Journal of Bone Oncology 19 (2019) 100263

Contents lists available at ScienceDirect

Journal of Bone Oncology

journal homepage: www.elsevier.com/locate/jbo

Research Paper

Radiomics signature extracted from diffusion-weighted magnetic resonance imaging predicts outcomes in osteosarcoma

Zhao Shuliang[a,b,1], Su Yi[c,1], Duan Jinghao[d], Qiu Qingtao[d], Ge Xingping[b], Wang Aijie[e], Yin Yong[d,*]

[a] School of Medicine, Shandong University, Ji'nan 250012, China
[b] Department of Radiotherapy, Yantaishan Hospital of Yantai, Yantai 264001, China
[c] Department of Radiotherapy, Yuhuangding Hospital of Yantai, Yantai 264001, China
[d] Department of Radiation Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Ji'nan 250117, China
[e] Department of CT/MR, Yantaishan Hospital, Yantai 264001, China

ARTICLE INFO

Keywords:
Osteosarcoma
Radiomics signature
DWI-MRI
Prognosis
Regression analysis

ABSTRACT

Objective: Osteosarcoma often requires multidisciplinary treatment including surgery, chemotherapy and radiotherapy. However, tumor behavior can vary widely among patients and selection of appropriate therapies in any individual patient remains a critical challenge. Radiomics seeks to quantify complex aspects of tumor images under the assumption that this information is related to tumor biology. This study tested the hypothesis that a radiomic signature extracted from Diffusion-weighted magnetic resonance images (DWI-MRI) can improve prediction of overall survival (OS) compared with clinical factors alone in localised osteosarcoma.

Materials/Methods: Pre-treatment DWI-MRI were collected from 112 patients (9–67 years of age) with histological-proven osteosarcoma that were treated with curative intent. The entire dataset was divided in two subsets: the training and validation cohorts containing 76 and 24% of the data respectively. Clinical data were extracted from our medical record. Two experienced radiotherapists evaluated DWI-MRIs for quality and segmented the tumor. A total of 103 radiomic features were calculated for each image. Least absolute shrinkage and selection operator (LASSO) regression was applied to select features. Association between the radiomics signature and OS was explored. Further validation of the radiomics signature as an independent biomarker was performed by using multivariate Cox regression. The Cox proportional-hazard regression model was also used to analyze the correlation between the prognostic factor and the survival for the clinical (C) model after the univariate analysis. Radiomics (R) model identified radiomics signature, which is the best predictor from the radiomic variable classes based on LASSO regression. Harrell's C-index was used to demonstrate the incremental value of the radiomics signature to the traditional clinical risk factors for the individualized prediction performance.

Results: Cox proportional-hazard regression model shows that: Tumor size, alkaline phosphatase (ALP) status before treatment and number of courses of chemotherapy were proven as the dependent clinical prognostic factors of osteosarcoma's overall survival time. The radiomics signature was significantly associated with OS, independent of clinical risk factors (radiomics signature: HR: 5.11, 95% CI: 2.85, 9.18, $P < 0.001$). Incorporating the radiomics signature into the coalition (C + R) model resulted in better performance ($P < .001$) for the estimation of OS (C-index: 0.813; 95% CI: 0.75, 0.89) than with the clinical (C) model (C-index: 0.764; 95% CI: 0.69, 0.85), or the single radiomics (R) model (C-index: 0.712; 95% CI: 0.65, 0.78).

Conclusion: This study shows that the radiomics signature extracted from pre-treatment DWI-MRI improve prediction of OS over clinical features alone. Combination of the radiomics signature and the traditional clinical risk factors performed better for individualized OS estimation in patients with osteosarcoma, which might enable a step forward precise medicine. This method may help better select patients most likely to benefit from intensified multimodality diagnosis and therapies. Future studies will focus on multi-center validation of an optimized model.

* Corresponding author.
E-mail address: yinyongsd@126.com (Y. Yin).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.jbo.2019.100263
Received 14 August 2019; Received in revised form 28 September 2019; Accepted 3 October 2019
Available online 04 October 2019

1. Introduction

Osteosarcoma is the most common primary malignant bone tumor and typically develops during puberty with tumors arising at sites of rapid bone growth, it accounts for approximately 19% of all bone cancers [1,2]. Osteosarcoma usually arises in the metaphysis of a long bone, most commonly around the knee. Involvement of the axial skeleton and craniofacial bones is primarily observed in adults [3]. Patients with clinically detectable lung metastasis account for about 20%–25% of total osteosarcomas diagnosed [4]. Despite significant advances in the detection and treatment of osteosarcoma over the past two decades, the 5-year survival of patients with a metastatic disease is around 20% [5]. Osteosarcoma often requires multidisciplinary treatment including surgery, chemotherapy and radiotherapy. Although the survival of patients with operable osteosarcoma is improved by chemotherapy [3,6], however, tumor behavior can vary widely among patients and selection of appropriate therapies in any individual patient, especially in children and adolescents, remains a critical challenge. An important component of management involves the assessment of treatment response to multidisciplinary treatment in ameliorating recurrence risk. Therefore, the identification of novel biomarkers of a metastatic and prognostic phenotype is essential for osteosarcoma therapy.

Although magnetic resonance imaging (MRI) is considered the best technique for the local staging of musculoskeletal neoplasms [7], it is relatively little used in the primary diagnosis of bone tumors and its capacity to evaluate the prognostic might be underestimated [8]. Presently, prognosis of non-metastatic osteosarcoma is determined largely based on histological response and completeness of resection [9]; magnetic resonance imaging serves an adjunct role in assessing treatment response, generally on the basis of imaging appearance on fluid sensitive and post-contrast imaging sequences with a focus on global tumor volume reduction. Metabolic imaging with FDG PET/CT is also sometimes employed to ascertain residual hypermetabolic tumor, but can result in substantial cumulative radiation doses in children [10]. One of the main challenges in assessing treatment response on conventional magnetic resonance imaging is that tumor volume may not significantly decrease following chemotherapy due to tumor necrosis, resulting in a stable or paradoxically increased tumor volume [11]. In addition, the osteoid component of osteogenic tumors such as conventional osteosarcoma does not appreciably change on fluid-sensitive and contrast-enhanced sequences [12]. Anatomic musculoskeletal MRI also does not provide information regarding viability of tumor tissue [13]. Diffusion-weighted imaging (DWI) MRI can capture changes at the cellular level thanks to differences in movement of water protons in the different tissue regions [14]. Therefore, DWI-MRI can provide information regarding tumor cellularity as a surrogate indicator of treatment response on the basis of a quantitative value [15].

However, the clinical utility of these factors was limited and unclear. Therefore, new tools are urgently needed to identify patients who are at risk of having a poor prognosis. For the past several years, as an emerging individualized precision medical technology, radiomics has applied advanced computational methodologies to transform the image data of the regions of interest into high dimensional feature data. Next, quantitative and high-throughput analysis of feature data is completed to probe tumor phenotype [16,17]. Radiomics utilizes noninvasive imaging to provide more comprehensive information about the entire tumor and can be used in diagnosis, prognosis and prediction [18].

Patients with overt metastases at diagnosis have a very different outlook to those with apparently localised disease. The value of quantitative imaging would be much more valuable in the majority of patients that have localised disease. Thus, in this study, we developed and validated multiparametric DWI-MRI based radiomics as a novel approach for providing individualized, pretreatment evaluation of overall survival in patients with localised osteosarcoma. In addition, we sought to reveal association between radiomics features and clinical data.

2. Materials and methods

2.1. Sample size

Small sample size will increase both the type-I (incorrectly detecting a difference) and type-II (not detecting an actual difference) error rates [19]. To generate accurate estimates of the impact of the depended variables, an adequate number of events per variable is required. For the training sample size, Chalkidou proposed that for linear models, like multiple regression, at least 10 to 15 observations per predictor variable is required to produce reasonably stable estimates [20]. In our study, eight features were selected for the final model and the minimum training data size was 80. While for the validation sample size, we performed a power calculation to estimate the sample size for our study [21] and found that the minimum sample size is 24. In our study, 112 patients (85 training data and 27 validation data) were analyzed, which were sufficient.

2.2. Study population

This study comprised an evaluation of the institutional database for medical records from January 2012 to December 2017 to identify patients with histologically confirmed osteosarcoma according to the World Health Organization (WHO) classification [22]. 112 consecutively selected patients with localised osteosarcoma were retrospectively analyzed. Ethical approval was obtained for this analysis. The entire dataset was divided in two cohorts: (I) the training set containing the data used to train the model (n = 85); (II) the validation set containing the data used to validate the model(n = 27); see the Statistical Analysis section. All patients underwent to a DWI-MRI acquisition before starting the treatment. All patients' data were anonymized before the analysis.

2.3. Follow-up

All patients were followed up every 3 months during the first 2 years, every 6 months in years 2–5, and annually thereafter. To provide an efficient tool that would allow final personalized treatment, we chose overall survival (OS) as the endpoint. We calculated OS from the first day of treatment to the death from osteosarcoma, and the date of any cause to death or the last follow-up visit (censored). Baseline clinical data in the training and validation cohort, including age, gender, Enneking stage, tumor size, Karnofsky performance status (KPS), location of tumors, alkaline phosphatase (ALP) and lactate dehydrogenase (LDH) status before treatment, local recurrence, number of courses of chemotherapy, chemotherapy regimen, pathological fracture, were obtained from the medical records (Table 1).

2.4. Image acquisition

DWI-MRI images without applying any preprocessing or normalization were acquired as Digital Imaging and Communications in Medicine (DICOM) using GE Discovery MR 750 3.0T with 8 channel phased array coil and spine array coil for signal reception. The data were acquired axially by means of echo planar imaging. DWI images were acquired using b-value (600 s/mm^2).

2.5. Manual tumor delineation

All tumor target areas were performed by two experienced radiotherapists, and each plan was verified by a senior radiotherapist with more than twenty years of experience. The region of interest (ROI) covered the gross tumor volume (GTV) and was delineated on each slice of primary osteosarcoma in MIM software (www.mimsoftware.com). The radiologists were blind to one another.

Table 1
Patient and tumor characteristics in the training and validation cohorts.

Characteristic	Training cohort (n = 85)	Validation cohort (n = 27)
Gender		
Male	49 (57.6%)	15 (55.6%)
Female	36 (42.4%)	12 (44.4%)
Age (y)		
Median(IQR)	18 (14–49)	17.5 (13.75–49)
≤15	33 (38.8%)	11 (40.7%)
>15	52 (61.2%)	16 (59.3%)
Enneking stage		
IA	8 (9.4%%	3 (11.1%)
IB	17 (20%)	6 (22.2%)
IIA	32 (36.6%)	10 (37%)
IIB	22 (25.9%)	5 (18.5%)
III	6 (7.1%)	3 (11.1%)
Tumor size (ml)		
≥150	23 (27%)	8 (29.6%)
<150	62 (73%)	19 (71.4%)
KPS		
≥80	57 (67%)	14 (51.9%)
<70	28 (33%)	13 (48.1%)
Pathological fracture		
Yes	14 (16%)	4 (14.8%)
No	71 (84%)	23 (85.2%)
Location of tumors		
Limb	61 (72%)	17 (63%)
Torso	24 (28%)	10 (37%)
ALP before treatment (U/L)		
Median (IQR)	129 (79–210)	125 (81–204)
>125	47 (55.3%)	18 (66.7%)
≤125	38 (44.7%)	9 (33.3%)
LDH before treatment (U/L)		
Median (IQR)	192 (168–235)	197 (162–241)
>245	19 (22.4%)	10 (37%)
≤245	66 (77.6%)	17 (63%)
Local recurrence		
Yes	27 (31.8%)	9 (33.3%)
No	58 (68.2%)	18 (66.7%)
Number of courses of chemotherapy		
≥6	33 (38.8%)	10 (37%)
<6	52 (61.2%)	17 (63%)
Chemotherapy regimen		
AP regimen	34 (40%)	12 (44.4%)
Others	51 (60%)	15 (55.6%)
Follow-up time (mo)		
Median(IQR)	45 (38–51)	43 (36–50)

Note. Unless otherwise specified, data are numbers of patients, with percentages in parentheses. No difference was found between the training data set and the validation data set in either the clinical characteristics or the follow-up data (P = .538–.982).
Abbreviations: IQR, inter-quartile range; KPS, Karnofsky Performance Status; ALP, alkaline phosphatase; LDH, lactate dehydrogenase; AP, cisplatin/doxorubicin

2.6. Quantitative imaging feature extraction

With the help of image quantification features, radiomics feature parameters can be used to quantitatively evaluate the differences between tumor phenotypes (Fig. 1). A total of 103 selected quantitative imaging features were extracted from the information contained in the voxels of the tumor region segmented by the three strategies (GraphCut,GrowCut,Manual Segmentation). This process was implemented in IBEX (Imaging Biomarker Explorer, MD Anderson cancer center, USA), an open-source, easy to use radiomic software [23]. These features were organized into three categories: (I) 49 first-order statistical features derived from the tumor intensity histogram reflect distribution of values of individual voxels without concern for spatial relationships; (II) 38 textural features describe spatial arrangement of voxels were calculated from different parent matrices, including gray-level co-occurrence matrix (GLCM), gray level run-length matrix (GLRLM), neighbor gray-tone difference matrix (NGTDM); (III) 16 shape based features provide the geometrical of tumor volume. The complete list of the main features and their texture interpretation is reported in Table 2. A more detailed description about their computation can be found in previous reviews [21,24].

2.7. Statistical analysis

The statistical analysis was performed with R software, version 3.4.3 (http://www.R-project.org. The survival R package was used for Kaplan-Meier survival analyses; the glmnet R package was used for the least absolute shrinkage and selection operator (LASSO) Cox regression model, which is suitable for the regression of high-dimensional data [25]; the Hmisc R package was used for comparisons between C-index), and X-tile software, version 3.6.1 (Yale University School of Medicine, New Haven, Conn) [26], SPSS version 20.0 (SPSS, Chicago, IL, USA). All statistical tests were two-sided, and P values of <0.05 were considered significant.

The differences in age, gender, Enneking stage, tumor size, KPS, location of tumors, ALP and LDH status before treatment, local recurrence, number of courses of chemotherapy, chemotherapy regimen, pathological fracture, mean follow-up time between the training and validation cohorts were assessed by using an independent samples t test, χ^2 test, or Mann-Whitney U test, where appropriate.

2.8. Construction of the radiomics score-based radiomics signature

According to the Harrell guideline, the number of events should exceed the number of included covariates by at least 10 times in a multivariate analysis [27]. To address this issue, LASSO regression model was used to select the most useful prognostic features in the training cohort. The selected imaging features were then combined into a radiomics signature (Fig. 2). A radiomics score (Rad-score) was computed for each patient through a linear combination of selected features weighted by their respective coefficients [28].

2.9. Validation of radiomics signature

The potential association of the radiomics signature with OS was first assessed in the training cohort and then validated in the validation cohort by using Kaplan-Meier survival analysis. With the help of X-tile software, the patients were classified into high-risk or low-risk groups according to the Rad-score. The difference in the survival curves of the high-risk and low-risk groups was evaluated by using a weighted log-rank test (the G-rho rank test, rho = 1) for a substantial increase in efficiency because the event rate was relatively low compared with the sample size [29].

Evaluation of the multi-feature-based radiomics signature as an independent biomarker was performed by integrating the selected clinical risk factors into the multivariable Cox proportional hazards model.

2.10. Clinical features selection

The factors of age, gender, Enneking stage, tumor size, KPS, location of tumors, ALP and LDH status before treatment, local recurrence, number of courses of chemotherapy, chemotherapy regimen, and pathological fracture were included in univariate analysis. Kaplan-Meier method was used to calculate the overall survival rate. The univariate analysis was used to determine the prognostic factors related with survival rate by log-rank test. The Cox proportional-hazard regression model was used to analyze the correlation between the prognostic factor and the survival [27].

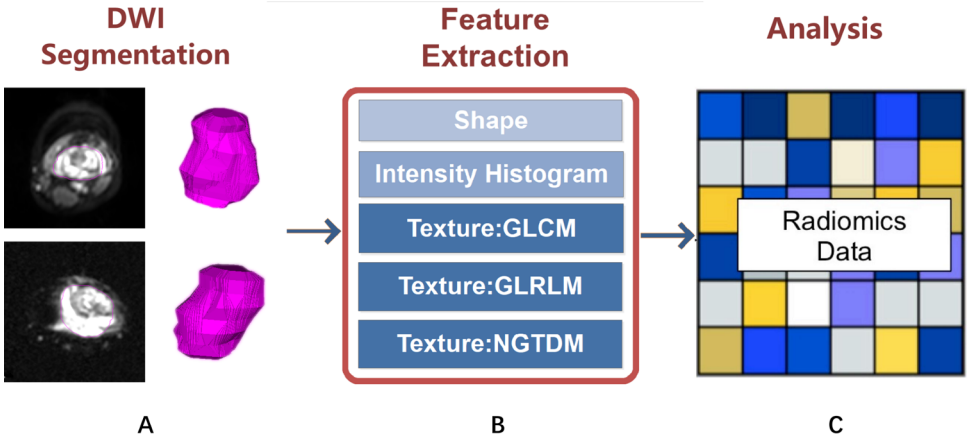

Fig. 1. Radiomics methods overview. (A), Image segmentation is performed on DWI-MRI images. Experienced radiologists contour the tumor areas on all MRI slices with planning software MIM. (B), Features are extracted from within the defined tumor contours on the MRI images, quantifying tumor intensity, shape, texture. (C) For the analysis, the radiomics features are compared with clinical data.
Abbreviations: DWI, Diffusion Weighted Imaging; GLCM, Gray-level co-occurrence matrix; GLRLM, Gray level run-length matrix; NGTDM, Neighbor gray-tone difference matrix.

Table 2
List of the main texture analysis methods.

Method	Texture interpretation	Main estimated features
Intensity Histogram	Global distribution of intensity values in terms of spread, symmetry, flatness, uniformity and randomness	Inter quartile range Quantile Percentile area Kurtosis Mean absolute deviation
GLCM	Spatial relationship between pixel in a specific direction, highlighting the properties of uniformity, homogeneity, randomness and linear dependency of the image	Inverse variance Entropy Dissimilarity Correlation Energy Contrast Homogeneity
GLRLM	Texture in a specific direction, where fine texture has more short runs whilst coarse texture presents more long runs with different intensity values	Short-run emphasis Run length non-uniformity Long-run emphasis Grey-level non-uniformity Run percentage
NGTDM	Spatial relationship among three or more pixels, closely approaching the human perception of the image	Busyness Coarseness Complexity Texture strength Contrast
Shape	Shape has 16 feature values that mainly describe the 3d size and Shape of the mask image.	Maximum 3D diameter Number of voxel Compactness Convex Mean breadth Mass Orientation Surface area density Volume

Abbreviations: GLCM, Gray-level co-occurrence matrix; GLRLM, Gray level run-length matrix; NGTDM, Neighbor gray-tone difference matrix.

2.11. Assessment of incremental value of radiomics signature in individual OS estimation

To demonstrate the incremental value of the radiomics signature to the traditional staging system and other clinical risk factors for individualized assessment of OS in patients with osteosarcoma, both a Rad-score and a clinical feature were presented in the validation cohort.

Compared with the incremental value of other clinical risk factors, the traditional clinical radiological characteristics were calibrated and evaluated. The potential of the radiomics signature was compared with that of the traditional assessment model. To quantify the discrimination performance, the Harrell concordance index (C-index) was measured, along with the concordance probability estimate considering the high degree of censoring in our data (1 indicates perfect concordance; 0.5 indicates no better concordance than chance) [30].

3. Results

3.1. Clinical Characteristics and OS

As of the last follow-up, the mean OS was 43.19 months, and the median OS was 45 months. The shortest OS was 12 months.

No difference was found in the factors of age, gender, Enneking stage, tumor size, KPS, location of tumors, ALP and LDH status before treatment, local recurrence, number of courses of chemotherapy, chemotherapy regimen, pathological fracture, mean follow-up time were not statistically different in the two groups between the training cohort and the validation cohort in either clinical characteristics or follow-up data (P = .538–.982; Table 1).

3.2. Radiomics feature selection and radiomics signature building

Among 254 originals parameters extracted from magnetic resonance images, 87 and 64 were excluded due to poor repeatability (ICC < 0.9) and strong regional volume correlation (r > 0.85), respectively. Finally, 103 independent radiomics features were incorporated into LASSO analysis. Among these features, we selected 8 features from DWI images after using glmnet R data package for LASSO logistic regression, and analyzed the correlation between parameters (Fig. 3). The results showed that although some parameters had a high positive correlation or negative correlation, on the whole, the correlation between parameters was low and had a high independence (r = 0.23 ± 0.65).

These characteristics are closely related to the overall survival time in the training cohort. To construct radiomics signature, these features were used for inclusion in the Rad-score prognostic model. Rad-score for each patient in the training cohort was shown in Fig. 4.

The texture features with a nonzero coefficient in the LASSO Cox regression model were as follows: GLCM-Inverse_Variance, GLRLM-Run_Length_Non_uniformity, GLRLM-Short_Run_Low_Gray_Level_Empha, IH-Quantile_0.025, SHAPE-Convex, SHAPE-Mass, SHAPE-Number_Of_Voxel, SHAPE-Orientation.

The radiomics signature was constructed, with a Rad-score calculated by using the following formula (Rad-score was computed for each patient through a linear combination of selected features weighted by their respective coefficients) :

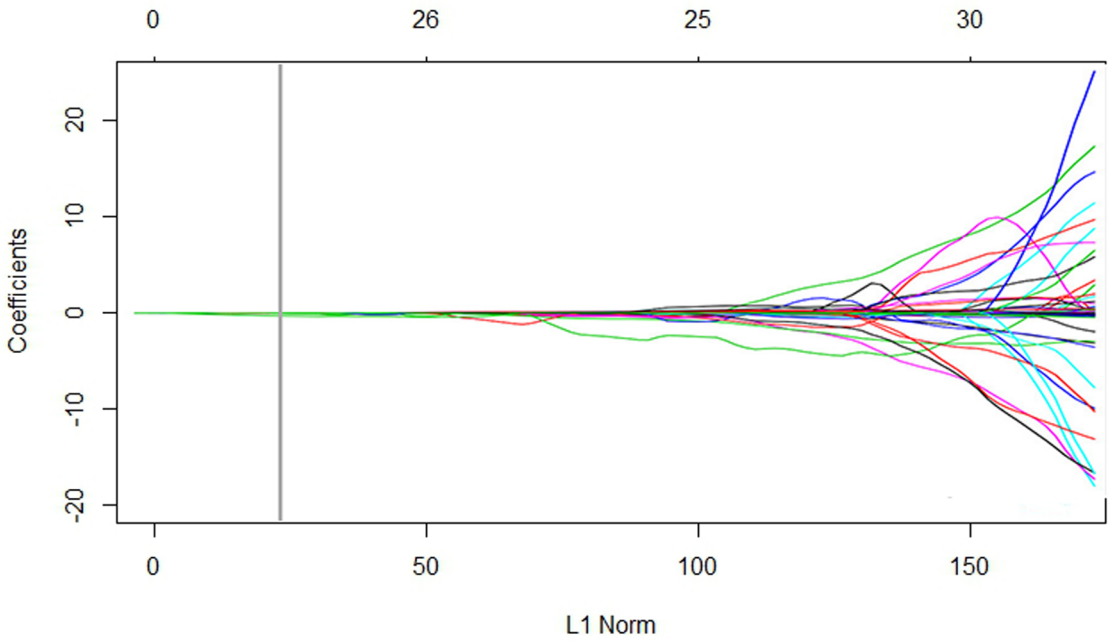

Fig. 2. R-plot algorithm for the relationship between the region-of-interest image filtration process and overall survival.

Rad-score = GLCM-InverseVariance × (–2.51323077) + GLRLM-Run_Length_Non_uniformity × (−0.10737204) + GLRLM-Short_Run_Low_Gray_Level_Empha × 0.00227 + IH-Quantile0.025 × (−0.74684231) + SHAPE-Convex × (−2.61241068) + SHAPE-Mass × 0.20896467 + SHAPE-Number_Of_Voxel × 2.35344525 + SHAPE-Orientation × (−0.97096879).

The optimum cutoff generated by the X-tile plot was −3.92. Accordingly, patients were classified into a high-risk group (Rad-score ≥ −3.92) and a low-risk group (Rad-score < −3.92) (Fig. 5).

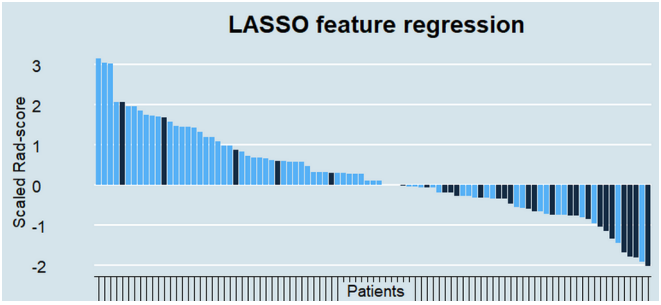

Fig. 4. Rad-score for each patient in the training cohort. Black bars show scores for patients who survived without disease relapse or were censored, while blue bars show scores for those who experienced relapse or died. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Validation of radiomics signature

The radiomics signature was associated with the OS in the training cohort (P = 0.0452; HR = 5.35, 95% confidence interval [CI]: 3.22,8.87), and this finding was confirmed in the validation cohort (*P* < 0.001; HR = 5.09; 95% CI: 3.12, 7.91). Patients with lower Rad-scores generally had better OS: When the patients were stratified on the basis of clinical risk factors, a significant association was found in one or more subgroups (Fig. 6).

3.4. Clinical features selection

After the univariate analysis in the clinical factors, number of courses of chemotherapy, KPS, tumor size, ALP status before treatment, and pathological fracture were selected to be applied in the clinical (C) model. Cox proportional-hazard regression model to analyze the correlation between the prognostic factor and the survival. The significant result as follow, ALP status before treatment: HR:1.96, 95% CI: 1.19, 3.12, P = .009; number of courses of chemotherapy: HR:1.89, 95% CI:

	GLCM-Inver	GLRLM-Ru	GLRLM-Sho	IH-Quantile	SHAPE-Conv	SHAPE-Mas	SHAPE-Num	SHAPE-Orienta
GLCM-Inverse_	1							
GLRLM-Run_Le	0.39138	1						
GLRLM-Short_R	-0.243538	0.479807	1					
IH-Quantile_0.0	0.044701	0.079095	0.051754	1				
SHAPE-Convex	0.508867	-0.124895	0.226216	-0.320145	1			
SHAPE-Mass	-0.288678	0.201061	0.345182	0.199872	0.688379	1		
SHAPE-Numbe	0.302144	-0.075293	-0.165787	0.175778	-0.297905	0.527851	1	
SHAPE-Orienta	0.132877	0.121676	0.460353	0.388616	0.164772	-0.1237302	0.080202	1

Fig. 3. The correlation between selected radiomics characteristic parameters. Green indicates that there is a positive correlation between the two parameters, while gray indicates that there is a negative correlation between the two parameters. The overall correlation between the radiomics parameters in the model was relatively low, with high independence: $r = 0.23 \pm 0.65$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

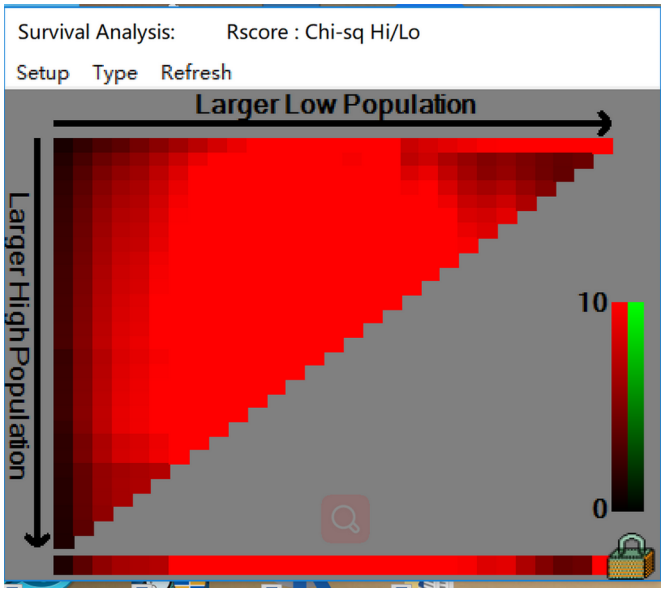

Fig. 5. X-Tile for OS: The optimum cutoff generated by the X-tile plot was −3.92

0.83, 4.29 P = .031; tumor size: HR: 3.56, 95% CI: 2.10, 6.23; P = .003). Therefore, in the clinical models, tumor size, ALP status before treatment and number of courses of chemotherapy were proven as the dependent prognostic factors of osteosarcoma's overall survival time (Table 3).

Then, a Cox regression analysis used in C+R (clinical + radiomics) model identified radiomics signature as independent risk factors (radiomics signature: HR: 5.11, 95% CI: 2.85, 9.18, $P < 0.001$).

3.5. Evaluation of incremental value of radiomics signature in individual OS performance

The discrimination performance of the radiomics signature improved when it was integrated into the coalition model along with the clinical risk factors (C-index for the C+R model:0.813; 95% CI: 0.75, 0.89). Compared with either clinical (C) model (C-index: 0.764; 95% CI: 0.69, 0.85) or the single radiomics (R) model (C-index: 0.712; 95% CI: 0.65, 0.78), the C+R model showed a better discrimination capability ($P < 0.01$ for each comparison, except the ALP + Rad-score vs. Rad-score group, $P = 0.607$ and tumor size + Rad-score vs. Rad-score group, $P = 0.191$) (Table 4).

Table 3
Risk Clinical Factors for overall survival in Osteosarcoma.

Clinical factors	β	*P*value	Hazard ratio	95% CI
KPS	−0.78	0.06	0.43	025–0.79
Number of courses of chemotherapy	0.64	0.03	1.89	0.83–4.29
Tumor size	1.29	0.003	3.56	2.10–6.23
ALP status before treatment	0.67	0.009	1.96	1.19–3.12
Pathological fracture	−0.11	0.92	0.94	0.43–2.11

Note. β is the regression coefficient.
Abbreviation: 95% CI, 95% confidence interval.

Table 4
Performance of models.

Model	C-Index	95% CI	*P*value
Radiomics signature	0.712	0.65, 0.78	<0.01
Clinical Characteristic	0.764	0.69, 0.85	<0.05
C+R	0.813	0.75, 0.89	<0.01

Abbreviations: C-index, concordance index; C, clinical model; R, radiomics model.

4. Discussion

This study developed and validated a diagnostic approach based on DWI-MRI radiomics model for individualized evaluation of OS before treatment in osteosarcoma. We showed that radiomics features complemented the clinical system, helping to provide better prognostic ability for the OS of osteosarcoma.

Our research shows that the difference of number of courses of chemotherapy, KPS, pathological fracture in the univariate analysis of patients with osteosarcoma was statistically significant, but there was no significant difference in the multiple factors analysis, which illustrates that one of the prognosis of the clinical factors as a single factors will be influenced by various other prognostic factors. Maybe there is a certain correlation in univariate analysis, but it has also lost its meaning in the multivariate analysis. Otherwise, tumor size, ALP status and number of courses of chemotherapy were proven as the dependent prognostic factors of osteosarcoma's overall survival time, this finding is consistent with findings of previous studies [3,31], but there are still some otherness, which may be caused by differences in the therapy method and the clinical stage.

The delineation of the target area is the most critical and technical part of radiology because the subsequent characteristic data are generated by the segmented volume, although the boundaries of many tumors are not clear [32,33]. Compared with CT or PET/CT, MRI provides better tissue contrast, it has multiplanar capacity, and exhibits

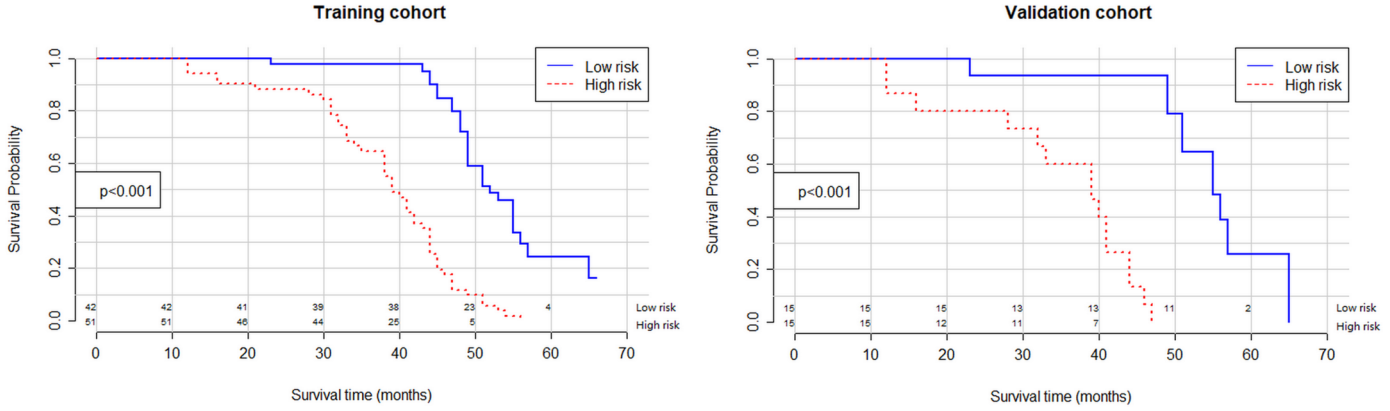

Fig. 6. Graphs show results of Kaplan-Meier survival analyses according to the radiomics signature for patients in the training data set (left) and those in the validation data set (right). A significant association of the radiomics signature with the overall survival time was shown in the training data set, which was then confirmed in the validation data set.

fewer artifacts from radiation and bone beam hardening, that allows tumor borders to be delineated more accurately [32,34]. DWI-MRI dependent upon tumor cellularity and perfusion, significant associations of key imaging features in the initial assessment of osseous sarcomas, which supports DWI as an alternative to gadolinium-based contrast-enhanced MRI [35].

For the construction of the radiomics signature, 103 candidate radiomics features were reduced to 8 potential predictors by examining the predictor-outcome association by shrinking the regression coefficients with the LASSO method. The identified signature consisted of the following features: GLCM-Inverse_Variance, GLRLM-Run_Length_Non_uniformity, GLRLM-Short_Run_Low_Gray_Level_Empha, IH-Quantile_0.025, SHAPE-Convex, SHAPE-Mass, SHAPE-Number_Of_Voxel, SHAPE-Orientation, which are consistent with results of recent studies of risk stratification [36]. Previous studies have supported the hypothesis that radiologic and phenotypic information of tumors can be inferred from imaging [37,38]. The current study suggests that the identified features may suggest survival outcomes, which supports the idea that radiomics signature capture intra-tumoral heterogeneity in a non-invasive manner and it can be associated with patient outcomes. Further research in radiomics is needed to establish a biological basis for texture heterogeneity to identify potential radiomics-biologics correlations. The multi-feature-based radiomics signature could be used to successfully estimate the survival time. The result is understandable because the potential nuances or confounding effects of other risk factors may exist in a single data set [39].

The radiomics signature of DWI-MRI successfully stratified patients into high-risk and low-risk groups, which were separated on the basis of the median Rad-score. It incorporates the radiomics signature and clinical risk factors into an easy-to-use model, which facilitates the individualized prediction of these patients with osteosarcoma. It shows that a radiomic signature extracted from pre-treatment DWI-MRI improves the prediction of OS over clinical features alone. The technology could help doctors better select patients who are most likely to benefit from multimodal treatment.

In addition, researchers naturally combine multiple disease manifestations of patients to estimate and determine follow-up treatment, rather than focusing on a single symptom, which emphasizes the necessity of multivariate estimation. Until now, the overall scientific quality and reporting of MRI radiomics studies is insufficient. Reporting of study objectives, blind assessment, sample size, and missing data are deemed to be necessary [40]. We still need to make scientific improvement to make them reproducible, and useful for clinical analysis and scientific categories. Studies have shown that in the absence of any combined model, it is difficult to assess the overall outcome of an individual patient's prognosis by using a single risk factor which is considered powerful. Recently, more findings prove that imaging is a useful tool for the combination of prognostic molecular signatures and clinical risk factors. 99m Tc-MIBI imaging is a useful tool for the evaluation of neoadjuvant chemotherapy in patients with osteosarcoma, and its dual mechanisms could be simultaneously used in predicting and evaluating tumor response to chemotherapy [41]. Gene pathways related to immune system regulation and extracellular signaling had the highest number of significant radiomic feature associations [42]. All these have proved that the multi-mode model combined with image data processing can be a good guide for the clinical prognosis and treatment of malignant tumors.

The limitations of this study are that the number of patients is low, and it was contributed by the retrospective nature of data collection. Although the preferred design should be a prospective longitudinal cohort study, this allows for better control of all relevant risk factors and outcomes, ensures that no bias is introduced, and minimizes the loss of follow-up [39]. The protracted length of a prospective longitudinal cohort study in Osteosarcoma (because of the long wait needed for survival outcomes) may make the research daunting [9,31]. However, we believe that our findings are sufficient to encourage larger, multicenter clinical studies in patients with osteosarcoma and even soft tissue sarcoma, they are expected to be evaluated by using radiomics to better determine indications for surgery and adjuvant chemotherapy.

5. Conclusion

It can be concluded from the results of this study that the union model (C + R) which incorporates both the radiomics signature and clinical risk factors is feasible, it provides valuable information to predict the overall survival time in patients with osteosarcoma, and improves the value of evaluation compared with the traditional clinical model. It is worth mentioning that the radiomics model alone also has a good prognostic ability, but in order to be more responsible for patients with osteosarcoma, it is recommended to use the combined model for prediction. This method may help better select patients most likely to benefit from intensified multimodality diagnosis and therapies. Optimization of the model is still needed with a study of larger cohorts and inclusion of other categories of features, other imaging modalities, and other "-omics" criteria.

Funding

This research was supported by the National Key Research and Development Project of China (No. 2017YFC0113202).

Compliance with ethical standards

All procedures performed in studies involving human participants were in accordance with the Institutional Review Board (IRB) and ethics committee of Shandong Cancer Hospital and Institute.

Declaration of Competing Interest

The authors declare no conflict of interest.

References

[1] C.M. Hattinger, M. P., S. Ferrari, P. Picci, M. Serra, Emerging drugs for high-grade osteosarcoma, Expert Opin. Emerg. Drugs 15 (4) (2010) 615–634.
[2] A.A. E., Perinatal factors associated with clinical presentation of osteosarcoma in children and adolescents, Pediatr. Blood Cancer 64 (6) (2017).
[3] Group, E.S.E.S.N.W., Bone sarcomas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up, Ann. Oncol. 25 (Suppl 3) (2014) iii113–23 https://doi.org/10.1093/annonc/mdu256.
[4] Z.Q. Bao, et al., Over-expression of Sox4 and beta-catenin is associated with a less favorable prognosis of osteosarcoma, J. Huazhong Univ. Sci. Technol. Med. Sci. 36 (2) (2016) 193–199 https://doi.org/10.1007/s11596-016-1565-z.
[5] O. Camuzard, et al., Role of autophagy in osteosarcoma, J. Bone Oncol. 16 (2019) 100235https://doi.org/10.1016/j.jbo.2019.100235.
[6] I.J. Lewis, et al., Improvement in histologic response but not survival in osteosarcoma patients treated with intensified chemotherapy: a randomized phase III trial of the European Osteosarcoma Intergroup, J. Natl. Cancer Inst. 99 (2) (2007) 112–128 https://doi.org/10.1093/jnci/djk015.
[7] D. Bisseret, et al., Periosteum: characteristic imaging findings with emphasis on radiologic-pathologic comparisons, Skeletal Radiol. 44 (3) (2015) 321–338 https://doi.org/10.1007/s00256-014-1976-5.
[8] J.L. de Sa Neto, et al., Diagnostic performance of magnetic resonance imaging in the assessment of periosteal reactions in bone sarcomas using conventional radiography as the reference, Radiol. Bras. 50 (3) (2017) 176–181 https://doi.org/10.1590/0100-3984.2015.0166.
[9] S.S. Bielack, B. K.-B., Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. 2002.
[10] S.C. Chawla, et al., Estimated cumulative radiation dose from PET/CT in children with malignancies: a 5-year retrospective review, Pediatr. Radiol. 40 (5) (2010) 681–686 https://doi.org/10.1007/s00247-009-1434-z.
[11] G. Lee, et al., Measurement variability in treatment response determination for non–small cell lung cancer: improvements using radiomics, J. Thorac. Imaging 34 (2019) 1.
[12] M. Uhl, et al., Evaluation of tumour necrosis during chemotherapy with diffusion-weighted MR imaging: preliminary results in osteosarcomas, Pediatr. Radiol. 36 (12) (2006) 1306–1311 https://doi.org/10.1007/s00247-006-0324-x.
[13] F.M. Costa, E.C. Ferreira, E.M. Vianna, Diffusion-weighted magnetic resonance imaging for the evaluation of musculoskeletal tumors, Magn. Reson. Imaging Clin. N. Am. 19 (1) (2011) 159–180 https://doi.org/10.1016/j.mric.2010.10.007.
[14] V.D.A. Corino, et al., Radiomic analysis of soft tissues sarcomas can distinguish

intermediate from high-grade lesions, J. Magn. Reson. Imaging (2017), https://doi.org/10.1002/jmri.25791.
[15] A.J. Degnan, C.Y. Chung, A.J. Shah, Quantitative diffusion-weighted magnetic resonance imaging assessment of chemotherapy treatment response of pediatric osteosarcoma and Ewing sarcoma malignant bone tumors, Clin. Imaging 47 (2018) 9–13 https://doi.org/10.1016/j.clinimag.2017.08.003.
[16] P. Lambin, et al., Radiomics: extracting more information from medical images using advanced feature analysis, Eur. J. Cancer 48 (4) (2012) 441–446 https://doi.org/10.1016/j.ejca.2011.11.036.
[17] H.J. Aerts, et al., Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach, Nat. Commun. 5 (2014) 4006 https://doi.org/10.1038/ncomms5006.
[18] H.J. Aerts, et al., Defining a radiomic response phenotype: a pilot study using targeted therapy in NSCLC, Sci. Rep. 6 (2016) 33860 https://doi.org/10.1038/srep33860.
[19] B. Zhang, et al., Radiomics features of multiparametric mri as novel prognostic factors in advanced nasopharyngeal carcinoma, Clin. Cancer Res. 23 (15) (2017) 4259–4269 https://doi.org/10.1158/1078-0432.CCR-16-2910.
[20] A. Chalkidou, M.J. O'Doherty, P.K. Marsden, False discovery rates in PET and CT studies with texture features: a systematic review, PLoS One 10 (5) (2015) e0124165https://doi.org/10.1371/journal.pone.0124165.
[21] S. Elisa, G. Rizzo, Texture analysis of medical images for radiotherapy applications, Br. Inst. Radiol. 90 (20) (2017) 20160642.
[22] V.Y J., Refinements in sarcoma classification in the current 2013 world health organization classification of tumours of soft tissue and bone, Surg. Oncol. Clin. N. Am. 4 (25) (2016).
[23] L. Zhang, et al., IBEX: an open infrastructure software platform to facilitate collaborative work in radiomics, Med. Phys. 42 (3) (2015) 1341–1353 https://doi.org/10.1118/1.4908210.
[24] F D., et al., Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? Insights Imaging 3 (6) (2012) 573–589.
[25] J. Gui, H. Li, Penalized Cox regression analysis in the high-dimensional and low-sample size settings, with applications to microarray gene expression data, Bioinformatics 21 (13) (2005) 3001–3008 https://doi.org/10.1093/bioinformatics/bti422.
[26] R.L. Camp, D.L. R., M. Dolled-Filhart, X-Tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization, Clin. Cancer Res. 10 (November 1) (2004) 7252–7259.
[27] F. Harrell, Regression Modeling Strategies with Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis, Springer Series in Statistics, 2015 ISBN 978-3-319-19424-0.
[28] Y. Huang, Z. L., Radiomics signature a potential biomarker for the prediction of disease-free survival in early-stage(I or II) non–small cell lung cancer, Radiology 281 (3) (2016) 947.
[29] S. Buyske, R. F., Z Ying, A class of weighted log-rank tests for survival data when the event is rare, J. Am. Stat. Assoc. 95 (449) (2000) 249–258.
[30] M.G.G. Heller, Concordance probability and discriminatory power in proportional hazards regression, Biometrika 92 (4) (2005) 965–970.
[31] B. Gaetano, et al., Prognostic factors for osteosarcoma of the extremity treated with neoadjuvant chemotherapy: 15-year experience in 789 patients treated at a single institution, Cancer 106 (5) (2010) 1154–1161.
[32] F. Tixiera, H. Um, Reliability of tumor segmentation in glioblastoma Impact on the robustness: impact on the robustness of MRI-radiomic features, J. Med. Imaging (Bellingham) (2019), https://doi.org/10.1002/mp.13624.
[33] L. Duron, et al., Gray-level discretization impacts reproducible MRI radiomics texture features, PLoS One 14 (3) (2019) e0213459https://doi.org/10.1371/journal.pone.0213459.
[34] S.B. Ginsburg, et al., Radiomic features for prostate cancer detection on MRI differ between the transition and peripheral zones: preliminary findings from a multi-institutional study, J. Magn. Reson. Imaging 46 (1) (2017) 184–193 https://doi.org/10.1002/jmri.25562.
[35] A.N. Alsharief, et al., Usefulness of diffusion-weighted MRI in the initial assessment of osseous sarcomas in children and adolescents, Pediatr. Radiol. 49 (9) (2019) 1201–1208 https://doi.org/10.1007/s00247-019-04436-y.
[36] Q. Qiu, et al., Reproducibility of radiomic features with GrowCut and GraphCut semiautomatic tumor segmentation in hepatocellular carcinoma, Transl. Cancer Res. 6 (5) (2017) 940–948 https://doi.org/10.21037/tcr.2017.09.47.
[37] A.M. Rutman, M.D. Kuo, Radiogenomics: creating a link between molecular diagnostics and diagnostic imaging, Eur. J. Radiol. 70 (2) (2009) 232–241 https://doi.org/10.1016/j.ejrad.2009.01.050.
[38] M. Artzi, et al., Differentiation between vasogenic edema and infiltrative tumor in patients with high-grade gliomas using texture patch-based analysis, J. Magn. Reson. Imaging (2018), https://doi.org/10.1002/jmri.25939.
[39] K.G. Moons, et al., Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration, Ann. Intern. Med. 162 (1) (2015) W1–73 https://doi.org/10.7326/M14-0698.
[40] J.E. Park, et al., Quality of science and reporting of radiomics in oncologic studies: room for improvement according to radiomics quality score and TRIPOD statement, Eur. Radiol. (2019), https://doi.org/10.1007/s00330-019-06360-z.
[41] C. Wu, Q. Wang, Y. Li, Prediction and evaluation of neoadjuvant chemotherapy using the dual mechanisms of (99m)Tc-MIBI scintigraphy in patients with osteosarcoma, J. Bone Oncol. 17 (2019) 100250https://doi.org/10.1016/j.jbo.2019.100250.
[42] A.C. Yeh, et al., Radiogenomics of breast cancer using dynamic contrast enhanced MRI and gene expression profiling, Cancer Imaging 19 (1) (2019) 48 https://doi.org/10.1186/s40644-019-0233-5.