criteria (ie, intermediate T2 signal and high DWI signal [higher than in the endometrium or lymph nodes]), they were classified as probably benign (ADC $>$ 0.905×10^{-3} mm²/sec) or highly suspicious (ADC $\leq$ 0.905×10^{-3} mm²/sec) in the most restricted portion. Finally, they were classified as highly suspicious in cases with peritoneal implants or enlarged lymph nodes.

The resulting algorithm achieved a sensitivity and specificity of 98% and 96%, respectively, in the training set, and was validated in two independent sets, including when the algorithm was applied by a reader with no previous experience in using it, with sensitivities and specificities of 88% and 100%, and 83% and 97%, respectively. There was also high interreader agreement ($\kappa = 0.86$).

Of note, perfusion acquisitions did not discriminate between atypical leiomyoma and sarcoma because even benign leiomyomas were highly vascular. By using this algorithm, we identified cases in which the radiologist can confidently diagnose the uterine mass as certainly benign (rate of malignancy, 0%; 0 of 96 sarcomas in total): when it manifests a focal or global low T2 signal intensity, or when the DWI signal is lower than that in the endometrium or lymph nodes.

Several studies have sought to differentiate benign leiomyomas from sarcomas, usually with patient numbers ranging from 17 to 70 patients (12,14,16–20). Populations included were diverse, in some cases including both endometrial and myometrial sarcomas, in some cases only leiomyosarcomas, and sometimes both typical and atypical fibroids.

Morphologic features such as tumor margins, hemorrhage, tumor enhancement, T2 and DWI qualitative signal, functional parameters such as ADC, and texture analysis have been explored, yielding performances ranging from 56% to 92% for sensitivity and 75% to 96% for specificity (21,22). However, none of the previous studies included a validation population. Cornfeld et al (23) tested the features proposed by Tanaka et al (17) and achieved a sensitivity of 17% versus 73% in the training set.

We chose to optimize diagnostic performance to diagnose invasive sarcomas most accurately but also to not overdiagnose atypical leiomyomas in young women. Our study tested the algorithm in two external data sets and evaluated interreader reproducibility and the effect of experience on performance of the algorithm. Our ADC cutoff was lower than previously

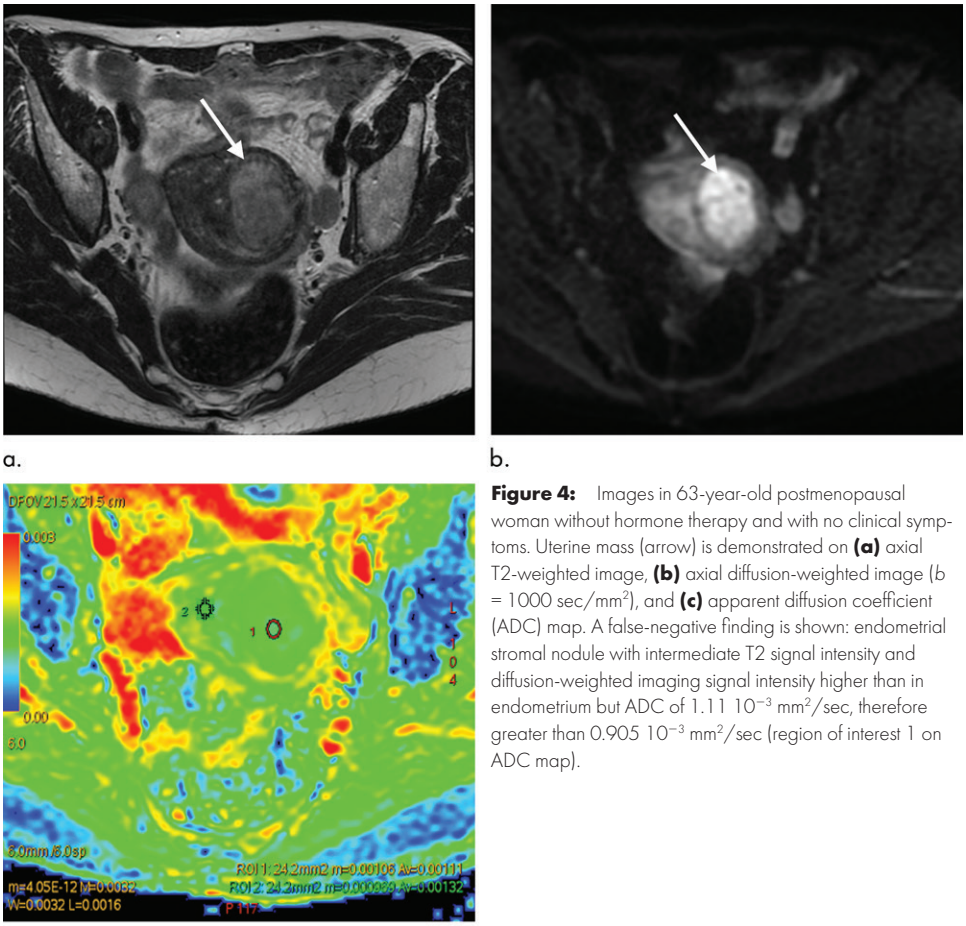

Figure 4: Images in 63-year-old postmenopausal woman without hormone therapy and with no clinical symptoms. Uterine mass (arrow) is demonstrated on **(a)** axial T2-weighted image, **(b)** axial diffusion-weighted image (*b* = 1000 sec/mm^2), and **(c)** apparent diffusion coefficient (ADC) map. A false-negative finding is shown: endometrial stromal nodule with intermediate T2 signal intensity and diffusion-weighted imaging signal intensity higher than in endometrium but ADC of 1.11 10^{-3} mm^2/sec, therefore greater than 0.905 10^{-3} mm^2/sec (region of interest 1 on ADC map).

Table 4: Contingency Table of Diagnostic Test

Reader Classification, Validation Set	Sarcoma or Tumor of Uncertain Malignant Potential	Leiomyoma	Total
Validation set 1			
Highly suspicious	14	0	14
Certainly or probably benign	2	26	28
Total	16	26	42
Validation set 2			
Highly suspicious	24	1	25
Certainly or probably benign	5	29	34
Total	29	30	59

Note.—Data are numbers of masses. Diagnostic test data are from 42 women in external validation set 1 (reader 1) and from 59 women in external validation set 2 with inexperienced reader (before consensus). Sarcomas or tumors of uncertain malignant potential were considered as malignant disease, and leiomyomas as absence of malignant disease. The test result was considered positive when the diagnostic algorithm classified a mass as highly suspicious, and negative if the algorithm classified the mass as certainly or probably benign.

reported values (14,16,18,19,24), which ranged from 1.06 to 1.23 × 10^{-3} mm^2/sec, with *b* values of 800 or 1000 sec/mm^2. In all these studies, ADC cut-off values were optimized to exclude sarcomas—that is, an ADC higher than the cut-off value had a negative predictive value of 100%. The drawback of this strategy is that some benign leiomyomas would be considered suspicious, therefore leading to unnecessary hysterectomy.

Two leiomyomas were misclassified (false-positive findings) in the training and validation sets, respectively. For one, no definitive histologic type was obtained with follow-up only. The other was confirmed as a cellular leiomyoma at pathology. This result is similar to that of another study in the literature (16). Of note, nine of 10 cellular leiomyomas were correctly identified by our algorithm. Regarding false-negative findings, an endometrial stromal nodule, a low-grade endometrial stromal sarcoma, and two smooth muscle tumors of uncertain malignant potential were misclassified in the three data sets. Reassuringly, these lesions have relatively better prognosis than more aggressive sarcomas, and all occurred in postmenopausal women, in whom hysterectomy is often performed even for benign lesions when symptomatic. Moreover, all of these were identified as probably benign (and not certainly benign), yielding a rate of malignancy for this category of 5.7% (four of 70).

Other clinical and MRI criteria previously described as predictive of malignancy were also found in our study, such as age, menopausal status, irregular margins with T2 sequences, and presence of intratumoral hemorrhage or intratumoral cystic alterations; however, there was overlap between benign and malignant lesions. These criteria may be useful when masses are nonevaluable or in the probably benign category of our algorithm, for which a follow-up could be recommended.

The presence of hemorrhage is potentially confounding. A low T2 solid component should not be confused with a low T2 signal intensity resulting from hemorrhagic necrosis, identified by its corresponding high T1 signal intensity and no enhancement after injection of contrast material. This criterion was the source of errors in interpretation by the inexperienced reader. Also, the DWI signal of the hemorrhagic portion is sometimes very high and "saturates" the DWI signal intensity because gray

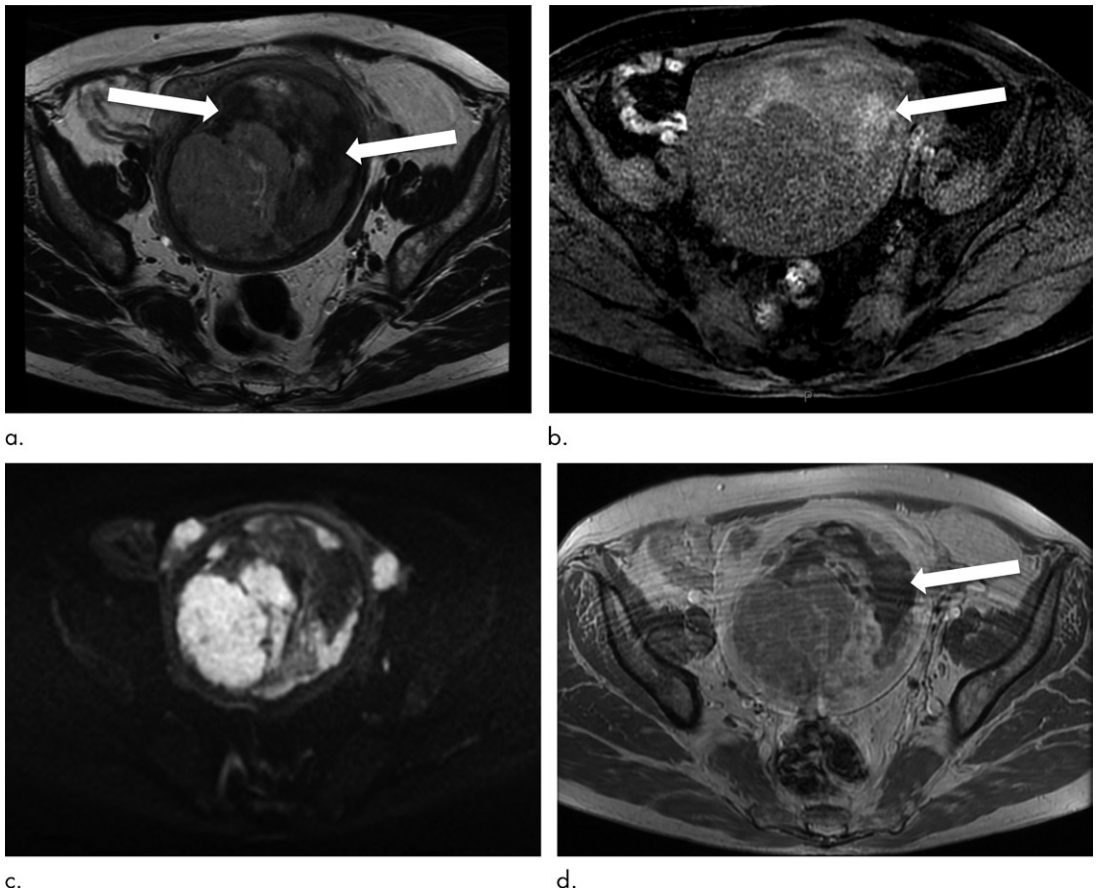

Figure 5: MRI scans in 52-year-old postmenopausal woman undergoing hormone therapy, with abnormal vaginal bleeding and palpation of pelvic mass. MRI scans show uterine mass on **(a)** axial T2-weighted image, **(b)** axial T1-weighted image with fat suppression, **(c)** axial diffusion-weighted image (b = 1000 sec/mm^2), and **(d)** axial T1-weighted image with fat suppression and gadolinium chelate injection. Figure represents example of leiomyosarcoma with low T2 signal intensity because of hemorrhage (arrows), which should not be confused with low T2 signal intensity in tissue associated with benign leiomyoma. Area manifesting low T2 signal intensity can be identified as hemorrhage because it manifests high T1 signal intensity and does not demonstrate enhancement after injection of contrast material.

levels are adjusted to the signals demonstrated in the image. This occurred in one sarcoma in the validation set, in which the DWI signal intensity of the soft-tissue component appeared relatively lower, making it difficult to interpret the signal relative to endometrium or lymph nodes. These misclassified cases highlight the limits of our algorithm, and we therefore suggest that the algorithm not be applied to extensively hemorrhagic masses.

Our study had several limitations. First, this study did not determine whether use of the MRI features provided any additional ability to identify sarcoma over and above clinical features such as age and menopausal status. Second, the retrospective nature of the study may have led to selection bias. However, the much higher percentage of sarcomas in our set compared with what would be naturally observed should have led to an overestimation of the false-negative findings. Because there were only four false-negative findings among the three populations (257 masses), we believe that real-world performance would be at least as good. Third, the MRI protocols changed during the study period, and some of the examinations were performed outside our institution. Different b values were used; however, no false-negative or false-positive findings were among the masses imaged by using a b value of 500 sec/mm^2. Fourth, our study included both endometrial and myometrial sarcomas. When women undergo MRI for characterization of a uterine mass, it is not always known beforehand whether the origin is endometrial or myometrial because it may be difficult to determine the tumor's origin when the masses are large. Fifth, we included women with multiple uterine masses but analyzed only the largest and most atypical mass on T2-weighted MRI scans and diffusion-weighted images. Of note, in 51% (26 of 51 women), a sarcoma developed in a uterus containing several masses. Finally, we performed the ADC measurement by using manually drawn regions of interest on the area with the most restricted diffusion on the ADC map, but this step led to interobserver variability when the algorithm was applied.

In conclusion, our study proposes an MRI-based algorithm for assessing atypical uterine masses in routine clinical practice, and diffusion-weighted imaging is an important element. However, training will be necessary to optimize performance in inexperienced readers.

Author contributions: Guarantors of integrity of entire study, C.A.W., A.S.J., L.S.F.; 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, C.A.W., C.B., C.C., M.A.L.B., I.T.N., A.B., L.S.F.; clinical studies, C.A.W., P.A.B., C.B., M.A.L.B., A.S.B., I.T.N., A.B., C.R., L.S.F.; experimental studies, C.A.W., M.A.L.B.; statistical analysis, C.A.W., A.S.J., M.A.L.B., C.R., L.S.F.; and manuscript editing, C.A.W., A.S.J., M.A.L.B., A.S.B., C.R., L.S.F.

Disclosures of Conflicts of Interest: **C.A.W.** disclosed no relevant relationships. **A.S.J.** disclosed no relevant relationships. **P.A.B.** disclosed no relevant relationships. **C.B.** disclosed no relevant relationships. **C.C.** disclosed no relevant relationships. **M.A.L.B.** disclosed no relevant relationships. **A.S.B.** disclosed no relevant relationships. **I.T.N.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: disclosed money paid to author for lectures from GE, Hologic, Canon, Siemens, Guerbet; disclosed money paid to author for travel expenses from GE. Other relationships: disclosed no relevant relationships. **A.B.** disclosed no relevant relationships. **C.R.** disclosed no relevant relationships. **L.S.F.** Activities related to the present article: disclosed no relevant relationships. Activities not related to the present article: disclosed money to author's institution for a grant from Invectys; disclosed money to author for lectures from GE Healthcare, Janssen, Novartis, Sanofi; disclosed money to author's institution for travel expenses from Guerbet. Other relationships: disclosed no relevant relationships.

References

1. Baird DD, Dunson DB, Hill MC, Cousins D, Schectman JM. High cumulative incidence of uterine leiomyoma in black and white women: ultrasound evidence. Am J Obstet Gynecol 2003;188(1):100–107.
2. Van den Bosch T, Coosemans A, Morina M, Timmerman D, Amant F. Screening for uterine tumours. Best Pract Res Clin Obstet Gynaecol 2012;26(2):257–266.
3. Ha HK, Jee MK, Lee HJ, et al. MR imaging analysis of heterogeneous leiomyomas of the uterus. Front Biosci 1997;2(6):f4–f12.
4. Huang YT, Huang YL, Ng KK, Lin G. Current Status of Magnetic Resonance Imaging in Patients with Malignant Uterine Neoplasms: A Review. Korean J Radiol 2019;20(1):18–33.
5. DeMulder D, Ascher SM. Uterine Leiomyosarcoma: Can MRI Differentiate Leiomyosarcoma From Benign Leiomyoma Before Treatment? AJR Am J Roentgenol 2018;211(6):1405–1415.
6. Tirumani SH, Ojili V, Shanbhogue AKP, Fasih N, Ryan JG, Reinhold C. Current concepts in the imaging of uterine sarcoma. Abdom Imaging 2013;38(2):397–411.
7. Zapletal E, Rodon N, Grabar N, Degoulet P. Methodology of integration of a clinical data warehouse with a clinical information system: the HEGP case. Stud Health Technol Inform 2010;160(Pt 1):193–197.
8. Bolan C, Caserta MP. MR imaging of atypical fibroids. Abdom Radiol (NY) 2016;41(12):2332–2349.
9. Kubik-Huch RA, Weston M, Nougaret S, et al. European Society of Urogenital Radiology (ESUR) Guidelines: MR Imaging of Leiomyomas. Eur Radiol 2018;28(8):3125–3137.
10. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377–381.
11. Oliva E, Carcangiu ML, Carinelli SG, et al. Mesenchymal tumors of the uterine corpus. In: Kurman RJ, Carcangiu ML, Herrington CS, Young RH, ed. WHO Classification of Tumours of Female Reproductive Organs. 4th ed. Vol. 6. Lyon: International Agency for Research on Cancer, 2014; 135-147.
12. Lakhman Y, Veeraraghavan H, Chaim J, et al. Differentiation of Uterine Leiomyosarcoma from Atypical Leiomyoma: Diagnostic Accuracy of Qualitative MR Imaging Features and Feasibility of Texture Analysis. Eur Radiol 2017;27(7):2903–2915.
13. Thomassin-Naggara I, Daraï E, Cuenod CA, Rouzier R, Callard P, Bazot M. Dynamic contrast-enhanced magnetic resonance imaging: a useful tool for characterizing ovarian epithelial tumors. J Magn Reson Imaging 2008;28(1):111–120.
14. Thomassin-Naggara I, Dechoux S, Bonneau C, et al. How to differentiate benign from malignant myometrial tumours using MR imaging. Eur Radiol 2013;23(8):2306–2314.
15. Therneau T, Atkinson B. port BR (producer of the initial R, maintainer 1999-2017). rpart: Recursive Partitioning and Regression Trees. https://CRAN.R-project.org/package=rpart. Published 2019. Accessed July 24, 2019.
16. Tamai K, Koyama T, Saga T, et al. The utility of diffusion-weighted MR imaging for differentiating uterine sarcomas from benign leiomyomas. Eur Radiol 2008;18(4):723–730.
17. Tanaka YO, Nishida M, Tsunoda H, Okamoto Y, Yoshikawa H. Smooth muscle tumors of uncertain malignant potential and leiomyosarcomas of the uterus: MR findings. J Magn Reson Imaging 2004;20(6):998–1007.
18. Namimoto T, Yamashita Y, Awai K, et al. Combined use of T2-weighted and diffusion-weighted 3-T MR imaging for differentiating uterine sarcomas from benign leiomyomas. Eur Radiol 2009;19(11):2756–2764.
19. Lin G, Yang LY, Huang YT, et al. Comparison of the diagnostic accuracy of contrast-enhanced MRI and diffusion-weighted MRI in the differentiation between uterine leiomyosarcoma / smooth muscle tumor with uncertain malignant potential and benign leiomyoma. J Magn Reson Imaging 2016;43(2):333–342.
20. Gerges L, Popiolek D, Rosenkrantz AB. Explorative Investigation of Whole-Lesion Histogram MRI Metrics for Differentiating Uterine Leiomyomas and Leiomyosarcomas. AJR Am J Roentgenol 2018;210(5):1172–1177.
21. Tong A, Kang SK, Huang C, Huang K, Slevin A, Hindman N. MRI screening for uterine leiomyosarcoma. J Magn Reson Imaging 2019;49(7):e282–e294.
22. Kim TH, Kim JW, Kim SY, Kim SH, Cho JY. What MRI features suspect malignant pure mesenchymal uterine tumors rather than uterine leiomyoma with cystic degeneration? J Gynecol Oncol 2018;29(3):e26.
23. Cornfeld D, Israel G, Martel M, Weinreb J, Schwartz P, McCarthy S. MRI appearance of mesenchymal tumors of the uterus. Eur J Radiol 2010;74(1):241–249.
24. Barral M, Placé V, Dautry R, et al. Magnetic resonance imaging features of uterine sarcoma and mimickers. Abdom Radiol (NY) 2017;42(6):1762–1772.

Radiology

Erratum

Originally published online:
https://doi.org/10.1148/radiol.2020191658

Diagnostic Algorithm to Differentiate Benign Atypical Leiomyomas from Malignant Uterine Sarcomas with Diffusion-weighted MRI

Cendos Abdel Wahab, Anne-Sophie Jannot, Pietro A. Bonaffini, Camille Bourillon, Caroline Cornou, Marie-Aude Lefrère-Belda, Anne-Sophie Bats, Isabelle Thomassin-Naggara, Alexandre Bellucci, Caroline Reinhold, Laure S. Fournier

Erratum in:
DOI:10.1148/radiol.2020209020

The last name for author Pietro A. Bonaffini was misspelled. It should appear as **Pietro A. Bonaffini.**