Clinical Radiology 74 (2019) 167.e1–167.e7

Contents lists available at ScienceDirect

Clinical Radiology

journal homepage: www.clinicalradiologyonline.net

A multiparametric MRI-based machine learning to distinguish between uterine sarcoma and benign leiomyoma: comparison with ^{18}F-FDG PET/CT

M. Nakagawa*, T. Nakaura, T. Namimoto, Y. Iyama, M. Kidoh, K. Hirata, Y. Nagayama, S. Oda, F. Sakamoto, S. Shiraishi, Y. Yamashita

Department of Diagnostic Radiology, Graduate School of Medical Sciences, Kumamoto University, Honjo 1-1-1, Tyuou-ku, Kumamoto, 860-0811, Japan

ARTICLE INFORMATION

Article history:
Received 24 January 2018
Accepted 18 October 2018

AIM: To compare the performance of machine learning using multiparametric magnetic resonance imaging (mp-MRI) and positron-emission tomography (PET) to distinguish between uterine sarcoma and leiomyoma.

MATERIALS AND METHODS: This retrospective study was approved by the institutional review board and informed consent was waived. Sixty-seven consecutive patients with uterine sarcoma or leiomyoma who underwent pelvic 3 T MRI and PET were included. Of 67 patients, 11 had uterine sarcomas and 56 had leiomyomas. Seven different parameters were measured in the tumours, from T2-weighted, T1-weighted, contrast-enhanced, and diffusion-weighted MRI, and PET. The areas under the receiver operating characteristic curves (AUC) with a leave-one-out cross-validation were used to compare the diagnostic performances of the univariate and multivariate logistic regression (LR) model with those of two board-certified radiologists.

RESULTS: The AUCs of the univariate models using MRI parameters (0.68–0.8) were inferior to that of the maximum standardised uptake value (SUVmax) of PET (0.85); however, the AUC of the multivariate LR model (0.92) was superior to that of SUVmax, and comparable to that of the board-certified radiologists (0.97 and 0.89).

CONCLUSION: The diagnostic performance of the machine learning using mp-MRI was superior to PET and comparable to that of experienced radiologists.

Introduction

Uterine sarcoma is a rare myometrial disease, accounting for 3–5% of all malignancies of the uterine body.[1] Tumour stage is the most important prognostic factor,[2] and early diagnosis of uterine sarcoma is important; however, it is generally difficult to differentiate early-stage uterine

* Guarantor and correspondent: M. Nakagawa, Department of Diagnostic Radiology, Graduate School of Medical Sciences, Kumamoto University, Honjo 1-1-1, Tyuou-ku, Kumamoto, 860-0811, Japan. Tel.: +81 96 344 2111.
E-mail address: mstknkgw.a.you.1@gmail.com (M. Nakagawa).

https://doi.org/10.1016/j.crad.2018.10.010
0009-9260/

sarcoma from leiomyoma, which is the most common disease of the uterine myometrium. This is because leiomyoma variants, such as cellular leiomyomas and myxoid leiomyomas, may mimic malignancy.[2]

Several recent studies have reported on the utility of magnetic resonance imaging (MRI) for the diagnosis of uterine tumours, including MRI with or without the addition of diffusion-weighted imaging (DWI), and 2-[^{18}F]-fluoro-2-deoxy-D-glucose (^{18}F-FDG) positron-emission tomography (PET) combined with computed tomography (CT).[3–7] Their results indicated that these techniques demonstrated a relatively good performance for the differentiation of uterine sarcoma from leiomyoma, although the diagnosis of uterine sarcoma generally depended on a subjective judgment, which was subject to high inter-reader variability. To the authors' knowledge, the utility of a diagnostic model that combines multiple MRI parameters to differentiate uterine sarcoma from uterine leiomyoma has not been evaluated. It was hypothesised that such an objective model might be useful in distinguishing uterine sarcoma from uterine leiomyoma.

The purpose of this study was therefore to develop and validate a predictive multiparametric MRI (mp-MRI) model to distinguish uterine sarcoma from uterine leiomyoma.

Materials and methods

Patients

This retrospective study was approved by the institutional review board and informed consent was waived. Between January 2007 and December 2015, 1,630 consecutive patients with uterine myometrial smooth muscle masses who underwent either pelvic MRI at 3 T or ^{18}F-FDG PET/CT were identified. Inclusion criteria were[1] patients had undergone contrast-enhanced MRI with DWI and ^{18}F-FDG PET/CT within 2 months before surgery at Kumamoto University Hospital,[2] the uterine myometrial smooth muscle masses were histopathologically proven uterine sarcomas or benign leiomyomas, or the lesions had not changed in size over an observation period of >6 months. The lesions with no change were considered to be benign leiomyomas. From the total of 1,630 patients identified, 1,367 had not undergone PET/CT, 154 had not undergone pelvic MRI, 12 patients had neither leiomyoma nor sarcoma identified at PET/CT, five patients had neither histopathologically proven leiomyoma nor sarcoma, one patient had not undergone DWI, 22 patients had not undergone contrast-enhanced MRI, and two patients had not undergone surgery or observation. These patients were therefore excluded from the study. A final total of 67 patients were enrolled in this study (Fig 1).

The patients had a mean ± SD age of 54.4±12.6 years (range 23–85 years). The histopathological findings in the 11 patients with uterine sarcomas were carcinosarcoma (n=4), leiomyosarcoma (n=3), endometrial stromal sarcoma (n=2), and adenosarcoma (n=2).

MRI examinations

All MRI examinations were acquired using a 3 T whole-body MRI system (Ingenia, Philips Medical Systems, Best, the Netherlands) equipped with a dual-source parallel radiofrequency transmission system in which a 32-channel phased-array surface coil served as the receiver coil. Patients received hyoscine butylbromide prior to imaging, providing that they had no contraindications. The MRI protocol included axial T2-weighted imaging (WI), axial T1WI, axial contrast-enhanced T1WI, and axial DWI with

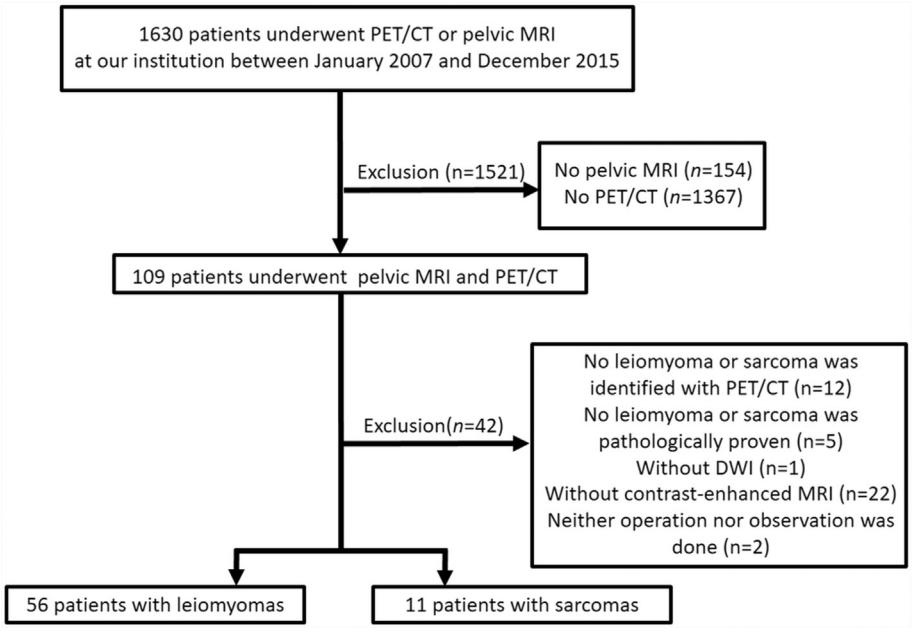

Figure 1 Flowchart showing the patient selection process.

two different b-values (0 and 1000). The apparent diffusion coefficient (ADC) values were calculated from b=0 and 1,000 s/mm^2. The imaging parameters were T1WI turbo spin echo: repetition time (TR)/echo time (TE)=500/10 ms, T2WI (TSE) TR/TE=3,700/100 ms, field of view: 250×250mm, matrix size: 576×461 (T2WI), 384×230 (T1WI), and 154×154 (DWI), 10° flip angle, and 5 mm section thickness. Patients were scanned in the supine position with their arms raised. MRI sequences were acquired in an axial orientation to the patients' body.

^{18}F-FDG PET examinations

All patients underwent PET/CT on a whole-body tomography scanner (Gemini-GXL 16, Philips). After fasting for at least 5 hours before tracer administration, each patient was injected intravenously with approximately 186–256 MBq of ^{18}F-FDG. Sixty minutes after ^{18}F-FDG administration, the whole-body emission scan began with 1.5 minutes in each region (one bed position: 18 cm). Each bed position was overlapped at 50% with the next bed position, so 3 minutes per bed position was performed for almost all of bed positions for image acquisition. Imaging parameters included PET: 144×144 matrix, and LOR-RAMLA reconstruction, CT: 120 kV, 80 mA, 16×1.5 mm collimation, 0.94 mm pitch, and 5 mm image reconstruction.

MRI image analysis

A radiologist (4 years of experience in MRI) then performed quantitative image analysis of the axial T2WI, axial T1WI, ADC maps, and axial contrast-enhanced T1WI images. For each patient, one axial section that included a uterine myometrial smooth muscle mass with greatest dimension was selected for analysis. The largest lesion was included for analysis in cases of patients with two or more lesions.

For all patients, the radiologist drew regions of interest (ROIs) in the gluteus maximus muscle on T2WI and measured their mean signal intensity (ROI_{muscle}). The mean signal intensity of the uterine myometrial smooth muscle mass was then measured from circular ROIs placed on selected slices. The entire tumour was outlined on the respective sections. The ROIs defined on the T2WI were transferred to the other images (contrast-enhanced T1WI and ADC maps) using copy and paste functions. The normalised T2WI signal intensity of the uterine myometrial smooth muscle mass was then calculated as the T2WI signal intensity/ROI_{muscle}. The normalised T1WI signal intensity and enhancement ratio were also calculated in the same way. All measurements were performed on a picture archiving and communication system (PACS) workstation (EV Insight; PSP, Tokyo, Japan).

PET/CT image analysis

A circular ROI was manually placed on the uterine myometrial smooth muscle mass image and the maximum standardised uptake value (SUVmax) was calculated for each lesion.

Statistical analysis

Continuous variables were expressed as mean ± SD. Statistical analyses were performed using the free statistical software program R (R, version 3.2.2; The R Project for Statistical Computing; http://www.r-project.org/). The age, tumour volume, normalised T1WI signal, normalised T2WI signal, enhancement ratio, mean and minimum ADC, and SUV_{max} were compared between uterine sarcoma and benign leiomyoma using the Mann–Whitney *U*-test. In this study, logistic regression (LR) was used as a machine learning algorithm. Single parametric LR models were then built for all variables associated with significant *p*-values. Two mp-MRI LR models were developed, one using mean ADC (mp-MRI model 1) and the other using minimum ADC (mp-MRI model 2), with all of the other significant variables included. These two separate models were used, as previous reports have suggested that mean and minimum ADC vary collinearly, and should not therefore be placed within the same multivariate analytic model. For all models, the Akaike information criterion (AIC), which is a measure of both training error and complexity, was calculated. This criterion was used as additional factors may result in a better mathematical fit to the data, but yield no additional biological information; this is due to overfitting to the training data.

The areas under the receiver operating characteristic (ROC) curves (AUCs) of the all LR models using a leave-one-patient-out cross-validation method to separate training from testing of the discriminant by the LR models in which the training with 66 patients and the validation with another patient were repeated 67 times with each different combination. The AUCs of the two radiologists (11 years and 20 years of experience) were also calculated. For MRI data, the criteria for sarcoma identification were as follows: heterogeneous hypointensity on T1WI, intermediate-to-high signal intensity on T2WI, heterogeneous enhancement after contrast medium administration, and hyperintensity on DWI with low ADC values. A value of $p<0.05$ was considered to indicate a statistically significant difference, and all interval estimations provided in this article are given with 95% confidence intervals (CIs).

Results

Univariate analysis

Table 1 shows the results of Welch's *t*-tests between uterine sarcoma and benign leiomyoma for all independent variables. There were no significant differences in age ($p=0.980$), tumour volume ($p=0.678$), normalised T1WI signal ($p=0.724$) or enhancement ratio ($p=0.^{18}8$) between uterine sarcoma and benign leiomyoma; however, the normalised T2WI signal was significantly higher in uterine sarcoma than in benign leiomyoma (2.19±0.65 versus 1.2±0.88, $p<0.001$), while the mean and minimum ADC were significantly lower in uterine sarcoma than in benign leiomyoma (mean ADC: 0.91±0.26 versus 1.22±0.31 s/mm^2,

Table 1
Comparison of each parameter between groups.

	Myoma (n=56)	Sarcoma (n=11)	p-Value
Age (year)	53.5±11.5	53.4±17.1	0.980
volume (mm^3)	464±1275	560±515	0.678
Standardised T1WI signal	0.87±0.27	0.89±0.17	0.724
Standardised T2WI signal	1.20±0.88	2.19±0.65	<0.001
Contrast enhancement ratio	1.83±0.71	1.53±0.66	0.188
Mean ADC value (s/mm^2)	1.22±0.31	0.91±0.26	0.004
Minimum ADC value (s/mm^2)	0.95±0.31	0.70±0.22	0.006
SUVmax	2.34±2.11	9.58±8.96	0.023

WI, weighted imaging; ADC, apparent diffusion coefficient; SUVmax, maximum standardised uptake value.

p=0.004; minimum ADC: 0.7±0.22 versus 0.95±0.31 s/mm^2, p=0.006, respectively). The SUV_{max} in uterine sarcoma was significantly higher than that in benign leiomyoma (9.58±8.96 versus 2.34±2.11, p=0.023).

Development of the LR models

In the present study, univariate LR models were developed for the normalised T2WI signal, mean ADC, minimum ADC, and SUV_{max}. Two multivariate LR models were also developed for normalised T2WI signal and mean ADC (mp-MRI model 1), and normalised T2WI signal and minimum ADC (mp-MRI model 2; Table 2). The AIC analysis of the univariate models revealed that the lowest AIC was with SUV_{max} (AIC=45.553), followed by T2WI (AIC=53.865), mean ADC (AIC=54.507), and minimum ADC (AIC=56.871); however, the multivariate LR models with mp-MRI revealed lower AIC values than obtained with SUV_{max} (mp-MRI model 1: 36.002, mp-MRI model 2: 38.218). The final model was developed using mean ADC and normalised T2WI signal.

The probability of a given ROI being uterine sarcoma was provided by the final LR model as follows:

$$Probabilty = \frac{e^{a*meanADC+b*T2WI+c}}{1+e^{a*meanADC+b*T2WI+c}} \quad (1)$$

where *meanADC* is the mean ADC, *T2WI* is the standardised T2WI signal, a is −6.82, b is 2.60, and c is −0.89.

Validation of LR models

Fig 2 shows the ROC curves for MRI, PET, and the two board-certified radiologists. The AUC of the LR model with SUV_{max} (AUC 0.85) was higher than the AUCs of the univariate LR models with mp-MRI: normalised T2WI (AUC 0.80), mean ADC (AUC 0.78), and minimum ADC (AUC 0.68); however, the AUC of the multivariate LR model (AUC 0.92) was superior to that of SUVmax, and comparable to that of the board certified radiologists (reader 1: 0.97; reader 2: 0.89). Figs 3–5 show representative cases.

Discussion

In the present study, the best mp-MRI LR model for determination of uterine sarcoma from uterine leiomyoma contained only the normalised T2WI signal and mean ADC value. The AUC of the multivariate LR model was superior to that of the SUVmax, and comparable to those of the board-certified radiologists.

Many quantitative MRI and PET parameters have been reported as useful tools for the differentiation of uterine leiomyosarcoma and benign leiomyoma; however, there is still debate over which technique offers superior performance in the differentiation of uterine leiomyosarcoma and benign leiomyoma. Lin *et al.* reported that contrast-enhanced MRI was found to show the most characteristic imaging features of leiomyosarcoma/uterine smooth

Table 2
Logistic regression models.

	T2WI model		Mean ADC model		Minimum ADC model		Maximum SUV model		Combination model 1		Combination model 2	
	OR (CI)	p	OR (CI)	p	OR (CI)	p	OR (CI)	p	OR (CI)	p	OR (CI)	p
(Intercept)	0.03		11.68		3.46		0.04		2.45		1.21	
Standardised T2WI signal	3.04 (1.51–6.87)	0.003							13.50 (3.58–103.93)	0.002	11.59 (3.41–68.25)	0.001
Mean ADC value			0.02 (0.00–0.29)	0.011					0.00 (0.00–0.04)	0.002		
Minimum ADC value					0.03 (0.00–0.44)	0.021					0.00 (0.00–0.03)	0.003
SUV max							1.47 (1.18–2.03)	0.005				
Observations	67		67		67		67		67		67	
Pseudo-R^2	R^2_{CS}=0.138 R^2_N=0.234 D=0.136		R^2_{CS}=0.130 R^2_N=0.220 D=0.156		R^2_{CS}=0.099 R^2_N=0.167 D=0.103		R^2_{CS}=0.239 R^2_N=0.404 D=0.356		R^2_{CS}=0.359 R^2_N=0.608 D=0.489		R^2_{CS}=0.338 R^2_N=0.572 D=0.460	
AIC	53.865		54.507		56.871		45.553		36.002		38.2[18]	
AUC	0.80		0.78		0.68		0.85		0.92			

OR, odds ratio, CI, 95% confidence interval, WI, weighted imaging; ADC, apparent diffusion coefficient; SUVmax, maximum standardised uptake value; R^2_{CS}, Cox and Snell pseudo-R^2; R^2_N, Nagelkerke pseudo-R^2, D, deviance.

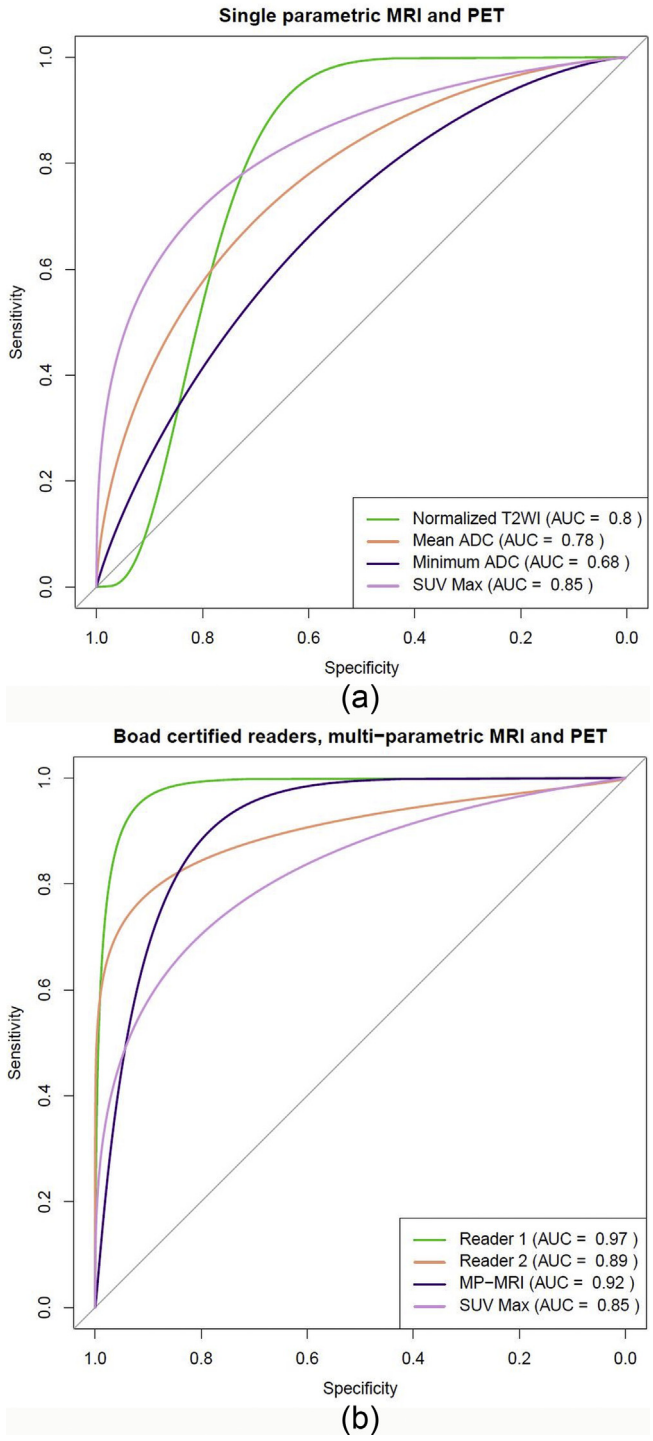

Figure 2 ROC analysis of MRI, PET, and board-certified radiologists. (a) ROC curves for the T2WI model (green line), mean ADC (yellow line), minimum ADC (deep blue line), and SUVmax (purple line). (b) ROC curves for the two radiologists (reader 1: green line, reader 2: orange line), mp-MRI (deep blue line), SUVmax (purple line).

muscle tumours with uncertain malignant potential, and yielded a diagnostic accuracy of 0.94, which was significantly higher than that achieved with DWI, T2WI, or T1WI criteria (0.52, 0.46, and 0.67 respectively, $p<0.05$ for all).[8]

Yoshida *et al.* reported that ^{18}F-FDG PET was marginally superior to MRI alone (T1WI and T2WI) according to ROC analysis ($p=0.05$),[3] and Nagamatsu *et al.* reported that the accuracy of an SUV cut-off value of 3 for the diagnosis of leiomyosarcoma was 79%[5]; however, the present authors are not aware of any previous reports that directly compared the diagnostic performance of MRI and PET in the differentiation of uterine leiomyosarcoma from benign leiomyoma. The present study suggests that for the differentiation of uterine leiomyosarcoma and benign leiomyoma, the diagnostic performance of a single MRI parameter is not as high as PET or the board-certified radiologists. This may be partly caused by the fact that atypical leiomyomas may present hyperintensities on T2WI and the fact that typical leiomyomas suffer from T2 black-out effect.[9]

Contrastingly, the present study suggested that the performance of mp-MRI was better than that of PET, and comparable to that of board-certified radiologists. This suggests that the combination of T2WI and ADC are important for the differentiation of uterine sarcoma from benign leiomyoma. This finding appears reasonable in consideration of previous reports. Namimoto *et al.* reported that for the differentiation of sarcomas from leiomyomas, a combination of ADC and contrast ratio on T2WI was significantly better than ADC or T2WI alone.[4] In clinical practice, there is usually no difficulty in diagnosing leiomyomas with low signal intensity on T2WI. A problem arises when lesions present hyperintensity on T2WI. Such lesions with hyperintensity on T2WI are not affected by the T2 black-out effect, and it is, therefore, reasonable for the AUC of the combined normalised T2WI and ADC to be high.

The AUC of the combined model (normalised T2WI and mean ADC) was comparable to those of the board-certified radiologists (0.92 and 0.97), which suggests it may have potential for clinical use. A substantial advantage of the LR model is that it allows the estimation of the diagnosis probability with a simple procedure.

There were some limitations to the present study. First, the number of patients included in this study was small. A large number of patients is needed to make a stronger LR model, although the present shows good diagnostic performance. A larger study should be performed prospectively to verify the present results. Second, some leiomyomas were not histopathologically proven; typical clinical and laboratory findings in combination with follow-up imaging findings were used as diagnostic criteria for these patients. Third, the follow-up interval was short because some patients included in the study had had a malignant tumour and died without a long enough follow-up interval. Fourth, there might be a selection bias because of patients in a single institute. These patients might not represent patients in the general population. Fifth, uterine sarcomas have some histological subtypes and there are differences in prognosis and imaging appearance between them. Lastly, volumetric analysis in multiple lesions was not performed, and important features associated with intra- and inter-tumour

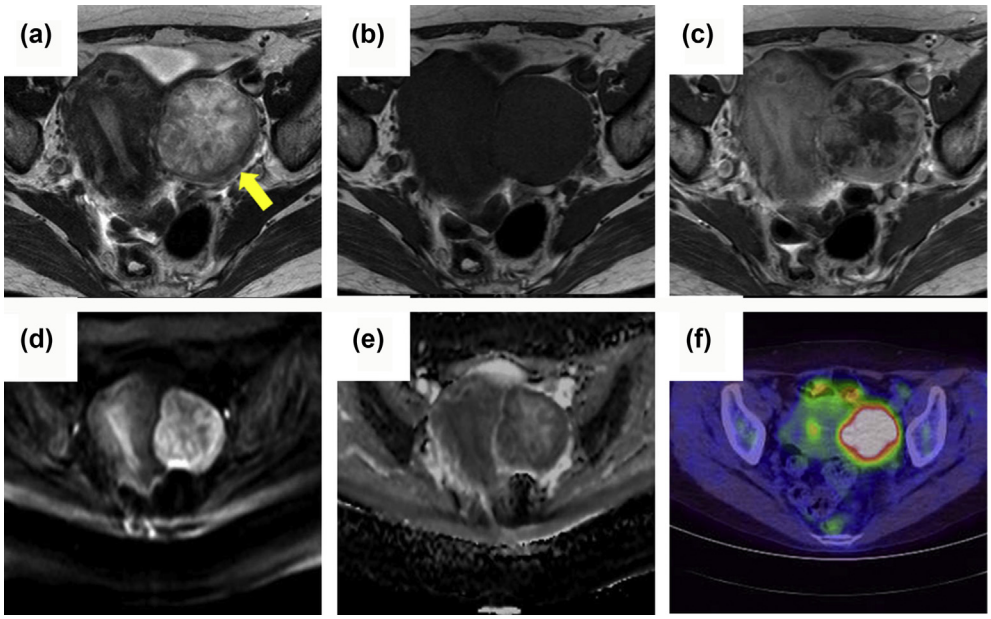

Figure 3 A 50-year-old woman with uterine leiomyosarcoma. Axial T2WI (a), axial T1WI (b), axial Gd-enhanced T1WI (c), axial DWI at b=1,000 (d), ADC (e), fusion of PET/CT (f). A heterogeneous hyperintense lesion is observed on T2WI and contrast-enhanced T1WI shows heterogeneous enhancement. The normalised T2WI signal is 2.42 (996/412), the mean ADC is 0.78, and the corresponding SUVmax of the lesion is 7.3. The probability of the lesion being uterine sarcoma can be calculated by the LR model as follows: Probability $= \frac{e^{-6.82\times0.78+2.60\times2.42-0.89}}{1+e^{-6.82\times0.78+2.60\times2.42-0.89}} = 52.1\%$.

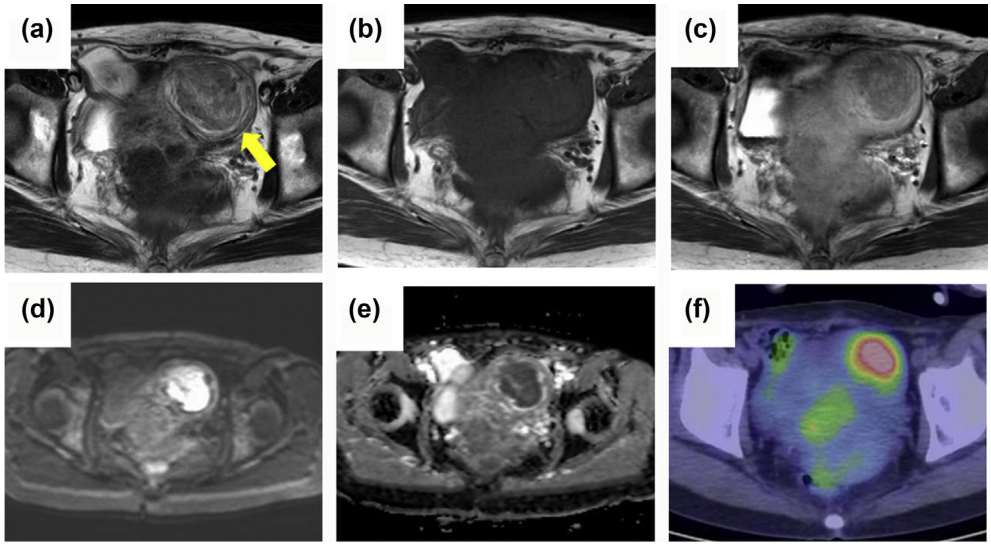

Figure 4 A 48-year-old woman with atypical leiomyoma. Axial T2WI (a), axial T1WI (b), axial Gd-enhanced T1WI (c), axial DWI at b=1,000 (d), ADC (e), fusion of PET/CT (f). A heterogeneous hyper-intense lesion is observed on T2WI. Contrast-enhanced T1WI demonstrates relatively homogeneous enhancement. The normalised T2WI signal is 2.74 (895/327), the mean ADC is 1.1, and the corresponding SUVmax of the lesion is 5.0. The probability of the lesion being uterine sarcoma can be calculated by the LR model as follows: Probability $= \frac{e^{-6.82\times1.1+2.60\times2.74-0.89}}{1+e^{-6.82\times1.1+2.60\times2.74-0.89}} = 21.9\%$.

heterogeneity can be missed with the present method; however, a single measurement method in the largest tumour might be useful in daily clinical use.

In conclusion, the diagnostic performance of the machine learning with mp-MRI was superior to the SUVmax of PET/CT, and comparable to that of experienced radiologists.

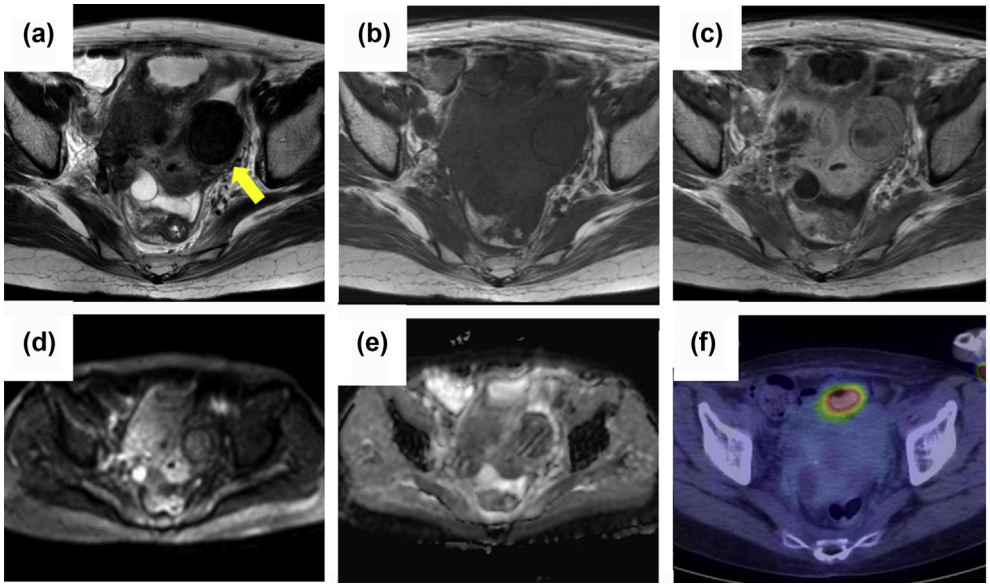

Figure 5 A 53-year-old woman with uterine leiomyoma. Axial T2WI (a), axial T1WI (b), axial Gd-enhanced T1WI (c), axial DWI at b=1,000 (d), ADC (e), fusion of PET/CT (f). A homogeneous hypo-intense lesion is observed on T2WI. Contrast-enhanced T1WI demonstrates relatively homogeneous enhancement. The normalised T2WI signal is 0.57 (217/379), the mean ADC is 1.58, and the corresponding SUVmax of the lesion is 1.1.

Conflict of interest

The authors declare no conflict of interest.

References

1. Nagai T, Takai Y, Akahori T, *et al*. Novel uterine sarcoma preoperative diagnosis score predicts the need for surgery in patients presenting with a uterine mass. *SpringerPlus* 2014;**3**:678.
2. Prat J, Mbatani. Uterine sarcomas. *Int J Gynaecol Obstet* 2015;**131** (Suppl. 2):S105–10.
3. Yoshida Y, Kurokawa T, Sawamura Y, *et al*. Comparison of ^{18}F-FDG PET and MRI in assessment of uterine smooth muscle tumors. *J Nucl Med* 2008;**49**(5):708–12.
4. 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):2756–64.
5. Nagamatsu A, Umesaki N, Li L, *et al*. Use of ^{18}F-fluorodeoxyglucose positron emission tomography for diagnosis of uterine sarcomas. *Oncol Rep* 2010;**23**(4):1069–76.
6. Dubreuil J, Tordo J, Rubello D, *et al*. Diffusion-weighted MRI and ^{18}F-FDG-PET/CT imaging: competition or synergy as diagnostic methods to manage sarcoma of the uterus? A systematic review of the literature. *Nucl Med Commun* 2017;**38**(1):84–90.
7. 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):2903–15.
8. 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):333–42.
9. Silvera S, Oppenheim C, Touze E, *et al*. Spontaneous intracerebral hematoma on diffusion-weighted images: influence of T2-shine-through and T2-blackout effects. *AJNR Am J Neuroradiol* 2005;**26**(2):236–41.