European Radiology (2019) 29:4485–4494
https://doi.org/10.1007/s00330-018-5845-9

MUSCULOSKELETAL

Conventional MR and diffusion-weighted imaging of musculoskeletal soft tissue malignancy: correlation with histologic grading

Avneesh Chhabra[1,2] · Oganes Ashikyan[1] · Chenelle Slepicka[1] · Nathan Dettori[1] · Helena Hwang[3] · Alexandra Callan[2] · Rohit R. Sharma[4] · Yin Xi[1]

Received: 20 August 2018 / Revised: 22 September 2018 / Accepted: 22 October 2018 / Published online: 3 December 2018

Abstract

Aim To evaluate proven soft tissue musculoskeletal malignancies blinded to their Fédération Nationale des Centres de Lutte Contre le Cancer histologic grades to identify the predictive values of conventional MR findings and best fit region of interest (ROI) apparent diffusion coefficient (ADC) measurements.

Materials and methods Fifty-one consecutive patients with different histologic grades were evaluated by four readers (R1–4) of different experience levels. Quantitatively, the maximum longitudinal size, tumor to muscle signal intensity ratios, and ADC measurements and, qualitatively, the spatial location of the tumor, its signal alterations, heterogeneity, intralesional hemorrhage or fat, and types of enhancement were assessed. Intraclass correlation, weighted kappa, ANOVA, and Fisher exact tests were used.

Results There were 22/51 (43%) men (mean age ± SD = 52 ± 16 years) and 29/51 (57%) women (mean age ± SD = 54± 17 years), with the majority of tumors 38/51 (75%) in the lower extremities. Histologic grades were I in 8/51 (16%), II in 17/51 (33%), and III in 26/51 (51%), respectively. The longitudinal dimensions were different among three grades ($p = 0.0015$), largest with grade I. More central enhancements and deep locations were seen in grade III tumors ($p = 0.0191$, 0.0246). The ADC mean was significantly lower in grade III than in grade I or II ($p < 0.0001$ and $p = 0.04$). The ADC min was significantly lower in grade III than in grade I ($p = 0.02$). Good to excellent agreements were seen for T1/T2 tumor/muscle ratios, longitudinal dimension, and ADC (ICC = 0.60–0.98).

Conclusion Longitudinal tumor dimension, central enhancement, and ADC values differentiate histology grades in musculoskeletal soft tissue malignancy with good to excellent inter-reader reliability.

Key Points

- *The longitudinal tumor dimension of grade III malignancy is smaller than that of grade I ($p < 0.0001$), and higher-grade tumors are located deeper ($p = 0.0246$).*
- *The ADC mean is significantly lower in grade III than in grade I or grade II ($p < 0.0001$ and $p = 0.04$).*
- *The ADC minimum is significantly lower in grade III than in grade I ($p = 0.02$).*

Keywords MRI · DWI · Sarcoma · Tumor grade · FNCLCC

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00330-018-5845-9) contains supplementary material, which is available to authorized users.

✉ Avneesh Chhabra
avneesh.chhabra@utsouthwestern.edu

[1] Radiology, UT Southwestern Medical Center, Dallas, TX 75022, USA

[2] Orthopedics, UT Southwestern Medical Center, Dallas, TX, USA

[3] Pathology, UT Southwestern Medical Center, Dallas, TX, USA

[4] Surgical Oncology, UT Southwestern Medical Center, Dallas, TX, USA

Abbreviations

ADC	Apparent diffusion coefficient
DWI	Diffusion-weighted imaging
FNCLCC	Fédération Nationale des Centres de Lutte Contre le Cancer
ROI	Region of interest

Introduction

Magnetic resonance (MR) imaging is the modality of choice for characterization of musculoskeletal soft tissue masses. In addition, MR is most helpful for evaluating the solid versus cystic nature of a soft tissue mass and delineating the extent of the malignancy for pre-operative planning and prognostication [1–4]. Histology obtained from percutaneous or surgical biopsy serves as the reference standard for the final diagnosis of soft tissue malignancy and its tumor grade [5]. The Fédération Nationale des Centres de Lutte Contre le Cancer (FNCLCC) grading system is commonly used in these patients for planning neoadjuvant chemo-radiation treatments, surgery, and prognostication [6, 7]. Moderate accuracy of conventional MR imaging has been shown in the differentiation of benign versus malignant soft tissue lesions and prediction of the tumor histology [8, 9]. Expert opinions or evaluation at a dedicated sarcoma center is of additional value for earlier and more accurate diagnosis of the malignancy as well as for facilitating expedited patient management [10–12]. Higher-grade tumors often require neoadjuvant radiation treatment and/or chemotherapy, and it can be helpful if the tumor grade can be predicted prospectively on imaging. Preliminary studies have been performed to predict the tumor histology and grading (low grade versus high grade) using conventional MR imaging [13, 14]. Various MR imaging findings such as infiltrating tumor margins, peritumoral edema or enhancement, rounded tumor margins, larger tumors, and intratumoral necrosis have been shown to be associated with malignancy [13, 14]. To identify higher grade of tumor, peritumoral enhancement was shown as the best predictor by Zhao et al [14]. However, there is a considerable overlap of these findings among various tumors and their histologic grades. Even chronic expanding hematomas have been shown to exhibit tumor-like enhancing nodularity on conventional MR imaging [15].

Diffusion-weighted imaging (DWI) is being used over the last few years in the domain of musculoskeletal soft tissue lesions to evaluate the tumor cellularity. DWI reflects the degree of free water diffusion within the tissues. Tumors or tissues with lower free water content, proteineous content, or high cellularity tend to restrict diffusion to a greater degree and vice versa. Thus, DWI provides functional information that can complement the structural or anatomic information obtained from the conventional MR imaging. The degree of diffusion is quantitated by the apparent diffusion coefficient (ADC) using mono-exponential or multi-exponential approaches. Simple mono-exponential ADC evaluation has been shown to be reliable and useful for the differentiation of benign versus malignant bone lesions [16], with the minimum ADC being reported to be the most accurate parameter. For soft tissue sarcoma heterogeneity characterization and tumor grading, different approaches have been tested recently, including volumetric segmentation of ADC [17] and positron emission tomography-MR imaging to correlate ADC with standardized glucose uptake values [18]. The selected observer-based method of ADC measurement has been shown to be comparable to the whole tumor volume measurement method and has been established as the most practical and fast method [19]. Major pitfalls of DWI interpretations include evaluation of such images in isolation without correlation with conventional anatomic MR images, and inadvertent inclusion of fatty, myxoid, and hemorrhagic areas in the ADC calculation [20]. Ideally bright areas on DWI and corresponding dark areas of ADC maps or vice versa are selected using the best-fit region of interest (ROI) for the optimal evaluation of tumor cellularity. Thus, the fibrous, fatty, infarcted, necrotic, myxoid, and calcified areas are avoided in the ADC calculation to estimate the real tumor cellular density. To our knowledge, systematic evaluation of soft tissue sarcomas with best-fit ADC ROI for tumor grading while avoiding fibrous, fatty, necrotic, myxoid, and calcified areas has not yet been performed.

The authors evaluated a consecutive series of histology-proven soft tissue malignancies blinded to their established different FNCLCC grades to identify the predictive values of various conventional MR imaging findings and the best-fit ROI ADC measurements. Inter-reader reliability analysis was performed. Our hypothesis was that best-fit ROI mean and minimum ADC measurements can separate the tumor grades.

Methods

This was a cross-sectional retrospective evaluation performed following local institutional IRB approval and regulations. The informed consent was waived.

Patient population

An institutional electronic search was performed for musculoskeletal extremity contrast-enhanced MR imaging studies performed for consecutive tumors over a 4-year period from 2013 to 2017 from PACS and sarcoma board records. Inclusion criteria included both genders, age between 18 and 100 years, upper or lower extremity malignant soft tissue tumors, established histology on biopsy or tumor resection, complete MR imaging protocol including conventional imaging sequences, contrast imaging and DWI prior to biopsy, and

surgery or neoadjuvant treatment. Exclusion criteria included benign tumors, metastasis, and lack of complete imaging protocol or final histology proof or tumor grading. We did not include gynecologic or retroperitoneal sarcomas.

The patient demographics including age and gender, location of tumor, and final histology grades were recorded. A sarcoma specialist pathologist (HH) was consulted, and the cases were re-reviewed to establish the missing FNCLCC grades from the final histology reports.

MR imaging

The MR imaging was performed on various scanners in the institute (Siemens, GE, and Philips) using a uniform protocol established for musculoskeletal tumor imaging using a flexible phased array body matrix coil. 27/51 scans were performed on 1.5-T scanners and 24 scans were performed on 3-T scanners. Conventional imaging (T1W, fsT2W in multiple planes), contrast-enhanced imaging (3D volumetric fat-suppressed isotropic voxel, pre- and post-contrast imaging), and DWI were obtained in all cases as outlined in the inclusion criteria. The parameters were as follows: T1-weighted (repetition time (TR)/echo time (TE), 600–715/9–17; section thickness, 4 mm; axial and sagittal planes), fat-suppressed (fs) T2-weighted (TR/TE, 3600–6000/60; section thickness, 4 mm; axial plane), unenhanced and gadolinium-enhanced fsT1W sequence (isotropic resolution, repetition time/echo time (TR/TE), 4.6–6.3/1.4–1.5; section thickness, 1.5 mm; modified Dixon fat suppression; coronal plane with axial and sagittal reconstructions; 0.1 mmol/kg gadolinium-based contrast agent). The axial DWI was performed before contrast administration using a single-shot echo-planar imaging (TR/TE, 8000–11,700/70–92; section thickness, 4 mm; *b* values, 50, 400, and 600–800 s/mm^2; flip angle, 90°; matrix, 128 × 128; fat suppression, spectral adiabatic inversion recovery; time of acquisition, 5 min 10 s). The ADC value was calculated using all *b* values and the ADC map was automatically generated from the scanner.

Data evaluation

Using conventional MR images and DWI, both quantitative and qualitative analyses were performed by multiple readers.

Quantitative analysis Four readers (a medical student, R1; a radiology resident, R2; and two experienced musculoskeletal fellowship-trained faculty, R3 and R4) evaluated the anonymized MR imaging data sets blinded to the histology and tumor grades as well as to each other's measurements. R1 and R2 were pre-trained on five cases by the faculty regarding conventional MR imaging and DWI measurements, following which both R1 and R2 independently recorded the maximum longitudinal size of the tumor on fsT2W images, tumor to adjacent non-involved skeletal muscle signal intensity ratios on T1W (weighted) and fat-suppressed T2W (fsT2W) images, and ADC measurements (mean and minimum) using the best-fit ROI encompassing the majority of the solid area on a single most representative slice. The slice with the lowest (darkest) ADC was chosen for analysis. The darkest area on ADC (at least 10 mm^2) corresponding to the brighter area on DWI and/or enhancing area on contrast-enhanced imaging was measured excluding any areas of hemorrhage and fat as correlated with the conventional MR images using a circular ROI tool. After three measurements, the mean and minimum ADCs representative of that lesion were recorded. R3 and R4 also measured the ADC (mean and minimum) on separate occasions using the same methodology. The tumor to muscle ratios were calculated to quantitate the signal intensity alterations, which are routinely evaluated by musculoskeletal radiology readers while assessing the soft tissue tumors.

Qualitative analysis Faculty readers R3 and R4 recorded the tumor signal intensity and tumor heterogeneity on T1W and fsT2W, intralesional hemorrhage or fat (including %), non-enhancing necrotic-cystic areas, and types of enhancement (central or peripheral). Central enhancement was defined as enhancement within the internal substance of the lesion including confluent enhancement, solid nodular enhancement, or enhancing thick (> 1 mm) septations in the central core of the tumor. Peripheral enhancement was recorded if the enhancement was limited to the rim of the lesion without a central enhancing component as above. Peritumoral edema and peritumoral enhancement were also recorded in consensus [14]. The tumors were considered superficial if localized to the skin and subcutaneous tissues. The spatial location of the tumor within the muscles, under the deep fascia, or any extension of a large mass into the deep muscle compartment or around the bones was recorded as deep location.

Statistical analysis

Intraclass correlation (ICC) and kappa (weighted kappa if applicable) were used to assess the reader agreements for continuous and categorical variables, respectively. Agreement was considered poor with ICC/kappa < 0.4; fair, 0.4–0.59; good, 0.6–0.74; and excellent, 0.75–1. Log transformation was used on continuous variables to adjust for right skewness. Quantile-quantile plots of the residuals were checked for normality assumption. No significant violation was observed. ANOVA were used to test the difference in the continuous variables between tumor grades on average. If the overall *p* value was significant, multiple comparisons were conducted with Tukey adjustment. Fisher's tests were used to test the association between tumor grades and the categorical variables. The Cochrane-Armitage test was used to evaluate differences of tumor grades with respect to the spatial location of the tumor as deep versus superficial, 1.5-T and 3-T scanners,

and the two DWI protocols with different diffusion moment values. Receiver operating curve (ROC) analyses were performed on measurements that were significant. The diagnostic performance of predicting tumor grades was evaluated with the area under the ROC curve (AUC, area under the curve) and corresponding 95% confidence intervals. Leave-one-out cross validation was used to reduce overfit. A p value of less than 0.05 was considered statistically significant. All analyses were done in SAS 9.4 (SAS Institute Inc.).

Results

Patient population

Among 83 patients with suspected soft tissue malignancies, 32 were excluded for reasons such as final pathology of non-malignant tumor [19], lack of final pathologic diagnosis [6], and no DWI [7]. Thus, 51 patients with 51 musculoskeletal soft tissue malignancies were included in the final inclusion sample.

Table 1 Soft tissue malignancy tumor types, histology grades, and locations

Study number	Tumor type	Tumor grade	Location
1	Angiosarcoma, epithelioid, high grade	III	R lower leg
2	Extraskeletal myxoid chondrosarcoma	III	L lower leg
3	Extraskeletal myxoid chondrosarcoma	I	L knee
4	Chondrosarcoma myxoid	II	R pelvis
5	Epithelioid sarcoma	II	R forearm
6	Fibromyxoid sarcoma	I	L thigh
7	Fibrosarcoma	I	R knee
8	Fibrous histiocytoma	III	L lower leg
9	Fibrous histiocytoma	III	L thigh
10	Leiomyosarcoma	II	R thigh
11	Leiomyosarcoma	II	L knee
12	Pleomorphic sarcoma undifferentiated	III	L knee
13	Liposarcoma	III	R forearm
14	Liposarcoma myxoid	II	R lower leg
15	Liposarcoma myxoid	I	R thigh
16	Liposarcoma myxoid	I	R lower leg
17	Liposarcoma myxoid	III	R lower leg
18	Liposarcoma myxoid	I	R thigh
19	Liposarcoma	I	L thigh
20	Lymphoma, B cell	III	L knee
21	Lymphoma, B cell	III	R forearm
22	Lymphoma, B cell	III	L shoulder
23	Malignant peripheral nerve sheath tumor	II	R thigh
24	Myxofibrosarcoma	II	L thigh
25	Myxofibrosarcoma	II	L thigh
26	Myxofibrosarcoma	II	R thigh
27	Myxofibrosarcoma	II	Pelvis
28	Pleomorphic sarcoma undifferentiated	II	L lower leg
29	Pleomorphic sarcoma undifferentiated—myxoid changes	II	R thigh
30	Pleomorphic sarcoma undifferentiated	III	R forearm
31	Pleomorphic sarcoma undifferentiated	III	R thigh
32	Pleomorphic sarcoma undifferentiated	III	L thigh
33	Pleomorphic sarcoma undifferentiated	III	L upper arm
34	Pleomorphic sarcoma undifferentiated	III	L thigh
35	Pleomorphic sarcoma undifferentiated	III	L thigh
36	Pleomorphic sarcoma undifferentiated	III	L thigh
37	Pleomorphic sarcoma undifferentiated	III	R lower leg
38	Pleomorphic sarcoma undifferentiated	III	L thigh
39	Pleomorphic sarcoma undifferentiated	II	L forearm
40	Pleomorphic sarcoma undifferentiated	III	L forearm
41	Synovial sarcoma	II	R thigh
42	Synovial sarcoma	III	R ankle
43	Synovial sarcoma	II	L thigh
44	Synovial sarcoma	II	R foot
45	Pleomorphic sarcoma undifferentiated with myxoid changes	III	R lower leg
46	Myxoid sarcoma	II	R chest wall
47	Pleomorphic sarcoma undifferentiated	III	R foot
48	Malignant peripheral nerve sheath tumor	I	R pelvis
49	Pleomorphic sarcoma undifferentiated	III	R arm
50	Small cell neuroendocrine tumor	III	R thigh
51	Pleomorphic sarcoma undifferentiated	III	R thigh

Table 2 Quantitative and qualitative MR imaging characteristics with respect to tumor grades. ANOVA were used to test the difference in the continuous variables between tumor grades on average. Fisher's tests were used to test the association between tumor grades and the categorical variables. The Cochrane-Armitage test was used to evaluate the differences of tumor grades with respect to the spatial location of the tumor as deep versus superficial

Feature	Grade I	Grade II	Grade III	*p* value (adjusted for age and gender)
T1 tumor/muscle signal ratio	1.29/0.72 (mean ± SD)	1.15/0.26	1.33/0.44	0.5264
fsT2 tumor/muscle signal ratio	4.02/1.98	3.81/1.21	4.23/2.49	0.9327
T1W				
Hypointense	1	1	1	0.57571
Isointense	2	6	5	
Hyperintense	5	10	20	
T1W				
Homogeneous	3	7	13	0.79957
Heterogeneous	5	10	13	
fsT2W				
Hypointense	1	0	0	0.12368
Isointense	0	0	0	
Hyperintense	1	3	10	
Mixed	6	14	16	
fsT2W				0.09372
Homogeneous	1	0	6	
Heterogeneous	7	17	20	
Non-enhancing cystic areas				0.14628
Present	3	4	14	
Absent	5	13	12	
Peritumoral edema	7/8	17/17	25/26	0.0985
Peritumoral enhancement	2/8	8/17	15/26	0.2529
Enhancement				0.01912
Peripheral	2	0	8	
Central or mixed	6	17	18	
Fat % in liposarcoma	2B, 1C, 1D	1B	1A	A—0–25%, B—25–50%, C—50–75%, D—75–100% of tumor volume
Hemorrhage	2/8	7/17	8/26	1.0
Largest tumor dimension	16.36 ± 9.42 cm	10.74 ± 6.09 cm	7.05 ± 4.82	0.0015
Deep tumor location	8/8	11/17	14/26	0.0246
ADC				
Mean	1.6 ± 0.59	1.2 ± 0.45	0.9 ± 0.34	0.0007
Minimum	1.21 ± 0.54	0.77 ± 0.34	0.66 ± 0.34	0.0209

The final diagnoses were established on percutaneous core biopsy (17/51) and surgical histopathology (34/51). All histology results containing tumor grades were obtained within 7.5 ± 7.6 weeks (range = 1–31 weeks). All MRI scans were performed before any neoadjuvant treatment was initiated. There were 22/51 (43%) men (mean age ± SD = 52 ± 16 years) and 29/51 (57%) women (mean age ± SD = 54 ± 17 years) with 38/51 (75%), 9/51 (18%), 1/51 (1%), and 3/51(6%) tumors in the lower extremities, upper extremities, chest wall, and pelvis, respectively. The most common histologies were pleomorphic sarcoma 17/51 (33%) and liposarcoma 7/51 (14%) (Table 1). The FNCLCC grades were I in 8/51 (16%), II in 17/51 (33%), and III in 26/51 (51%), respectively. Among the 51 malignancies, only 3/51 showed multifocal appearance, and in such cases, the largest tumor was evaluated for the analysis purpose. The various descriptive statistics of MR imaging characteristics with respect to tumor grades are outlined in Table 2.

Table 3 Adjusted *p* values for pairwise comparisons (Tukey adjustment)

Quantitative MRI and DWI characteristics	Tumor grade		
	1 vs II	I vs III	II vs III
ADC mean	0.16	< 0.0001	0.04
ADC min	0.17	0.02	0.45
Largest dimension (cm)	0.18	< 0.0001	0.07

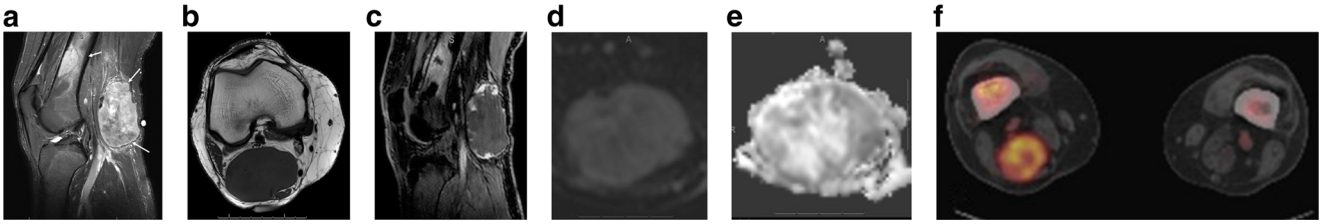

Fig. 1 Grade 1 tumor. **a** Sagittal fat-suppressed T2W image shows a 7.4-cm heterogeneous popliteal fossa mass with mixed signal intensity and peritumoral edema (small arrows). Also note a metastatic lesion in the distal femur (large arrow). **b** Axial T1W image shows homogeneous signal without hemorrhage. **c** 3D fat-suppressed post-contrast sagittal image shows a peripheral nodular enhancement. **d**, **e** Axial DWI and ADC images showed mean and minimum ADC of 1.2 and 0.8, respectively. Diagnosis on histology was fibrosarcoma, sclerosing epithelioid type and FNCLCC grade I. **f** F18 FDG image shows high glucose uptake in the popliteal mass lesion and additional femoral lesion, a metastatic lesion

Conventional MR imaging analysis

The longitudinal tumor dimensions were different among the three grades ($p = 0.0015$), with the larger lesions seen with grade I malignancy. On adjusted pairwise comparison, the difference persisted in grade I vs grade III tumors (Table 3). The largest dimension was significantly lower in grade III than in grade I ($p < 0.0001$); however, no significant difference was found between grade I and II or grade III and II ($p = 0.18$ and $p = 0.07$). The tumor to skeletal muscle signal intensity ratios were not different among the groups ($p = 0.5$–1.0). None of the other conventional MRI characteristics including signal intensity alterations, heterogeneity, or cystic areas were significant. Among the 51 malignancies, 33 were located in the deep compartments and 18 in the superficial compartment. Higher-grade tumors were more likely to be deep ($p = 0.0246$). On contrast imaging, there was significantly more central enhancement in grade III tumors as compared to the other grades ($p = 0.0191$). Intralesional hemorrhage was seen in 17/51 (33.3%) tumors. Intralesional fat was seen in 6/7 liposarcomas and was absent in the grade III tumors while grade I tumors expectedly showed a higher percentage of fat. None of the other malignancies showed intralesional fat. There was no significant difference in peritumoral edema and enhancement among the different grades ($p = 0.0985$, 0.2529). Interestingly, one patient had a soft tissue malignancy in the lower limb with peritumoral edema and no peritumoral enhancement, and additional bone metastasis in the same field of view (Fig. 1). It turned out to be a grade I lesion. On further F18 FDG PET imaging, multiple additional metastases were seen in other bones.

DWI analysis

Among the 51 scans performed with DWI (19 with b values of 0 and 600 and 32 with b values of 50, 400, and 800), there was not enough evidence to suggest a trend in tumor grade between the b values ($p = 0.5022$) or between the scanners ($p = 0.5601$). The DWI images showed a mild wrap artifact at the terminal field of views in 7/51 cases and a mild susceptibility artifact in 6/51 cases. None of these artifacts overlapped the tumor in question or hindered its ADC calculation. The ADC mean was significantly lower in grade III than in grade I or grade II ($p < 0.0001$ and $p = 0.04$) on average; no significant difference was found in the ADC mean between grades I and II on average ($p = 0.16$) on adjusted pairwise comparison. The ADC min was significantly lower in grade III than in grade I ($p = 0.02$) on average; on adjusted pairwise comparison, however, no significant difference was found in ADC min between grades I and II or grades III and II on average ($p = 0.17$ and $p = 0.45$) (Figs. 2 and 3). The ADC values were different among myxoid (ADC mean = 0.92 ± 0.30, ADC min = 0.67 ± 0.30) and non-myxoid (ADC mean = 1.57 ± 0.56, ADC min = 1.06 ± 0.52) tumors on an average with p values of 0.0006 and 0.0141, respectively.

Diagnostic performance

The largest dimension combined with central enhancement provided the highest AUC but was not significantly higher than

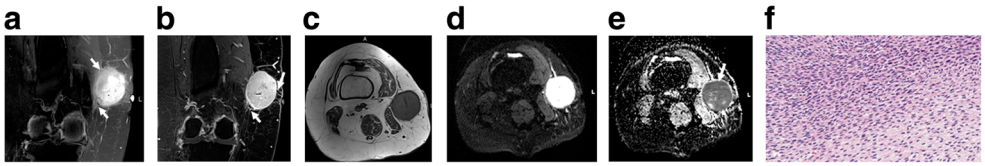

Fig. 2 Grade 2 tumor. **a** Coronal fat-suppressed T2W image shows a 5.1-cm heterogeneous thigh mass with hyperintense signal intensity and peritumoral edema (small arrows). **b** 3D fat-suppressed post-contrast coronal image shows central and minimal peritumoral enhancement (small arrows). **c** Axial T1W image shows homogeneous signal hyperintense to muscle without hemorrhage. **d**, **e** Axial DWI and ADC (arrow) images show mean and minimum ADC of 0.9 and 0.6, respectively. Diagnosis on histology was MPNST and FNCLCC grade II. **f** F18 FDG image shows high glucose uptake in the popliteal mass lesion and additional femoral lesion, a metastatic lesion

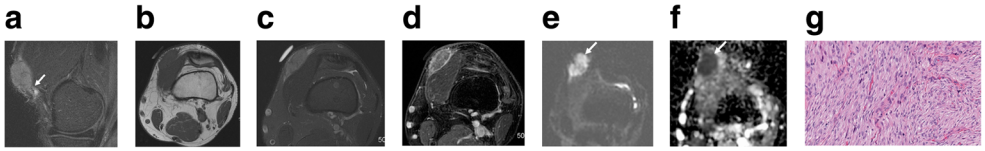

Fig. 3 Grade 3 tumor. **a**, **c** Sagittal and axial fat-suppressed T2W images show a 2.9-cm mildly heterogeneous thigh mass with hyperintense signal intensity and peritumoral edema (small arrow). **b** Axial T1W image shows homogeneous signal isointense to muscle without hemorrhage. **d** 3D fat-suppressed post-contrast axial image shows central and peripheral enhancement as well as minimal peritumoral enhancement. **e**, **f** Axial DWI and ADC images show mean and minimum ADC of 1.0 and 0.8, respectively. Diagnosis on histology was pleomorphic undifferentiated sarcoma, FNCLCC grade 3. **g** The tumor consists of pleomorphic spindled cells with brisk mitoses (H&E stain, × 200)

any of the univariate models with ADC min, ADC mean, or largest dimension or combined with central enhancement (Table 4). With equal weights on false positives and false negatives, the ADC mean cutoff was 0.88 for grade III versus grade I and II with a sensitivity of 62% and specificity of 80%. After excluding all myxoid and chondroid lesions, the cutoff for the ADC mean was 0.74 with a sensitivity of 48% and specificity of 85%. When differentiating grade I from higher grades (grade II and III), the ADC mean cutoff was 0.98 with a sensitivity of 56% and specificity of 88% (Fig. 4). After excluding all myxoid and chondroid lesions, a similar cutoff for the ADC mean at 0.98 showed a sensitivity of 66% and specificity of 75%.

Inter-reader analysis

The interobserver agreement between the four readers was fair to excellent based on the kappa values as illustrated in Table 5. The best agreements (good to excellent) were seen with respect to T1 and T2 tumor/muscle ratios, the largest tumor dimension, cystic areas, and ADC values (ICC, 0.60–0.98).

Discussion

The final diagnosis and grading of sarcoma is currently established by histology. This study supports that DWI can be helpful in the grading of musculoskeletal soft tissue malignancies and ADC values decrease with the increasing grades of these malignancies. The ADC mean was significantly lower in grade III than in grade I or grade II. On adjusted pairwise comparisons, however, no significant difference was found in the ADC mean between grades I and II ($p = 0.16$). Though there are likely histology differences not captured by diffusion metrics, with a larger sample, grade I and II tumors might be separated due to the expected larger power of the study. These results agree with those reported for other malignancies such as prostate cancer, glioma, and hepatocellular carcinoma, where the ADC has been shown to be of value in predicting the tumor grading [21–23]. Thus, DWI not only is useful for differentiating benign from malignant musculoskeletal lesions as previously shown in literature [16, 24, 25] but also provides an insight into the tumor grading. There was marginal (not statistically significant) superiority of the combination of conventional MR and contrast imaging over DWI. However, DWI can be performed without contrast imaging and its accuracy (AUC, 0.72) was fairly close to the combination of other MRI features (AUC, 0.74). We did not assess perfusion effects in DWI analysis as done by Rijswijk et al [26], which may further improve the AUC of ADC calculation, but it requires changing the protocol to many more *b* values and it would be a subject of a future study. Furthermore, DWI can add to the specificity of conventional MRI when contrast-enhanced examination is contraindicated, for instance in pregnant patients, very low GFR, and contrast allergy. DWI has already been shown to be useful for assessing treatment response in musculoskeletal bone tumors, with successfully treated tumors showing good response and higher minimum ADC values and ratios compared to the non-responsive tumors [27, 28]. Higher ADC values have been reported in the myxoid lesions and cystic lesions [29–31] as also seen in our study. The ADC numbers for myxoid and non-myxoid lesions compared very closely to those reported by Maeda et al [29].

Among the conventional MR imaging findings, the longitudinal tumor dimensions were different among the three

Table 4 Area under the ROC curve (AUC) using the Delong method (33)

Model	AUC	95% confidence interval		*p* value comparing to largest dimension + central enhancement
ADC min	0.6246	0.4652	0.784	0.2462
ADC mean	0.7262	0.5859	0.8664	0.8906
Largest dimension (cm)	0.7046	0.5594	0.8499	0.4411
ADC min + enh	0.66	0.5061	0.8139	0.3776
ADC mean + enh	0.7185	0.5761	0.8608	0.814
Largest dimension (cm) + enh	0.74	0.5977	0.8823	–

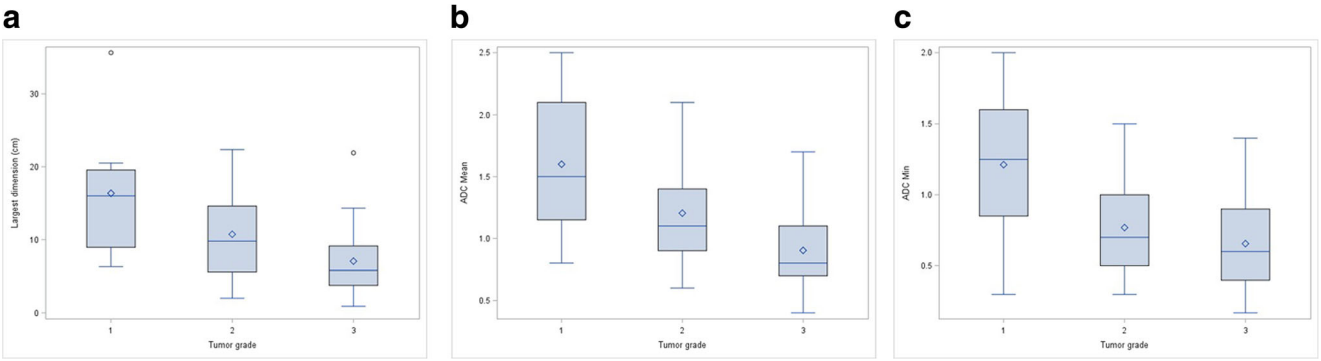

Fig. 4 Box plots showing the significant differences among different tumor grades with respect to the longitudinal largest dimension (**a**), ADC mean (**b**), and ADC minimum (**c**)

grades ($p = 0.0015$), with the larger lesions seen with grade I malignancy. On adjusted pairwise comparison, the difference again persisted among grade I vs grade III tumors. This finding disagrees with the previous report by Zhao et al [14] where high-grade tumors were larger than the low-grade tumors. It might be reflective of the slow growth and late presentation expected with such lower-grade lesions especially grade I liposarcomas. None of the other non-contrast quantitative or qualitative findings were, however, found to be different among the various grades of these tumors including peritumoral edema and enhancement ($p = 0.0985$, 0.2529). On contrast imaging, there was significantly more central enhancement in grade III tumors compared to other grades ($p = 0.0191$). This result has not been reported before. Gruber et al [13] reported the presence of heterogeneous enhancement in malignant tumors, and Zhao et al [14] observed a trend towards a higher percentage of tumor enhancement with increasing tumor grade.

Good to excellent agreement was seen with ADC values (ICC = 0.60–0.98) validating the inter-reader reliability among four different experience levels of readers as has also been shown in other studies using different methods of ADC measurements [16–19]. In future, DWI or contrast images can be used with artificial intelligence approaches to build automated predictive models for prospective tumor grading, which could differentiate intermediate-grade from high-grade tumors and aid in patient management [32].

The limitations included different sample sizes of various tumor grades; however, we included a consecutive series of cases with all possible grades of tumors reflecting our tertiary care practice. The tumor grades were also confirmed by an expert sarcoma pathologist for accuracy. Since soft tissue sarcoma occurs with less frequency, we could not assess a larger cohort of same histology and its different grades. The imaging was performed on different scanners; however, a uniform imaging protocol was used during the study period. As compared to the previous study [14], where heterogeneous protocols with and without fat suppression were assessed, in this study, all patients had T2W imaging with fat suppression and T1W imaging; therefore, the signal intensity ratios were assessed in all lesions. Finally, as per the study design and reflecting our current practice, the readers did not evaluate DWI in isolation without anatomic and contrast MR imaging; however, they were blinded to the final tumor histology and grading. DWI in the domain of musculoskeletal imaging is currently limited to a few centers in the country. In future, a larger prospective or multi-institutional study can be performed using similar MR platforms and same histologic types of sarcomas can be

Table 5 Inter-reader agreements with respect to conventional MR and DWI interpretations

Features	ICC/kappa	95% CI	
T1 tumor/muscle signal	0.88	0.79	0.93
T2 FS tumor/muscle signal ratio	0.85	0.76	0.91
ADC mean	0.74	0.58	0.85
ADC min	0.60	0.38	0.75
Largest dimension (cm)	0.98	0.96	0.99
T1 signal—hypointense, isointense, hyperintense (weighted kappa)	0.30	0.15	0.45
T1—homogeneous, heterogeneous	0.46	0.23	0.69
T2 signal—hypointense, isointense, hyperintense, mixed	0.43	0.24	0.61
T2—homogeneous, heterogeneous	0.56	0.19	0.93
Non-enhancing cystic/necrotic T2 bright areas—present, absent	0.65	0.44	0.87
Enhancement—peripheral, central	0.58	0.30	0.86

grouped with larger sample sizes for generation of more uniform data and better results can be generated with wider applicability.

To conclude, diffusion-weighted imaging is helpful in tumor grading of soft tissue malignancies with good to excellent inter-reader reliability. Longitudinal tumor size, deep tumor location, and central contrast enhancement are also useful conventional MR imaging markers of higher-grade soft tissue malignancies.

Funding The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Avneesh Chhabra, MD.

Conflict of interest The authors of this manuscript declare relationships with the following companies: AC: consultant, ICON Medical; royalties: Jaypee, Wolters. OA, CS, ND, HW, AC, RS, and YX: no disclosures.

The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry One of the authors (YX) has significant statistical expertise.

Informed consent Written informed consent was waived by the Institutional Review Board.

Ethical approval Institutional Review Board approval was obtained.

Methodology
- Retrospective
- Cross sectional
- Performed at one institution

References

1. Murphey MD, Kransdorf MJ, Smith SE (1999) Imaging of soft tissue neoplasms in the adult: malignant tumors. Semin Musculoskelet Radiol 3:39–58. https://doi.org/10.1055/s-2008-1080050
2. Kransdorf MJ, Bancroft LW, Peterson JJ, Murphey MD, Foster WC, Temple HT (2002) Imaging of fatty tumors: distinction of lipoma and well-differentiated liposarcoma. Radiology 224:99–104. https://doi.org/10.1148/radiol.2241011113
3. Murphey MD, Gibson MS, Jennings BT, Crespo-Rodríguez AM, Fanburg-Smith J, Gajewski DA (2006) From the archives of the AFIP: imaging of synovial sarcoma with radiologic-pathologic correlation. Radiographics 26:1543–1565. https://doi.org/10.1148/rg.265065084
4. Walker L, Thompson D, Easton D et al (2006) A prospective study of neurofibromatosis type 1 cancer incidence in the UK. Br J Cancer 95:233–238. https://doi.org/10.1038/sj.bjc.6603227
5. Coran A, Ortolan P, Attar S et al (2017) Magnetic resonance imaging assessment of lipomatous soft-tissue tumors. In Vivo 31:387–395
6. Guillou L, Coindre JM, Bonichon F et al (1997) Comparative study of the National Cancer Institute and French Federation of Cancer Centers Sarcoma Group grading systems in a population of 410 adult patients with soft tissue sarcoma. J Clin Oncol 15:350–362. https://doi.org/10.1200/JCO.1997.15.1.350
7. Oliveira AM, Nascimento AG (2001) Grading in soft tissue tumors: principles and problems. Skeletal Radiol 30:543–559. https://doi.org/10.1007/s002560100408
8. Gielen JL, De Schepper AM, Vanhoenacker F et al (2004) Accuracy of MRI in characterization of soft tissue tumors and tumor-like lesions. A prospective study in 548 patients. Eur Radiol 14:2320–2330. https://doi.org/10.1007/s00330-004-2431-0
9. Berquist TH, Ehman RL, King BF, Hodgman CG, Ilstrup DM (1990) Value of MR imaging in differentiating benign from malignant soft-tissue masses: study of 95 lesions. AJR Am J Roentgenol 155:1251–1255. https://doi.org/10.2214/ajr.155.6.2122675
10. Vanhoenacker FM, Van Looveren K, Trap K et al (2012) Grading and characterization of soft tissue tumors on magnetic resonance imaging: the value of an expert second opinion report. Insights Imaging 3:131–138. https://doi.org/10.1007/s13244-012-0151-6
11. Mesko NW, Wilson RJ, Lawrenz JM et al (2018) Pre-operative evaluation prior to soft tissue sarcoma excision—why can't we get it right? Eur J Surg Oncol 44:243–250. https://doi.org/10.1016/j.ejso.2017.11.001
12. Dyrop HB, Vedsted P, Rædkjær M, Safwat A, Keller J (2017) Imaging investigations before referral to a sarcoma center delay the final diagnosis of musculoskeletal sarcoma. Acta Orthop 88:211–216. https://doi.org/10.1080/17453674.2016.1278113
13. Gruber L, Gruber H, Luger AK, Glodny B, Henninger B, Loizides A (2017) Diagnostic hierarchy of radiological features in soft tissue tumours and proposition of a simple diagnostic algorithm to estimate malignant potential of an unknown mass. Eur J Radiol 95:102–110. https://doi.org/10.1016/j.ejrad.2017.07.020
14. Zhao F, Ahlawat S, Farahani SJ et al (2014) Can MR imaging be used to predict tumor grade in soft-tissue sarcoma? Radiology 272:192–201. https://doi.org/10.1148/radiol.14131871
15. Jahed K, Khazai B, Umpierrez M, Subhawong TK, Singer AD (2018) Pitfalls in soft tissue sarcoma imaging: chronic expanding hematomas. Skeletal Radiol 47:119–124. https://doi.org/10.1007/s00256-017-2770-y
16. Ahlawat S, Khandheria P, Subhawong TK, Fayad LM (2015) Differentiation of benign and malignant skeletal lesions with quantitative diffusion weighted MRI at 3T. Eur J Radiol 84:1091–1097. https://doi.org/10.1016/j.ejrad.2015.02.019
17. Singer AD, Pattany PM, Fayad LM, Tresley J, Subhawong TK (2016) Volumetric segmentation of ADC maps and utility of standard deviation as measure of tumor heterogeneity in soft tissue tumors. Clin Imaging 40:386–391. https://doi.org/10.1016/j.clinimag.2015.11.017
18. Sagiyama K, Watanabe Y, Kamei R et al (2017) Multiparametric voxel-based analyses of standardized uptake values and apparent diffusion coefficients of soft-tissue tumours with a positron emission tomography/magnetic resonance system: preliminary results. Eur Radiol 27:5024–5033. https://doi.org/10.1007/s00330-017-4912-y
19. Ahlawat S, Khandheria P, Del Grande F et al (2016) Interobserver variability of selective region-of-interest measurement protocols for quantitative diffusion weighted imaging in soft tissue masses: comparison with whole tumor volume measurements. J Magn Reson Imaging 43:446–454. https://doi.org/10.1002/jmri.24994
20. Fisher S, Wadhwa V, Manthuruthil C, Cheng J, Chhabra A (2016) Clinical impact of magnetic resonance neurography in patients with brachial plexus neuropathies. Br J Radiol 89:20160503. https://doi.org/10.1259/bjr.20160503
21. Barbieri S, Brönnimann M, Boxler S, Vermathen P, Thoeny HC (2017) Differentiation of prostate cancer lesions with high and with

low Gleason score by diffusion-weighted MRI. Eur Radiol 27: 1547–1555. https://doi.org/10.1007/s00330-016-4449-5

22. Li X, Zhang K, Shi Y, Wang F, Meng X (2016) Correlations between the minimum and mean apparent diffusion coefficient values of hepatocellular carcinoma and tumor grade. J Magn Reson Imaging 44:1442–1447. https://doi.org/10.1002/jmri.25323
23. Kang Y, Choi SH, Kim YJ et al (2011) Gliomas: histogram analysis of apparent diffusion coefficient maps with standard- or high-b-value diffusion-weighted MR imaging—correlation with tumor grade. Radiology 261:882–890. https://doi.org/10.1148/radiol.11110686
24. Teixeira PA, Gay F, Chen B et al (2016) Diffusion-weighted magnetic resonance imaging for the initial characterization of non-fatty soft tissue tumors: correlation between T2 signal intensity and ADC values. Skeletal Radiol 45:263–271. https://doi.org/10.1007/s00256-015-2302-6
25. Pekcevik Y, Kahya MO, Kaya A (2015) Characterization of soft tissue tumors by diffusion-weighted imaging. Iran J Radiol 12: e15478. https://doi.org/10.5812/iranjradiol.15478v2
26. van Rijswijk CS, Kunz P, Hogendoorn PC, Taminiau AH, Doornbos J, Bloem JL (2002) Diffusion-weighted MRI in the characterization of soft-tissue tumors. J Magn Reson Imaging 15:302–307
27. Degnan AJ, Chung CY, Shah AJ (2018) Quantitative diffusion-weighted magnetic resonance imaging assessment of chemotherapy treatment response of pediatric osteosarcoma and Ewing sarcoma malignant bone tumors. Clin Imaging 47:9–13. https://doi.org/10.1016/j.clinimag.2017.08.003
28. Oka K, Yakushiji T, Sato H, Hirai T, Yamashita Y, Mizuta H (2010) The value of diffusion-weighted imaging for monitoring the chemotherapeutic response of osteosarcoma: a comparison between average apparent diffusion coefficient and minimum apparent diffusion coefficient. Skeletal Radiol 39:141–146. https://doi.org/10.1007/s00256-009-0830-7
29. Maeda M, Matsumine A, Kato H et al (2007) Soft-tissue tumors evaluated by line-scan diffusion-weighted imaging: influence of myxoid matrix on the apparent diffusion coefficient. J Magn Reson Imaging 25:1199–1204. https://doi.org/10.1002/jmri.20931
30. Subhawong TK, Durand DJ, Thawait GK, Jacobs MA, Fayad LM (2013) Characterization of soft tissue masses: can quantitative diffusion weighted imaging reliably distinguish cysts from solid masses? Skeletal Radiol 42:1583–1592. https://doi.org/10.1007/s00256-013-1703-7
31. Del Grande F, Ahlawat S, Subhawong T, Fayad LM (2017) Characterization of indeterminate soft tissue masses referred for biopsy: what is the added value of contrast imaging at 3.0 tesla? J Magn Reson Imaging 45:390–400
32. Corino VDA, Montin E, Messina A et al (2018) Radiomic analysis of soft tissues sarcomas can distinguish intermediate from high-grade lesions. J Magn Reson Imaging 47:829–840. https://doi.org/10.1002/jmri.25791