Research Paper

Survival Prediction in High-grade Osteosarcoma Using Radiomics of Diagnostic Computed Tomography

Yan Wu [a,d,1], Lei Xu [b,c,1], Pengfei Yang [b,c], Nong Lin [a], Xin Huang [a], Weibo Pan [a], Hengyuan Li [a], Peng Lin [a], Binghao Li [a], Varitsara Bunpetch [d], Chen Luo [b,c], Yangkang Jiang [b,c], Disheng Yang [a], Mi Huang [e], Tianye Niu [b,c,*], Zhaoming Ye [a,**]

[a] Department of Orthopaedics, The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, Zhejiang 310009, China
[b] Institute of Translational Medicine, Zhejiang University School of Medicine, Hangzhou, Zhejiang 310020, China
[c] Department of Radiology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang 310016, China
[d] Dr. Li Dak Sum & Yip Yio Chin Center for Stem Cell and Regenerative Medicine, Zhejiang University-University of Edinburgh Institute, Zhejiang University, Hangzhou, Zhejiang 310058, China
[e] Department of Radiation Oncology, Duke University Medical Center, Durham, NC 27708, USA

ABSTRACT

The poor 5-year survival rate in high-grade osteosarcoma (HOS) has not been increased significantly over the past 30 years. This work aimed to develop a radiomics nomogram for survival prediction at the time of diagnosis in HOS. In this retrospective study, an initial cohort of 102 HOS patients, diagnosed from January 2008 to March 2011, was used as the training cohort. Radiomics features were extracted from the pretreatment diagnostic computed tomography images. A radiomics signature was constructed with the lasso algorithm; then, a radiomics score was calculated to reflect survival probability by using the radiomics signature for each patient. A radiomics nomogram was developed by incorporating the radiomics score and clinical factors. A clinical model was constructed by using clinical factors only. The models were validated in an independent cohort comprising 48 patients diagnosed from April 2011 to April 2012. The performance of the nomogram was assessed with respect to its calibration, discrimination, and clinical usefulness. Kaplan–Meier survival analysis was performed.

The radiomics nomogram showed better calibration and classification capacity than the clinical model with AUC 0.86 vs. 0.79 for the training cohort, and 0.84 vs. 0.73 for the validation cohort. Decision curve analysis demonstrated the clinical usefulness of the radiomics nomogram. A significant difference (p-value <.05; log-rank test) was observed between the survival curves of the nomogram-predicted survival and non-survival groups. The radiomics nomogram may assist clinicians in tailoring appropriate therapy.

1. Introduction

Osteosarcoma is the most common primary bone malignancy, with an age-standardized incidence rate of 2.9 per 1 million men and 2.2 per 1 million women [1]. Nearly 90% of cases are classified as high-grade osteosarcoma (HOS) at the time of diagnosis [2]. Although the implementation of neoadjuvant and adjuvant chemotherapies and limb salvage surgeries has gradually increased the survival rate of HOS, the overall survival rate has not increased significantly over the past 30 years [3,4]. The 5-year overall survival rate for HOS ranges from 45% to 75% [5].

Although aggressive treatment plans, including multi-cycle treatments and adjuvant chemotherapies, are beneficial for patients who are likely to exhibit poor survival, not all HOS patients benefit from these treatments [6,7]. If patients with poor survival could be identified preoperatively, personalized treatment plans could be helpful for decision support for these patients. Therefore, there is a critical need to identify patients who are more likely to experience poor survival and thus benefit from additional therapy. Several clinical factors, such as age [8], tumor volume [9], stage [5,10], histologic subtype [11] and pathological fractures [12] have been associated with treatment outcome [13,14]. Nevertheless, a preoperative prognostic model for survival prediction has not yet been constructed. To address this issue, we built a reliable model to predict 5-year survival status at the time of diagnosis in HOS.

Recently, rapid developments in diagnostic imaging have become essential in the context of osteosarcoma decision-making in clinical practice; this especially includes computed tomography (CT) and magnetic resonance imaging (MRI). Notably, CT images can be used to

determine tumor size, location, and migration status [15]. Radiomics is an emerging field that converts medical images into a high-dimensional mineable feature space via high-throughput quantitative feature extraction [16,17]. Previous radiomics studies have shown that objective and quantitative radiomics features might serve as prognostic imaging biomarkers [18]. In the past 5 years, radiomics has been used in multiple aspects of the clinical assessment of tumors, including detection, diagnosis, curative effect, and prognosis [19–22]. These studies demonstrated the feasibility of developing a nomogram with radiomics features to predict 5-year survival status for patients with HOS.

Hence, this study aimed to develop and validate a survival prediction nomogram that incorporates both a radiomics signature and clinical risk factors at the time of diagnosis for individualized prediction of survival in patients with HOS. In addition, we compared prediction performance between the nomogram and a model built with clinical factors alone. To the best of our knowledge, this is the first study that used radiomics to model survival prediction at the time of diagnosis in HOS, based on CT images.

2. Material and Methods

2.1. Patients

This retrospective study was approved by the Institutional Review Board (IRB) of our institution, which waived the requirement for signed informed consent forms. A total of 150 patients with HOS, diagnosed from January 2008 to April 2012, were enrolled in this study, in accordance with the following inclusion and exclusion criteria. The inclusion criteria were as follows: (a) patients with HOS diagnosed by multidisciplinary teams; (b) an open biopsy or CT-guided core needle biopsy, pathologically evaluated by specialized sarcoma pathologists; (c) a standard CT scan performed at the time of diagnosis; and (d) clinical characteristics available. The exclusion criteria were as follows: (a) patients who underwent chemotherapy treatment before undergoing a CT scan in our institution; (b) patients suffering from other synchronous cancers; (c) incomplete or indeterminate clinical characteristics; and (d) death by a cause other than osteosarcoma. Supplementary Data I shows the patient recruitment pathway.

Treatment options included preoperative neoadjuvant and/or adjuvant chemotherapy and surgery with the aim of achieving a wide excision (either by limb-salvage or amputation surgery). Margins were defined on the basis of Enneking's criteria [23]. Chemotherapeutic regimens included cisplatin, doxorubicin, and high-dose methotrexate. Most patients received neoadjuvant chemotherapy and additional adjuvant chemotherapy after surgery. The total duration of chemotherapy was at least 6–8 months.

The overall patient population was divided into two cohorts on the basis of diagnosis time: training cohort and independent validation cohort. The training cohort was used for construction of the prediction model. This cohort included 102 patients (49 males and 53 females, 8–54 years of age) who were diagnosed between January 2008 and March 2011. The independent validation cohort consisted of 48 patients (25 males and 23 females, 8–47 years of age) who were diagnosed between April 2011 and April 2012; this cohort was used to test the prediction power of the model. Patients who survived ≥5 years after treatment were classified within the survival group, whereas those who died within 5 years of the operation were classified within the non-survival group. The power test was performed to evaluate the reliability of this study by using sample size and 5-year survival rates in both training and independent validation cohorts [24,25].

Baseline clinical data, including age, gender, tumor anatomic site, tumor stage (local/metastatic) [10], and the presence of pathological fracture (no/yes), were obtained from the Electronic Medical Record System (EMRS). All CT images were collected from the Picture Archiving and Communication System (PACS). The tumor stage and presence of pathological fracture at the time of diagnosis were determined by

consensus of three experienced radiologists. The follow-up time in our study comprised every 6 weeks in the first and second years after treatment, every 3 months in the third and fourth years, and every 6 months after the fourth year. All data were collected and evaluated in April 2017 with a minimum follow-up of 5 years for all included patients.

2.2. CT Image Acquisition, Region of Interest Segmentation and Radiomics Feature Extraction

CT image acquisition is described in Supplementary Data II. Image resampling and gray level quantization were performed prior to feature extraction. ITK-SNAP software was used for three-dimensional regions of interest (ROI) segmentation [26]. Texture features were extracted by using in-house developed software in MATLAB 2015b (MathWorks, Natick, MA, USA) [27,28]. The feature pool extracted based on the 3-dimension region of interest (ROI) comprised 4 groups: (i) 6 histogram statistics features; (ii) 7 shape features; (iii) 53 texture features; and (iv) 408 wavelet features. Supplementary Data III describes the detailed features and references. The ROI was segmented by 3 orthopedists with 6 years (Orthopedist-1), 4 years (Orthopedist-2), and 4 years (Orthopedist-3) of experience in orthopedic CT interpretation. The patients in the training cohort were segmented separately by Orthopedist-1 and Orthopedist-2. The feature set based on the segmentation of Orthopedist-1 was used for model training. The feature set based on the segmentation of Orthopedist-2 was used to test the reproducibility and stability of each of the features. Patients in the validation cohort were segmented by Orthopedist-3 to test the prediction power of the trained model.

2.3. Feature Selection, Radiomics Signature Building and Validation

Feature selection was performed in 3 steps to select the optimal survival-related features via the training cohort. Firstly, the reproducibility and stability of each feature was determined by calculating the correlation coefficient between feature sets, based on the segmentations of Orthopedist-1 and Orthopedist-2. We only retained stable features with an intra-class correlation coefficient >0.8. Secondly, we used the Spearman rank correlation test to investigate the internal linear correlation between individual features. Redundant features (a linear correlation coefficient >0.95, determined by the Spearman test) were removed [28]. Finally, the least absolute shrinkage and selection operator (LASSO) logistic regression algorithm, which is applicable for high-dimensional data reduction [29], was performed for optimal feature selection. LASSO regression reduced the coefficients for survival-unrelated variables to zero; variables with non-zero coefficients were retained. To select optimal parameters in LASSO regression, we performed 100 iterations of 10-fold cross-validation with binomial deviance minimization criteria from the training cohort [30]. Binomial deviance was used as the loss function in the model training process; the model with the minimum binomial deviance was selected. Then, a radiomics score calculation formula, defined as the radiomics signature, was generated by a linear combination of selected features multiplied by LASSO coefficients. The radiomics signature was a prediction model constructed by using the selected features; the radiomics score was calculated to reflect survival probability by using the radiomics signature for each patient. The performance of the radiomics signature was assessed by its discrimination in both training and validation cohorts, which measured how well the model could distinguish patients in the survival or non-survival groups [31]. Discrimination was demonstrated by a receiver operating characteristic (ROC) curve and the associated area under the ROC curve (AUC); sensitivity and specificity were also calculated.

2.4. Development of the Radiomics Nomogram and Clinical Model

Considering the potential prediction value of the clinical characteristics, a multivariable logistic regression analysis was developed by

combining radiomics score and clinical characteristics based on the training cohort. The clinical characteristics included age, gender, tumor anatomic site, stage, and presence of pathological fracture. The Variance Inflation Factor (VIF) was calculated to detect the severity of multicollinearity among variables in the multiple logistic regression model [22]. The VIF was the ratio between the variance in the multiple variables mode and the variance in a model with one variable [19]; if VIF was >10, then multicollinearity was high [32]. Then, a radiomics nomogram was developed based on the multivariate logistic regression model [33]. To compare the prediction performance of the radiomics nomogram and clinical characteristics, a clinical model was constructed separately using multivariable logistic regression analysis based on clinical factors alone.

2.5. Validation and Assessment of the Radiomics Nomogram

The performance of the radiomics nomogram was tested in both training and validation cohorts with respect to discrimination, calibration, and clinical usefulness [31]. The AUC was measured to quantify discrimination performance; calibration measured the model's ability to generate predictions that were generally close to the average observed outcome. The calibration curve, representing the agreement between predicted survival probability and average actual survival probability, was plotted to assess the calibration of the radiomics nomogram. The Hosmer-Lemeshow test was applied to evaluate the goodness-of-fit of the nomogram [33]. Decision curve analysis (DCA) was used to evaluate whether the nomogram was sufficiently robust for clinical practice [34]. The net benefit was derived by calculating the difference between the true positive rate and weighted false positive rate across different threshold probabilities in the validation cohort. The "decision curve" was plotted against the threshold probability. In addition, Kaplan-Meier survival curves were plotted by using predictive survival status as the prediction factor. The patients were classified into two groups on the basis of the prediction factor: predictive survival group and predictive non-survival group. The log-rank test was used to assess the difference between survival curves from the nomogram-predicted survival and non-survival groups (p-value <.05, 95% confidence interval [CI]).

2.6. Statistical Analysis

Univariate analysis for clinical factors used the Mann-Whitney U test or Chi-square test, as appropriate. Correlation was assessed by using Spearman's correlation rank test. Reported significance levels were two-sided, p-value <.05. The LASSO logistic regression model was used with penalty parameter (λ) tuning, which was conducted with 10-fold cross-validation based on the binomial deviance minimum criteria. The backward search method was used with the minimum criteria of Akaike's information criterion (AIC) as the stopping rule for the multivariate logistic regression model [35,36]. A detailed description of AIC is presented in Supplementary Data IV.

All statistical analyses were performed with R software (version 3.4.1; http://www.Rproject.org), MedCalc Statistical Software (version 15.2.2; https://www.medcalc.org), and PASS software (version 11.0.7; https://www.ncss.com). Kaplan-Meier analysis was performed by using MedCalc. The power test was calculated by using the PASS software. LASSO logistic regression analysis was performed by using the "glmnet" package. Nomograms and calibration plots were performed with the "rms" package. The Hosmer-Lemeshow test was performed by using the "generalhoslem" package. VIFs were calculated by using the "car" package. AUC analysis was conducted by using the "pROC" package. DCA was performed by using the "dca.R" function. The two-sided statistical significance level used in this study was p-value <.05.

3. Results

The schematic depiction for this study is presented in Fig. 1.

3.1. Clinical Characteristics

Clinical characteristics in the training and validation cohorts are summarized in Table 1. No significant difference was observed in survival rate (p-value = .6090, Chi-square test) between the two cohorts. Survival rates in the training and validation groups were 46.07% (47/102) and 52.08% (25/48), respectively. A power of 0.9238 indicated that the sample size for training and validation in this study was sufficient.

3.2. Construction and Validation of Radiomics Signature

A total of 474 features were extracted based on CT images for each patient. The heatmap showing the distribution of these radiomics features is shown in Supplementary Data V [19]. After removing survival-unrelated and redundant features, 76 features remained. Subsequently, 5 potential predictors (1:20 ratio) of the 76 features were selected via LASSO regression, based on the training cohort. Parameter (λ) selection in the LASSO model and LASSO coefficients are shown in Supplementary Data VI. The radiomics score calculation formula is presented in Supplementary Data VI, along with the selected features.

Patients in the survival group generally had higher scores than patients in the non-survival group. A significant difference in radiomics score was observed between patients in the survival and non-survival groups (0.5483 ± 0.1256 vs. 0.3860 ± 0.1678; p-value <.0001, Mann-Whitney U test) in the training cohort. The AUC for the training cohort was 0.79 (95% CI, 0.70–0.86). The performance of the radiomics signature was confirmed by testing in the validation cohort. A significant difference in radiomics score was observed in the validation cohort (0.4935 ± 0.1564 vs. 0.3326 ± 0.2005; p-value = .0031, Mann-Whitney U test), with an AUC of 0.76 (95% CI, 0.61–0.87). Distributions of radiomics scores for the training and validation cohorts are shown in Fig. 2.

3.3. Development and Validation of the Radiomics Nomogram and Clinical Model

A radiomics nomogram was developed by combining the radiomics signature, stage, and tumor volume (Fig. 3a). No severe collinearity was observed in our regression model. The VIFs for radiomics score, stage, and tumor volume were 1.0799, 1.0850, and 2.9030, respectively. The clinical model was constructed with clinical characteristics of stage and tumor volume by using multivariable logistic regression analysis based on the training cohort.

The calibration curve showed good agreement between the predicted probability of survival and actual probability of survival. A nonsignificant statistic (p-value = .8185, Hosmer-Lemeshow test) for the training cohort suggested no departure from the ideal fit. The AUC for the nomogram was 0.86 (95% CI, 0.77–0.92) in the training cohort. For the clinical model, an AUC of 0.79 (95% CI, 0.70–0.86) was observed. Good calibration was also observed for the probability of the 5-year survival rate in the dependent validation cohort. A nonsignificant statistic (p-value = .6431, Hosmer-Lemeshow test) showed agreement between the actual and observed survival probability with the nomogram for the dependent validation cohort. The AUC for the nomogram was 0.84 (95% CI, 0.71–0.93). The AUC for the clinical model was 0.73 (95% CI, 0.58–0.85). ROC curves for the radiomics nomogram and clinical model are presented in Fig. 3b and c for the training and validation cohorts, respectively. Detailed performances of the radiomics signature, radiomics nomogram, and clinical model are summarized in Table 2.

Calibration curves for the training and independent validation cohorts are shown in Fig. 4a and b. DCAs for the radiomics nomogram in the training and independent validation cohorts are presented in Fig. 4c and d. For both training and validation cohorts, DCA curves showed that the radiomics nomogram gained more net benefits than the treat-all-patients strategy, the treat-none strategy, and the clinical model.
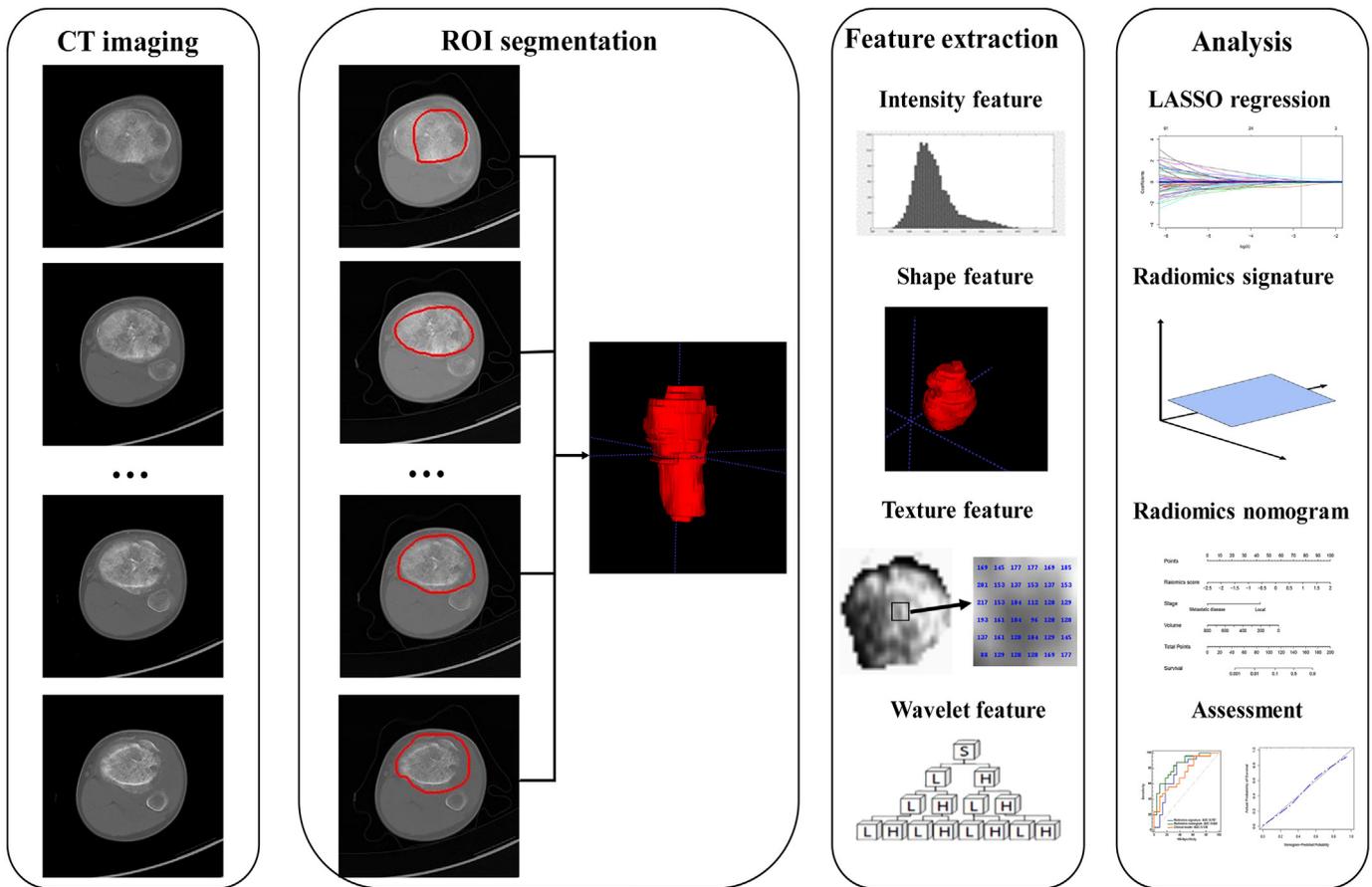
Fig. 1. Radiomics schematic for this study.

For the patients in the non-survival group, mean survival time was 31.7 months and median survival time was 28.6 months. Kaplan-Meier curves (Fig. 5a and b) showed a significant difference between the nomogram-predicted survival and non-survival groups (p-value <.05, log-rank test) in both training and validation cohorts.

4. Discussion

Preoperative prediction of 5-year survival is important for treatment planning. Previous studies solely analyzed the association between clinical factors and survival status [5,8–12]; they did not investigate the

Table 1
Characteristics at time of diagnosis in patients with high-grade osteosarcoma primary.

	Training cohort (n = 102)			Independent validation cohort (n = 48)		
Characteristic	Survival (n = 47)	Non-survival (n = 55)	p-value	Survival (n = 25)	Non-survival (n = 23)	p-value
Age (years)			0.4739			0.4371
<15 years	19	24		10	9	
≥15 years	28	31		15	14	
Gender			0.1190	14:11	11:12	0.7817
Male: Female	27:20	22:33				
Tumor volume			**0.0261**			0.0662
Median	70.94	161.34		108.35	140.28	
Range	9.31, 521.98	13.06, 782.12		21.05, 277.01	83.95, 395.01	
Location			0.9762			0.7776
Distal femur	29	33		10	12	
Lower extremities	11	13		6	6	
Pelvis	1	2		2	2	
Proximal tibia	6	7		7	3	
Stage at diagnosis			**0.0024**			**0.0083**
Local: Metastatic	46:1	41:14		24:1	14:9	
Pathological fracture			0.0630			0.8456
No: Yes	42:5	40:15		18:7	15:8	
Radiomics score			**<0.0001**			**0.0031**
Median	0.5419	0.3640		0.4922	0.2992	
Range	0.2485, 0.8367	0.0771, 0.7587		0.2280, 0.8479	0.0622, 0.7993	

Individual clinical factors were analyzed using Mann-Whitney U test or Chi-square test, as appropriate.
p-value <.05 indicates the significant difference.
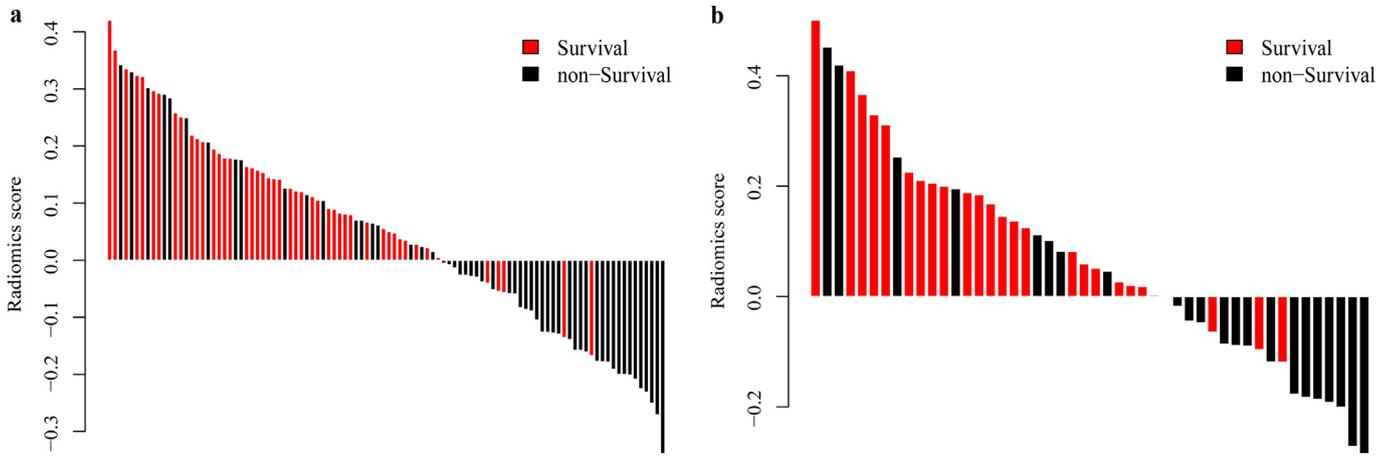
Fig. 2. Radiomics score for each patient in training (a) and validation (b) cohorts. Red bars indicate scores for patients in the survival group, while black bars indicate scores for patients in the non-survival group. Patients with radiomics scores >0 were predicted to be in the survival group, while patients with radiomics scores <0 were predicted to be in the non-survival group.
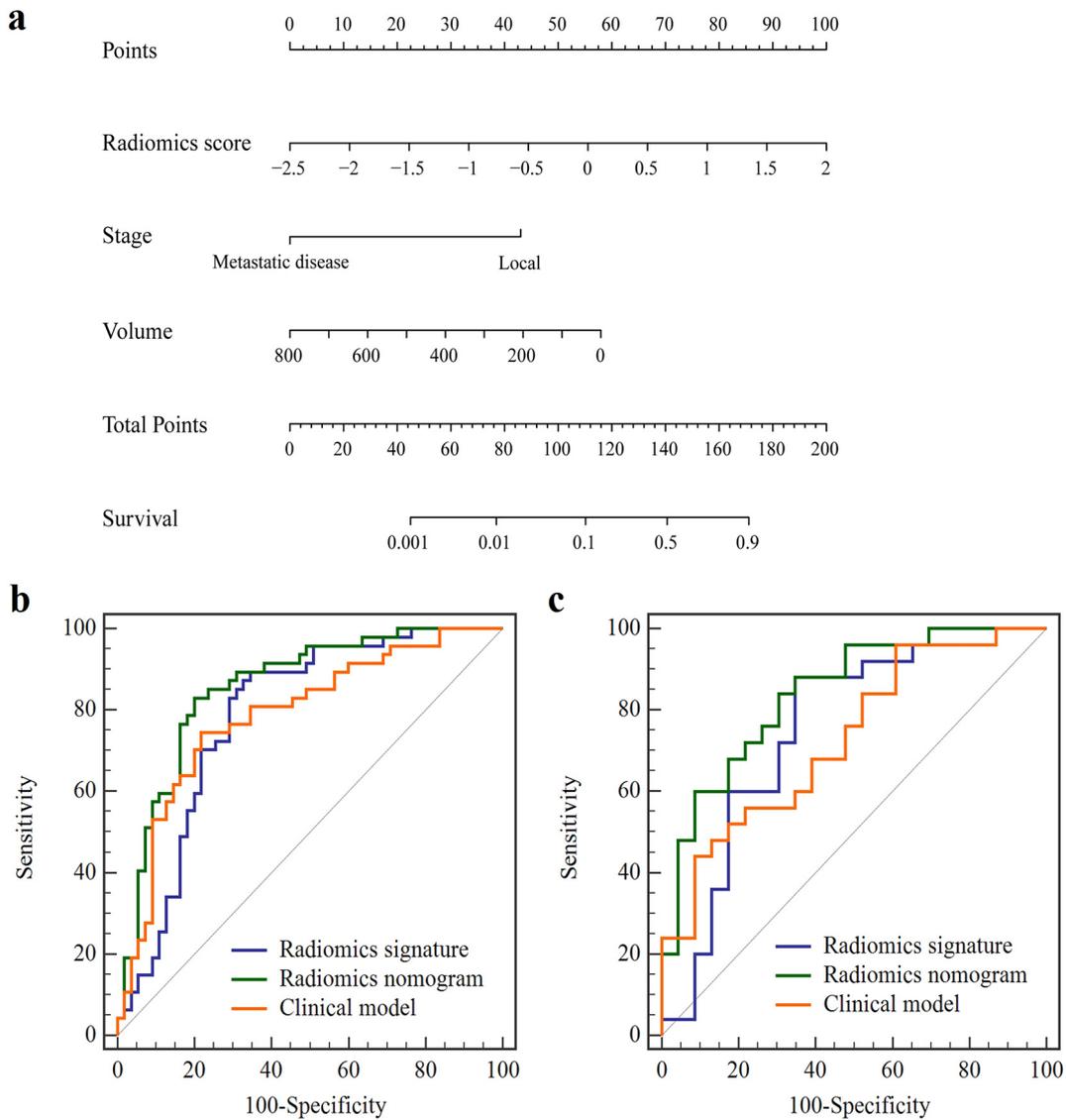
Fig. 3. The radiomics nomogram, combining radiomics signature, stage, and tumor volume, developed in the training cohort (a). Radiomics score ranges from −2.5 to 2. Radiomics scores < −2.5 were set to −2.5. Radiomics scores >2 were set to 2. Tumor volume ranges from 0 to 800 cm³. Volumes >800 cm³ were set to 800 cm³. ROC curves for the radiomics nomogram, radiomics signature, and clinical model in training (b) and validation (c) cohorts.

Table 2
Model performance.

Model	Training cohort			Independent validation cohort		
	Sensitivity	Specificity	AUC (95% CI)	Sensitivity	Specificity	AUC (95% CI)
Radiomics signature	89.36%	65.45%	0.79 (0.70 to 0.86)	88.00%	65.22%	0.76 (0.61 to 0.87)
Radiomics nomogram	82.98%	80.00%	0.86 (0.77 to 0.92)	84.00%	69.57%	0.84 (0.71 to 0.93)
Clinical model	74.47%	78.18%	0.79 (0.70 to 0.86)	44.00%	91.30%	0.73 (0.58 to 0.85)

AUC: area under the curve; CI: confidence interval.

underlying association between 5-year survival status and imaging features by using radiomics method. In this study, we developed and validated a radiomics nomogram combining radiomics score and clinical factors to evaluate survival status preoperatively; this nomogram showed better performance than the clinical model.

We chose to include axial HOS in this study because these patients were systematically referred to our multidisciplinary center, thereby reducing the risk of bias associated with investigating osteosarcoma throughout the body. Currently, it remains unclear whether there is a significant difference in survival between patients with axial and appendix tumors. Some studies suggest a potential for confounding bias when comparing these two types of HOSs [37] because of the large number of overweight, elderly patients with axial tumors [38], as well as the possibility that axial tumors might be more malignant [39]. In contrast, other studies have found no statistically significant differences in survival rates when comparing axial and appendicular HOSs [40].

Among the 5 clinical characteristics at diagnosis, the tumor stage of the osteosarcoma was included in the nomogram. The evaluation of patients with both localized and metastatic disease was a key point for 5-year survival rates for patients with osteosarcoma. A previous European Intergroup Osteosarcoma study on osteosarcoma reported a 5-year survival-rate of 56% in a cohort of 1067 localized high-grade appendicular osteosarcomas in patients ≤40 years of age. Among HOSs, we found a 48% 5-year survival-rate in our study, which was lower than in the European Intergroup Osteosarcoma's study; this may have occurred because our cases included patients of all ages.

In agreement with our findings, tumor volume has previously been suggested as a prognostic factor by several studies, which reported
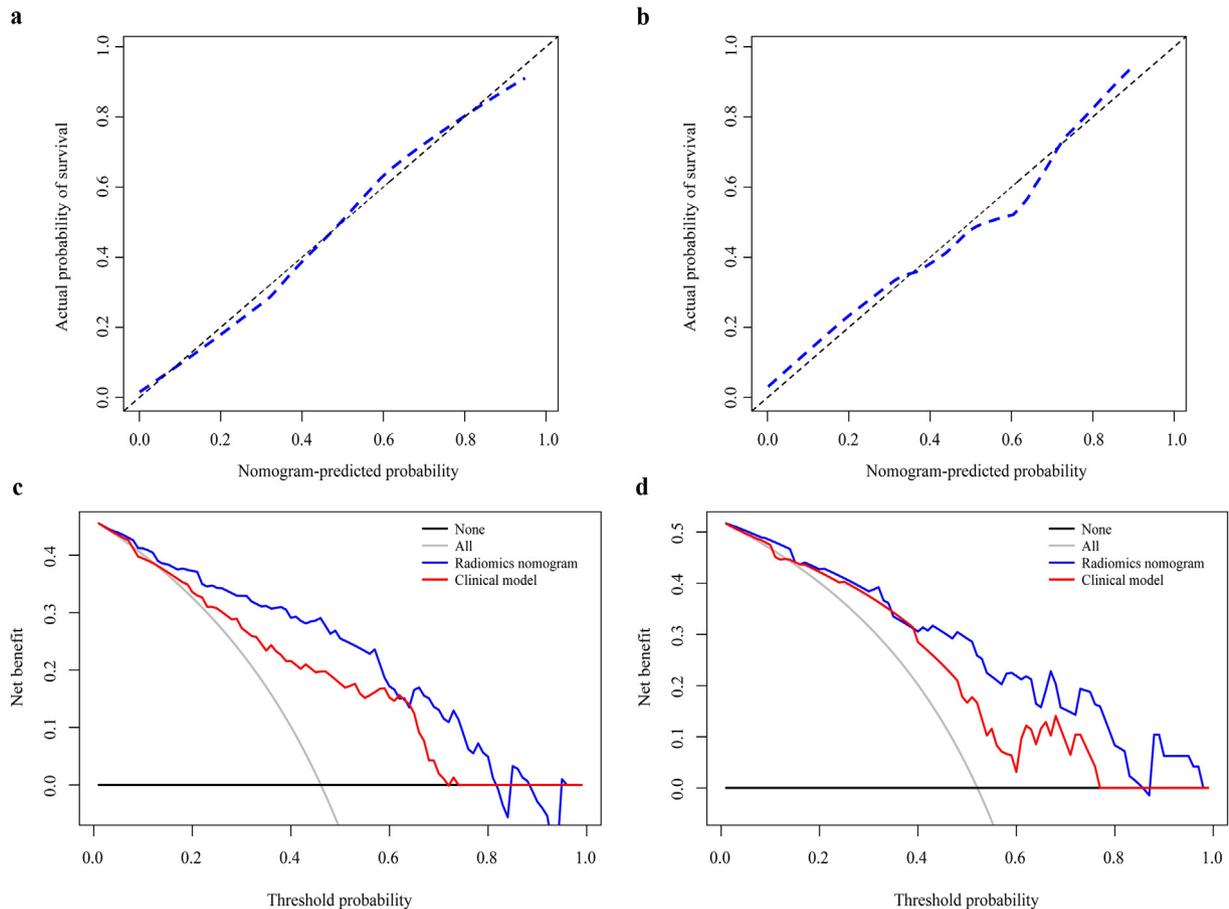
Fig. 4. Calibration curves for the radiomics nomogram in training (a) and validation (b) cohorts. The y-axis indicates the actual probability of survival; x-axis indicates the predicted probability of survival. The 45-degree black line represents the ideal prediction; red line represents the performance of the radiomics nomogram. As the red line approaches the ideal prediction line, the predictive accuracy of the nomogram increases. DCA for the radiomics nomogram and clinical model in both training (c) and independent validation cohorts (d). The y-axis indicates the net benefit; x-axis indicates threshold probability. The blue line represents net benefit of the radiomics nomogram; red line represents net benefit of the clinical model. The black line represents the hypothesis that all patients die within 5 years; gray line represents the hypothesis that no patient dies within 5 years.
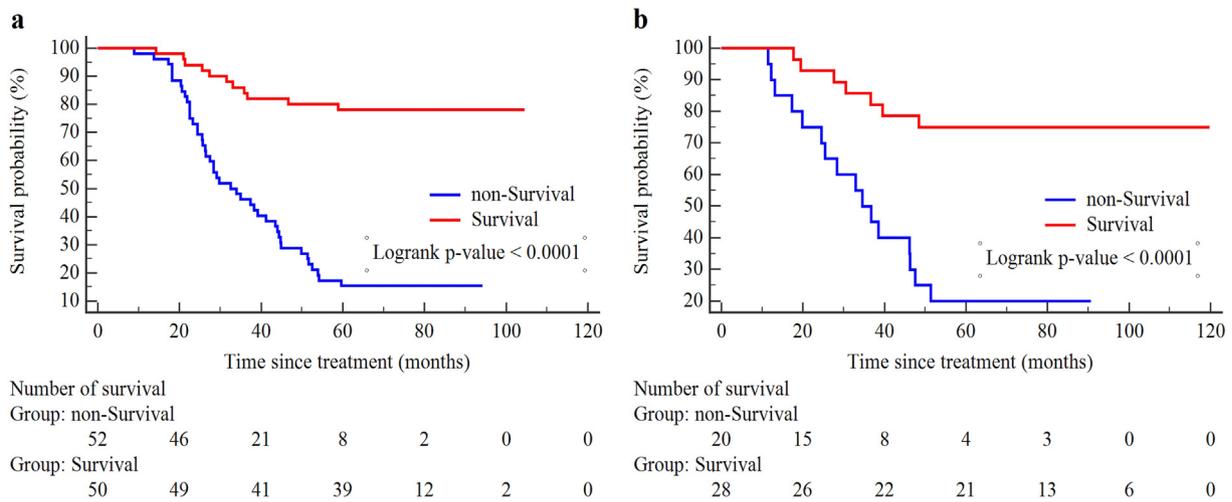
Fig. 5. Kaplan-Meier curves for patients in the nomogram-predicted survival and non-survival groups from training (a) and validation (b) cohorts.

that a volume > 150 mL led to poor prognosis and lower survival rate [41]. Other studies have shown that primary tumor size can be closely related to the anatomic site of the tumor; axial tumors tend to be larger when diagnosed, which may be related to a lower survival rate [13]. Determining tumor size has often been a difficult task because of the performance of imaging devices and the heterogeneity of tumor sizes among different bone types. Recent studies have described new methods for determining tumor size to predict prognosis in HOSs [14]. In our study, tumor volume was assessed to equalize the sites of tumors. Tumor volume was significantly different between the survival and non-survival groups in both training and validation cohorts; it could be regarded as a qualified component of the early prediction model of osteosarcoma. We suspect that this parameter may reflect the heterogeneity of osteosarcoma. Our research results constitute an improvement in the prediction accuracy of the model by combining radiomics signature and clinical factors; this is sufficiently effective to be an independent component of the nomogram.

Current literature has not yet reached a consensus regarding prognostic factors for survival in treating this difficult disease; therefore, our radiomics-based prediction nomogram for early survival prediction of osteosarcoma may allow physicians to provide a more appropriate treatment strategy for each patient. Furthermore, because our radiomics nomogram can enable personalized survival prediction for each patient, a reasonable follow-up interval could be arranged, thereby avoiding unnecessary medical resources and expenses.

Importantly, there were some limitations in our study. Firstly, the database in this study was retrospectively collected from a single center. Due to the low incidence rate of osteosarcoma, the sample size was not large. A power of 0.9238 shows that the sample size (training cohort: 102 patients; validation cohort: 48 patients) was sufficient for robust analysis, based on power analysis of the training and validation sample cohorts. Multi-center validation with a larger sample size is essential for acquiring high-level evidence for future clinical application. Secondly, since survival potential may correlate with the location of the osteosarcoma tumor, but the corresponding mechanism remains unclear, combined statistical analysis of both axial and appendicular HOS in this study may have resulted in bias. In the pre-processing of CT images from osteosarcoma tumors in different sites, tumor volumes were corrected to suppress possible biases and extend the limited database when only one site was analyzed. In future work, a series of prediction models for different sites will be developed; a dedicated prediction model will, theoretically, improve prediction accuracy. Thirdly, genetic markers were not considered in our study because they have not been shown to be closely correlated with osteosarcoma prognosis in clinical practice. Lastly, only CT images were used in this study. In surgical planning, MRI images are indispensable because of their brilliant resolution for soft tissue. We plan to develop an additional model combining CT and MRI image features.

In conclusion, we have developed a non-invasive predictive tool, combining radiomics features and clinical risk factors, to predict survival period at the time of diagnosis in HOS cases. The developed predictive tool can also provide a basis for clinical doctors to make decisions for personalized diagnosis and treatment. Multi-center retrospective validation studies, including prospective randomized clinical trials, should be performed to obtain high-level evidence for future clinical application.

Conflicts of Interest

We declare that we have no conflicts of interest.

Author Contributions

Conception and design: Y·W, L·X, P·Y, T·N, Z·Y.
Collection and assembly of data: Y·W, N·L, X·H, W·P, H·L, P·L, B·L, D·Y.
Data analysis and interpretation: Y·W, L·X, P·Y, C·L, Y·J, T·N, Z·Y.
Manuscript writing: Y·W, L·X, P·Y, V·B, M·H, T·N.
Final approval of manuscript: All authors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ebiom.2018.07.006.

References

[1] Whelan J, Mctiernan A, Cooper N, Wong YK, Francis M, Vernon S, et al. Incidence and survival of malignant bone sarcomas in England 1979-2007. Int J Cancer 2012;131 (4):E508–17.

[2] Zaikova O, Sundby Hall K, Styring E, Eriksson M, Trovik CS, Bergh P, et al. Referral patterns, treatment and outcome of high-grade malignant bone sarcoma in Scandi-navia—SSG Central Register 25 years' experience. J Surg Oncol 2016;112(8):853–60.

[3] Mirabello L, Troisi R. Sa. Osteosarcoma incidence and survival rates from 1973 to 2004: data from the Surveillance, Epidemiology, and End Results Program. Cancer 2010;115(7):1531–43.

[4] Venkatramani R, Murray J, Helman L, Meyer W, Hicks MJ, Krance R, et al. Risk-Based Therapy for Localized Osteosarcoma. Pediatr Blood Cancer 2016;63(3):412–7.

[5] Durnali A, Alkis N, Cangur S, Yukruk FA, Inal A, Tokluoglu S, et al. Prognostic factors for teenage and adult patients with high-grade osteosarcoma: an analysis of 240 patients. Med Oncol 2013;30(3):624.

[6] Bacci G, Bertoni F, Longhi A, Ferrari S, Forni C, Biagini R, et al. Neoadjuvant chemo-therapy for high-grade central osteosarcoma of the extremity. Cancer 2003;97 (12):3068–75.

[7] Briccoli A, Rocca M, Salone M, Guzzardella GA, Balladelli A, Bacci G. High grade oste-osarcoma of the extremities metastatic to the lung: long-term results in 323 patients treated combining surgery and chemotherapy, 1985–2005. Surg Oncol 2010;19(4): 193–9.

[8] Bielack SS, Kempf-Bielack B, Gn Delling, Exner GU, Flege S, Helmke K, et al. Prognos-tic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. J Clin Oncol 2002;20(3):776–90.

[9] Bieling P, Rehan N, Winkler P, Helmke K, Maas R, Fuchs N, et al. Tumor size and prognosis in aggressively treated osteosarcoma. J Clin Oncol 1996;14(3):848–58.

[10] Enneking WF, Spanier SS, Goodman MA. The classic: A system for the surgical staging of musculoskeletal sarcoma. Clin Orthop Relat Res 2003;415:4–18.

[11] Dae-Geun Jeon MD, Chang-Bae Kong MD, Wan HC, Song WS, Sang HC, Choi SW, et al. Examination of the cutoff value of postchemotherapy increase in tumor volume as a predictor of subsequent oncologic events in stage IIB osteosarcoma. J Surg Oncol 2014;109(3):275–9.

[12] Scully SP, Ghert MA, Zurakowski D, Thompson RC, Gebhardt MC. Pathologic fracture in osteosarcoma : prognostic importance and treatment implications. J Bone Joint Surg Am 2002;84-A(1):49.

[13] Anderson ME. Update on Survival in Osteosarcoma. Orthop Clin N Am 2016;47(1): 283.

[14] Kim SH, Shin KH, Park EH, Yong JC, Park BK, Suh JS, et al. A new relative tumor sizing method in epi-metaphyseal osteosarcoma. BMC Cancer 2015;15(1):284.

[15] James PO, Connor. Rethinking the role of clinical imaging. Elife 2017;6.

[16] Katsila T, Liontos M, Patrinos GP, Bamias A, Kardamakis D. The New Age of-omics in Urothelial Cancer–Re-wording Its Diagnosis and Treatment. EBioMedicine 2018;28: 43–50.

[17] Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48(4):441–6.

[18] O'Connor JP, Aboagye EO, Adams JE, Aerts HJ, Barrington SF, Beer AJ, et al. Imaging biomarker roadmap for cancer studies. Nat Rev Clin Oncol 2017;14(3):169.

[19] Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Cavalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5:4006.

[20] Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. Sci Found China 2016;34(4):2157–64.

[21] Lee G, Lee HY, Park H, Schiebler ML, van Beek EJ, Ohno Y, et al. Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical manage-ment: state of the art. Eur J Radiol 2017;86:297–307.

[22] Wu S, Zheng J, Li Y, Yu H, Shi S, Xie W, et al. A Radiomics Nomogram for the Preop-erative Prediction of Lymph Node Metastasis in Bladder Cancer. Clin Cancer Res 2017;23(22):6904–11.

[23] Enneking WF, Spanier SS, Goodman MA. A system for the surgical staging of muscu-loskeletal sarcoma. Clin Orthop Relat Res 1980(153):106–20.

[24] Chow S-C, Shao J, Wang H, Lokhnygina Y. Sample size calculations in clinical re-search. Chapman and Hall/CRC; 2017.

[25] Zhang B, Tian J, Dong D, Gu D, Dong Y, Zhang L, et al. Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma. Clin Cancer Res 2017;23(15):4259–69.

[26] Yushkevich PA, Piven J, Hazlett HC, Smith RG, Ho S, Gee JC, et al. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 2006;31(3):1116–28.

[27] Valliä˙Res M, Freeman CR, Skamene SR, El NI. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarco-mas of the extremities. Phys Med Biol 2015;60(14):5471–96.

[28] Wu J, Aguilera T, Shultz D, Gudur M, Rubin DL, Jr LB, et al. Early-Stage Non-Small Cell Lung Cancer: Quantitative Imaging Characteristics of (18)F Fluorodeoxyglucose PET/ CT Allow Prediction of Distant Metastasis. Radiology 2016;281(1):151829.

[29] Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc B Methodol 1996:267–88.

[30] Roosen CB, Hastie TJ. Logist Response Projection Pursuit 1993;76(376):817–23.

[31] Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: The Hosmer-Lemeshow test revisited. Crit Care Med 2007;35(9): 2052–6.

[32] Kutner MH, Nachtsheim C, Neter J. Applied linear regression models. McGraw-Hill/ Irwin; 2004.

[33] Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol 2008;26(8):1364–70.

[34] Vickers AJ, Elkin EB. Decision curve analysis: A novel method for evaluating predic-tion models. Med Decis Making Int J Soc Med Decis Making 2006;26(6):565.

[35] Bozdogan H. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika 1987;52(3):345–70.

[36] Sauerbrei W, Royston P, Binder H. Selection of important variables and determina-tion of functional form for continuous predictors in multivariable model building. Stat Med 2007;26(30):5512–28.

[37] Whelan JS, Jinks RC, Mctiernan A, Sydes MR, Hook JM, Trani L, et al. Survival from high-grade localised extremity osteosarcoma: combined results and prognostic fac-tors from three European Osteosarcoma Intergroup randomised controlled trials. Ann Oncol 2012;23(6):1607–16.

[38] Luetke A, Meyers PA, Lewis I, Juergens H. Osteosarcoma treatment - where do we stand? A state of the art review. Cancer Treat Rev 2014;40(4):523–32.

[39] Meazza C, Luksch R, Daolio P, Podda M, Luzzati A, Gronchi A, et al. Axial skeletal os-teosarcoma: a 25-year monoinstitutional experience in children and adolescents. Med Oncol 2014;31(4):875.

[40] Aggerholmpedersen N, Marettynielsen K, Keller J, Baerentzen S, Schrøder H, Jørgensen PH, et al. The importance of standardized treatment in high-grade osteosarcoma: 30 years of experience from a hospital-based database. Acta Oncol 2015;54(1):17.

[41] Ferrari S, Bertoni F, Mercuri M, Picci P, Giacomini S, Longhi A, et al. Predictive factors of disease-free survival for non-metastatic osteosarcoma of the extremity: An analysis of 300 patients treated at the Rizzoli Institute. Ann Oncol 2001;12(8): 1145–50.