Skeletal Radiology
https://doi.org/10.1007/s00256-024-04837-7

SCIENTIFIC ARTICLE

Predicting progression-free survival in sarcoma using MRI-based automatic segmentation models and radiomics nomograms: a preliminary multicenter study

Nana Zhu[1,2] · Feige Niu[1,2] · Shuxuan Fan[3] · Xianghong Meng[2] · Yongcheng Hu[2] · Jun Han[2,4] · Zhi Wang[2]

Received: 1 September 2024 / Revised: 17 November 2024 / Accepted: 18 November 2024

Abstract

Objectives Some sarcomas are highly malignant, associated with high recurrence despite treatment. This multicenter study aimed to develop and validate a radiomics signature to estimate sarcoma progression-free survival (PFS).

Materials and methods The study retrospectively enrolled 202 consecutive patients with pathologically diagnosed sarcoma, who had pre-treatment axial fat-suppressed T2-weighted images (FS-T2WI), and included them in the ROI-Net model for training. Among them, 120 patients were included in the radiomics analysis, all of whom had pre-treatment axial T1-weighted and transverse FS-T2WI images, and were randomly divided into a development group ($n=96$) and a validation group ($n=24$). In the development cohort, Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression was used to develop the radiomics features for PFS prediction. By combining significant clinical features with radiomics features, a nomogram was constructed using Cox regression.

Results The proposed ROI-Net framework achieved a Dice coefficient of 0.820 (0.791–0.848). The radiomics signature based on 21 features could distinguish high-risk patients with poor PFS. Univariate Cox analysis revealed that peritumoral edema, metastases, and the radiomics score were associated with poor PFS and were included in the construction of the nomogram. The Radiomics-T1WI-Clinical model exhibited the best performance, with AUC values of 0.947, 0.907, and 0.924 at 300 days, 600 days, and 900 days, respectively.

Conclusion The proposed ROI-Net framework demonstrated high consistency between its segmentation results and expert annotations. The radiomics features and the combined nomogram have the potential to aid in predicting PFS for patients with sarcoma.

Keywords Sarcoma · Deep learning · Recurrence · Tumor segmentation · Radiomics

Abbreviations

PFS	Progression-free survival
ROI	Region of interest
ICC	Intra- and inter-class correlation coefficients
FS-T2WI	Fat-suppressed T2-weighted images
T1WI	T1-weighted imaging
ER	Early recurrence
3D-ROI	Three-dimensional regions of interest
LASSO	Least absolute shrinkage and selection operator
timeROC	Time-dependent receiver operating characteristic
LoG	Gaussian-Laplacian

✉ Zhi Wang
15002213919@163.com

[1] Graduate School, Tianjin Medical University, Tianjin, China

[2] The Department of Radiology, Tianjin University Tianjin Hospital, 406 Jiefang Southern Road, Tianjin 300211, China

[3] The Department of Radiology, Tianjin Medical University Cancer Hospital, Tianjin, China

[4] Graduate School, Tianjin University, Tianjin, China

Introduction

Sarcomas are tumors that originate in mesenchymal tissues, and are classified into two major categories: primary malignant bone tumors and soft tissue sarcomas. Although the disease location differs, the histological origin is the same [1]. Many sarcomas occur in both bone and soft tissue, and the incidence of both is relatively low, accounting for approximately 1% and 15% of all malignant tumors in adults and

Published online: 04 December 2024

children, respectively, so the efficiency of modeling alone is insufficient. Some are highly malignant, with a very poor prognosis despite treatment and a very high recurrence rate. There is similarity in the treatment methods. In this study, bone and soft tissue sarcomas were merged and modeled. Bone and soft tissue sarcomas usually require comprehensive treatment, including surgery, chemotherapy, and radiotherapy; adjuvant radiotherapy and chemotherapy are considered to be auxiliary interventions to improve prognosis [2]. However, despite optimal local treatment, nearly 40% of patients with bone or soft tissue sarcoma experience metastatic disease or local recurrence during follow-up [3]. Patients with metastatic disease have a poor prognosis, with 5-year event-free survival rates ranging from 20 to 30% and less than 20% for those with recurrent disease [4, 5]. The overall median survival for patients who develop local recurrence or distant metastases after initial treatment is poor, ranging from 12 to 18 months [6, 7].

A randomized, prospective study of 141 patients with soft tissue sarcomas of the extremities evaluated the effect of postoperative external radiotherapy on local recurrence (LR) after limb salvage resection. The results showed that the 10-year local recurrence rate of these patients was 0%, while that of patients without postoperative radiotherapy was 22% [8]. Such lesions may benefit from radiotherapy, requiring reliable prognostic markers to identify patients at high risk for recurrence to optimize personalized treatment.

Radiomics has recently gained substantial attention, and many biomarkers have been identified for diagnosis, staging, prognosis, gene expression, and prediction of therapeutic response [9–12]. It has high throughput to extract tumor features that cannot be observed by human eyes from conventional images [11]. However, large-sample joint model studies exploring radiomics to predict post-treatment recurrence or metastasis in patients with sarcoma are currently limited.

Tumor delineation is essential when establishing radiomics models. Currently, the most widely used methods are manual and semi-automatic segmentation. However, both require expert intervention and are prone to subjectivity and variability. Experts may produce different segmentations for the same image, leading to inter-observer and intra-observer variability [13]. Additionally, manually delineating the region of interest (ROI) is a labor-intensive process, particularly for large datasets or when high precision is required, limiting its scalability in clinical practice and research [14].

In this study, we have designed and proposed a lightweight deep learning architecture specifically for the delineation of ROI, named ROI-Net, which aims to facilitate the automation, acceleration, and enhancement of accuracy in the process of tumor delineation. Subsequently, we established a combined clinical radiomics model to accurately predict postoperative survival and better quantify the risk of recurrence in patients with bone and soft tissue sarcomas. This model helps formulate surgical plans for patients with varying risk levels, thus aiding personalized treatment.

Materials and methods

Patients

This retrospective study was approved by the hospital ethics review committees of all participating agencies and exempted from the requirement for written informed consent. The research design is shown in Fig. 1. We searched the data of 202 patients with sarcoma, with complete imaging records, from January 2019 to December 2023 in Tianjin Hospital (72 males 65 females; mean age 43.5 years) and Tianjin Cancer Hospital (30 males 35 females; mean age 46.7 years).

The construction of the ROI-Net model was based on all 202 patients (all with pre-treatment axial fat-suppressed T2-weighted images [FS-T2WI]). Of these, 162 (~80%) and 40 (~20%) cases were used as training and validation sets, respectively.

For the construction of the imaging model, we included 120 patients (80 from xxx hospital and 40 from xxx hospital) who met the inclusion criteria of radiological analysis (see electronic supplementary materials). All patients received pre-treatment axial T1-weighted imaging (T1WI) and FS-T2WI. The imaging scanning parameters are provided in the supplementary material. The training and verification sets were randomly divided 96:24 to ensure the stability and reproducibility of the model.

All patients followed the National Comprehensive Cancer Network guidelines [15] for neoadjuvant chemotherapy and surgical resection, and were followed up for at least 1 year after treatment. Follow-ups with CT or MRI scans were conducted every 2–3 months to screen for recurrence. If extrapulmonary metastasis was suspected, additional CT or MRI was routinely performed. Early recurrence (ER) and distant metastasis were confirmed through definitive pathology or reliable imaging. Patients with local ER and/or distant metastasis within 1 year were the ER group, while those without were the non-ER group. The follow-up period was from the first diagnosis of sarcoma to the next recurrence or no recurrence in September 2023. The average follow-up time was 25.8 months, with a median of 24.5 months.

MRI segmentations and image preprocessing

All three-dimensional ROIs were segmented using the uAI Research Portal platform (Shanghai Lianying Intelligent Medical Technology Co., Ltd.; http://urp.united-imaging.com:8080/#/).

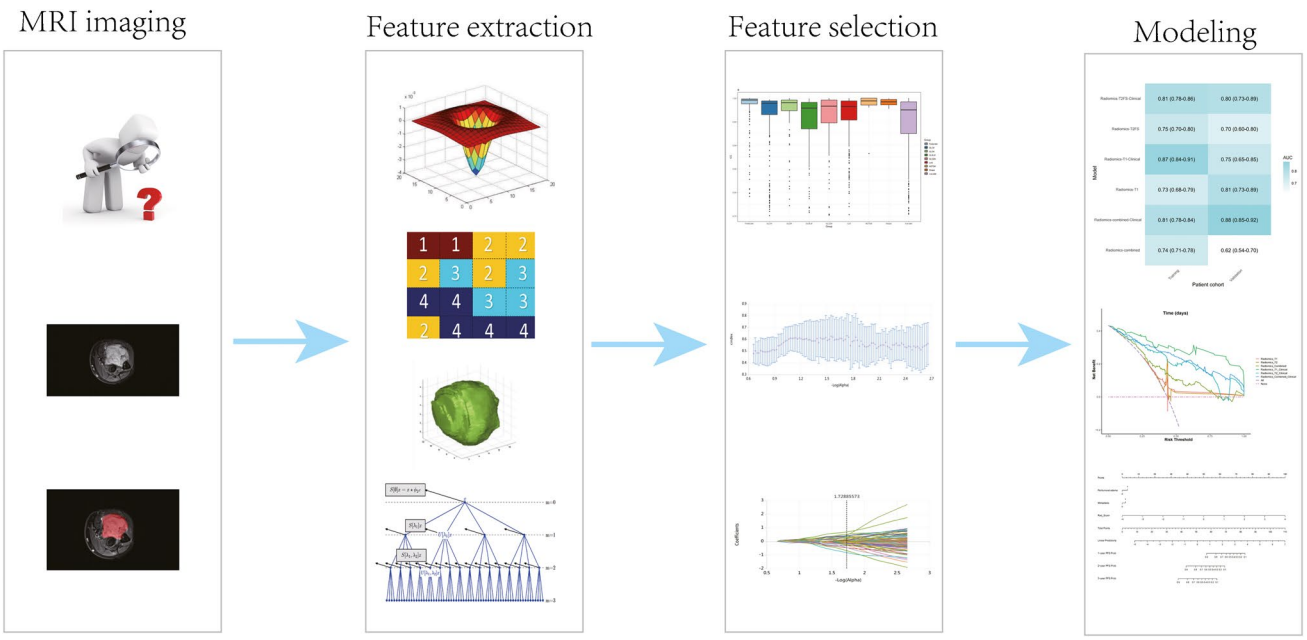

Fig. 1 Schematic workflow of the study

ROI-Net segmentations

Reader A (A med tech with 2 years of radiology research experience) manually delineated the ROI on each transverse slice of the T2FS sequence for all 202 patients; these were automatically converted into 3D-ROIs. The 3D-ROI segmentation encompassed the entire primary tumor, avoiding any apparent peritumoral edema. Subsequently, an experienced radiologist, Reader C (He had fellowship training and 33 years' experience in musculoskeletal imaging) refined and approved the segmentations.

Radiomics segmentations

The ROIs delineated by Reader A on the T2FS images of 120 patients were copied using ITK-SNAP and transferred to another sequence of interest (T1WI) as 3D ROIs, with visual inspection for consistency. Similarly, Reader B (B radiologist with 15 years' experience in musculoskeletal imaging) manually segmented the ROIs for a randomly selected subset of images, and Reader A repeated the segmentation after one month. The intra- and inter-class correlation coefficients (ICC) were calculated, based on 30 randomly selected patients. Features with an ICC > 0.75 were retained for further analysis.

Image preprocessing

Before image analysis, we applied a three-step preprocessing approach to all the MRI images to reduce noise and variability, thereby enhancing the stability and reliability of subsequent analyses: intensity discretization, intensity normalization, and isotropic resampling.

ROI-Net configuration information

Image preprocessing Given the heterogeneous features of multicenter images, data preprocessing is crucial for normalization. The Adaptive Normalizer preprocessed images to achieve mean–variance normalization of data features.

Configuration of target spacing In the coarse-level model (low resolution), a large target spacing of $5 \times 5 \times 5$ mm^3 was adopted to capture global positional information, with a cropping size of $64 \times 64 \times 64$.

Image resampling strategy Nearest neighbor interpolation was used for image resampling to the target spacing.

Optimizer The Adam optimizer dynamically adjusted the learning rate for each parameter by computing the first (mean) and second (average of the squared gradients) moments of the gradients. It adapts to sparse gradients and non-stationary objective functions by considering both the gradient average and variance.

Learning rate The learning rate was set to 0.0001 to fine-tune the network. As the network converged, the learning rate decreased from a larger to a smaller initial value.

Sampling schemes Global sampling schemes randomly generated center points throughout the foreground space to locate a target ROI.

Loss functions Segmentation accuracy was quantified using the Dice coefficient, ranging from 0 to 1; a value of 1 indicated a perfect overlap between the segmented result and its ground truth.

Network component In our framework, V-Net was the key component for multiscale organ segmentation. V-Net employs an encoder-decoder structure in which the input image undergoes convolutions and pooling operations in the encoder for feature extraction and down-sampling. In the decoder, transposed convolutions and skip connections were used to gradually restore the spatial resolution of the image and generate segmentation results. Skip connections preserved detailed information by directly passing features from the encoder to corresponding layers in the decoder.

Radiomics signature construction

The uAI Research Portal software (Version 730) was used for feature selection, where each radiomics feature was *z*-score-normalized to achieve a normal distribution. Subsequently, a three-step procedure was developed to rank and select the most informative features. First, we assessed the stability of the selected features using ICC. Features with ICC > 0.75 were retained. Second, we applied univariate Cox regression analysis to identify features associated with recurrence, considering adjusted p-values < 0.05 as statistically significant. Third, we employed Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression analysis to select the most robust and prognostic radiomics features.

We balanced the model's fit and complexity by adjusting the alpha value to improve its generalizability. The regularization parameter (lambda) in the LASSO model was determined through fivefold cross-validation. Features with non-zero coefficients selected by the optimal lambda were used to compute the radiomics score (rad-score), which is the weighted sum of the radiomics features and their corresponding LASSO coefficients. All patients were categorized into high-score and low-score groups using the median radiomics score as the cutoff.

Predictive model building

Three radiomics models (Radiomics-T1, Radiomics-T2FS, and Radiomics-combined) were developed using their respective features as inputs. To further establish a clinical radiomics-combined model, baseline clinical data, including age, sex, and tumor location, were obtained from medical records. Qualitative MRI features, including tumor size, peritumoral edema, tumor border, signal heterogeneity on fat-suppressed images, and metastases, were also collected. Clinical features and rad-scores significantly associated with progression-free survival (PFS) were selected for model construction using univariate Cox regression analysis across the entire cohort. This led to the establishment of clinical radiomics feature models: Radiomics-T1-Clinical, Radiomics-T2FS-Clinical, and Radiomics-combined-Clinical.

Harrell's concordance index (C-index) was measured to quantify discriminative performance, with values close to 1 indicating perfect consistency and 0.5 indicating consistency no better than chance. Kaplan–Meier curves and log-rank tests were used to plot and compare survival curves. Time-dependent receiver operating characteristic (timeROC) analysis was used to evaluate predictive ability over time. Forest plots were constructed using the Cox regression results to visualize risk factors. Decision curve analysis was used to estimate radiomics forest plots.

Statistical analysis

The segmentation results from the ROI-Net model based on deep learning were generated using the uAI Research Portal platform. Statistical analyses were conducted using R software (version 3.6.1). Clinical and MRI morphological features were evaluated using univariate analysis. The Student's *t*-test or Mann–Whitney *U* test were applied to determine significant inter-group differences for ordinal variables or continuous data, while the chi-square test or Fisher's exact test were used for categorical variables. Statistical significance was set at $p < 0.05$. Multivariate Cox model construction and LASSO-Cox regression analysis were implemented using the "survival" and "glmnet" packages. Nomograms were built, and the performance of predictive models over time was evaluated using the "nomogram" and "timeROC" packages. A decision curve analysis was conducted using the "ggDCA" package for statistical assessment.

Results

Patient demographics

Among the 202 patients trained with ROI-Net, there were 115 men (56.9%) and 87 women (43.1%). Primary malignant bone tumors accounted for 104 cases (51.5%), whereas soft tissue sarcomas accounted for 98 (48.5%). Diameters ranged from 2.3 to 32.1 cm, with an average of 10.71 cm.

Among the 120 patients included in the radiomics analysis, there were 63 men and 57 women (52.5% vs 47.5%, $p > 0.05$), with an average age of 44.57 ± 22.24 years. All had tumors larger than 2.0 cm, with 52 (43.3%) having tumors larger than 10.0 cm. Among these, 62 cases were

primary malignant bone tumors (51.7%), and 58 were soft tissue sarcomas (48.3%). For histology, see supplementary materials.

The clinicopathological data and semantic features of patients with sarcoma in the recurrence and non-recurrence groups are shown in Table 1. Patients with recurrence had significantly more prominent peritumoral edema ($p < 0.004$) and metastases ($p < 0.002$). However, no significant differences were observed regarding sex, age, tumor size, primary tissue of origin, tumor border, or signal heterogeneity ($p > 0.05$).

Evaluation of segmentation results for ROI-Net

Figure 2 illustrates the trend of the loss function during the training steps. Initially, the training loss decreased rapidly and stabilized, indicating continued learning and gradual model convergence.

The model was saved every 100 steps during training, and its performance was evaluated on the test set. Table 2 presents the performances of several key checkpoint models. Performance metrics steadily improved as the number of training steps increased and stabilized between 500 and 900 steps, with the optimal model achieving peak performance at 800 steps. Additionally, the Dice coefficient peaked at

Table 1 Baseline characteristics of patients

Characteristics	Training dataset			Validation dataset		
	ER ($n=40$)	Non-ER($n=56$)	*P*-value	ER ($n=10$)	Non-ER($n=14$)	*P*-value
Clinical characteristics						
Gender [No. (%)]			0.704[a]			1.000[a]
Male	19	30		6	8	
Female	21	26		4	6	
Age (mean ± SD, years)	46.05 ± 24.98	46.32 ± 23.47	0.907[b]	44.82 ± 19.05	41.08 ± 21.44	0.646[b]
Primary tissue			0.917[a]			0.945[a]
Bone	22	29		4	7	
Soft tissue	18	27		6	7	
MRI qualitative characteristics						
Tumor size (cm)			1.000[a]			0.884[a]
≤ 10	20	28		6	10	
> 10	20	28		4	4	
Peritumoral edema			0.004[a]			0.00007[a]
Yes	7	1		10	0	
No	23	55		0	14	
Tumor boundary			0.283[a]			0.048[a]
Clear with capsule/hardened edges	24	28		11	19	
Clear	6	16		2	1	
Unclear	10	12		2	6	
STIR homogeneity			0.438[a]			1.000[a]
Homogenous	13	13		3	3	
Inhomogenous	27	43		7	11	
Metastasis			0.002[a]			0.318[a]
Yes	8	0		2	0	
No	32	56		8	14	
Bone tumor location			0.742[a]			0.528[a]
Femur	13	14		3	3	
Tibia and fibula	4	7		1	3	
Others	5	8		0	1	
Soft tissue sarcoma location			0.303[a]			0.929[a]
Thigh	3	8		1	2	
Calf	12	16		3	4	
Others	5	3		1	1	

[a]: Chi-square test; [b]: Mann-Whitney U test

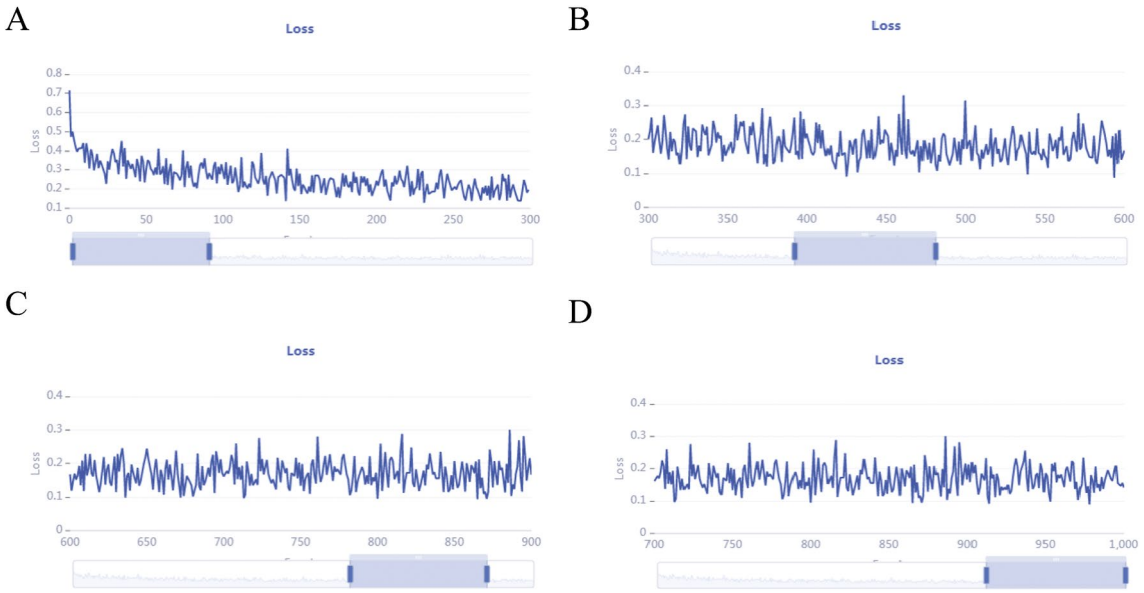

Fig. 2 Loss function curve for training steps at 300 (**A**), 600 (**B**), 900 (**C**), and 1000 (**D**); the model continues to learn and converges with increasing numbers of steps

Table 2 Prediction performance of models with different steps based on loss function. *Std* standard deviation

Training steps	Loss	Dice (95%CI)	Std
200	0.242	0.682 (0.602–0.761)	0.217
300	0.201	0.776 (0.741–0.810)	0.089
400	0.217	0.811 (0.777–0.845)	0.092
500	0.315	0.819 (0.789–0.849)	0.083
600	0.167	0.818 (0.790–0.847)	0.077
700	0.161	0.818 (0.790–0.846)	0.077
800	0.094	0.820 (0.791–0.848)	0.077
900	0.166	0.819 (0.790–0.847)	0.078
1000	0.141	0.812 (0.786–0.837)	0.065

0.820 (0.791–0.848), indicating high consistency between the segmentation results and expert annotations.

Radiomics signature construction

After reducing the radiomics features owing to segmentation variability, we developed Radiomics-T1, Radiomics-T2FS, and Radiomics-combined models using 2040, 1636, and 2169 features, respectively. During the training of these models, we employed univariate Cox regression and LASSO-Cox regression analyses to select 20, 21, and 12 features with non-zero coefficients, respectively. These features included one first-order statistical feature, one shape feature, one gray-level co-occurrence matrix feature, two gray-level run length matrix features, eight gray-level size zone matrix features, four gray-level dependence matrix features, one neighboring gray-tone difference matrix feature, 12 Gaussian-Laplacian (LoG) features, and 23 wavelet features. These features were used to establish radiomics predictors of PFS. All patients were categorized into high (≥ median rad-score) or low (< median rad-score) groups based on the median rad-score threshold.

Development of the radiomics and clinical radiomics models

Figure 3 presents the C-index and 95% confidence intervals (95% CI) for Radiomics-T2FS, Radiomics-T1, and Radiomics-combined. Using univariate Cox regression analysis, two significant MRI qualitative features (peritumoral edema and metastases) and rad-scores were identified as significantly associated with recurrence. Combining these risk factors, a clinical radiomics nomogram was constructed to predict 1-year, 3-year, and 5-year PFS (Fig. 2). Compared with radiomics models based solely on MRI sequences, this nomogram demonstrated superior performance. Kaplan–Meier analysis showed that patients in the high-score group had worse PFS than those in the low-score group in both the development and validation cohorts (Fig. 4).

TimeROC analysis demonstrated that the nomogram model outperformed the radiomics features (Fig. 5b). Radiomics-T1-Clinical exhibited the best performance in the training set, whereas the combined model of Radiomics-T2FS-Clinical and Radiomics-combined-Clinical showed good performance in the training set and good reproducibility and optimal performance in the validation set. Across the entire cohort, the AUCs at 300, 600, and 900 days for the

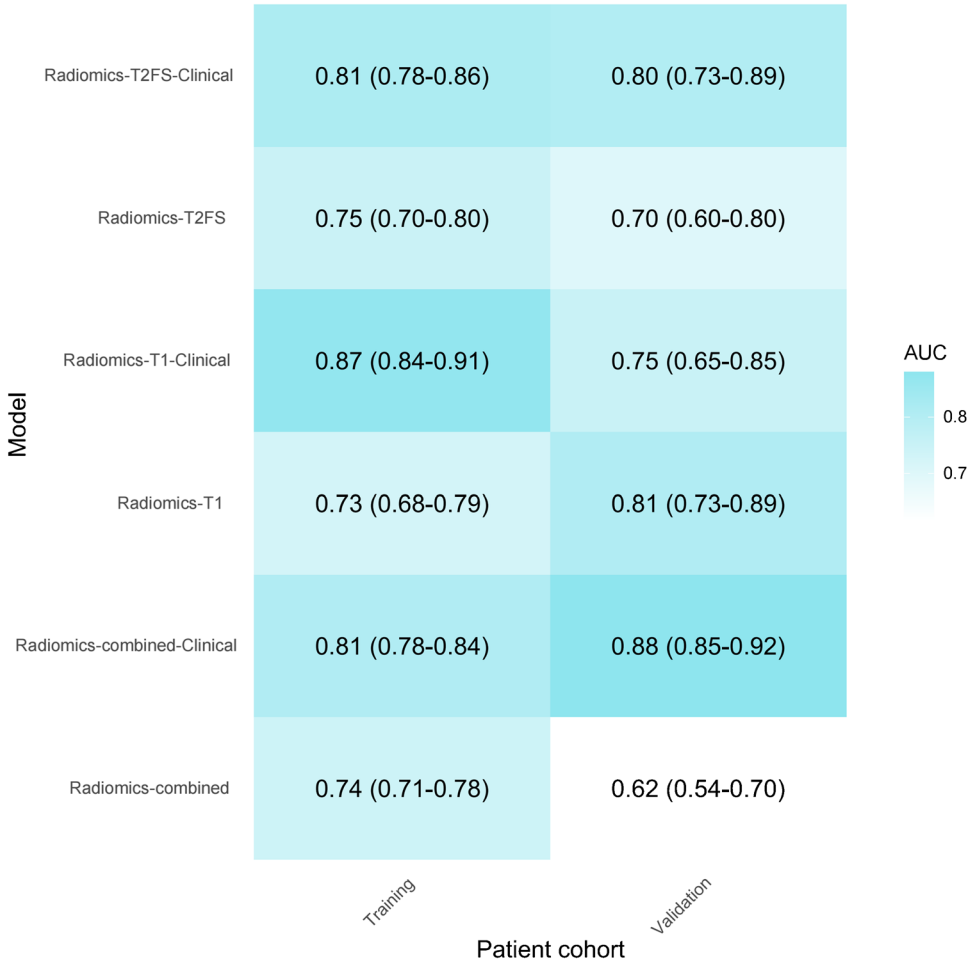

Fig. 3 C-index of different radiomics models after Cox regression analysis. C-index: concordance index

Radiomics-T1-Clinical, Radiomics-T2FS-Clinical, and Radiomics-combined-Clinical groups were 0.947, 0.907, and 0.924; 0.913, 0.853, and 0.869; and 0.918, 0.866, and 0.889, respectively (Table 3). Decision curve analysis showed that the combined nomogram model was more advantageous than the radiomics model for PFS prediction (Fig. 5c).

Discussion

The proposed ROI-Net framework initially realized automatic, accurate, and efficient tumor delineation. Quantitative radiology, based on preoperative MRI, was used to develop imaging features related to recurrence in patients with sarcoma. These preoperative radiological features combined with clinical baseline characteristics improved predictive performance and effectively divided patients into low-risk and high-risk groups. Therefore, the prediction model we proposed can help identify high-risk patients with poor PFS, thus contributing to personalized management. Including preoperative assessment of PFS can help doctors assess the complexity and risk of surgery, to develop a more reasonable surgical strategy. For example, patients with a shorter expected PFS may require more aggressive surgical strategies, such as expanding the scope of resection, to reduce the risk of recurrence. Additionally, preoperative evaluation enables the choice of appropriate adjuvant therapy, such as chemotherapy, radiotherapy, or immunotherapy. These adjuvant treatments can further reduce the risk of recurrence and improve survival rates.

Sarcoma represents a heterogeneous group of malignant tumors with high risk of recurrence and poor prognosis. Existing predictive models have focused on differentiating between benign and malignant cases and determining malignancy grading [16–19]. First, Liu et al.'s model [20] combines different clinical predictors, the MRI-based line chart C-index ≥ 0.721, and the median AUC ≥ 0.746. In the model of Chen et al. [21] results in the training queue were C-index: 0.907 CI 0.838, 95% CI 0.653–0.977 and that in the external verification queue were C-index: 0.811 Personality 95% CI 0.653–0.970. However, this study, based on the Radiomics-T1WI-Clinical model, showed the best performance, with AUC values of 0.947, 0.907, and 0.924 at 300, 600, and 900 days, respectively. Second, Liu [20] and Chen

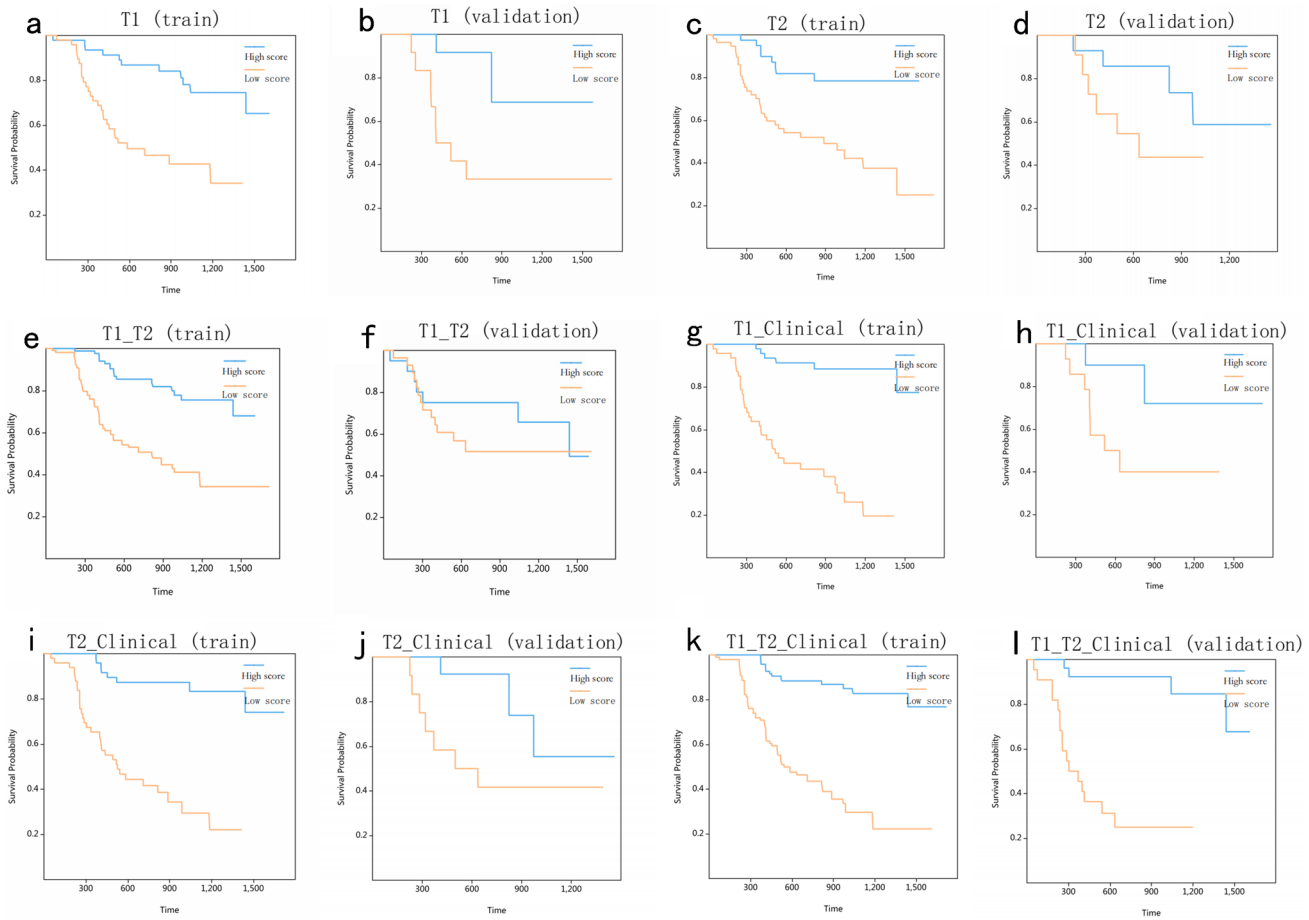

Fig. 4 Kaplan–Meier curves showing that the high radiomics score group has significantly worse progression-free survival in both the development and validation cohorts. **a**, **c**, **e**, **g**, **i**, and **k** are the training sets for Radiomics-T1, Radiomic-T2FS, Radiomics-combined, Radiomics-T1-Clinical, Radiomics-T2FS-Clinical, and Radiomics-combined-Clinical, respectively. **b**, **d**, **f**, **h**, **j**, and **i** are the validation sets for Radiomics-T1, Radiomic-T2FS, Radiomics-combined, Radiomics-T1-Clinical, Radiomics-T2FS-Clinical, and Radiomics-combined-Clinical, respectively

et al.'s model [21] combines postoperative clinical predictors (such as the staging system and Ki67) and enhanced sequence-based imaging features (joint invasion and perivascular involvement). However, our study is a multi-sequence, basic MRI plain scan sequence that does not require the use of contrast media, has no risk of allergy, reduces complications, and controls cost-effectiveness. Additionally, the clinical and radiological features included in our study are preoperative, allowing doctors to more accurately judge severity of patients and tumor characteristics, to develop more personalized treatment plans.

To address the limitations of manual contouring in radiomics analysis, such as time consumption, and low reproducibility and consistency, we developed ROI-Net, a deep learning-based framework for accurate and efficient automatic tumor contouring. This AI-accelerated workflow offers two potential advantages: (1) automatic segmentation significantly reduces the subjectivity and variability associated with manual segmentation, enhances consistency and accuracy in image analysis, and ensures reproducible and precise decision-making. (2) Segmentation time is restricted to 15 s, much shorter than the 5–10 min or longer required for manual outlining. Therefore, clinicians spend less time on automatic contouring and manual editing, significantly reducing the time required for image analysis. This allows for faster results and accelerates diagnostic and treatment decisions [22]. Additionally, automatic segmentation technology can effectively handle large volumes of image data, making it suitable for large-scale research and clinical applications, thereby promoting the widespread use of radiomics in medical research [13].

However, we also need to consider the results of this study more carefully. First of all, the automatic segmentation model based on deep learning relies on a large amount of training data to learn the segmentation features of tumors. However, the sample size of this study was relatively small,

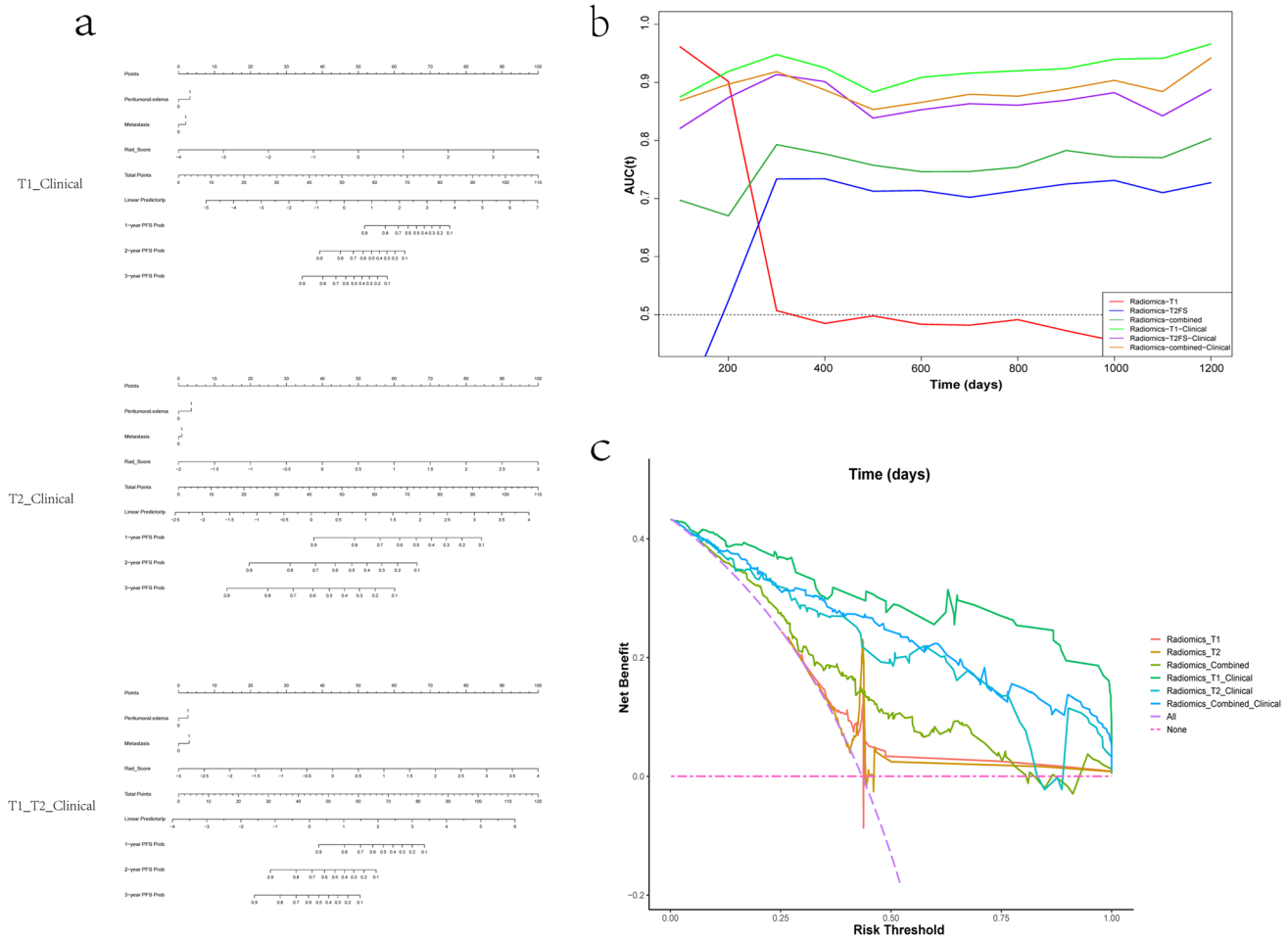

Fig. 5 Nomogram models established by combining clinical features and radiomics signatures of different sequences for predicting 1-, 3-, and 5-year progression-free survival (**a**). Comparison of timeAUC between the nomogram and other models in predicting progression-free survival (**b**). Decision curve analyses for predicting progression-free survival (**c**). timeAUC, time-dependent area under the curve

and the model may not have fully learned the complex characteristics of the tumors. Additionally, the tumor is highly heterogeneous in terms of location, shape, size, and contrast with other tissues. Therefore, for complex tumors, such as tumors with blurred boundaries, irregular shapes, or multiple sub-regions, higher segmentation accuracy is required. Figure 4 in the supplementary material shows the high accuracy of the segmentation model and the poor segmentation because of the mixed tumor signals and irregular shape. Therefore, future research needs to expand the sample size and integrate information from different imaging modes, such as T1-weighted MRI, T2-weighted MRI, and FLAIR MRI, to improve the accuracy of segmentation. Second, radiomics still lack a standardized process and repeatable research. There may be differences in image acquisition equipment, parameter settings, and acquisition techniques among different medical institutions, which limits the clinical applicability of the imaging model. In order to ensure data consistency and improve the accuracy of feature extraction and the generalizability of the model, firstly, the intensity values of the image are discretized using a fixed bin width of 10. This step converts the continuous intensity values in the image to discrete values for further processing. Second, through the resampling technology, the spatial resolution of all image data can be unified to eliminate the data inconsistency caused by equipment differences. Finally, the normalization process can scale the pixel values to a unified range, such as [0,1] or [− 1,1], thus eliminating the data inconsistency caused by the difference in the pixel value range. Third, because the incidence of sarcoma is low, the sample size of patients with complete concurrent image data and clinical features is small for model training, so this study included multiple histological types of sarcoma to collect enough data for modeling. This may have led to reduced model specificity and over-fitting. However, in view of the same histological origin of bone and soft tissue sarcomas and some similarities in treatment, we combined them to model. In future studies, we may group bone and soft tissue

Table 3 Comparison performance of models in predicting progression-free survival. *T* time (day)

Model	T = 100	T = 200	T = 300	T = 400	T = 500	T = 600	T = 700	T = 800	T = 900	T = 1000	T = 1100	T = 1200
Radiomics-T1	0.961	0.901	0.507	0.485	0.498	0.483	0.482	0.491	0.472	0.455	0.389	0.404
Radiomics-T2FS	0.325	0.524	0.733	0.734	0.713	0.714	0.702	0.714	0.725	0.731	0.710	0.727
Radiomics-combined	0.697	0.670	0.793	0.777	0.757	0.747	0.746	0.754	0.783	0.771	0.770	0.803
Radiomics-T1-Clinical	0.875	0.919	0.947	0.925	0.883	0.907	0.916	0.920	0.924	0.940	0.941	0.966
Radiomics-T2FS-Clinical	0.821	0.874	0.913	0.901	0.839	0.853	0.863	0.861	0.869	0.882	0.842	0.888
Radiomics-combined-Clinical	0.869	0.897	0.918	0.887	0.853	0.866	0.880	0.876	0.889	0.903	0.884	0.942

sarcomas into groups and expand the sample size based on functional imaging or enhanced scan sequences to obtain more dimensional tumor information training and prospectively verify the universality of the model in a multicenter population. Additionally, it can be developed into software; clinicians would only need to input the patient's specific scanning sequence and clinical characteristics, and then the corresponding score can be generated through the model, and the relapse-free survival status of preoperative patients can be obtained and inform the treatment plan.

The radiomics features developed in this study primarily included the LoG and wavelet features. The LoG filter is an edge detection filter that removes noise through Gaussian smoothing, while the wavelet filters decompose the original image in different directions, revealing spatial heterogeneity across multiple dimensions. Both techniques enhance texture details to help unmask potential tumor heterogeneity. Wavelet radiomics features assist in distinguishing between different types of tissues or tumors based on texture. For example, wavelet features have been used in gliomas to differentiate between low- and high-grade tumors by analyzing texture variations at different scales [23]. Elizabeth et al. [24] discussed the prognostic significance of LoG features related to overall survival in patients with lung cancer undergoing stereotactic radiotherapy. Similarly, wavelet radiomics features were found to have significant prognostic implications for distant metastasis. Li et al. combined LoG and wavelet features to construct a radiomics model for predicting prostate cancer versus benign prostatic hyperplasia, aiding clinicians in making better treatment decisions and reducing unnecessary biopsies [25].

Our research has some limitations. First, the sample size was relatively small, and multiple histological types of sarcomas were included; this is due to strict inclusion and exclusion criteria and the incidence of sarcomas. Second, this is a multicenter retrospective study, so selection bias could not be avoided even with strict exclusion criteria. With multicenter image acquisition, the acquisition protocols differ slightly. Therefore, we use the three-step method of image preprocessing to reduce the difference and enhance the stability of the line map. However, once radiomics methods are eventually applied to routine clinical settings in multiple institutions, this variability may be an advantage. Third, our study only used MRI image features, but CT images or other enhanced scanning and functional sequences may improve the radiomics features. In future research, it is necessary to compare the performance of related prognostic biomarkers based on multiple scans or other sequences with that of plain scan MRI sequences only. Finally, external verification in a larger multicenter population will help to prove the universality of our research results.

We proposed the first automatic sarcoma segmentation model trained on a multicenter dataset. This model can

quickly and accurately delineate tumor boundaries, reducing physician workloads and improving efficiency. Radiomics features can effectively predict the PFS of patients with sarcoma and provide valuable prognostic insights. Based on patient-specific clinical and radiomics features, this model serves as a useful predictive tool for developing personalized follow-up plans, detecting recurrence or progression early, and devising individualized treatment strategies based on risk stratification.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00256-024-04837-7.

Data availability The data that support the findings of this study is available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Declarations

Ethical approval This study was approved by the Ethics Committee (Ethical Review of Medicine 2024, Case No. 120). All authors listed meet the authorship criteria according to the latest guidelines of the International Committee of Medical Journal Editors, and all authors are in agreement with the manuscript.

Competing interests The authors declare no competing interests.

References

1. Skubitz KM, D'Adamo DR. Sarcoma. Mayo Clin Proc. 2007;82:1409–32.
2. Crago AM, Brennan MF. Principles in management of soft tissue sarcoma. Adv Surg. 2015;49:107–22.
3. Callegaro D, Miceli R, Bonvalot S, et al. Development and external validation of a dynamic prognostic nomogram for primary extremity soft tissue sarcoma survivors. EClinicalMedicine. 2019;17:100215.
4. Anderson JL, Denny CT, Tap WD, Federman N. Pediatric sarcomas: translating molecular pathogenesis of disease to novel therapeutic possibilities. Pediatr Res. 2012;72:112–21.
5. Hingorani P, Janeway K, Crompton BD, et al. Current state of pediatric sarcoma biology and opportunities for future discovery: a report from the sarcoma translational research workshop. Cancer Genet. 2016;209:182–94.
6. Italiano A, Mathoulin-Pelissier S, Cesne AL, et al. Trends in survival for patients with metastatic soft-tissue sarcoma. Cancer. 2011;117:1049–54.
7. Gronchi A, Miceli R, Colombo C, et al. Primary extremity soft tissue sarcomas: outcome improvement over time at a single institution. Ann Oncol. 2011;22:1675–81.
8. Yang JC, Chang AE, Baker AR, et al. Randomized prospective study of the benefit of adjuvant radiation therapy in the treatment of soft tissue sarcomas of the extremity. J Clin Oncol. 1998;16:197–203.
9. Varghese BA, Cen SY, Hwang DH, Duddalwar VA. Texture analysis of imaging: what radiologists need to know. AJR Am J Roentgenol. 2019;212:520–8.
10. Varghese BA, Hwang D, Cen SY, et al. Reliability of CT-based texture features: Phantom study. J Appl Clin Med Phys. 2019;20:155–63.
11. Aerts HJ, Velazquez ER, Leijenaar RT, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. 2014;5:4006.
12. Wang T, Deng J, She Y, et al. Radiomics signature predicts the recurrence-free survival in stage I non-small cell lung cancer. Ann Thorac Surg. 2020;109:1741–9.
13. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60–88.
14. Cardenas CE, Yang J, Anderson BM, Court LE, Brock KB. Advances in auto-segmentation. Semin Radiat Oncol. 2019;29:185–97.
15. Pluchino LA, D'Amico TA. National Comprehensive Cancer Network Guidelines: who makes them? What are they? Why are they important? Ann Thorac Surg. 2020;110:1789–95.
16. Xu W, Hao D, Hou F, Zhang D, Wang H. Soft tissue sarcoma: preoperative MRI-based radiomics and machine learning may be accurate predictors of histopathologic grade. AJR Am J Roentgenol. 2020;215:963–9.
17. Navarro F, Dapper H, Asadpour R, et al. Development and external validation of deep-learning-based tumor grading models in soft-tissue sarcoma patients using MR imaging. Cancers. 2021;13(12):2866. https://doi.org/10.3390/cancers13122866.
18. Vos M, Starmans MPA, Timbergen MJM, et al. Radiomics approach to distinguish between well differentiated liposarcomas and lipomas on MRI. Br J Surg. 2019;106:1800–9.
19. Fields BKK, Demirjian NL, Hwang DH, et al. Whole-tumor 3D volumetric MRI-based radiomics approach for distinguishing between benign and malignant soft tissue tumors. Eur Radiol. 2021;31:8522–35.
20. Liu S, Sun W, Yang S, et al. Deep learning radiomic nomogram to predict recurrence in soft tissue sarcoma: a multi-institutional study. Eur Radiol. 2022;32:793–805.
21. Chen H, Liu J, Cheng Z, et al. Development and external validation of an MRI-based radiomics nomogram for pretreatment prediction for early relapse in osteosarcoma: a retrospective multicenter study. Eur J Radiol. 2020;129:109066.
22. Zunair H, Ben Hamza A. Sharp U-Net: depthwise convolutional network for biomedical image segmentation. Comput Biol Med. 2021;136:104699.
23. Çinarer G, Emiroğlu BG, Yurttakal AH. Prediction of glioma grades using deep learning with wavelet radiomic features. Appl Sci. 2020;10:6296.
24. Huynh E, Coroller TP, Narayan V, et al. CT-based radiomic analysis of stereotactic body radiation therapy patients with lung cancer. Radiother Oncol. 2016;120:258–66.
25. Li C, Deng M, Zhong X, et al. Multi-view radiomics and deep learning modeling for prostate cancer detection based on multiparametric MRI. Front Oncol. 2023;13:1198899.