Radiomics Analysis of Multiparametric MRI for Prediction of Synchronous Lung Metastases in Osteosarcoma

Zhendong Luo†, Jing Li†, YuTing Liao, RengYi Liu, Xinping Shen* and Weiguo Chen*

1 Department of Radiology, The University of Hong Kong - Shenzhen Hospital, Shenzhen, China, 2 Department of Radiology, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou, China, 3 Department of Pharmaceuticals Diagnosis, GE Healthcare, Shanghai, China, 4 Department of Radiology, Nanfang Hospital, Southern Medical University, Guangzhou, China

Purpose: To establish and verify a predictive model involving multiparameter MRI and clinical manifestations for predicting synchronous lung metastases (SLM) in osteosarcoma.

Materials and Methods: Seventy-eight consecutive patients with osteosarcoma (training dataset, n = 54; validation dataset, n = 24) were enrolled in our study. MRI features were extracted from the T1-weighted image (T1WI), T2-weighted image (T2WI), and contrast-enhanced T1-weighted image (CE-T1WI) of each patient. Least absolute shrinkage and selection operator (LASSO) regression and multifactor logistic regression were performed to select key features and build radiomics models in conjunction with logistic regression (LR) and support vector machine (SVM) classifiers. Eight individual models based on T1WI, T2WI, CE-T1WI, T1WI+T2WI, T1WI+CE-T1WI, T2WI+CE-T1WI, T1WI+T2WI+CE-T1WI, and clinical features, as well as two combined models, were built. The area under the receiver operating characteristic curve (AUC), sensitivity and specificity were employed to assess the different models.

Results: Tumor size was the most significant univariate clinical indicator (1). The AUC values of the LR predictive model based on T1WI, T2WI, CE-T1WI, T1WI+T2WI, T1WI+CE-T1WI, T2WI+CE-T1WI, and CE-T1WI were 0.686, 0.85, 0.87, 0.879, 0.736, 0.85, and 0.914, respectively (2). The AUC values of the SVM predictive model based on T1WI, T2WI, CE-T1WI, T1WI+T2WI, T1WI+CE-T1WI, T2WI+CE-T1WI, and T1WI+T2WI+CE-T1WI were 0.629, 0.829, 0.771, 0.879, 0.643, 0.829, and 0.929, respectively (3). The AUC values of the clinical, combined 1 (clinical and LR-radiomics) and combined 2 (clinical and SVM-radiomics) predictive models were 0.779, 0.957, and 0.943, respectively.

Conclusion: The combined model exhibited good performance in predicting osteosarcoma SLM and may be helpful in clinical decision-making.

Keywords: radiomics, predictive value of tests, magnetic resonance imaging, osteosarcoma, metastasis
INTRODUCTION

Osteosarcoma is a highly prevalent primary bone malignancy. Fortunately, complete ablation of nonmetastatic high-grade osteosarcoma is possible in 60–70% of cases when treated with adjuvant and neoadjuvant multiagent chemotherapies in addition to surgery (1). However, the prognoses of osteosarcoma patients with distant metastasis remain poor. Among all forms of metastasis, lung metastasis is the most common, occurring in over 80% of patients. Approximately 20% of osteosarcoma patients also exhibit metastasis at initial diagnosis (synchronous metastases) (2, 3). The primary tumor is more resistant to chemotherapy in patients with synchronous metastases than in patients with localized disease at presentation (4). Both the number of nodules and lobes are strong indicators of survival (5). At present, the best indicators of survival are tumor grade, tumor size, and distal metastases, which can be detected from biopsies and microscopic evaluations (6). Predicting individual and early metastases is essential to osteosarcoma management, as it informs treatment strategies and increases survival rates. Chest computerized tomography (CT) has been the most commonly used imaging modality for the detection of lung nodules. Although there have been great advancements in imaging technology, particularly in enhancing the sensitivity of detection, the specificity of the data remains insufficient. Metastases cannot be properly distinguished from benign tissue (5). When nodules are detected at diagnosis, it is usually assumed that these nodules represent metastatic disease. However, not all pulmonary nodules that develop during tumor therapy are malignant, which poses additional challenges for physicians. Hence, the goal of this study was to evaluate the diagnostic abilities of two distinct classifiers (logistic regression (LR) and support vector machine (SVM)) and radiomics features retrieved from different magnetic resonance imaging (MRI) parameters, including T1-weighted imaging (T1WI), T2-weighted imaging (T2WI) and contrast-enhanced T1-weighted imaging (CE-T1WI), and the combinations of two and three of these parameters. We also developed and validated combined models according to multiparametric MRI and clinical features to predict synchronous lung metastases (SLM) in osteosarcoma.

MATERIALS AND METHODS

Our retrospective investigation was approved by the Institutional Review Board. Participant informed consent was waived due to the retrospective nature of the study.

Patient Selection

Overall, 360 patients who received MRI evaluations between January 1, 2014, and December 30, 2020, were recruited for this study. The following patients were included in the study: (i) patients with no history of surgical or medical treatment administered for suspected osteosarcoma; (ii) patients who underwent multiparametric MRI, including T1WI, T2WI, and CE-T1WI, prior to treatment; (iii) patients with a osteosarcoma diagnosis confirmed by surgical resection or CT/ultrasound-guided needle biopsy and histopathological results; and (iv) patients diagnosed with SLM according to follow-up chest CT or confirmed by pathology. The patients in this study had lung nodules, and the possibility of viral, bacterial or fungal infection was ruled out. Three criteria were used to identify SLM lung nodules on follow-up chest CT according to previous studies (7, 8): first, the presence of multiple round nodules with or without changes in size or number; second, nodule size ≥5 mm and the presence of calcifications or ossification that remained stable or increased in size relative to the initial chest CT; third, changes in size or morphology during chemotherapy. Fifteen and 18 patients with SLM were diagnosed by biopsy and follow-up chest CT, respectively. The following patients were excluded from this study: (i) patients who received biopsy and locoregional therapy before MRI; (ii) patients with low-quality images rendering analysis difficult (such as images with metallic artefacts or motion artefacts); and (iii) patients with missing images or relevant sequences. A schematic diagram of our patient selection process is provided in Figure 1. After the application of these criteria, 78 patients were eligible for this study. The clinical characteristics of the 78 osteosarcoma patients divided into non-SLM and SLM groups are shown in Table 1. We next arbitrarily divided the patients into two populations: 54 patients were placed in the training cohort (TC) and 24 were placed in the validation cohort (VC) based on the seed point set obtained from programming. The clinical characteristics of the 78 osteosarcoma patients in the TC and VC are shown in Table 2, and a further breakdown of the clinical characteristics of these cohorts in terms of SLM and lack of SLM of osteosarcoma are summarized in Table 3.

Patient clinical features, such as age, sex, tumor size, pathological type, tumor location, bone destruction type, and alkaline phosphatase (ALP) and lactate dehydrogenase (LDH) levels, were recorded.

MR Imaging

All MR imaging was conducted with 1.5- or 3.0-T superconducting magnet systems. The imaging sequences included axial T1WI, T2WI and CE-T1WI. The detailed scan parameters of the four MRI scanners are described in Table 4. Gadolinium contrast agent was intravenously administered via a weight-based dosing protocol (0.1 mmol/kg) at an injection rate of 2.5 mL/s. All the MR data were obtained from the picture archiving and communication system (PACS) of our institutes and stored in Digital Imaging and Communications in Medicine (DICOM) format for additional analyses.

Preprocessing of MR Images

All the images were exported to ITK-SNAP software (version 3.8.0, http://www.itksnap.org/) for segmentation before radiomics analysis. Lesion segmentation was performed by a radiologist with over 5 years of MRI diagnostic experience, and proper segmentation was further confirmed by a separate radiologist with over 10 years of MRI diagnostic experience. If disagreements arose about a specific image segmentation, a revision was made by two radiologists after discussion. The
segmentation for 24 randomly selected patients was then repeated by another radiologist (over 10 years of experience). A separate region of interest (ROI) was manually selected for all sequences on each axial T1WI, T2WI and CE-T1WI slice. All the images were acquired without fat suppression. Delineation of the ROI, including the entire tumor and necrotic areas, cyst degeneration, hemorrhage, periosteal reactions, and peritumoral oedema, was carried out on the images from each sequence. Figure 2 shows an example of a segmented MRI image. Image intensity normalization was performed before feature extraction, including image gray normalization to uniform grayscale of 0–255 and resampling to 1 mm × 1 mm ×
1 mm voxel size using linear interpolation by AK software (Analysis Kit; GE Healthcare).

**Radiomic Feature Extraction**

In total, 944 radiomic features, quantifying phenotypic differences on the basis of shape (reflecting the size and shape of tumors), first-order (measuring the signal intensity of different tumors), and texture features (representing the relationship between each tumor voxel and its surrounding environments) (9), were automatically extracted from each segmented region of interest by using in-house software written in Python (Pyradiomics version: V 3.0; https://github.com/Radiomics/pyradiomics) (10). All the features were calculated in 3D directions within the whole-tumor volume and normalized by transforming the data into standardized intensity ranges (z-score transformation). Intraclass correlation coefficients (ICCs) based on a multiple-rating, consistency, 2-way random-effects model were calculated to assess the stability and reproducibility of radiomic features. For both tumor ROIs, only features with an ICC > 0.75 were considered to suggest good agreement and retained for further radiomic feature selection.

**Radiomics Feature Selection**

Radiomics features were automatically calculated with the noncommercial Analysis Kit (A.K. GE Healthcare). First, we performed least absolute shrinkage and selection operator (LASSO) regression on all features to grossly choose attributes with discriminative ability. The goal was to reduce certain attribute coefficients to zero by regulating parameter $\lambda$. Subsequently, the area under the receiver operating characteristic (ROC) curve (AUC) could be determined versus log($\lambda$) by employing tenfold cross-validation. The advantage of the LASSO technique is that it can analyse a massive amount of radiomics characteristics from low numbers of samples. Second, we applied multivariate logistic regression to select the most predictive features.

**Machine Learning Model**

This study used two machine learning classifiers: LR and SVM. An SVM model was generated based on the established optimal feature subsets of the TC dataset. The kernel, gamma, degree, coef, and C parameters were set to ‘rbf’, 0.0, 3, 0.0, and 1.0, respectively.

The individual sequence models were constructed by T1WI, T2WI and CE-T1WI. Next, four combined models were generated via a combination of features of dissimilar sequences, namely, T1WI+T2WI, T1WI+CE-T1WI, T2WI+CE-T1WI, and T1WI+T2WI+CE-T1WI. Clinical features were analysed by univariate analysis, and variables for which $P < 0.05$ were entered into the clinical model.

Two combined models were constructed by combining the best LR and SVM radiomics models with clinical features. The models were conditioned with the TC using the repeated 10-fold cross-validation technique, and their performance was assessed in the VC.

The radiomics framework of our study is shown in Figure 3.

**Statistical Analysis**

The t-test or Mann-Whitney U-test was employed for the comparison of continuous variables, whereas the chi-squared test or Fisher’s exact test was employed for the comparison of intergroup categorical variables. All the statistical analyses were two-sided, and a Bonferroni-corrected $P$ value was employed to determine the feature significance of multiple comparisons. ROC curves were generated to assess the performance of the machine learning models, and the sensitivity, specificity and AUC values were calculated. The AUCs of each two models (clinical, radiomics and combined models) in the two cohorts were compared by using the DeLong test. All the data analyses were performed in R 3.5.1 and Python 3.5.6. A two-tailed $P$ value < 0.05 was set as the significance threshold.

**TABLE 1 | Clinical characteristics of 78 cases of osteosarcoma.**

| Characteristic       | Non-SLM | SLM     | $P$ value |
|----------------------|---------|---------|-----------|
| Sex                  |         |         |           |
| Female               | 15 (33.33%) | 14 (42.42%) | 0.412     |
| Male                 | 30 (66.67%) | 19 (57.58%) |           |
| Pathology            |         |         |           |
| Osteoblastic         | 34 (75.56%) | 25 (75.76%) | 0.984     |
| Others               | 11 (24.44%) | 8 (24.24%) |           |
| Location             |         |         |           |
| Femur                | 29 (64.44%) | 17 (51.52%) | 0.486     |
| Tibia                | 8 (17.78%) | 7 (21.21%) |           |
| Others               | 8 (17.78%) | 9 (27.27%) |           |
| Bone destruction     |         |         |           |
| Mx                   | 21 (46.67%) | 14 (42.42%) | 0.067     |
| Osteolytic           | 22 (48.89%) | 12 (36.36%) |           |
| Osteoblastic         | 2 (4.44%) | 7 (21.21%) |           |
| Age (years)          | 19.49 ± 13.86 | 16.45 ± 7.53 | 0.268     |
| Tumor size (cm)      | 6.31 ± 1.32 | 8.09 ± 2.39 | <0.001*   |
| ALP (IU/L)           | 758.49 ± 2286.19 | 913.30 ± 1659.41 | 0.742     |
| LDH (IU/L)           | 256.81 ± 105.03 | 347.63 ± 312.71 | 0.073     |

SLM, synchronous lung metastases; ALP, alkaline phosphatase; LDH, lactate dehydrogenase.

* $p < 0.05$. 
RESULTS

Clinical Characteristics of the Patients

In total, 78 osteosarcoma patients (49 males, 29 females; between 15-83 years of age) were recruited. Based on our univariate analysis, the tumor size was markedly different between the two groups (P < 0.05) (Table 1). No obvious differences were observed in age, sex, pathological type, tumor location, bone destruction type, and ALP or LDH levels between the SLM and non-SLM groups. Moreover, no marked differences were observed between the TC and VC (Table 2). In addition, the clinical features were not markedly different between the SLM and non-SLM cohorts, except for age in the VC and tumor size in the TC (Table 3).

Performance of the Radiomics Models

Overall, 944 radiomics features were obtained from each of the T1WI, T2WI and CE-T1WI images. A total of 702, 839 and 835 radiomics features from T1WI, T2WI and CE-T1WI were included, respectively, with ICC greater than 0.75. The radiomics features with the largest differences between the models are summarized in Table 5.

In terms of a distinct sequence in the LR classifier, CE-T1WI features displayed a stronger predictive performance (AUC = 0.87, 95% CI, 0.655-0.965) than T2WI (AUC = 0.85, 95% CI, 0.699-0.981) and T1WI (AUC = 0.686, 95% CI, 0.488-0.873) features in the VC. In terms of combined features, T1WI+T2WI+CE-T1WI had a higher performance (AUC = 0.914, 95% CI, 0.776-0.998) than T1WI+CE-T1WI (AUC = 0.736, 95% CI, 0.533-0.902), T2WI+CE-T1WI (AUC =...
0.85, 95% CI, 0.699-0.981) and T1WI+T2WI (AUC = 0.879, 95% CI, 0.746-0.993) (Tables 6, 7 and Figures 4, 5).

Delong-test results in Table 8 showed that there were significant differences between predictive performance of T1WI-radiomic model and that of T1WI+T2WI+CE-T1WI radiomic model in both cohorts.

In terms of a distinct sequence in the SVM classifier, T2WI features were more enhanced (AUC = 0.829, 95% CI, 0.621-0.950), compared to CE-T1WI (AUC = 0.771, 95% CI, 0.556-0.916) and T1WI (AUC = 0.629, 95% CI, 0.409-0.815) features in the VC. In terms of the combined features, T1WI+T2WI+CE-T1WI had a higher performance (AUC = 0.929, 95% CI, 0.746-0.993) than T1WI+CE-T1WI (AUC = 0.643, 95% CI, 0.423-0.826), T2WI+CE-T1WI (AUC = 0.829, 95% CI, 0.621-0.950) and T1WI+T2WI (AUC = 0.879, 95% CI, 0.681-0.975) (Tables 9, 10 and Figures 6, 7).

Delong-test results in Table 11 showed that there were significant differences between predictive performance of T1WI-radiomic model and that of T1WI+T2WI+CE-T1WI radiomic model in validation cohort.

Based on our univariate analysis, marked differences were observed in tumor size between the non-SLM and SLM sets (P < 0.05). Thus, the clinical model was built using tumor size alone, and this model performed well in the TC (AUC = 0.75, 95% CI,
0.613-0.858) and VC (AUC = 0.779, 95% CI, 0.564-0.921). When tumor size was combined with radiomics features, the combined model achieved enhanced prediction compared to the clinical model. The first combined model involving LR-radiomics features + clinical features had an AUC of 0.938 (95% CI, 0.838-0.986) in the TC and 0.957 (95% CI, 0.787-0.999) in the VC. The second combined model based on SVM-radiomics features + clinical features had an AUC of 0.944 (95% CI, 0.845-0.988) in the TC and 0.943 (95% CI, 0.766-0.997) in the VC (Tables 12, 13 and Figures 8, 9).

Delong-test results in Table 14 indicated that there were insignificant differences between predictive performance of combined mode and that of clinical model, LR-Radiomic model and SVM-Radiomic model in both cohorts. The combined model 1 achieved equivalent AUCs compared to combined model 2. Additionally, the combined model 1 performed better than the clinical model as indicated by AUCs of borderline statistical significance (p = 0.0619) in the validation cohort.

**DISCUSSION**

According to our univariate analysis, tumor size was the most reliable indicator of SLM in osteosarcoma patients, which is in accordance with prior findings (11–14). Huang et al (11) performed a retrospective study examining the characteristics of 1057 osteosarcoma patients. These authors reported that large tumors (>5 cm) were at a substantially elevated risk of resulting in lung metastases in osteosarcoma patients. Munajat et al. (13) also examined the correlation between lung metastasis and tumor volume in a population of 70 osteosarcoma patients. These authors reported that 33 patients (47%), who mostly exhibited larger tumor volumes, also showed signs of lung metastasis. However, in contrast to our study, these authors primarily concentrated on lung metastasis without distinguishing synchronous and metachronous metastases. In our study, we compared tumor size (in terms of diameter) between the non-SLM (6.31 ± 1.32 cm) and SLM (8.09 ± 2.39 cm) groups and found that the tumor size in the SLM group was significantly larger than that in the non-SLM group.

Tumor heterogeneity strongly modulates tumor invasion and prognosis, and a radiomics profile can specifically reflect the complicated histopathology of tumor (15, 16). Several MRI radiomics studies conducted on osteosarcoma were recently reported. Chen H et al. (17, 18) proved that a radiomics signature based on MRI was useful for predicting the response to neoadjuvant chemotherapy and early relapse. Zhao SL et al. (19) showed that a radiomics signature extracted from diffusion-weighted imaging (DWI)-MRI prior to treatment improved the estimation of osteosarcoma.

T1WI, T2WI, and CE-T1WI are the most commonly used MRI sequences for bone tumor. T1WI can be used to observe anatomical structures, but sometimes it is difficult to distinguish soft tissue masses from muscle tissues. T2WI can accurately determine tumor margins and reveal, to a certain extent, the total lesion cell density, whereas CE-T1WI can reveal lesion vascularity, establish the degree of malignancy, and identify
necrosis and solid components of tumor. Current research has shown that multiple MRI sequences, such as DWI, T2WI/FS-T2WI and CE-T1WI images, can enhance tumor information extraction and thereby augment specificity, and such sequences for bone and soft tissue tumor radiomics analysis yielded favorable outcomes (20–23). In addition, several studies revealed that classifiers originating from varying classifier families exhibit differing performances for different forms of tumors (24–28). We used two well-known machine learning classifiers in this study. The first is LR, which is a machine learning stratification algorithm used for the prediction of the class probability of a given categorical dependent variable. The second is SVM, which generates a decision margin between two classes to facilitate label estimation from one or more feature vectors. Seven radiomics models using these three sequences alone and combined were established in our study (29). Differences in tumor vessel morphology affect tumor vascular permeability (30). Increased vessel permeability may accelerate cancer metastasis and spread; in the absence of blood vessels, tumors cannot develop beyond a critical volume or invade other organs (31). Enhancing MRI in neoplasms represents regions of admixed vascularity and necrosis, in which contrast permeability is elevated owing to damaged vascular integrity (32). Among the seven radiomics modes, the LR and SVM classifiers had excellent performance (AUC=0.914 and 0.929) in the model combining T1WI, T2WI and CE-T1WI. The combination of T1WI, T2WI, and CE-T1WI allowed the detection of morphological information and indirectly reflected

| Model      | Radiomics features                       | Coef.  |
|------------|------------------------------------------|--------|
| T1WI       | Intercept                                | -0.4597|
|            | T1WI_wavelet-LLL_glcmm_Correlation       | 1.3060 |
|            | T1WI_wavelet-LLL_gldm_GrayLevelNonUniformity | 0.8114 |
| T2WI       | Intercept                                | -1.6745|
|            | T2WI_wavelet-LLL_glcmm_Correlation       | 3.4374 |
|            | T2WI_wavelet-HHL_firstorder_Mean         | -2.7065|
|            | T2WI_wavelet-HHL_gldm_MCC               | -3.5169|
|            | T2WI_wavelet-LHL_gldm_LargeDependenceHighGrayLevelEmphasis | 2.1317 |
| CE-T1WI    | Intercept                                | -1.2416|
|            | CE-T1WI_wavelet-LLL_glcmm_Correlation   | 1.6253 |
|            | CE-T1WI_wavelet-LHL_firstorder_Mean     | -1.4719|
|            | CE-T1WI_wavelet-HHL_firstorder_Skewness | -1.7320|
|            | CE-T1WI_wavelet-HHL_gldm_MCC            | -1.3359|
|            | CE-T1WI_wavelet-LHH_firstorder_Kurtosis  | 1.1688 |
| T1WI+T2WI  | Intercept                                | -1.0635|
|            | T2WI_wavelet-LLL_glcmm_Correlation       | 2.2998 |
|            | T2WI_wavelet-HHL_firstorder_Mean         | -1.5630|
|            | T2WI_log-sigma-1-0-mm-3D_ngtdm_Busyness  | 1.1760 |
|            | T2WI_wavelet-HHL_gldm_SmallDependenceLowGrayLevelEmphasis | -1.2490 |
| T1WI+CE-T1WI | Intercept                               | -0.8489|
|            | T1WI_wavelet-LLL_glcmm_Correlation       | 1.2206 |
|            | T1WI_wavelet-LHL_firstorder_Mean         | -1.6295|
|            | T1WI_wavelet-HHL_firstorder_Skewness     | -1.1276|
| T2WI+CE-T1WI | Intercept                               | -1.6745|
|            | T2WI_wavelet-LLL_glcmm_Correlation       | 3.4374 |
|            | T2WI_wavelet-HHL_firstorder_Mean         | -2.7065|
|            | T2WI_wavelet-HHL_gldm_MCC               | -3.5169|
|            | T2WI_wavelet-LHL_gldm_LargeDependenceHighGrayLevelEmphasis | 2.1317 |
| T1WI+T2WI+CE-T1WI | Intercept               | -1.2077|
|            | T2WI_wavelet-LLL_glcmm_Correlation       | 2.4347 |
|            | T2WI_wavelet-HHL_firstorder_Mean         | -1.6936|
|            | T2WI_wavelet-HHL_gldm_MCC               | -1.5491|

| Classifiers | Model       | AUC   | 95% CI        | Sensitivity | Specificity |
|-------------|-------------|-------|---------------|-------------|-------------|
| LR          | T1WI        | 0.795 | 0.663 - 0.893 | 0.565       | 0.806       |
|             | T2WI        | 0.951 | 0.855 - 0.991 | 0.826       | 0.884       |
|             | CE-T1WI     | 0.909 | 0.799 - 0.970 | 0.870f      | 0.871       |
|             | T1WI+T2WI   | 0.937 | 0.836 - 0.985 | 0.783       | 0.903       |
|             | T1WI+CE-T1WI| 0.846 | 0.722 - 0.930 | 0.739       | 0.774       |
|             | T2WI+CE-T1WI| 0.951 | 0.855 - 0.991 | 0.826       | 0.864       |
|             | T1WI+T2WI+CE-T1WI | 0.940 | 0.840 - 0.986 | 0.913       | 0.903       |

T1WI, T1-weighted imaging; T2WI, T2-weighted imaging; AUC, area under curve; 95% CI, 95% confidence interval; LR, logistic regression.
the permeability of tissue microvessels. Although the SVM classifier in the T1WI, T2WI and CE-T1WI combined model performed the best (AUC = 0.929), no obvious difference was observed between the two classifiers of each radiomics model.

In addition, among all radiomics models, those including the T2WI parameter (T2WI, T2WI+CE-T1WI, T1WI+T2WI, and T1WI+T2WI+CE-T1WI) exhibited excellent performance (AUC=0.829-0.929). We found that all of the final features that included the T2WI parameter after selection contained ‘GLCM_Correlation’ and ‘firstorder_Mean’ features, which were high-dimensional features that could not be readily interpreted by humans and included comprehensive tumor information. Among these features, the mean, which is a first-order feature, assesses the average grey level intensity within a specified area of interest. The grey level cooccurrence matrix (GLCM) is a second-order feature and is a summary of the frequency of the various combinations of pixel brightness values that occur between neighboring voxels in a given image. GLCM represents the similarity of voxel values along a given direction, whereas homogeneity represents regional grey level uniformity.
and correlation establishes the consistency of image texture (33, 34). In prior studies, these features established tumor heterogeneity and were correlated with the histopathological characteristics and prognosis of numerous tumors, such as osteosarcoma, rectal cancer, thymic tumors, and breast cancer (18, 25, 35–37).

Furthermore, we analysed seven discrete radiomics models, two combined models and one clinical model in osteosarcoma patients. The AUC of the clinical model (0.779) was lower than that of the radiomics model. The prediction ability of the combined model was markedly enhanced relative to that of other models, namely, the clinical model and radiomics model, which involved multiple and single sequences. As clinical information may take into account only some aspects of tumors, multiparametric MRI may better reflect of all tumor information (38). Hence, once the clinical and radiomics characteristics were combined, the performance greatly improved. Based on our data, machine learning analysis involving multiparametric MRI radiomics characteristics can accurately and efficiently predict SLM in osteosarcoma.

Our work encountered certain limitations. First, this work was retrospective in nature. Radiomic features are heavily influenced by differences in the acquisition and reconstruction settings. In the present study, although the most commonly used MRI sequences (T1WI, T2WI and CE-T1WI) were selected, MRI image data were acquired from four distinct scanners, which can also influence the acquired characteristics. Device inconsistency within a dataset is a challenging issue, particularly in retrospective analyses. Second, the sample size was quite small, and all the obtained images were collected over several years. We eliminated non-long bone extremity osteosarcoma and patients who did not receive MRI and chest CT prior to surgery, which accounted for the majority of the enrolled patients. In addition, few osteosarcomas simultaneously met both the axial plane and multiparametric MRI requirements. Due to our strict criteria for patient eligibility, it was difficult to gather large datasets. Despite a statistically insufficient sample size, the results of this study may allow for the improvement of future clinical studies with limited sample sizes. A large sample population with multicentre validation is warranted to achieve high-level evidence for future studies.

### Table 9 | The ROC curve of different models of SVM-classifier in the training cohort.

| Classifiers | Model         | AUC 95% CI | Sensitivity | Specificity |
|-------------|---------------|------------|-------------|-------------|
| SVM         | T1WI          | 0.829      | 0.702 - 0.918 | 0.957       | 0.677 |
|             | T2WI          | 0.973      | 0.888 - 0.998 | 1.000       | 0.938 |
|             | CE-T1WI       | 0.935      | 0.834 - 0.984 | 1.000       | 0.871 |
|             | T1WI+T2WI     | 0.930      | 0.826 - 0.981 | 0.957       | 0.839 |
|             | T1WI+CE-T1WI  | 0.885      | 0.769 - 0.956 | 0.783       | 0.871 |
|             | T2WI+CE-T1WI  | 0.973      | 0.888 - 0.998 | 1.000       | 0.839 |
|             | T1WI+T2WI+CE-T1WI | 0.938 | 0.838 - 0.986 | 0.957       | 0.903 |

### Table 10 | The ROC curve of different models of SVM-classifier in the validation cohort.

| Classifiers | Model         | AUC 95% CI | Sensitivity | Specificity |
|-------------|---------------|------------|-------------|-------------|
| SVM         | T1WI          | 0.629      | 0.409 - 0.815 | 1.000       | 0.429 |
|             | T2WI          | 0.829      | 0.621 - 0.950 | 0.800       | 0.786 |
|             | CE-T1WI       | 0.771      | 0.556 - 0.916 | 1.000       | 0.500 |
|             | T1WI+T2WI     | 0.879      | 0.681 - 0.975 | 0.800       | 0.867 |
|             | T1WI+CE-T1WI  | 0.643      | 0.423 - 0.826 | 0.800       | 0.500 |
|             | T2WI+CE-T1WI  | 0.829      | 0.621 - 0.950 | 0.800       | 0.786 |
|             | T1WI+T2WI+CE-T1WI | 0.929 | 0.746 - 0.993 | 0.900       | 0.857 |
FIGURE 6 | SVM-classifier in the training cohort.

FIGURE 7 | SVM-classifier in the validation cohort. Figures 4–7 The ROC curve of different models and classifier in the training and validation cohorts.

TABLE 11 | Delong Test between each two models of SVM-classifier in the training and validation cohorts.

| Radiomic model | T1WI | T2WI | CE-T1WI | T1WI+T2WI | T1WI+CE-T1WI | T2WI+CE-T1WI | T1WI+T2WI+CE-T1WI |
|----------------|------|------|---------|-----------|--------------|--------------|-------------------|
| T1WI           | –    | 0.0084* | 0.1247 | 0.1305    | 0.2795       | 0.0084*       | 0.0956            |
| T2WI           | 0.0928 | –    | 0.2589 | 0.1822    | 0.0449*      | 1.0000       | 0.2545            |
| CE-T1WI        | 0.2519 | 0.4948 | –       | 0.8594    | 0.3360       | 0.2589       | 0.9250            |
| T1WI+T2WI      | 0.0306* | 0.4472 | 0.2551  | –         | 0.3718       | 0.1822       | 0.5620            |
| T1WI+CE-T1WI   | 0.8881 | 0.1217 | 0.1746  | 0.0203*   | –            | 0.0449*       | 0.2863            |
| T2WI+CE-T1WI   | 0.0928 | 1.0000 | 0.4948  | 0.4472    | 0.1217       | –            | 0.2545            |
| T1WI+T2WI+CE-T1WI | 0.0079* | 0.1404 | 0.0912  | 0.3870    | 0.0125*      | 0.1404       | –                 |

Training cohort Validation cohort

T1WI, T1-weighted imaging; T2WI, T2-weighted imaging; CE, contrast-enhanced. *p < 0.05.

TABLE 12 | The ROC curve of clinical features, radiomic, clinical features + radiomic model in the training cohort.

| Model                        | AUC   | 95% CI         | Sensitivity | Specificity |
|------------------------------|-------|----------------|-------------|-------------|
| Clinical model               | 0.750 | 0.613 - 0.858  | 0.696       | 0.839       |
| LR-radiomic                  | 0.940 | 0.840 - 0.986  | 0.913       | 0.903       |
| SVM-radiomic                 | 0.938 | 0.838 - 0.986  | 0.957       | 0.903       |
| Combined 1                   | 0.938 | 0.838 - 0.986  | 0.957       | 0.903       |
| (clinical+LR-radiomic)       |       |                |             |             |
| Combined 2                   | 0.944 | 0.845 - 0.988  | 0.956       | 0.900       |
| (clinical+SVM-radiomic)      |       |                |             |             |

AUC, area under curve; LR, logistic regression; 95% CI, 95% confidence interval; SVM, support vector machine.
**TABLE 13** | The ROC curve of clinical features, radiomic, clinical features + radiomic model in the validation cohort.

| Model                          | AUC    | 95% CI       | Sensitivity | Specificity |
|-------------------------------|--------|--------------|-------------|-------------|
| Clinical model                | 0.779  | 0.564 - 0.921| 0.600       | 0.929       |
| LR-radiomic                   | 0.914  | 0.776 - 0.998| 0.700       | 0.929       |
| SVM-radiomic                  | 0.929  | 0.746 - 0.993| 0.900       | 0.857       |
| Combined 1 (clinical+LR-radiomic) | 0.957  | 0.787 - 0.999| 1.000       | 0.857       |
| Combined 2 (clinical+SVM-radiomic) | 0.943  | 0.766 - 0.997| 0.846       | 0.929       |

AUC, area under curve; LR, logistic regression; 95% CI, 95% confidence interval; SVM, support vector machine.

**FIGURE 8** | The training cohort.

**FIGURE 9** | The validation cohort. **Figures 8, 9** The ROC curve of clinical features, radiomic, clinical features + radiomic model in the training and validation cohorts.

**TABLE 14** | Delong Test between each two models (clinical, radiomic, clinical features + radiomic model) in the training and validation cohort.

| Model                          | Clinical | LR-Radiomic | SVM-Radiomic | Clinical +LR-Radiomic | Clinical +SVM-Radiomic |
|-------------------------------|----------|-------------|--------------|-----------------------|------------------------|
| Clinical                      | –        | 0.0140*     | 0.0161*      | 0.0138*               | 0.0085*                |
| LR-Radiomic                   | 0.0966   | –           | 0.01065      | 0.7927                | 0.9065                |
| SVM-Radiomic                  | 0.1361   | 0.5247      | –            | 1.0000                | 0.8908                |
| Combined 1                    | 0.0619   | 0.7807      | 0.2536       | –                     | 0.8798                |
| Combined 2                    | 0.0787   | 0.8381      | 0.6251       | 0.6265                | –                     |

Training cohort Validation cohort

T1WI, T1-weighted imaging; T2WI, T2-weighted imaging; CE, contrast-enhanced; LR, logistic regression; SVM, support vector machine. *p < 0.05.
clinical application. Third, we compared only T1WI, T2WI and CE-T1WI sequences of MRI. The most commonly used clinical examination for osteosarcoma is X-ray, and the effectiveness of X-ray radiomics needs to be investigated in future studies. Additional MRI functional data must be included in future evaluations to enhance the accuracy and clinical value of our model.

In conclusion, we established a noninvasive predictive tool involving radiomics and clinical characteristics to predict SLM in osteosarcoma patients. The LR and SVM classifiers exhibited an elevated degree of diagnostic performance while employing a combination of characteristics for distinguishing SLM and lack of SLM in osteosarcoma patients. Among all radiomics models, those including the T2WI parameter exhibited good predictive performance for the prediction of osteosarcoma SLM. The constructed model involving the combination of clinical and radiomics characteristics is more effective in evaluating osteosarcoma SLM relative to the clinical model and radiomics model, and the constructed model can provide a new basis for early clinical intervention in metastasis.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

REFERENCES

1. Ritter J, Bielack S. Osteosarcoma. J Am Oncol Off J Eur Soc Med Oncol (2010) (Supplement 7):320–5. doi: 10.1093/anonic/mdq276

2. Meyers P, Heller G, Healey J, Huvos A, Applewhite A, Sun M, et al. Osteogenic Sarcoma With Clinically Detectable Metastasis at Initial Presentation. J Clin Oncol Off J Am Soc Clin Oncol (1993) 11(3):449–53. doi: 10.1200/jco.1993.11.3.449

3. Kager L, Zoubek A, Pötschger U, Kastner U, Flege S, Kempf-Bielack B, et al. Primary Metastatic Osteosarcoma: Presentation and Outcome of Patients Treated on Neoadjuvant Cooperative Osteosarcoma Study Group Protocols. J Clin Oncol Off J Am Soc Clin Oncol (2003) 21(10):2011–8. doi: 10.1200/jco.2003.08.132

4. Bacci G, Picci P, Brizzi A, Avella M, Ferrari S, Femino F, et al. Osteosarcoma of the Extremity Metastatic at Presentation: Results Achieved in 26 Patients Treated With Combined Therapy (Primary Chemotherapy Followed by Simultaneous Resection of the Primary and Metastatic Lesions). Tumori (1992) 78(3):200–6. doi: 10.1177/030089169207800311

5. Kaste S, Pratt C, Cain A, Jones-Wallace D, Rao B. Metastases Detected at the Time of Diagnosis of Primary Pediatric Extremity Osteosarcoma: Results of a Multicenter Study. J Bone Oncol (2016) 5:200306. doi: 10.1016/j.jbono.2019.05.003

6. Sheen H, Kim W, Byun B, Kong C, Song W, Cho W, et al. Metastases and Overall Survival After Neoadjuvant Chemotherapy and Surgery for Patients With Nonmetastatic Osteosarcoma: A Multi-Institutional Study. Cancer (2015) 121(21):3844–52. doi: 10.1002/cncr.289575

7. Brader P, Abramson S, Price A, Ishill N, Emily Z, Moskowitz C, et al. Do Characteristics of Pulmonary Nodules on Computed Tomography in Children With Known Osteosarcoma Help Distinguish Whether the Nodules Are Malignant or Benign? J Pediatr Surg (2011) 46(4):729–35. doi: 10.1016/j.jpedsurg.2010.11.027

8. Chiesa A, Spinnato P, Miceli M, Facchini G. Radiologic Assessment of Osteosarcoma Lung Metastases: State of the Art and Recent Advances. Cells (2021) 10(3):553. doi: 10.3390/cells10030553

ETHICS STATEMENT

Written informed consent was not obtained from the individual(s), nor the minor(s)’ legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

ZL, JL, XS, and WC contributed to the design and implementation of the concept. ZL and JL contributed equally to this work. JL and ZL contributed in collecting and reviewing patients’ clinical and imaging data. RL and ZL contributed in segmenting the lesions. YL and RL contributed in the statistical analysis of the data. ZL, YL, and RL contributed in building models. All authors contributed to the writing and reviewing of the paper. All authors read and approved the final manuscript.

FUNDING

High Level-Hospital Program, Health Commission of Guangdong Province, China (No: HKUSZH201901026).

9. Fang J, Zhang B, Wang S, Jin Y, Wang F, Ding Y, et al. Association of MRI-Derived Radiomic Biomarker With Disease-Free Survival in Patients With Early-Stage Cervical Cancer. Theranostics (2020) 10(5):2284–92. doi: 10.7150/thno.37429

10. van Griethuysen I, Fedorov A, Parmar C, Honsy A, Aucoin N, Narayan V, et al. Computational Radiomics System to Decode the Radiographic Phenotype. Cancer Res (2017) 77(21):e104–107. doi: 10.1158/0008-5472.Can-17-0339

11. Huang X, Zhao J, Bai J, Shen H, Zhang B, Deng L, et al. Risk and Clinicopathological Features of Osteosarcoma Metastasis to the Lung: A Population-Based Study. J Bone Oncol (2019) 16:100230. doi: 10.1016/j.jbono.2019.100230

12. Kim S, Shin K, Kim H, Cho Y, Noh J, Suh J, et al. Postoperative Nomogram to Predict the Probability of Metastasis in Enneking Stage IIB Extremity Osteosarcoma. BMC Cancer (2014) 14:666. doi: 10.1186/1471-2407-14-666

13. Munajat I, Zulmi W, Norazman M, Wan Faisham W. Tumour Volume and Nodules Are Malignant or Benign? J Pediatr Surg (2011) 46(4):729–35. doi: 10.1016/j.jpedsurg.2010.11.027

14. Wang S, Zheng S, Hu K, Sun H, Zhang J, Rong G, et al. A Predictive Model to Estimate the Pretest Probability of Metastasis in Patients With Osteosarcoma. Medicine (2017) 96(3):e5909. doi: 10.1097/MD.0000000000005909

15. Ogura K, Fujiwara T, Yasunaga H, Matsui H, Jeon D, Cho W, et al. Development and External Validation of Nomograms Predicting Distant Metastases and Overall Survival After Neoadjuvant Chemotherapy and Surgery for Patients With Nonmetastatic Osteosarcoma: A Multi-Institutional Study. Cancer (2015) 121(21):3844–52. doi: 10.1002/cncr.289575

16. Li H, Zhu Y, Burnsire E, Drukker K, Hoadley K, Fan C, et al. MR Imaging Radiomics Signatures for Predicting the Risk of Breast Cancer Recurrence as Given by Research Versions of MammaPrint, Oncotype DX, and PAM50 Gene Assays. Radiology (2016) 281(2):382–91. doi: 10.1148/radiol.2016152110

17. Chen H, Zhang X, Wang X, Quan X, Deng Y, Lu M, et al. MRI-Based Radiomics Signature for Pretreatment Prediction of Pathological Response to Neoadjuvant Chemotherapy in Osteosarcoma: A Multicenter Study. Eur Radiol (2021) 31(10):7913–24. doi: 10.1007/s00330-021-07748-6
18. Chen H, Liu J, Cheng Z, Lu X, Wang X, Lu M, et al. Development and External Validation of an MRI-Based Radiomics Nomogram for Pretreatment Prediction of Early Benign vs. Malignant Osteosarcoma: A Retrospective Multicenter Study. *Eur J Radiol* (2020) 129:109066. doi: 10.1016/j.ejrad.2020.109066

19. Zhao S, Su Y, Duan J, Qiu Q, Ge X, Wang A, et al. Radiomics Signature Extracted From Diffusion-Weighted Magnetic Resonance Imaging Predicts Outcomes in Osteosarcoma. *J Bone Oncol* (2019) 19:100263. doi: 10.1016/j.jbo.2019.100263

20. Dai Y, Yin P, Mao N, Sun C, Wu J, Cheng G, et al. Differentiation of Pelvic Osteosarcoma and Ewing Sarcoma Using Radiomic Analysis Based on T2-Weighted Images and Contrast-Enhanced T1-Weighted Images. *BioMed Res Int* (2020) 2020:9078603. doi: 10.1155/2020/9078603

21. Yin P, Mao N, Liu X, Sun C, Wang S, Chen L, et al. Can Clinical Radiomics Nomogram Based on 3D Multiparametric MRI Features and Clinical Characteristics Estimate Early Recurrence of Pelvic Chondrosarcoma? *J Magn Reson Imaging* JMRI (2020) 51(2):435–45. doi: 10.1002/jmri.26834

22. Zhong J, Hu Y, Si L, Jia G, Xing Y, Zhang H, et al. A Systematic Review of Radiomics in Osteosarcoma: Utilizing Radiomics Quality Score as a Tool Promoting Clinical Translation. *Eur Radiol* (2021) 31(3):1526–35. doi: 10.1007/s00330-020-07221-w

23. Wang H, Zhang J, Bao S, Liu J, Hou F, Huang Y, et al. Preoperative MRI-Based Radiomic Machine-Learning Nomogram May Accurately Distinguish Between Benign and Malignant Soft-Tissue Lesions: A Two-Center Study. *J Magn Reson Imaging* JMRI (2020) 52(3):873–82. doi: 10.1002/jmri.27111

24. Viswanath S, Chirra P, Yim M, Rofsky N, Purysoy A, Rosen M, et al. Comparing Radiomic Classifiers and Classifier Ensembles for Detection of Peripheral Zone Prostate Tumors on T2-Weighted MRI: A Multi-Site Study. *BMJ Med Imaging* (2019) 19(1):22. doi: 10.1186/s12880-019-0308-6

25. Hu J, Zhao Y, Li M, Liu Y, Wang F, Weng Q, et al. Machine-Learning-Based Computed Tomography Radiomic Analysis for Histologic Subtype Classification of Thymic Epithelial Tumours. *Eur J Radiol* (2020) 126:108929. doi: 10.1016/j.ejrad.2020.108929

26. Parmar C, Grossmann P, Rietveld D, Rietbergen M, Lambin P, Aerts H. Radiomic Machine-Learning Classifiers for Prognostic Biomarkers of Head and Neck Cancer. *Front Oncol* (2015) 5:272. doi: 10.3389/fonc.2015.00272

27. Delzetto D, Magnuson S, Peter T, Smith M, Smith B. Machine Learning and Feature Selection Methods for Disease Classification With Application to Lung Cancer Screening Image Data. *Front Oncol* (2019) 9:1393. doi: 10.3389/fonc.2019.01393

28. Zhang Y, Zhu Y, Shi X, Tao J, Cui J, Dai Y, et al. Soft Tissue Sarcomas: Preoperative Predictive Histopathological Grading Based on Radiomics of MRI. *Acad Radiol* (2019) 26(9):1262–8. doi: 10.1016/j.acra.2018.09.025

29. Nazari M, Shir I, Hajanfar G, Oveis N, Abdollahi H, Drevband M, et al. Noninvasive Fuhrman Grading of Clear Cell Renal Cell Carcinoma Using Computed Tomography Radiomic Features and Machine Learning. *La Radiol Med* (2020) 125(8):754–62. doi: 10.1007/s11547-020-01169-z

30. Roberts W, Delaat J, Nagane M, Huang S, Cavenee W, Palade G. Host Microvasculature Influence on Tumor Vascular Morphology and Endothelial Gene Expression. *Am J Pathol* (1998) 153(4):1239–48. doi: 10.1016/s0002-9440(10)65668-4

31. Carmeliet P, Jain R. Angiogenesis in Cancer and Other Diseases. *Nat Rev Cancer* (2000) 4(7):249–57. doi: 10.1038/35025220

32. Chan A, Fox J, Perez Johnston R, Brouwer L, Grizzard J, et al. Late Gadolinium Enhancement Cardiac Magnetic Resonance Tissue Characterization for Cancer-Associated Cardiac Masses: Metabolic and Prognostic Manifestations in Relation to Whole-Body Positron Emission Tomography. *J Am Heart Assoc* (2019) 8(10):e011709. doi: 10.1161/jaha.118.011709

33. Zhang L, Fried D, Fave X, Hunter L, Yan J, Court L. IBEX: An Open Infrastructure Software Platform to Facilitate Collaborative Work in Radiomics. *Med Phys* (2015) 42(3):1341–53. doi: 10.1118/1.4908210

34. Giannini V, Mazzetti S, Bertotto I, Chiarenza C, Cauda S, Delmastro E, et al. Predicting Locally Advanced Rectal Cancer Response to Neoadjuvant Therapy With F-FDG PET and MRI Radiomics Features. *Eur J Nucl Med Mol Imaging* (2019) 46(4):878–88. doi: 10.1007/s00259-018-4250-6

35. Nie K, Shi L, Chen Q, Hu X, Jabbour S, Yue N, et al. Rectal Cancer: Assessment of Neoadjuvant Chemoradiation Outcome Based on Radiomics of Multiparametric MRI. *Clin Cancer Res* (2016) 22(21):5256–64. doi: 10.1158/1078-0432.Ccr-15-2997

36. Liu S, Zheng H, Pan X, Chen L, Shi M, Guan Y, et al. Texture Analysis of CT Imaging for Assessment of Esophageal Squamous Cancer Aggressiveness. *J Thorac Dis* (2017) 9(11):4724–32. doi: 10.21037/td.2017.06.46

37. Xiong Q, Zhou X, Liu Z, Lei C, Yang C, Yang M, et al. Multiparametric MRI-Based Radiomics Analysis for Prediction of Breast Cancers Insensitive to Neoadjuvant Chemotherapy. *Clin Trans Oncol* (2020) 22(1):50–9. doi: 10.1007/s12094-019-02109-8

38. Gillies R, Kinahan P, Hricak H. Radiomics: Images Are More Than Pictures, They Are Data. *Radiology* (2016) 278(2):563–77. doi: 10.1148/radiol.2015151169

**Conflict of Interest:** Author YL was employed by GE Healthcare China.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher’s Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Luo, Li, Liu, Shen and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.