Construction and Validation of Nomogram for Prediction of Metastasis in Osteosarcoma Patients

weilong xu

Guangzhou University of Chinese Medicine

wei niu (✉ niuwei@139.com)

Guangdong Provincial Hospital of Traditional Chinese Medicine

Research Article

Keywords: Osteosarcoma, Metastasis, Nomogram, SEER Program

Posted Date: March 10th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-293568/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Osteosarcoma is one of the most common malignant tumors of the bone with poor prognosis. The present study was aimed to establish and validate nomogram to predict the risk of metastasis for individual osteosarcoma patients.

Methods: 2195 osteosarcoma patients were retrospectively collected from Surveillance, Epidemiology, and End Results (SEER) database. The least absolute shrinkage and selection operator regression model was applied to promote feature selection for the metastasis signature model. Univariate and multivariate logistic analysis were used to discern independent predictive factors. These predictive factors were included in the nomogram to evaluate metastasis probability in osteosarcoma patients. Nomogram was validated internally and externally.

Results: We randomly assigned 2195 osteosarcoma patients into the training (n = 1464) and validation (n = 731) subgroups. Age, sex, tumor grade, tumor size and use of surgery were identified as the predictive factors for the risk of metastasis (all $p < 0.05$) and further incorporated to develop the nomogram. Predictive accuracy (AUC) for metastasis prediction were 0.7427067 in the training subgroup and 0.6798374 in the external validation subgroup, respectively. And the decision curve analysis also proved the value of the model.

Conclusions: We have constructed new high-performance nomogram to assist clinicians in identifying osteosarcoma patients with high risk of metastasis.

Background

Osteosarcoma, the most common primary bone malignant tumor in children and adolescents and the third most frequent bone cancer in adults [1], originates from interstitial cells. Osteosarcoma, characterized by proliferating tumor cells, directly form immature bone or bone-like tissue [2]. It occurs mostly in long bones of the limbs. Common sites of osteosarcoma include humerus, tibia and femur, and less commonly pelvis, skull and jaw [3]. Even though osteosarcoma mainly occurs in adolescence aged between 10 to 25 years old, there is a second incidence peak among individuals aged over 60 years [4]. Osteosarcoma has a high metastatic rate, and over 20% of osteosarcoma patients have lung metastases at the time of initial diagnosis [5]. The 5-year survival rate of osteosarcoma patients with metastasis (20%) is significantly lower than that of patients with focal lesions (65%), and most patients died of osteosarcoma with metastasis [6]. Consequently, it is of importance and significance to identify predictive factors for the metastasis risk of osteosarcoma. Single predictive factor is limited to perform an accurate prediction of metastasis risk for individualized osteosarcoma patients. Therefore statistical prediction model integrating the cumulative effects of individual predictive factors should be established for more accurate predictions.

In this era of personalized medicine, nomogram, a sample statistical predictive tool, have obtained popularity in clinics. As a pictorial representation of a complicated mathematical formula [7], it uses

clinical and biologic variables, such as patient sex and tumor size, to graphically depict a statistical predictive model and estimate the probability of clinical event for a given individual [8]. Nomogram is able to assist clinicians in predicting the risk of metastasis for an individual patient with tumors, which has been reported in breast cancer, papillary thyroid cancer and squamous non-small cell lung cancer [9–11].

Therefore, the current study aimed to establish and validate available nomogram to evaluate the metastasis risk in osteosarcoma patient based on the data from Surveillance, Epidemiology, and End Results (SEER[12, 13]) database. In order to construct nomogram predicting the metastasis risk of osteosarcoma, we used univariate and multivariate logistic [14] analysis to discern independent predictive factors. Age, sex, tumor grade, tumor size and use of surgery were identified as the predictive factors. Then we validated our nomogram in validation data sets. Nomogram constructed in present study might serve as a new and alternative tool to predict metastasis risk in osteosarcoma and provide valuable information for surgical planning and clinical management.

Materials And Methods

Patient data

All the osteosarcoma patient's data were downloaded from SEER database. The SEER*Stat software (Version 8.3.5; NCI, Bethesda, USA) was applied to acquired osteosarcoma patient's information. Osteosarcoma patients who met the following criteria were recruit in the present study: (1) positive histological confirmation of osteosarcoma; (2) diagnosed with osteosarcoma as primary malignancy between 1983 and 2014 (International classification of diseases for oncology: 9180, 9181, 9182, 9183, 9184, 9185, 9186, 9187, 9192, 9193, 9194, 9200) [15]; (3) extremity site (short and long bones of lower and upper extremity) or axial site (skull, spine, ribs and pelvis); (4) known survival months and cause of death after diagnosis. And osteosarcoma patients with unknown surgical stage or unknown use of surgery or unknown tumor size were excluded in present study. Data on age, sex, marital status, race, histologic type, tumor site, tumor grade, tumor size and use of surgery were applied in further analysis to identify metastasis risk factors.

Data analysis

The receiver operating characteristic (ROC) curve, which illustrates sensitivity against false positive rate, has been frequently used to obtain optimal tumor size cutoff values in diagnostic tests. ROC is also a tool applied to describe the discrimination accuracy of prediction model or diagnostic test [16].

The least absolute shrinkage and selection operator (LASSO) method is effective learning method for feature selection of high dimensional data. This analysis method can impose constraint during parameter estimation to penalize the magnitude of the variables' coefficients [17]. In present study, this method was applied to optimize the predictive features in metastasis risk factors. Features with nonzero

coefficients were chosen in LASSO regression model. These predictive features were applied in the univariate logistic analysis for the risk of metastasis.

We randomly assigned all selected osteosarcoma patients (n = 2195) into a training subgroup (n = 1464) and a validation subgroup (n = 731). Further, we used Chi-square test to compare the demographic baseline and clinical characteristics of the patients between the training and validation subgroup. Significant variables were further assessed using univariate and multivariate logistic regression model. And odds ratio of each variable with corresponding 95% confidence interval was calculated regarding the risk of metastasis in osteosarcoma patients.

Construction of the nomogram

Based on the multivariable analysis results, metastasis nomogram was formulated. The calibration of this metastasis nomogram was evaluated by calibration curves. Predictive discrimination of nomogram was evaluated internally and externally via ROC curve. The minimum AUC value was 0.5 which indicated a random chance to correctly discriminate outcome for the constructed signature, and 1.0 indicated a perfect discrimination. Decision curve analysis (DCA) was a proposed novel method to compare the potential net benefit of predictive signatures. Instead of the AUC value in ROC curve which only demonstrated discriminative accuracy of predictive signature, DCA could visualize clinical consequences for treatment strategy and quantify net benefits at various threshold probabilities to evaluate clinical availability of metastasis risk nomogram.

Statistical analysis

Chi-square test, and univariate and multivariate logistic analysis were performed via SPSS 22.0 (SPSS Inc, Chicago, IL, USA). Construction and validation of the nomogram were conducted on R version 3.5.3 (<http://www.rproject.org>) with “rms” package. LASSO regression analysis was performed in R with “glmnet” package. ROC curves were conducted with “ROCR” package. DCA analysis was performed in R with “rmda” package. All P values were two-sided and statistical significance was set at $P < 0.05$.

Results

Demographic baseline and clinical characteristics

According to exclusion and inclusion criteria, entire 2195 osteosarcoma patients were identified from the SEER database (Fig 1). An optimal cutoff of tumor size (9.5cm) was defined by ROC analysis using clinical date of all the osteosarcoma patients. Based on ROC analysis result, AUC for this test was 0.650 (95% CI 0.630–0.670). The tumor size with largest Youden value was considered as the optimal cutoff. An optimal tumor size cutoff value 9.5cm was obtained with a diagnostic sensitivity of 60.24% (95% CI 55.4–65.0), a specificity of 62.31% (95% CI 60.0–64.6) and Youden value 0.2255 (Fig 2).

Table 1 showed demographic characteristics of all osteosarcoma patients. Of the 2195 patients with osteosarcoma, 659 patients (45.0%) were female and 805 patients (55.0%) were male. 1682 patients

(76.6%) were diagnosed when they were unmarried, while the others (23.4%) were married. Among those patients, 19.2% of the tumors were located at the axial bone, and the rest (80.8%) were in the extremity bone. 420 patients (19.1%) had distant metastasis disease, while the rest 1775 patients (80.9%) hadn't. Randomly, all of osteosarcoma patients were assigned into the training subgroup (n=1464) to establish and validate nomogram internally and validation subgroup (n=731) to validate nomogram externally. We performed Chi-square analysis in this table and found that patient sex, age, marital status, race, histologic subtype, tumor size, tumor site, tumor grade and use of surgery had no significant differences between validation and training subgroups (Table 1).

Feature selection

All 9 features were selected as potential predictors with nonzero coefficients (Fig 3A and 3B) in LASSO regression model. These clinical features included patient sex, age, marital status, race, tumor site, tumor grade, tumor size, histologic subtype and use of surgery.

Predictive factors for the risk of metastasis

With regard to the training subgroup, we used univariate logistic regression to analyses patient sex, age, marital status, race, tumor site, tumor grade, tumor size, histologic subtype and use of surgery. As the result, these variables, including sex, age, use of surgery, tumor size and tumor grade, were found to be connected with metastasis risk in osteosarcoma patients in the univariate analysis ($p < 0.05$) (Table 2). While tumor site, histologic subtype, marital status and race were excluded ($p > 0.05$) (Table 2). Furthermore, we performed multivariate logistic analysis to control the confounding variable. Sex, age, tumor grade, tumor size and use of surgery were recognized as independent predictive factors for the risk of metastasis in osteosarcoma patients.

Construction and validation of nomogram for metastasis risk in osteosarcoma patients

Independent predictive factors including sex, age, tumor size, tumor grade and use of surgery were included to construct nomogram predicting the risk of metastasis in osteosarcoma patients (Fig 4). The nomogram gave every prognostic factor a distinctive score on the nomogram point scale (Table 3). We also validated the nomogram internally and externally. Based on the information of an individual patient, we could use the nomogram by adding up the points of each predictive factor and correlating the total points with prediction in the nomogram. The calibration curve displayed perfect consistency between actual probability and predicted probability (Fig 5A). Fig 6 displayed the predictive accuracy of nomogram. AUC of the present model was 0.7427067. And AUC value of external validation subgroup was 0.6798374. These AUC results confirmed that this prognostic nomogram was reasonably accurate.

Clinical use

DCA for the nomogram was presented in Fig 5B. According to threshold probability in this analysis, the decision analysis curve was applied to assess the clinical application of metastasis prediction model. Decision curve analysis result showed that when the probability of metastasis generated by the

nomogram is less than 3% or more than 76%, decisions based on the nomogram would be meaningless. In other words, the nomogram provided additional values relative to the treat-all patients scheme or the treat-none scheme in these threshold probabilities. This analysis result suggested that this model was extremely useful for clinical determinations.

Discussion

Varieties of predictive factors have influence on the metastasis of osteosarcoma. Utility of single predictive factor to predict metastasis risk for individual osteosarcoma patients is limited. Nomogram, which have been widely used to predict the risk of metastasis of tumors for an individual patient[9–11, 18, 19], can incorporate all kind of predictive factors and calculate their cumulative effect to predict the probability of tumor patients with metastasis. Taking advantage of the SEER database, covering about 28% of overall U.S population, we constructed nomogram to predict the risk of metastasis in osteosarcoma patients.

Utilizing LASSO, univariate and multivariate logistic analysis, we identified the independent predictive factors for the risk of metastasis in osteosarcoma patients. LASSO, originally proposed for the linear regression model, was a popular high dimensional data analysis method which could applied for variable selection [20]. This method could improve both interpretation and prediction accuracy. In this current study, several clinicopathological characteristics such as age, sex, tumor grade, tumor size and use of surgery might become independent predictive factors for the risk of metastasis for osteosarcoma patients via multivariate analysis. We found that age was connected with the risk of metastasis for osteosarcoma patients. With the patient's age increasing, the risk become higher (except age < 20 group). According to Yusuke Tsuda's research, patients older than 65 years in Japan also showed a significantly higher proportion of tumors with metastasis at diagnosis [21]. In general, osteosarcoma is recognized as a disease of adolescents aged between 10 to 25 years old. This might result in higher point scale of age < 20 group when compared with young and middle-aged groups (20 < age < 60). There is a second incidence peak among individuals aged over 60 years in osteosarcoma [4]. With early diagnosis and treatment, elderly people is more likely to be diagnosed with metastasizing osteosarcoma. Consequently, we considered more studies should be conducted to prove this conclusion. We also identified sex as an independent predictive factor of metastasis for osteosarcoma patients, with male having a higher risk in comparison to female. There are some similar results observed in previous studies [22]. Estrada-Villaseor et al reported males have a higher frequency of metastasis in osteosarcoma in contrast to females [22]. And this result is also report in metastasis research of other tumor (Thyroid Cancer) [23]. In our study, use of surgery was an independent predictive factor of metastasis in osteosarcoma patients. Compared with osteosarcoma patients without surgery, the patients, who adopted surgery treatment after being diagnosed osteosarcoma, had a lower risk of metastasis. Surgery treatment increases disease-free survival probabilities according the guideline of treatment for osteosarcoma by European Society for Medical Oncology [24].

We also found that tumor grade was a definitive predictive factor for metastases in osteosarcoma patients, which is reported by previous study [25]. With regard to tumor size, we revealed that tumor size was independent predictive factor for the metastases of osteosarcoma, and Bielack [24] and Kim [26] et al consider tumor size as a definitive predictive factor of metastasis for osteosarcoma patients. In our study, we found that the cutoff of 9.5 cm in maximal tumor diameter was a predictive factor of metastasis. Conspicuously, osteosarcoma patient with maximal tumor diameter over 9.5 cm has a high probability of metastasis.

Based on the independent predictive factors identified by multivariate logistic analysis, we developed nomogram to predict the risk of metastasis for osteosarcoma patients. Utilizing this nomogram, we were able to easily and accurately predict the probability of metastasis in an individual patient. For instance, a 60-year-old female patient newly diagnosed with primary osteosarcoma of 12cm in size and with IV-grade. To use the nomogram (Fig. 4), we needed to take the following three steps. Firstly, we drew vertical lines from each independent predictive factor to the points scale to gain the matching points. Then, we added up each matching points of the independent predictive factors, and she got 21.5 points in total. Lastly, a vertical line was drawn from the total points scale to the risk scale obtaining the estimated risk of metastasis of osteosarcoma. Hence, her estimated risk of metastasis was 69%.

But our nomogram had some limitations. First, our nomogram was constructed only by clinical factors (without molecular markers). Information about several important clinicopathological parameters or molecular markers such as chemotherapy regimens, tumor volume, tumor location, and histopathology were also not complete. And the option of distant metastasis is not explicitly indicated metastasis sites. Specific metastasis sites data of every osteosarcoma was also not complete. Due to the unavailability of these variables in the SEER database, these variables could not be incorporated in our study. Their effects on metastasis in osteosarcoma patients deserve further exploration. Second, we could not use our nomogram to predict the time on which metastasis occurs for the reason that it was on the basis on logistic regression rather than Cox regression. Third, excluding patients with unknown important clinical information may cause bias. Fourth, our training set and validation set were quite small, so the generalizability of our nomogram ought to be verified in a larger population. More clinical data are needed to verify the results of this study, so as to further improve the credibility of the results.

Conclusions

In a brief, age, sex, tumor grade, tumor size and use of surgery were identified as independent predictive factors. Including those identified predictive factors, we developed and validated nomogram to estimate the risk of metastasis for individual osteosarcoma patients. The newly developed nomogram was highly sensitive and accurate. Utilization of our nomogram could help clinicians estimate individual patients' risk of metastasis of osteosarcoma.

Abbreviations

Surveillance, Epidemiology, and End Results (SEER)

receiver operating characteristic (ROC)

least absolute shrinkage and selection operator (LASSO)

Decision curve analysis (DCA)

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable

Availability of data and materials

The data for constructing model and external validation were obtained from the SEER database. All data generated or analyzed during this study are included in this published article.

Competing interests

All authors declare that they have no conflict of interest to state.

Funding

Scientific research project of health system Shenzhen Pingshan District Grant/Award Number 201844

Authors' contributions

Conception and design of the research: Wei Niu. Acquisition, analysis, and interpretation of data: Weilong Xu and Wei Niu. Drafting the manuscript: Weilong Xu and Wei Niu.

Acknowledgements

Not applicable.

References

1. Czarnecka AM, Synoradzki K, Firlej W, Bartnik E, Sobczuk P, Fiedorowicz M, Grieb P, Rutkowski P: **Molecular Biology of Osteosarcoma.** *Cancers* 2020, **12**(8).
2. Isakoff MS, Bielack SS, Meltzer P, Gorlick R: **Osteosarcoma: Current Treatment and a Collaborative Pathway to Success.** *J Clin Oncol* 2015, **33**(27):3029-3035.

3. Bielack SS, Kempf-Bielack B, Delling G, Exner GU, Flege S, Helmke K, Kotz R, Salzer-Kuntschik M, Werner M, Winkelmann W *et al*: **Prognostic 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):776-790.
4. Marina N, Gebhardt M, Teot L, Gorlick R: **Biology and therapeutic advances for pediatric osteosarcoma.** *The oncologist* 2004, **9**(4):422-441.
5. Bacci G, Rocca M, Salone M, Balladelli A, Ferrari S, Palmerini E, Forni C, Briccoli A: **High grade osteosarcoma of the extremities with lung metastases at presentation: treatment with neoadjuvant chemotherapy and simultaneous resection of primary and metastatic lesions.** *J Surg Oncol* 2008, **98**(6):415-420.
6. Bielack SS, Carrle D, Harges J, Schuck A, Paulussen M: **Bone tumors in adolescents and young adults.** *Current treatment options in oncology* 2008, **9**(1):67-80.
7. Grimes DA: **The nomogram epidemic: resurgence of a medical relic.** *Annals of internal medicine* 2008, **149**(4):273-275.
8. Kattan MW, Eastham JA, Stapleton AM, Wheeler TM, Scardino PT: **A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer.** *Journal of the National Cancer Institute* 1998, **90**(10):766-771.
9. Boutros C, Mazouni C, Lerebours F, Stevens D, Lei X, Gonzalez-Angulo AM, Delaloge S: **A preoperative nomogram to predict the risk of synchronous distant metastases at diagnosis of primary breast cancer.** *Br J Cancer* 2015, **112**(6):992-997.
10. Thompson AM, Turner RM, Hayen A, Aniss A, Jalaty S, Learoyd DL, Sidhu S, Delbridge L, Yeh MW, Clifton-Bligh R *et al*: **A preoperative nomogram for the prediction of ipsilateral central compartment lymph node metastases in papillary thyroid cancer.** *Thyroid : official journal of the American Thyroid Association* 2014, **24**(4):675-682.
11. Jiang L, Jiang S, Lin Y, Yang H, Xie Z, Lin Y, Long H: **Nomogram to Predict Occult N2 Lymph Nodes Metastases in Patients With Squamous Nonsmall Cell Lung Cancer.** *Medicine (Baltimore)* 2015, **94**(46):e2054.
12. Kuo TM, Mobley LR: **How generalizable are the SEER registries to the cancer populations of the USA?** *Cancer causes & control : CCC* 2016, **27**(9):1117-1126.
13. Noone AM, Lund JL, Mariotto A, Cronin K, McNeel T, Deapen D, Warren JL: **Comparison of SEER Treatment Data With Medicare Claims.** *Med Care* 2016, **54**(9):e55-64.
14. Cho SH: **Using multilevel analysis in patient and organizational outcomes research.** *Nursing research* 2003, **52**(1):61-65.
15. Yang QK, Lai QY, Wang Y, Wang Y, Yao ZX, Zhang XJ: **Establishment and validation of prognostic nomograms to predict overall survival and cancer-specific survival for patients with osteosarcoma.** *Neoplasma* 2020.
16. Obuchowski NA, Bullen JA: **Receiver operating characteristic (ROC) curves: review of methods with applications in diagnostic medicine.** *Physics in medicine and biology* 2018, **63**(7):07tr01.

17. Hu X, Shen F, Zhao Z, Qu X, Ye J: **An individualized gait pattern prediction model based on the least absolute shrinkage and selection operator regression.** *Journal of biomechanics* 2020, **112**:110052.
18. Iasonos A, Schrag D, Raj GV, Panageas KS: **How to build and interpret a nomogram for cancer prognosis.** *J Clin Oncol* 2008, **26**(8):1364-1370.
19. Dong D, Tang L, Li ZY, Fang MJ, Gao JB, Shan XH, Ying XJ, Sun YS, Fu J, Wang XX *et al*: **Development and validation of an individualized nomogram to identify occult peritoneal metastasis in patients with advanced gastric cancer.** *Annals of oncology : official journal of the European Society for Medical Oncology* 2019, **30**(3):431-438.
20. Vasquez MM, Hu C, Roe DJ, Halonen M, Guerra S: **Measurement error correction in the least absolute shrinkage and selection operator model when validation data are available.** *Statistical methods in medical research* 2019, **28**(3):670-680.
21. Tsuda Y, Ogura K, Shinoda Y, Kobayashi H, Tanaka S, Kawai A: **The outcomes and prognostic factors in patients with osteosarcoma according to age: a Japanese nationwide study with focusing on the age differences.** *BMC Cancer* 2018, **18**(1):614.
22. Estrada-Villaseor E, Escamilla-Urbe R, De la Garza-Montano P, Dominguez-Rubio R, Martinez-Lopez V, Avila-Luna A, Alfaro-Rodriguez A, Ruvalcaba-Paredes EK, Garciadiego-Cazares D, Bandala C: **Association of Metastasis with Clinicopathological Data in Mexican Patients with Osteosarcoma, Giant Cell Tumor of Bone and Chondrosarcoma.** *Asian Pacific journal of cancer prevention : APJCP* 2015, **16**(17):7689-7694.
23. Kao YH, Gan HK, Zaheer S, Lam WW, Loke KS, Wong WY, Ng DC, Goh AS: **Gender, Race, and Age at Diagnosis as Risk Factors for Metastasis or Recurrence among 1,657 Thyroid Cancer Patients Treated with Radioiodine across 40 Years in Singapore.** *Oncology research and treatment* 2015, **38**(12):679-682.
24. Bielack S, Carrle D, Casali PG, Group EGW: **Osteosarcoma: ESMO clinical recommendations for diagnosis, treatment and follow-up.** *Ann Oncol* 2009, **20 Suppl 4**:137-139.
25. Wang S, Zheng S, Hu K, Sun H, Zhang J, Rong G, Gao J, Ding N, Gui B: **A predictive model to estimate the pretest probability of metastasis in patients with osteosarcoma.** *Medicine* 2017, **96**(3):e5909.
26. Kim MS, Lee SY, Cho WH, Song WS, Koh JS, Lee JA, Yoo JY, Shin DS, Jeon DG: **An examination of the efficacy of the 8 cm maximal tumor diameter cutoff for the subdivision of AJCC stage II osteosarcoma patients.** *J Surg Oncol* 2008, **98**(6):427-431.

Tables

Table 1: Clinical characteristics and baseline demographic of osteosarcoma patients

Variables	Training subgroup (n=1464)		Validation subgroup (n=731)		Total (n=2195)		P
Sex, n, %							0.668
Female	659	45.0%	322	44.0%	981	44.7%	
Male	805	55.0%	409	56.0%	1214	55.3%	
Age, n, %							0.877
0-20	801	54.7%	414	56.6%	1215	55.4%	
21-40	352	24.0%	168	23.0%	520	23.7%	
41-60	202	13,8%	96	13.1%	298	13.6%	
61-80	90	6.1%	46	6.3%	136	6.2%	
>80	19	1.3%	7	1.0%	26	1.2%	
Tumor site, n, %							0.927
Axial	280	19.1%	141	19.3%	421	19.2%	
Extremity	1184	80.9%	590	80.7%	1774	80.8%	
Histology, n, %							0.725
9180	905	61.8%	450	61.6%	1355	61.7%	
9181	220	15.0%	109	14.9%	329	15.0%	
9182	101	6.9%	42	5.7%	143	6.5%	
9183	52	3.6%	29	4.0%	81	3.7%	
9184	9	0.6%	6	0.8%	15	0.7%	
9185	16	1.1%	5	0.7%	21	1.0%	
9186	43	2.9%	20	2.7%	63	2.9%	
9187	1	0.1%	3	0.4%	4	0.2%	
9192	93	6.4%	50	6.8%	143	6.5%	
9193	18	1.2%	13	1.8%	31	1.4%	
9194	6	0.4%	4	0.5%	10	0.5%	
Tumor stage, n, %							0.068
Non-metastasis	1168	79.8%	607	83.0%	1775	80.9%	
Distant metastasis	296	20.2%	124	17.0%	420	19.1%	

Surgery, n, %						0.262
No	133	9.1%	56	7.7%	189	8.6%
Yes	1331	90.9%	675	92.3%	2006	91.4%
Size, n, %						0.408
<9.5	822	56.1%	424	58.0%	1246	56.8%
≥9.5	642	43.9%	307	42.0%	949	43.2%
grade, n, %						0.148
I	80	5.5%	58	7.9%	138	6.3%
II	129	8.8%	62	8.5%	191	8.7%
III	399	27.3%	202	27.6%	601	27.4%
IV	856	58.5%	409	56.0%	1265	57.6%
Marital						0.169
Unmarried	1109	75.8%	573	78.4%	1682	76.6%
Married	355	24.2%	158	21.6%	513	23.4%
Race						0.505
Black	231	15.8%	119	16.3%	350	15.9%
Other	137	9.4%	79	10.8%	216	9.8%
White	1096	74.9%	533	72.9%	1629	74.2%

Abbreviation: 9180, conventional osteosarcoma; 9181, chondroblastic osteosarcoma; 9182, fibroblastic osteosarcoma; 9183, telangiectatic osteosarcoma; 9184, osteosarcoma in Paget disease of bone; 9185, small cell osteosarcoma; 9186, central osteosarcoma; 9187, intraosseous well differentiated osteosarcoma; 9192, parosteal osteosarcoma; 9193, periosteal osteosarcoma; 9194, high grade surface osteosarcoma.

Table 2: Univariable and multivariable logistic regression results for all variables

Variables	Univariable Analysis		Multivariable Analysis	
	OR (95%CI)	P	OR (95%CI)	P
Sex				
Female	1		1	
Male	1.473 (1.134-1.914)	0.004	1.345 (1.009-1.792)	0.043
Age				
0-20	1			
21-40	0.615 (0.437-0.865)	0.005	0.634 (0.434-0.924)	0.018
41-60	0.709 (0.471-1.067)	0.099	0.749 (0.479-1.169)	0.203
61-80	1.907 (1.197-3.040)	0.007	1.480 (0.855-2.563)	0.162
>80	2.640 (1.046-6.665)	0.040	1.939 (0.691-5.444)	0.209
Tumor site				
Axial	1			
Extremity	0.989 (0.716-1.368)	0.949		
Histology				
9193	1			
9180	2.447 (0.558-10.730)	0.235		
9181	1.832 (0.405-8.283)	0.431		
9182	1.395 (0.291-6.700)	0.677		
9183	1.455 (0.279-7.587)	0.657		
9184	4.000 (0.530-30.162)	0.179		
9185	3.636 (0.595-22.234)	0.162		
9186	1.297 (2.236-7.132)	0.765		
9187	0.000 (-)	1.000		
9192	0.267 (0.041-1.724)	0.165		
9194	1.600 (0.119-21.586)	0.723		
Surgery				
No	1		1	
Yes	0.128 (0.088-0.187)	<0.001	0.129 (0.086-0.193)	<0.001

Size				
<9.5	1		1	
≥9.5	2.788 (2.140-3.633)	<0.001	2.598 (1.952-3.458)	<0.001
Grade				
I	1		1	
II	1.949 (0.606-6.265)	0.263	1.887 (0.569-6.253)	0.299
III	5.614 (2.000-15.760)	0.001	3.567 (1.239-10.269)	0.018
IV	5.384 (1.945-14.905)	0.001	3.244 (1.144-9.199)	0.027
Marital				
Unmarried	1			
Married	0.852 (0.628-1.156)	0.304		
Race				
Black	1			
Other	0.769 (0.448-1.320)	0.340		
White	0.919 (0.651-1.300)	0.635		

Abbreviation: 9180, conventional osteosarcoma; 9181, chondroblastic osteosarcoma; 9182, fibroblastic osteosarcoma; 9183, telangiectatic osteosarcoma; 9184, osteosarcoma in Paget disease of bone; 9185, small cell osteosarcoma; 9186, central osteosarcoma; 9187, intraosseous well differentiated osteosarcoma; 9192, parosteal osteosarcoma; 9193, periosteal osteosarcoma; 9194, high grade surface osteosarcoma.

Table 3: Detailed scores of prognostic factors in the nomograms

Characteristic	Nomogram
Age	
<20	2.2
21-40	0
41-60	0.8
61-80	4.1
>80	5.5
Sex	
Female	0
Male	1.4
Surgery	
No	10.0
Yes	0
Size	
<9.5	0
≥9.5	4.7
Grade	
I	0
II	3.1
III	6.2
IV	5.7

Figures

Figure 1

Flow chart.

AUC=0.65

Figure 2

ROC curve for representative cut-off levels of tumor size. Tumor size levels were plotted with true positives on the vertical axis (sensitivity) and false positives (1-specificity) on the horizontal axis.

Figure 3

Clinical feature selection via the least absolute shrinkage and selection operator (LASSO) regression model. Notes: (A) Optimal parameter selection via LASSO regression model used fivefold cross-validation. The partial likelihood deviance curve was drawn versus $\log(\lambda)$. Dotted vertical lines were plotted at the optimal values via minimum criteria and the 1 standard error (SE) of minimum criteria (the 1-SE criteria). (B) LASSO coefficient profiles of 9 features. A coefficient profile plot was produced against $\log(\lambda)$ sequence. Vertical line was plotted at selected value selected via fivefold cross-validation where optimal λ led to all 9 features with nonzero coefficients.

Figure 4

Nomograms to predict metastasis risk for patients with osteosarcoma. Notes: Vertical line between points scale and each variable was showed to acquire points of each variable. And predicted metastasis risk was calculated and evaluated based on the total points which were drew by a vertical line from Total Points scale.

Figure 5

Decision curve and calibration curves analysis of predictive nomogram evaluation in the subgroup. Notes: (A) Calibration curves of the metastasis nomogram. The x-axis shows the predicted metastasis risk. And the y-axis shows actual diagnosed metastasis risk. The diagonal dotted line shows perfect prediction of ideal model. The solid line displays nomogram performance, of which a closer fit to diagonal dotted line displays a better prediction. (B) Decision curve analysis result of nomogram. And the y-axis evaluates the net benefit. The dotted line displays metastasis nomogram. The gray line displays the hypothesis that all the patients have metastasizing osteosarcoma. The black line displays the hypothesis that no patients are metastasizing osteosarcoma.

Figure 6

Predictive accuracy of nomogram to predict metastasis. Notes: (A) Predictive accuracy of nomogram in training subgroup. (B) Predictive accuracy of nomogram in validation subgroup