



Machine Learning in the Differentiation of Soft Tissue Neoplasms: Comparison of Fat-Suppressed T2WI and Apparent Diffusion Coefficient (ADC) Features-Based Models

Peian Hu¹ · Lei Chen² · Zhengrong Zhou^{2,3,4}

Received: 12 May 2021 / Revised: 18 August 2021 / Accepted: 22 August 2021 / Published online: 20 September 2021
© The Author(s) 2021

Abstract

Machine learning has been widely used in the characterization of tumors recently. This article aims to explore the feasibility of the whole tumor fat-suppressed (FS) T2WI and ADC features-based least absolute shrinkage and selection operator (LASSO)-logistic predictive models in the differentiation of soft tissue neoplasms (STN). The clinical and MR findings of 160 cases with 161 histologically proven STN were reviewed, retrospectively, 75 with diffusion-weighted imaging (DWI with b values of 50, 400, and 800 s/mm²). They were divided into benign and malignant groups and further divided into training (70%) and validation (30%) cohorts. The MR FS T2WI and ADC features-based LASSO-logistic models were built and compared. The AUC of the FS T2WI features-based LASSO-logistic regression model for benign and malignant prediction was 0.65 and 0.75 for the training and validation cohorts. The model's sensitivity, specificity, and accuracy of the validation cohort were 55%, 96%, and 76.6%. While the AUC of the ADC features-based model was 0.932 and 0.955 for the training and validation cohorts. The model's sensitivity, specificity, and accuracy were 83.3%, 100%, and 91.7%. The performances of these models were also validated by decision curve analysis (DCA). The AUC of the whole tumor ADC features-based LASSO-logistic regression predictive model was larger than that of FS T2WI features ($p = 0.017$). The whole tumor fat-suppressed T2WI and ADC features-based LASSO-logistic predictive models both can serve as useful tools in the differentiation of STN. ADC features-based LASSO-logistic regression predictive model did better than that of FS T2WI features.

Keywords Soft tissue neoplasms (STN) · Least absolute shrinkage and selection operator (LASSO) · Texture analysis (TA) · Apparent diffusion coefficient (ADC) · Diffusion MR weighted imaging

✉ Zhengrong Zhou
zhouzr_16@126.com

Peian Hu
pahu16@fudan.edu.cn

Lei Chen
mac_chenlei@126.com

¹ Department of Radiology, Children's Hospital of Fudan University, National Children's Medical Center, No.399, Wanyuan Road, Shanghai 201102, China

² Department of Radiology, Minhang Branch, Fudan University Shanghai Cancer Center, No. 106, Ruiji Road, Shanghai 201100, China

³ Department of Radiology, Fudan University Shanghai Cancer Center, Fudan University, No. 270, Dongan Road, Shanghai 200032, China

⁴ Department of Oncology, Shanghai Medical College, Fudan University, No. 270, Dongan Road, Shanghai 200032, China

Introduction

Soft tissue neoplasms (STN), a group of heterogeneous tumors, are derived from blood vessels, lymphatic vessels, nerves, muscles, or other connective tissue [1]. STNs are commonly seen with complicated components and classified as benign, intermediate (metastatic or recurrent occasionally), and malignant subtypes by the WHO [1]. Except for a few tumors with characteristic imaging features, a definite histological diagnosis is usually challenging on imaging. A better prognosis can be achieved for most benign and intermediate STN. Soft tissue sarcoma (STS) represents about 1% of all malignancy; it recurs and metastasizes commonly with a poor prognosis [2].

MR imaging is the preferred method for detecting and staging of STN [3–6]. Conventional MR assessment of STN mainly focused on the morphologic findings, such as the tumor's size ($>$ or ≤ 5 cm), contour (round or lobulated),

margins (well- or ill-defined), heterogeneity of masses, and involvement of adjacent vital structures (bone/neurovascular bundle) [3–5, 7, 8]. Several studies were designed to explore the effectiveness of conventional MR in the differentiation of STN. The reported diagnostic accuracy ranged from 50 to 90% [3, 5, 6, 8, 9]. An overlap of the radiological features between benign and malignant tumors was frequently seen. Gadolinium (Gd)-based enhanced MR scan helped differentiate cystic from solid masses [10]. Additionally, the knowledge of prevalence and presentation of onset can serve as a supplement of morphological features in the differentiation of STN [3].

Surgical excision was the first-choice treatment for STN. Although the role of chemotherapy was controversial [11], a few subtypes of sarcomas were sensitive to chemotherapy, such as rhabdomyosarcoma (embryonal and alveolar subtypes), Ewing sarcoma family of tumors, round cell liposarcoma, desmoplastic small round cell tumor, and synovial sarcoma [11].

Diffusion magnetic resonance weighted imaging (DWI) based on the Brownian motion of water molecules can reflect the tissue microstructures [12]. The apparent diffusion coefficient (ADC) is a widely used quantitative parameter. Low ADC values mean highly cellular density and/or restricted microenvironments, while acellular regions are found with elevated ADC values [12–15]. Muscular sarcomas were reported with a broad range of ADC values [16]. Some researchers thought that ADC value was a reliable quantitative parameter in the differentiation of STN [13, 14, 17].

Texture analysis (TA) is a method to evaluate the tumor by extracting and using features that were invisible to the naked eye. Texture analysis was employed to differentiate tumors or tumors with different grades be employed to differentiate tumors or tumors with different grades [18–20] but scarcely did they focus on the application of TA based on FS T2WI and ADC mapping in the differentiation of STN. Machine learning, as the intersection of statistics and computer science, has been gradually applied in the medical field recently [21]. It mainly focused on how computers learn from big data and included many algorithmic models, such as the least absolute shrinkage and selection operator (LASSO), support vector machine (SVM), random forest, and decision tree [22–24]. LASSO was commonly used and robust. It overcame the shortcomings of multiple regression in high-dimensional data and was beneficial in feature selection [23–25].

We supposed that the TA of the whole tumor FS T2WI and ADC features-based LASSO-logistic regression predictive models can be used in the characterization of STN precisely. An then to assess the effectiveness of these two models in the characterization of STN, we retrospectively collected and reviewed the clinical and imaging findings of 160 patients with 161 histologically proven STN (75 of them with DWI).

Methods

Study Population

This retrospective study was approved by our institutional review board, and informed consent was waived. Between July 1, 2015, and December 31, 2015, the imaging features and clinical findings of patients with suspected soft tissue neoplasms were collected and reviewed retrospectively. The inclusion criteria were as follows: STN were all histologically proven (surgery or biopsy), and all the patients underwent an MR scan. The suspected STN that were not histologically proven or without MR scans were excluded.

At last, 160 cases (161 histologically proven masses) with MR scans were collected and reviewed, and 75 of them with diffusion-weighted imaging (DWI, with b values 50, 400, and 800 s/mm²). The 38 soft tissue sarcoma (STS) cases with DWI were divided three times, into the chemosensitive and non-chemosensitive groups [11]; the small round cell and non-small round cell sarcoma groups; and the rhabdomyosarcoma and non-rhabdomyosarcoma groups.

Demographic and Clinical Data

The demographic and clinical data were reviewed, including the age of onset, gender, main manifestations, tumor locations, and histological results. The locations were recorded as the head and neck, trunk, retroperitoneum, and extremities, respectively.

Imaging Acquisition

All the patients underwent conventional MR and/or DWI (with b values of 50, 400, and 800 s/mm²). Axial FS T2WI imaging and/or ADC mapping was used for whole tumor 3D volume segmentation and feature extraction (Figs. 1–3):

- The scanned FS T2WI parameters: TR 3,500–4,000 ms, TE 100–110 ms, ETL 15, matrix 512×512, the number of excitation 2, the slice thickness 5 mm, the gap of slice 1 mm, and FOV 250–350 mm
- T1WI: axial FSE/TSE sequences, TR 410–500 ms, TE 15 ms, matrix 512×512, the number of excitation 2, slice thickness 5 mm, and the gap of slice 1 mm
- T2WI: coronal or sagittal TSE/FSE, TR 3,500–4,000 ms, TE 100–110 ms, the number of excitation 2, the slice thickness 5 mm, and the gap of slice 1 mm

DWI was performed before enhanced T1WI. DWI was acquired by using the single-shot echo-planar imaging pulse (SS-SE-EPI)-DWI sequence in free breathing with parallel

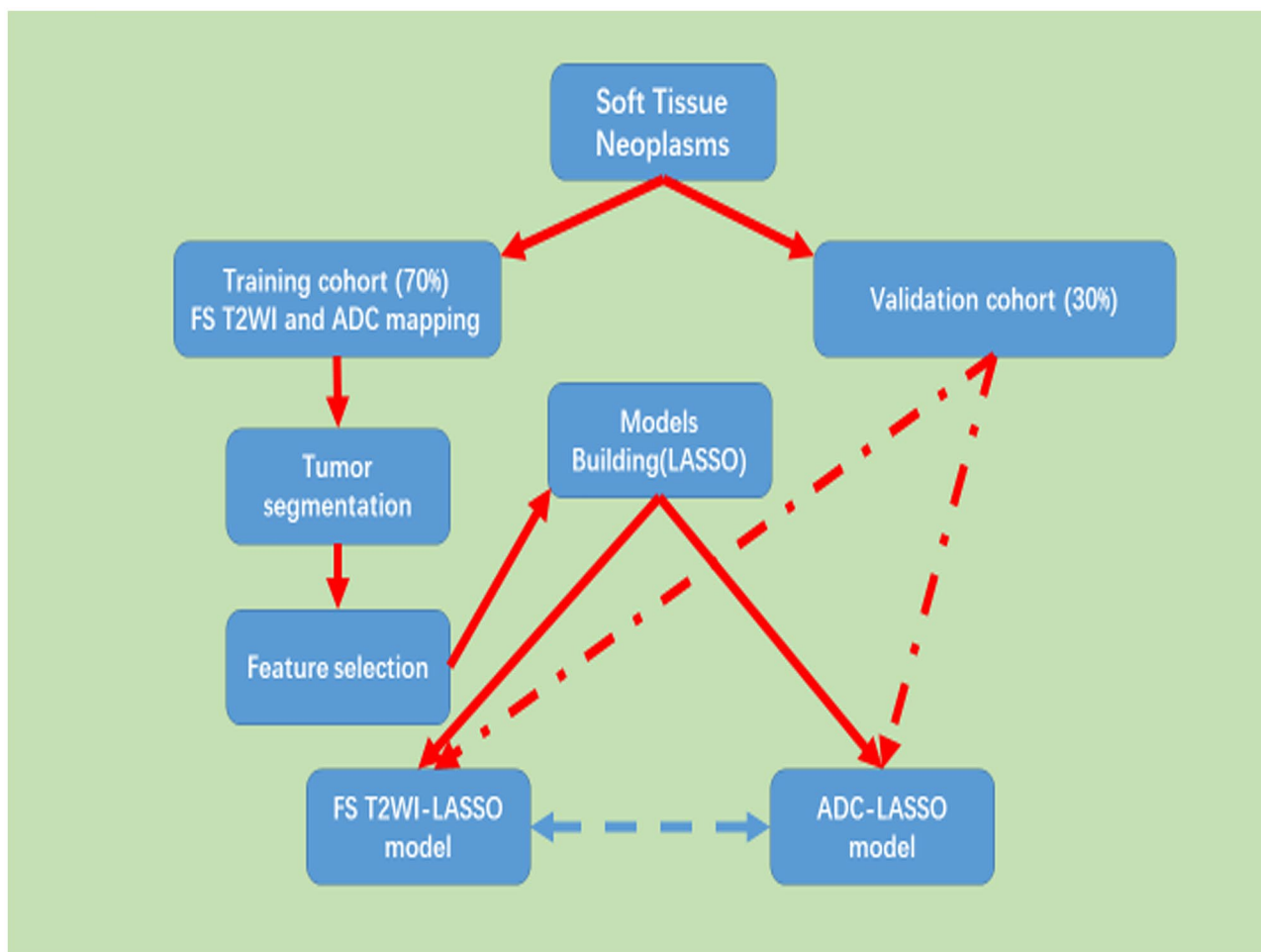


Fig. 1 Flowchart of radiomics and machine learning

imaging, with b values of 50, 400, and 800 s/mm². Other scanning parameter was the same as that described above. The ADC mapping was generated using the mono-exponential decay mode.

Subsequently, all patients underwent enhanced T1-weighted imaging after the intravenous injection of 0.1 mmol/kg contrast medium (Magnevist, Bayer Schering Pharma, Berlin, Germany) at a flow rate of 2–3 ml/s.

Tumor Segmentation and the Extraction of FS T2WI and ADC Features

LIFEx v4.00 software (<https://www.lifexsoft.org/>) was employed for tumor segmentation and feature extraction.

Tumor segmentation was done by a radiologist with 12 years of experience on MR interpretations of STN (Figs. 2 and 3). Conventional MR images were referred to during selection of the region of interest (ROI).

The ROIs were manually selected using LIFEx v4.00 software, to cover the whole tumor. The steps of texture

feature extraction were as follows: ROI selection (3D model), spatial resampling (1 mm × 1 mm × 1 mm), intensity discretization (number of Gray-level, 64), and intensity rescaling (relative, mean ± 3SD). The ROIs were measured twice at a 1-year interval.

These features were extracted (Table 1), including first-order features (the shape (Sphericity, Compacity, Volume_ml, and Volume_vx), histogram (HISTO_Skewness, HISTO_Kurtosis, HISTO_Entropy_log10, HISTO_Entropy_log2, and HISTO_Energy) (Figs. 2b and 3a), and second-order features (GLCM (GLCM, Gray-level co-occurrence matrix) (GLCM_Homogeneity, GLCM_Energy, GLCM_Contrast, GLCM_Correlation, GLCM_Entropy_log10, GLCM_Entropy_log2, and GLCM_Dissimilarity), GLRLM (GLRLM, Gray-level run length matrix) (GLRLM_SRE, GLRLM_LRE, GLRLM_LGRE, GLRLM_HGRE, GLRLM_SRLGE, GLRLM_SRHGE, GLRLM_LRLGE, GLRLM_LRHGE, GLRLM_GLNU, GLRLM_RLNU, and GLRLM_RP), NGLDM (NGLDM, Neighboring Gray-level dependence matrix) (NGLDM_Coarseness, NGLDM_Contrast, and NGLDM_Busyness), and

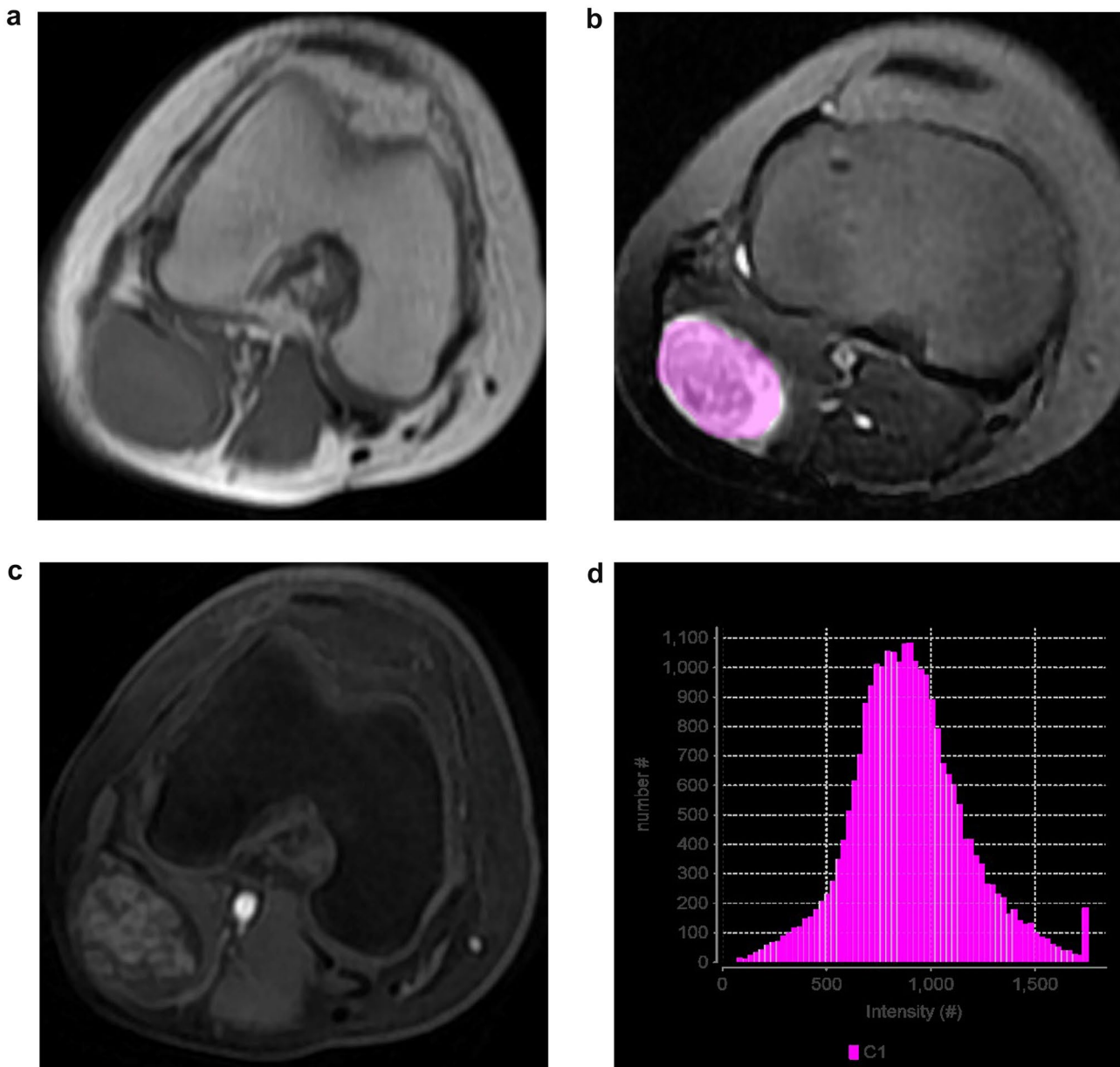


Fig. 2 A 19-year-old female with a schwannoma in the right popliteal fossa. **a–c** showed iso-signal intensity on MR T1WI (**a**), heterogeneous hyper-signal intensity (SI) on MR FS (fat-suppressed) T2WI (**b**),

and moderate heterogeneous enhancement (**c**). **d** the FS T2WI SI distribution of the whole tumor

GLZLM (GLZLM, Gray-level zone length matrix) (GLZLM_SZE, GLZLM_LZE, GLZLM_LGZE, GLZLM_HGZE, GLZLM_SZLGE, GLZLM_SZHGE, GLZLM_LZLGE, GLZLM_LZHGE, GLZLM_GLNU, GLZLM_ZLNU, and GLZLM_ZP)).

The Construction and Validation of the Predictive Model

These cases were randomly divided into training (70%) and validation (30%) cohorts. The texture features of the training

cohort were used for constructing the predictive model, and the features of the validation cohort were used for validation.

The inter-observer correlation coefficient (ICC) was used to evaluate the repeatability of these features. In order to handle high-dimensional data better and select features, the LASSO algorithm was employed. LASSO-logistic regression with tenfold cross-validation and 1 standard error rule was used to reduce data dimensions, select features, and build a predictive model. The receiver operating characteristic (ROC) and DCA were used to validate the effectiveness of the model.

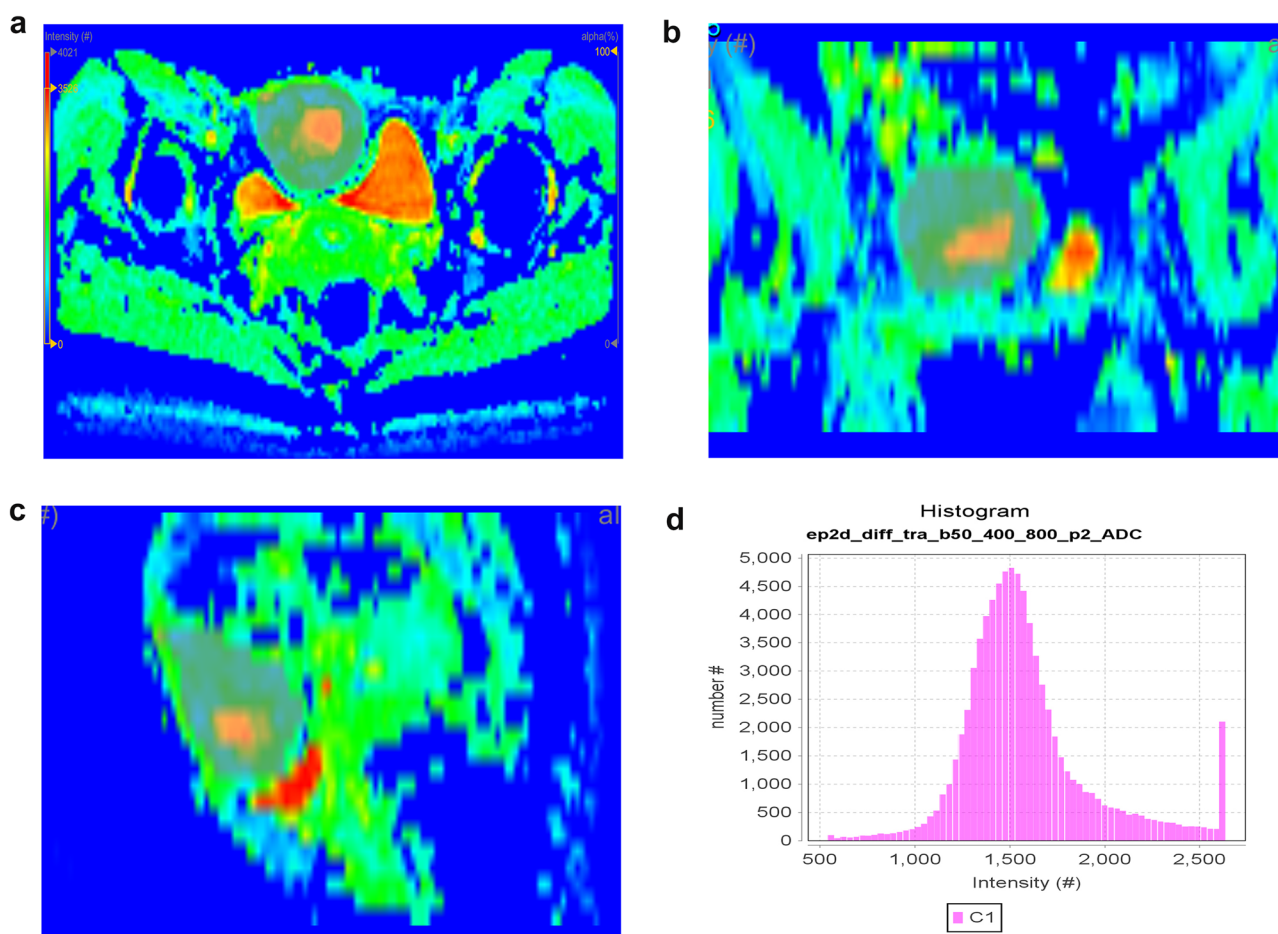


Fig. 3 A 29-year-old female with fibrosarcoma in the anterior abdominal wall showed 3D manually selected ROI on color ADC map (a–c) and histogram distribution of whole tumor ADC values (d)

Table 1 The histological type and numbers (cases) of two groups

Group	Histological type and number
Benign	Hemangioma 22, schwannoma 19, desmoid tumor 8, lymphangioma 5, benign soft tissue neoplasms 5, inflammatory myofibroblastic tumor 5, kaposiform hemangioendothelioma 3, intermediate tumor 3, lipoblastoma 2, leiomyoma 2, nodular fasciitis 2, giant cell tumor of tendon sheath 2, angiomyxoma 1, dermatofibrosarcoma protuberans 1, fibroma 1, myxoma 1, hemangiopericytoma 1, fibrous hamartoma 1
Malignant	Rhabdomyosarcoma 23, malignant fibroblast/myofibroblast origin tumor 12, leiomyosarcoma 9, malignant mesenchymal tumor with uncertain origin 7, GIST 7, MPNST 4, liposarcoma 4, malignant solitary fibrous tumor 3, pleomorphic undifferentiated sarcoma 2, PNET 1, malignant rhabdomyoid tumor 1, malignant hemangiopericytoma 1, mesenchymal chondrosarcoma 1, clear cell sarcoma 1, malignant mucinous tumor 1

Statistical Analysis

The R (version 3.6.0, <https://www.r-project.org/>), SPSS 20.0, and MedCalc statistical software were employed for data analysis. Kolmogorov–Smirnov test was employed for testing normal distribution. Independent student’s *t* test was employed to analyze the differences in texture features. ROC curves were generated to determine the cut-off values. The AUCs were calculated and further compared by the Delong test. The DCA was done by R software. The glmnet and pROC packages of R software were employed. Values of *p* < 0.05 were considered statistically significant.

Results

Demographic and Clinical Data

There were 84 masses in the benign and 77 in the malignant group (Table 1). And there were 37 benign and 38 malignant STNs with DWI.

The gender ratio (female:male) was 77:83. The ages ranged from 1 month to 82 years old, and the median age was 29.5 years old. Thirty-three were in the head and neck region, 93 arise in the trunk (7 in retroperitoneal space), and 35 arisen in the extremity (21 in the lower, 14 in the upper).

In the malignant group, metastases were found in 23.4% (18/77) cases; 13 pulmonary metastasis, 3 liver metastasis, 3 lymph nodule metastasis, and 3 intraperitoneal dissemination were found.

There were 38 cases with STS that underwent DWI; 17 chemosensitive and 21 non-chemosensitive sarcomas, 13 small and 25 non-small round cell sarcomas, and 17 rhabdomyosarcomas and 21 non-rhabdomyosarcomas were enrolled.

Most of them complained of enlarging, pain, or painless masses. The other manifestations were the Kasabach–Merritt phenomenon (KMP), proteinuria (1 case), and yellowish skin.

The Differences of MR FS T2WI and ADC Features Between Benign and Malignant Groups

The ICC of texture features ranged from 0.81 to 0.94, showing good repeatability.

There were 14 MR FS T2WI features with significant differences between benign and malignant STN ($p < 0.05$) (Table 2). And there were 12 features between benign and malignant tumors ($p < 0.05$) with significant differences, including mean ADC, max ADC, STD value, and HISTO-skewness values (Table 2).

The mean ADC value ($1668.93 \pm 406.34 \mu\text{m}^2/\text{s}$), max ADC value ($3167.38 \pm 711.68 \mu\text{m}^2/\text{s}$), and STD value (374.54 ± 110.21) of benign tumors were higher than those of malignant ones ($1173 \pm 289.65 \mu\text{m}^2/\text{s}$, $2746.72 \pm 823.17 \mu\text{m}^2/\text{s}$, 293.32 ± 102.60 , $p \leq 0.021$). Compared with benign tumors (-0.16 ± 0.58), malignant tumors had a higher HISTO-skewness value (0.52 ± 0.60 , $p = 0.000$).

The cut-off values of the mean ADC value, max ADC value, STD value, and HISTO-skewness value for

differentiation were $1388.55 \times \mu\text{m}^2/\text{s}$, $2568.88 \times \mu\text{m}^2/\text{s}$, $309.02 \times \mu\text{m}^2/\text{s}$, and 0.29, respectively. The AUC of the mean ADC value (0.868, 95% CI, 0.771–0.935) was bigger than that of max ADC value (0.653, 95% CI, 0.534–0.759) and STD value (0.717, 95% CI, 0.601–0.815) ($p < 0.0001$, $p = 0.0026$). There were no significant differences between the AUC of mean ADC value and that of HISTO-skewness (0.791, 95% CI, 0.682–0.876) ($p = 0.1624$).

The features between chemosensitive and non-chemosensitive sarcomas, between small round and non-small round cell sarcomas, and between rhabdomyosarcomas and non-rhabdomyosarcomas were not significantly different ($p > 0.05$).

The Construction and Validation of FS T2WI and ADC Features-Based Predictive Models

LASSO algorithm with tenfold cross-validation was employed for reducing data dimensions and feature selection.

The whole tumor 3D MR FS T2WI features of the training cohort (114 cases) were used to build predictive models. The deviance of classification was minimized when the λ (lambda) was 0.134 (Fig. 4). And only one feature, GLZLM_ZP, was selected. The LASSO-logistic regression predictive model was built and the linear regression equation was $Y_{benign/malignant} = -0.0713 - 0.2472 \times (GLZLM_ZP)$.

The AUC of the ROC curve was 0.65 for the training cohort. The AUC of the ROC curve was 0.75 for the validation cohort (Fig. 5a), and the sensitivity, specificity, and accuracy were 55%, 96%, and 76.6%, respectively.

The deviance of classification was minimized when the lambda (λ) was 0.038 (Fig. 4). Meanwhile, six features (mean ADC value, HISTO-skewness value, GLCM_Contrast, GLCM_Energy, GLRLM_LRHGE, and GLZLM_ZP) were selected. The LASSO-logistic regression model was constructed, and the regression equation was $Y_{benign/malignant} = -0.0615 - 1.9112 \times (\text{mean ADC value}) + 0.0959 \times (\text{HISTO_Skewness value}) + 0.0534 \times (\text{GLCM_Energy}) - 0.0811 \times (\text{GLCM_Contrast}) + 0.7319 \times (\text{GLRLM_LRHGE}) - 0.151 \times (\text{GLZLM_ZP})$.

Table 2 Differences of MR FS T2WI and ADC features between benign and malignant groups

Sequences	Features ($p < 0.05$)	
MR FS T2WI	First-order	SHAPE_VolumemL, SHAPE_Volume#vx
	Second-order	GLCM_Homogeneity, GLCM_Contrast, GLCM_Correlation, GLCM_Dissimilarity, GLRLM_GLNU, GLRLM_RLNU, GLRLM_RP, NGLDM_Contrast, NGLDM_Busyness, GLZLM_SZE, GLZLM_SZHGE, GLZLM_ZP
ADC	First-order	Mean value, max value, STD value, HISTO_Skewness
	Second-order	GLCM_Energy, GLCM_Contrast, GLCM_Dissimilarity, GLRLM_LRE, GLRLM_LRHGE, NGLDM_Contrast, GLZLM_HGZE, GLZLM_ZP

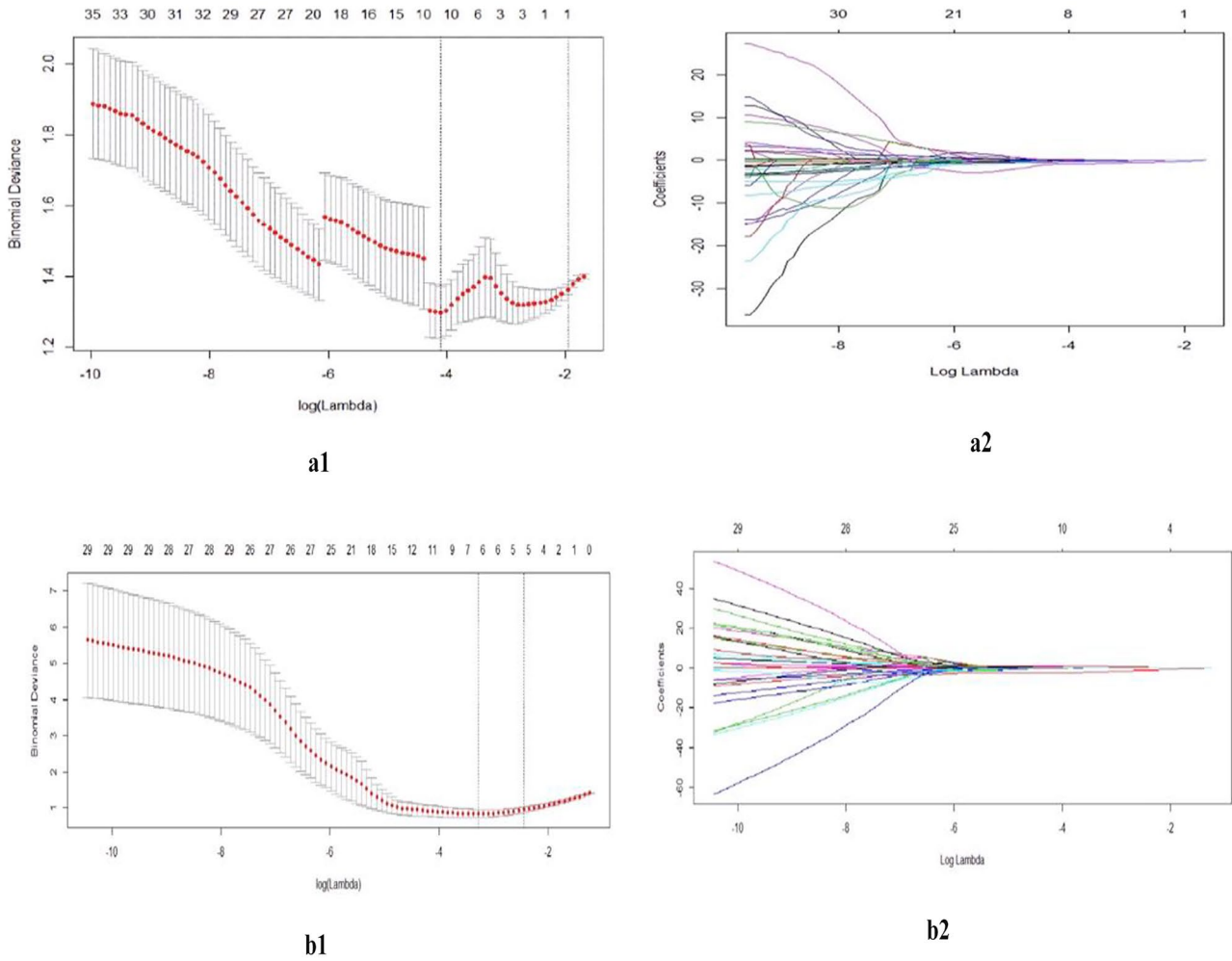
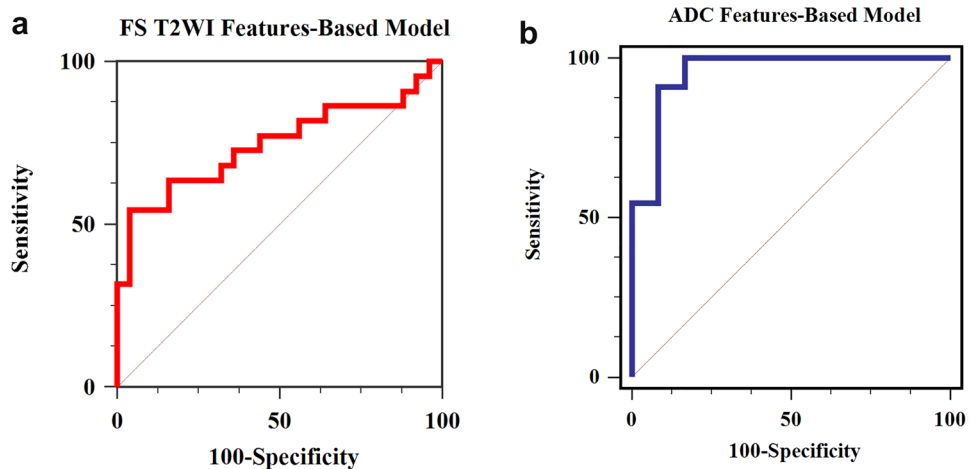


Fig. 4 Feature selection using the LASSO-logistic algorithm using tenfold cross-validation and 1 standard error rule. The optimal tuning value (**a1, b1**) was selected for benign and malignant STN prediction and (**a2, b2**) the corresponding features

The AUC was 0.932 for the training set. The AUC was 0.955 for the validation set (Fig. 5b), and the sensitivity, specificity, and accuracy were 83.3%, 100%, and 91.7% respectively.

The effectiveness of the predicted models was also validated by DCA (Fig. 6). DCA of FS T2WI and ADC features-based predictive models showed that these two

Fig. 5 The ROC curve of benign and malignant STN differentiation in validation cohorts, the ADC features-based LASSO-logistic regression predictive model did better than that of the FS T2WI features-based model ($z=2.386$, $p=0.017$). **a** FS T2WI features-based predictive model (AUC, 0.75) **b** ADC features-based predictive model (AUC, 0.955)



models provided greater benefit for benign and malignant characterization.

The Comparison of FS T2WI and ADC Features-Based Predictive Models

The ROCs of validation cohorts were used for the comparison of FS T2WI and ADC features-based predictive models. The ADC features-based LASSO-logistic regression predictive model did better than that of the FS T2WI in the differentiation of STN ($z=2.386$, $p=0.017$).

Discussion

The whole tumor ADC value was not helpful in the differentiation of chemosensitive and non-chemosensitive sarcomas, small round and non-small round sarcomas, or

rhabdomyosarcomas and non-rhabdomyosarcomas. The mean ADC value did better than max ADC value and STD value in the differentiation of STN. The HISTO-skewness value can be served as another useful feature in the differentiation. Machine learning of the whole tumor FS T2WI and ADC values did facilitate the differentiation of benign and malignant STN. And ADC features-based LASSO-logistic regression predictive model did better than that of FS T2WI features.

Texture analysis by extracting indiscernible radiomic features was useful in analyzing tumor heterogeneity. The utilization of images can be maximized without adding scan sequences [20]. Corino VDA et al. found that MR radiomic features can be used to distinguish intermediate soft tissue sarcomas from high-grade ones accurately [26]. The accuracy and AUC were 0.90 and 0.85 and 0.88 and 0.87 for the validation and test sets. Although we found the FS T2WI features-based model with high specificity (96%), the sensitivity was low (55%). The ADC features-based model can achieve high effectiveness. The sensitivity, specificity, and accuracy were 83.3%, 100%, and 91.7%, respectively. And ADC features-based LASSO-logistic regression predictive model did better than that of FS T2WI features.

ADC value was affected by ROI position and selected b values. We selected the whole tumor as ROI to avoid the selected bias. The quantitative parameter we measured showed good repeatability. Similar to literature [14], we chose three b values (50, 400, and 800 s/mm^2). For $b=50 s/mm^2$, it was less affected by microvascular perfusion than $b=0 s/mm^2$, and the selection of 800 s/mm^2 was to ensure enough signal-to-noise ratio (SNR).

DWI reflecting water molecule diffusion is useful in the detection and differentiation of tumors and facilitates the therapeutic assessment [13, 27–32]. Some benign STN resembled malignant ones on conventional MR sequences and were usually misdiagnosed [31, 33]. Most researchers thought the mean ADC and minimal ADC values help in the differentiation of STN [34, 35]. The mean ADC value of volumetric quantification had a high inter-observer agreement and reflected tumors' heterogeneity [36]. Although Van Rijswijk CSP et al. (37) harbored different opinions, they thought that malignant ones had significantly lower true diffusion coefficients. We found the mean ADC and HISTO-skewness values were valuable in the characterization of STN and did better than minimal ADC values. And it was tested by the LASSO-logistic model. The HISTO-skewness value can be served as another useful feature in differentiation, which was not mentioned previously. Benign STN often exhibited a negatively skewed distribution due to their low cell density and large extracellular space, while the malignant ones showed a positively skewed distribution. TA of ADC mapping can acquire more quantitative or semi-quantitative features for the differentiation of STN.

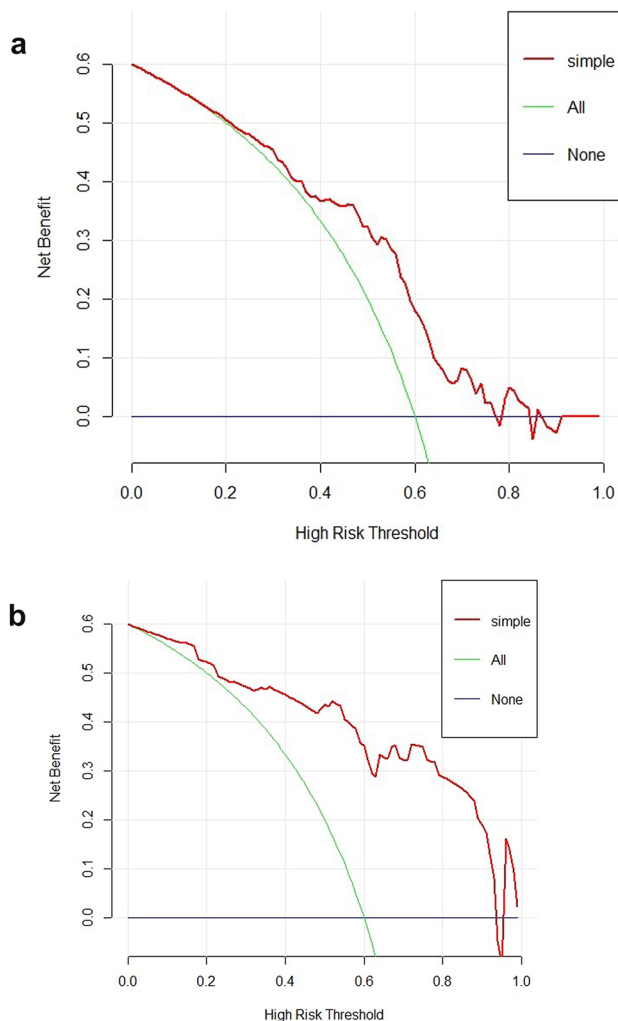


Fig. 6 DCA showed the predictive models provided greater benefit for benign and malignant characterization. **a** FS T2WI features-based predictive model. **b** ADC features-based predictive model

Several limitations should be mentioned. Selective bias could not be avoided; these patients were relatively younger, and the rhabdomyosarcoma was the most common malignancy. The sample size of intermediate tumors was relatively small. Those tumors seldom metastasize or recur and therefore were classified as benign. The value of texture analysis in the differentiation of STN should be explored at different anatomic sites. Considering the sample size, we did not compare the efficacy of different machine learning models. Moreover, the point-to-point radiological and histological correlation couldn't be done, due to the retrospective property.

Conclusion

ADC features of the whole tumor couldn't differentiate chemosensitive from non-chemosensitive sarcomas, small round from non-small round sarcomas, or rhabdomyosarcomas from non-rhabdomyosarcomas. The mean ADC and HISTO-skewness values did help in differentiating benign from malignant STN.

The ADC features-based LASSO-logistic predictive model did better than the FS T2WI features-based model in the characterization of STN.

Funding This research was supported in part by grants from the Science and Technology Council of Shanghai (grant no. 15ZR1408000, grant no. 18. no.12140901302, and grant no. 18140901200).

Availability of Data and Material The raw data can be made available.

Code Availability The R, SPSS 20.0, and MedCalc statistical software were used.

Declarations

Ethics Approval This retrospective study was approved by our institutional review board, and informed consent was waived.

Consent for Publication All authors have agreed to publish this article.

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not

permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Fletcher CDM, Bridge JA, Hogendoorn PCW, et al: WHO Classification of Tumours of Soft Tissue and Bone. 4th ed. Lyon: IARC Press, 2013.
2. Toro JR, Travis LB, Wu HJ, Zhu K, Fletcher CDM, Devesa SS: Incidence patterns of soft tissue sarcomas, regardless of primary site, in the surveillance, epidemiology and end results program, 1978–2001: An analysis of 26,758 cases. *Int J Cancer*. 2006;119(12):2922-30.
3. Berquist TH, Ehman RL, King BF, Hodgman CG, Ilstrup DM: Value of MR imaging in differentiating benign from malignant soft-tissue masses: study of 95 lesions. *AJR Am J Roentgenol*. 1990 Dec;155(6):1251-5.
4. Schlesinger AE, Hernandez RJ: Diseases of the musculoskeletal system in children: imaging with CT, sonography, and MR. *AJR Am J Roentgenol*. 1992; 158: 729-41.
5. Ma LD, Frassica FJ, Scott WW Jr, Fishman EK, Zerhouni EA: Differentiation of benign and malignant musculoskeletal tumors: potential pitfalls with MR imaging. *Radiographics*. 1995; 15:349-66.
6. Ma LD, McCarthy EF, Bluemke DA, Frassica FJ: Differentiation of benign from malignant musculoskeletal lesions using MR imaging: pitfalls in MR evaluation of lesions with a cystic appearance. *AJR Am J Roentgenol*. 1998; 170:1251-8.
7. van der Woude HJ, Verstraete KL, Hogendoorn PC, Taminiou AH, Hermans J, Bloem JL: Musculoskeletal tumors: does fast dynamic contrast-enhanced subtraction MR imaging contribute to the characterization? *Radiology*. 1998; 208:821-828.
8. Gielen JL, De Schepper AM, Vanhoenacker F, Parizel PM, Wang XL, Sciort R, et al: Accuracy of MRI in characterization of soft tissue neoplasms and tumor-like lesions. A prospective study in 548 patients. *Eur Radiol*. 2004; 14:2320–30.
9. Moulton JS, Blebea JS, Dunco DM, Braley SE, Bisset GS 3rd, Emery KH: MR imaging of soft-tissue masses: diagnostic efficacy and value of distinguishing between benign and malignant lesions. *AJR Am J Roentgenol*. 1995;164:1191-9.
10. May DA, Good RB, Smith DK, Parsons TW: MR imaging of musculoskeletal tumors and tumor mimickers with intravenous gadolinium: experience with 242 patients. *Skeletal Radiol*. 1997; 26:2-15.
11. Dangoor A, Seddon B, Gerrand C, Grimer R, Whelan J, Judson I: UK guidelines for the management of soft tissue sarcomas. *Clin Sarcoma Res*. 2016;15:6:20.
12. Subhawong TK, Jacobs MA, Fayad LM: Insights into quantitative diffusion-weighted MRI for musculoskeletal tumor imaging. *AJR Am J Roentgenol*. 2014; 203:560-72.
13. Subhawong TK, Durand DJ, Thawait GK, Jacobs MA, Fayad LM: Characterization of soft tissue masses: can quantitative diffusion weighted imaging reliably distinguish cysts from solid masses? *Skeletal Radiol*. 2013;42:1583-92.
14. Subhawong TK, Jacobs MA, Fayad LM: Diffusion-weighted MR imaging for characterizing musculoskeletal lesions. *Radiographics*. 2014;34:1163-77.
15. Mannelli L, Nougaret S, Vargas HA, Do RK: Advances in diffusion-weighted imaging. *Radiol Clin North Am*. 2015;53:569-81.
16. Surov A, Nagata S, Razek AA, Tirumani SH, Wienke A, Kahn T: Comparison of ADC values in different malignancies of the

- skeletal musculature: a multicentric analysis. *Skeletal Radiol.* 2015;44:995-1000.
17. Zou Y, Wang QD, Zong M, Zou YF, Shi HB: Apparent diffusion coefficient measurements with diffusion-weighted imaging for differential diagnosis of soft-tissue tumor. *J Cancer Res Ther.* 2016;12:864-70.
 18. Umanodan T, Fukukura Y, Kumagae Y, Shindo T, Nakajo M, Takumi K, et al: ADC histogram analysis for adrenal tumor histogram analysis of apparent diffusion coefficient in differentiating adrenal adenoma from pheochromocytoma. *J Magn Reson Imaging.* 2017;45(4):1195-1203.
 19. Hoffman DH, Ream JM, Hajdu CH, Rosenkrantz AB: Utility of whole-lesion ADC histogram metrics for assessing the malignant potential of pancreatic intraductal papillary mucinous neoplasms (IPMNs). *Abdom Radiol.* 2017;42(4):1222-1228.
 20. Vendrami CL, Velichko YS, Miller FH, Chatterjee A, Villavicencio CP, Yaghmai V, et al: Differentiation of Papillary Renal Cell Carcinoma Subtypes on MRI: Qualitative and Texture Analysis. *AJR Am J Roentgenol.* 2018; 211(6):1234-1245.
 21. Deo RC: Machine Learning in Medicine. *Circulation.* 2015;132(20):1920-1930.
 22. Richter AN, Khoshgoftaar TM: A review of statistical and machine learning methods for modeling cancer risk using structured clinical data. *Artif Intell Med.* 2018;90:1-14.
 23. Tang G, Qi L, Sun Z, Liu J, Lv Z, Chen L, et al: Evaluation and analysis of incidence and risk factors of lower extremity venous thrombosis after urologic surgeries: A prospective two-center cohort study using LASSO-logistic regression. *Int J Surg.* 2021;89:105948.
 24. Ueno D, Kawabe H, Yamasaki S, Demura T, Kato K: Feature selection for RNA cleavage efficiency at specific sites using the LASSO regression model in *Arabidopsis thaliana*. *BMC Bioinformatics.* 2021;22(1).
 25. Tishbirani R: Regression shrinkage and selection via the Lasso. *J R Stat Soc Ser B (Methodological).* 1996;58:267–88.
 26. Corino VDA, Montin E, Messina A, Casali PG, Gronchi A, Marchianò A, et al: Radiomic analysis of soft tissues sarcomas can distinguish intermediate from high-grade lesions. *J Magn Reson Imaging.* 2018;47(3):829-840.
 27. Dallaudière B, Lecouvet F, Vande Berg B, Omoumi P, Perlepe V, Cerny M, et al: Diffusion-weighted MR imaging in musculoskeletal diseases: current concepts. *Diagn Interv Imaging.* 2015; 96:327-40.
 28. Koh DM, Collins DJ: Diffusion-weighted MRI in the body: applications and challenges in oncology. *AJR Am J Roentgenol.* 2007;188:1622-35.
 29. Schnapauff D, Zeile M, Niederhagen MB, Fleige B, Tunn PU, Hamm B, et al: Diffusion-weighted echo-planar magnetic resonance imaging for the assessment of tumor cellularity in patients with soft-tissue sarcomas. *J Magn Reson Imaging.* 2009;29:1355-9.
 30. Türkbey B, Aras Ö, Karabulut N, Turgut AT, Akpınar E, Alibek S, et al: Diffusion-weighted MRI for detecting and monitoring cancer: a review of current applications in body imaging. *Diagn Interv Radiol.* 2012;18:46-59.
 31. Khoo MM, Tyler PA, Saifuddin A, Padhani AR: Diffusion-weighted imaging (DWI) in musculoskeletal MRI: a critical review. *Skeletal Radiol.* 2011; 40:665-81.
 32. Genovese E, Cani A, Rizzo S, Angeretti MG, Leonardi A, Fugazzola C: Comparison between MRI with spin-echo echo-planar diffusion-weighted sequence (DWI) and histology in the diagnosis of soft-tissue tumours. *Radiol Med.* 2011; 116:644-56.
 33. Oka K, Yakushiji T, Sato H, Fujimoto T, Hirai T, Yamashita Y, et al: Usefulness of diffusion-weighted imaging for differentiating between desmoid tumors and malignant soft tissue tumors. *J Magn Reson Imaging.* 2011; 33:189-93.
 34. Bonarelli C, Teixeira PA, Hossu G, Meyer JB, Chen B, Gay F, et al: Impact of ROI Positioning and Lesion Morphology on Apparent Diffusion Coefficient Analysis for the Differentiation Between Benign and Malignant Nonfatty Soft-Tissue Lesions. *AJR Am J Roentgenol.* 2015;205:W106-13.
 35. Lee SY, Jee WH, Jung JY, Park MY, Kim SK, Jung CK, et al: Differentiation of malignant from benign soft tissue tumours: use of additive qualitative and quantitative diffusion-weighted MR imaging to standard MR imaging at 3.0 T. *Eur Radiol.* 2016;26:743–54.
 36. Singer AD, Pattany PM, Fayad LM, Tresley J, Subhawong TK. Volumetric segmentation of ADC maps and utility of standard deviation as measure of tumor heterogeneity in soft tissue tumors. *Clin Imaging.* 2016;40:386-91.
 37. van Rijswijk CSP, Kunz P, Hogendoorn PCW, Taminiau AEHM, Doornbos J, Bloem JL. Diffusion-weighted MRI in the characterization of soft-tissue tumors. *J Magn Reson Imaging.* 2002; 15(3):302-7.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.