

# 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

Koichi Ogura, MD<sup>1,2</sup>; Tomohiro Fujiwara, MD<sup>1,3</sup>; Hideo Yasunaga, MD, PhD<sup>4</sup>; Hiroki Matsui, MPH<sup>4</sup>; Dae-Geun Jeon, MD<sup>5</sup>; Wan Hyeong Cho, MD<sup>5</sup>; Hiroaki Hiraga, MD<sup>6</sup>; Takeshi Ishii, MD<sup>7</sup>; Tsukasa Yonemoto, MD<sup>7</sup>; Hiroto Kamoda, MD<sup>7</sup>; Toshifumi Ozaki, MD<sup>8</sup>; Eiji Kozawa, MD<sup>8</sup>; Yoshihiro Nishida, MD<sup>8</sup>; Hideo Morioka, MD<sup>9</sup>; Toru Hiruma, MD<sup>10</sup>; Shigeki Kakunaga, MD<sup>11</sup>; Takafumi Ueda, MD<sup>11</sup>; Yusuke Tsuda, MD<sup>2</sup>; Hirotaka Kawano, MD<sup>2</sup>; and Akira Kawai, MD<sup>1</sup>

**BACKGROUND:** In this era of individualized cancer treatment, data that could be applied to predicting the survival of patients with osteosarcoma are still limited because of the rarity of the disease and the difficulty in accumulating a sufficient number of patients. Therefore, a multi-institutional collaboration was implemented to develop and externally validate nomograms that would predict metastasis-free survival (MFS) and overall survival (OAS) for patients with nonmetastatic osteosarcoma. **METHODS:** This study retrospectively examined 1070 patients treated with neoadjuvant chemotherapy and surgery for nonmetastatic osteosarcoma. Data from Japanese patients (n = 557) were used to develop multivariate nomograms based on Cox regression. Six clinical and pathologic variables were built into nomograms estimating the probability of MFS and OAS 3 and 5 years after diagnosis. The model was internally validated for discrimination and calibration with bootstrap resampling and was externally validated with an independent patient cohort from Korea (n = 513). **RESULTS:** A patient's age, tumor site, and histologic response were found to have a stronger influence on MFS and OAS in the model than sex, tumor size, or pathologic fracture. The nomograms and calibration plots based on these results well predicted the probability of MFS (concordance index, 0.631) and OAS (concordance index, 0.679). The concordance indices for external validation were 0.682 for MFS and 0.665 for OAS. **CONCLUSIONS:** The nomograms were externally validated and verified to be useful for the prediction of MFS and OAS and for the assessment of the postoperative prognosis. They can be used for counseling patients and for establishing appropriate surveillance strategies after surgery. *Cancer* 2015;121:3844-52. © 2015 The Authors. *Cancer* published by Wiley Periodicals, Inc. on behalf of *American Cancer Society*. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

**KEYWORDS:** external validation, nomogram, osteosarcoma, prognosis, survival.

## INTRODUCTION

Osteosarcoma, the most common sarcoma of bone, is characterized by the production of immature osteoid by the tumor cells. The long-term survival of osteosarcoma patients has improved dramatically because of the introduction of effective systemic chemotherapy.<sup>1-4</sup> Recent studies have demonstrated a more favorable outcome, especially for patients with nonmetastatic osteosarcoma.<sup>5,6</sup> Five-year survival rates of 60% to 80% have been reported for patients receiving intensive multidrug chemotherapy and surgery for aggressive local control.<sup>5-11</sup> However, despite the recent improvement in survival, a substantial number of patients still develop metastases or suffer tumor-related death, so more accurate identification of patient subgroups with different levels of risk is substantially important.

Clinical staging, such as staging using the American Joint Committee on Cancer classification,<sup>12</sup> which is based on tumor grade and size, skip metastases, and nodal or distant metastasis at the time of diagnosis is reportedly correlated with

**Corresponding author:** Akira Kawai, MD, Department of Musculoskeletal Oncology, National Cancer Center Hospital, 5-1-1 Tsukiji, Chuo-Ku, Tokyo 104-0045, Japan; Fax: (011) 81-3-3542-3815; akawai@ncc.go.jp

<sup>1</sup>Department of Musculoskeletal Oncology, National Cancer Center Hospital, Tokyo, Japan; <sup>2</sup>Department of Orthopedic Surgery, The University of Tokyo Hospital, Tokyo, Japan; <sup>3</sup>Department of Orthopedic Surgery, Okayama University, Okayama, Japan; <sup>4</sup>Department of Health Economics and Epidemiology Research, School of Public Health, The University of Tokyo, Tokyo, Japan; <sup>5</sup>Department of Orthopedic Surgery, Korea Cancer Center Hospital, Seoul, Korea; <sup>6</sup>Department of Orthopedic Surgery, Hokkaido Cancer Center, Hokkaido, Japan; <sup>7</sup>Department of Orthopedic Surgery, Chiba Cancer Center, Chiba, Japan; <sup>8</sup>Department of Orthopedic Surgery, Nagoya University, Aichi, Japan; <sup>9</sup>Department of Orthopedic Surgery, Keio University, Tokyo, Japan; <sup>10</sup>Department of Orthopedic Surgery, Kanagawa Cancer Center, Kanagawa, Japan; <sup>11</sup>Department of Orthopedic Surgery, Osaka National Hospital, Osaka, Japan

Additional supporting information may be found in the online version of this article.

**DOI:** 10.1002/cncr.29575, **Received:** March 20, 2015; **Revised:** June 16, 2015; **Accepted:** June 19, 2015, **Published online** July 20, 2015 in Wiley Online Library (wileyonlinelibrary.com)

survival for patients with bone sarcomas.<sup>13</sup> However, in osteosarcoma, other prognostic factors, including patient age,<sup>14-16</sup> tumor location,<sup>5,16</sup> pathologic fracture,<sup>17</sup> and responses to preoperative chemotherapy, have been reported to correlate strongly with survival.<sup>5,16,18</sup> Therefore, for the prediction of survival after treatment, it would be desirable to consider these factors, and a comprehensive and easy-to-use tool that incorporated all of them and could be used to assess individualized survival would be valuable for decision making.

A nomogram is a graphical depiction of a prediction model that can be used as a statistically based tool for assessing the overall probability of a specific outcome for any individualized patient. In the era of individualized cancer treatment, the use of such a tool for the estimation of survival would be particularly helpful for physicians when they are determining an appropriate interval for follow-up or imaging. To our knowledge, however, data that could be used for developing a nomogram for osteosarcoma patients are still limited.<sup>16,19</sup> In addition, no externally validated nomogram has been reported to date because of the rarity of the disease and the difficulty in collecting a sufficient number of patients.

Therefore, in the current study, we collected data from a large nationwide cohort of patients with osteosarcoma in Japan to develop nomograms for predicting metastasis-free survival (MFS) and overall survival (OAS) as part of a nationwide research project for the development of a prognostic nomogram for musculoskeletal tumors. In addition, we implemented a multi-institutional collaboration to validate them externally with data from 1 institution in Korea.

## MATERIALS AND METHODS

### *Eligibility*

This study was designed as a multi-institutional, cooperative, retrospective study using 11 major referral centers in Japan as a training set (National Cancer Center Hospital, Hokkaido Cancer Center, Chiba Cancer Center, Okayama University, Nagoya University, Keio University, Kanagawa Cancer Center, Osaka National Hospital, Osaka Medical Center for Cancer and Cardiovascular Diseases, Osaka University, and The University of Tokyo Hospital) and 1 major referral center in Korea as a validation set (Korea Cancer Center Hospital). This study was approved by the institutional review board of each hospital.

The inclusion criteria for patients were as follows: 1) a histological diagnosis of high-grade osteosarcoma of the extremity or trunk, 2) no distant metastasis at presentation,

3) being scheduled for neoadjuvant chemotherapy and surgery, and 4) a follow-up period of more than 3 years for survivors. Patients with osteosarcoma of the craniofacial bone, extraskelatal osteosarcoma, or secondary osteosarcoma were excluded. Clinical data, treatment modalities, and treatment outcomes were reviewed retrospectively by reference to the medical records at each institution.

### *Prognostic Variables*

The following data were extracted from 203 patients treated at 3 institutions (National Cancer Center Hospital, Chiba Cancer Center, and The University of Tokyo Hospital): patient age at diagnosis, sex, primary tumor site, maximal tumor size, presence or absence of a pathologic fracture before surgery (ie, at diagnosis and during preoperative chemotherapy), type of surgery (limb salvage surgery, amputation, or rotation plasty), surgical margin (positive or negative), and histologic response to preoperative chemotherapy. The primary tumor sites were classified into 6 categories according to the number of cases: distal extremity, distal femur, proximal tibia/fibula, proximal humerus, proximal femur, and trunk. A distal extremity was defined as any location distal to the knee or elbow joint with the exception of the proximal tibia/fibula (ie, radius, ulna, distal tibia, or hand and foot). The trunk included the pelvic bone, spine, rib, clavicle, sternum, and scapula. A histologic response was assessed by an examination of the resected specimen after surgery, and histologic responses were categorized into 4 groups according to the degree of tumor necrosis after preoperative chemotherapy, as described previously: grade 1 (<50%), grade 2 (50%-89%), grade 3 (90%-99%), and grade 4 (100%).<sup>1</sup> According to the results of the initial analysis, we excluded 2 variables from the analysis: type of surgery and surgical margin. The type of surgery was excluded because no significant difference was seen between the amputation and limb salvage surgery groups ( $P = .120$ ). Surgical margins did not affect MFS and OAS, perhaps because of the extremely small number ( $n = 2$ ) of margin-positive cases (MFS,  $P = .616$ ; OAS,  $P = .059$ ). As a result, data for 6 prognostic variables (age, sex, site, size, pathologic fracture, and histologic response to preoperative chemotherapy) were collected from the remaining 8 institutions.

### *Statistical Analysis*

The primary endpoints were the occurrence of 1) distant metastasis and 2) tumor-related death. MFS was defined as the period from the date of diagnosis until the appearance of distant metastasis or until the last follow-up for patients without metastasis. OAS was defined as the

**TABLE 1.** Descriptive Characteristics of the Study Population

Characteristic	Entire Cohort (n = 1070)	Training Set: Japanese Cohort (n = 557)	Validation Set: Korean Cohort (n = 513)	P
Age, mean (SD), y	20.2 (12.9)	20.6 (13.5)	18.5 (10.4)	.006
Age, No. (%)				
≤12 y	265 (24.8)	136 (24.4)	129 (25.1)	.026
13-19 y	482 (45.0)	236 (42.4)	246 (48.0)	
20-39 y	242 (22.6)	131 (23.5)	111 (21.6)	
≥40 y	81 (7.6)	54 (9.7)	27 (5.3)	
Sex, No. (%)				
Male	677 (63.3)	344 (61.8)	333 (64.9)	.285
Female	393 (36.7)	213 (38.2)	180 (35.1)	
Tumor size, mean (SD), cm	9.8 (4.1)	9.7 (4.1)	9.9 (4.1)	.329
Tumor size, No. (%)				
≤8 cm	450 (42.1)	246 (44.2)	204 (39.8)	.145
>8 cm	620 (57.9)	311 (55.8)	309 (60.2)	
Location, No. (%)				
Distal extremity	70 (6.5)	42 (7.5)	28 (5.5)	.520
Distal femur	476 (44.5)	251 (45.1)	225 (43.9)	
Proximal tibia/fibula	304 (28.4)	151 (27.1)	153 (29.8)	
Proximal humerus	96 (9.0)	47 (8.4)	49 (9.6)	
Proximal femur	77 (7.2)	38 (6.8)	39 (7.6)	
Trunk	47 (4.4)	28 (5.0)	19 (3.7)	
Pathologic fracture, No. (%)				
No	978 (91.4)	516 (92.6)	462 (90.1)	.132
Yes	92 (8.6)	41 (7.4)	51 (9.9)	
Histologic response (tumor necrotic rate), No. (%)				
Grade 1 (<50%)	244 (22.8)	144 (25.9)	100 (19.5)	.013
Grade 2 (50%-89%)	357 (33.4)	168 (30.2)	189 (36.8)	
Grade 3 (90%-99%)	322 (30.1)	176 (31.6)	146 (28.5)	
Grade 4 (100%)	147 (13.7)	69 (12.4)	78 (15.2)	

Abbreviation: SD, standard deviation.

period from the date of diagnosis until tumor-related death or until the last follow-up for survivors. Patients without distant metastasis or tumor-related death or patients who died without metastasis or because of other causes were censored at the last follow-up. OAS and MFS were estimated with the Kaplan-Meier method, and comparisons were assessed with the log-rank test. Multivariate analysis was conducted with the Cox proportional hazards model. In univariate and multivariate analyses, differences were considered to be statistically significant at  $P < .05$ .

### Nomograms

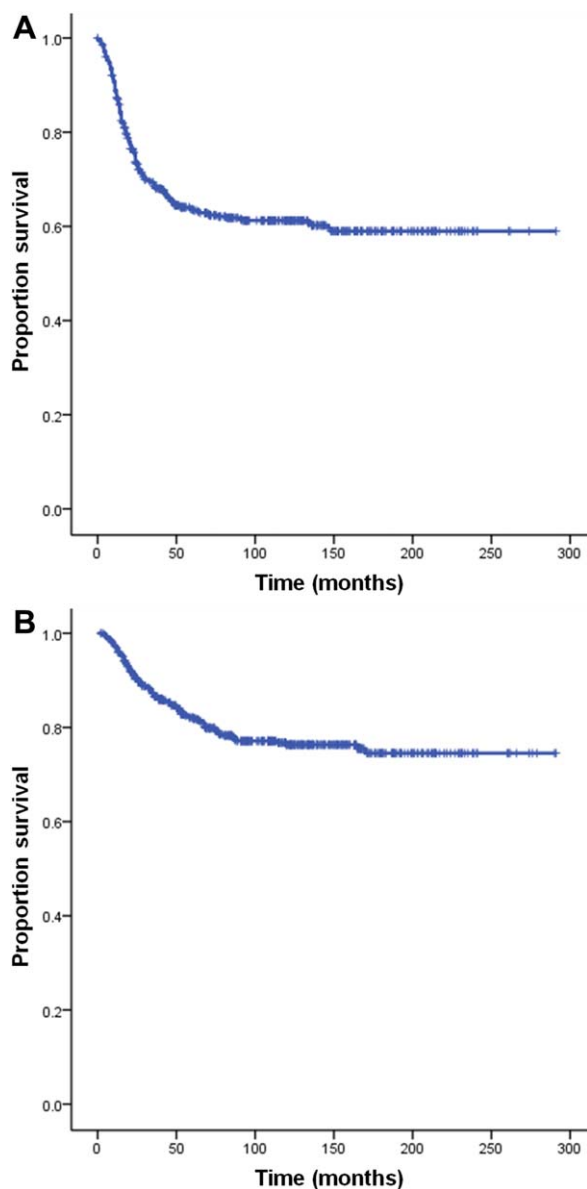
For the development of the multivariate nomograms based on Cox regression, data from 11 major referral centers in Japan (training set) were used ( $n = 557$ ). The final regression model was chosen on the basis of the clinical and statistical significance of the predictors. The 3- and 5-year predicted probabilities of each endpoint were calculated for each patient with the Cox regression model underlying the nomogram. The discriminative ability of the nomogram was assessed with the concordance index for purposes of comparison with the literature. The predictive model was internally validated with 200 bootstrap samples to prevent overfitting and obtain a relatively

unbiased estimate. Calibration of the nomogram was assessed via the plotting of the observed probabilities against the 3- and 5-year nomogram-predicted MFS and OAS. The model was externally validated with data from 1 institution from Korea (validation set;  $n = 513$ ). The model performance for predicting outcomes was evaluated by the calculation of the concordance index also in calibration and external validation. In external validation, predictive accuracy was also assessed.

All statistical analyses were conducted with IBM SPSS 19.0 (IBM SPSS, Armonk, NY), and the nomogram was built with R 3.0.1 (R Foundation for Statistical Computing, Vienna, Austria) with the rms library.

### RESULTS

During the whole period of 1990-2010, we identified 1270 eligible patients who met the inclusion criteria. Among these patients, we excluded those for whom any data were missing with respect to the age at diagnosis, sex, tumor site, maximum tumor size, presence or absence of any pathologic fracture before surgery, and histologic response to preoperative chemotherapy; for example, patients who underwent reconstruction with a devitalized autograft and whose histologic response to preoperative



**Figure 1.** Kaplan-Meier curves for (A) metastasis-free survival and (B) overall survival for all patients.

chemotherapy was not able to be evaluated were excluded. Consequently, we included 1070 patients (557 in the training set and 513 in the validation set) who had a complete set of data for prognostic variables. Among these patients, the chemotherapy regimen in the Japanese cohort was based on a protocol composed of methotrexate, doxorubicin, and cisplatin (MAP) or MAP plus ifosfamide.<sup>20,21</sup> The majority of the patients in the Korean cohort were treated with MAP.<sup>22</sup>

The clinicopathologic characteristics of the training set and validation set are summarized in Table 1. The

mean follow-up period was 94 months (range, 12-291 months). Distant metastasis and tumor-related death occurred in 206 patients and 114 patients, respectively. The 3- and 5-year MFS rates for all patients were 69% and 64% (Fig. 1A), respectively. The 3- and 5-year OAS rates for the patients overall were 87% and 82% (Fig. 1B), respectively. The univariate associations of the various factors with MFS and OAS rates determined via Kaplan-Meier plots are shown in Supporting Figures 1 and 2, respectively (see online supporting information).

Table 2 shows the univariate and multivariate Cox proportional hazards models for MFS and OAS. Univariate analyses showed that age, tumor size, tumor site, pathologic fracture, and degree of tumor necrosis were significantly associated with an increased risk of distant metastasis. Moreover, univariate analyses showed that patient age, sex, tumor size, tumor site, pathologic fracture, and degree of tumor necrosis were significantly associated with an increased risk of tumor-related death. Multivariate analyses demonstrated significant associations between distant metastasis and a patient age of 13 to 19 years (hazard ratio [HR], 1.51; 95% confidence interval [CI], 1.03-2.21;  $P = .035$ ) or  $\geq 40$  years (HR, 2.12; 95% CI, 1.28-3.52;  $P = .004$ ), a tumor size  $> 8$  cm (HR, 1.53; 95% CI, 1.14-2.06;  $P = .005$ ), a tumor located in the proximal humerus (HR, 2.84; 95% CI, 1.31-6.14;  $P = .008$ ), the presence of pathologic fracture (HR, 1.66; 95% CI, 1.05-2.63;  $P = .032$ ), and the degree of tumor necrosis (50%-89% [HR, 0.71; 95% CI, 0.51-0.99;  $P = .045$ ], 90%-99% [HR, 0.53; 95% CI, 0.37-0.77;  $P = .001$ ], and 100% [HR, 0.35; 95% CI, 0.19-0.62;  $P < .001$ ]). Moreover, multivariate analyses demonstrated significant associations between tumor-related death and a patient age of 20 to 39 years (HR, 1.96; 95% CI, 1.06-3.64;  $P = .032$ ) or  $\geq 40$  years (HR, 3.50; 95% CI, 1.77-6.92;  $P < .001$ ), female sex (HR, 0.62; 95% CI, 0.41-0.93;  $P = .022$ ), a tumor size  $> 8$  cm (HR, 1.59; 95% CI, 1.05-2.39;  $P = .028$ ), the tumor site (distal femur [HR, 8.39; 95% CI, 1.15-61.06;  $P = .036$ ], proximal tibia/fibula [HR, 10.75; 95% CI, 1.46-78.90;  $P = .020$ ], proximal humerus [HR, 9.45; 95% CI, 1.20-74.15;  $P = .033$ ], proximal femur [HR, 14.15; 95% CI, 1.84-108.72;  $P = .011$ ], and trunk [HR, 13.15; 95% CI, 1.63-105.93;  $P = .015$ ]), the presence of pathologic fracture (HR, 2.55; 95% CI, 1.46-4.43;  $P = .001$ ), and the degree of tumor necrosis (50%-89% [HR, 0.57; 95% CI, 0.36-0.90;  $P = .015$ ], 90%-99% [HR, 0.48; 95% CI, 0.29-0.78;  $P = .003$ ], and 100% [HR, 0.28; 95% CI, 0.12-0.63;  $P = .002$ ]).

**TABLE 2.** Cox Proportional Hazards Models for Metastasis-Free Survival and Overall Survival

	Metastasis-Free Survival				Overall Survival			
	Univariate Analysis		Multivariate Analysis		Univariate Analysis		Multivariate Analysis	
	Hazard Ratio (95% CI)	<i>P</i>	Hazard Ratio (95% CI)	<i>P</i>	Hazard Ratio (95% CI)	<i>P</i>	Hazard Ratio (95% CI)	<i>P</i>
Age								
≤12 y	Reference		Reference		Reference		Reference	
13-19 y	1.51 (1.04-2.21)	.031	1.51 (1.03-2.21)	.035	1.74 (1.00-3.02)	.049	1.72 (0.98-3.00)	.058
20-39 y	1.30 (0.85-2.00)	.229	1.28 (0.82-1.98)	.277	1.85 (1.01-3.38)	.046	1.96 (1.06-3.64)	.032
≥40 y	2.55 (1.57-4.16)	<.001	2.12 (1.28-3.52)	.004	4.31 (2.26-8.24)	<.001	3.50 (1.77-6.92)	<.001
Sex								
Male	Reference		Reference		Reference		Reference	
Female	0.83 (0.62-1.10)	.198	0.85 (0.64-1.14)	.271	0.60 (0.40-0.90)	.013	0.62 (0.41-0.93)	.022
Tumor size								
≤8 cm	Reference		Reference		Reference		Reference	
>8 cm	1.64 (1.23-2.18)	.001	1.53 (1.14-2.06)	.005	1.69 (1.14-2.48)	.008	1.59 (1.05-2.39)	.028
Tumor site								
Distal extremity	Reference		Reference		Reference		Reference	
Distal femur	1.98 (1.00-3.92)	.05	1.80 (0.90-3.58)	.095	9.27 (1.28-67.10)	.028	8.39 (1.15-61.06)	.036
Proximal tibia/fibula	1.74 (0.857-3.54)	.125	1.72 (0.84-3.50)	.138	10.21 (1.40-74.64)	.022	10.75 (1.46-78.90)	.020
Proximal humerus	3.13 (1.45-6.73)	.004	2.84 (1.31-6.14)	.008	10.36 (1.33-80.95)	.026	9.45 (1.20-74.15)	.033
Proximal femur	2.31 (1.02-5.24)	.044	1.90 (0.83-4.33)	.127	17.89 (2.34-136.77)	.005	14.15 (1.84-108.72)	.011
Trunk	2.98 (1.25-7.06)	.013	2.27 (0.95-5.43)	.065	19.76 (2.47-158.07)	.005	13.15 (1.63-105.93)	.015
Pathologic fracture								
No	Reference		Reference		Reference		Reference	
Yes	1.81 (1.15-2.84)	.01	1.66 (1.05-2.63)	.032	2.38 (1.40-4.04)	.004	2.55 (1.46-4.43)	.001
Histologic response (tumor necrotic rate)								
Grade 1 (<50%)	Reference		Reference		Reference		Reference	
Grade 2 (50%-89%)	0.71 (0.51-0.99)	.046	0.71 (0.51-0.99)	.045	0.55 (0.35-0.85)	.008	0.57 (0.36-0.90)	.015
Grade 3 (90%-99%)	0.50 (0.35-0.71)	<.001	0.53 (0.37-0.77)	.001	0.41 (0.25-0.65)	<.001	0.48 (0.29-0.78)	.003
Grade 4 (100%)	0.33 (0.18-0.58)	<.001	0.35 (0.19-0.62)	<.001	0.25 (0.11-0.55)	.001	0.28 (0.12-0.63)	.002

Abbreviation: CI, confidence interval.

On the basis of these results, we developed nomograms that visually depicted the multivariate impact of each variable with the Cox regression model. The nomograms were able to predict MFS (Fig. 2A) and OAS (Fig. 2B) 3 or 5 years after the date of diagnosis.

The calibration plots for the probabilities of MFS (concordance index, 0.631) and OAS (concordance index, 0.679) are shown in Supporting Figure 3 (see online supporting information). The calibration plots for the probabilities of MFS and OAS were good for external validation and are shown in Figure 3 (concordance index, 0.682 and 0.665, respectively).

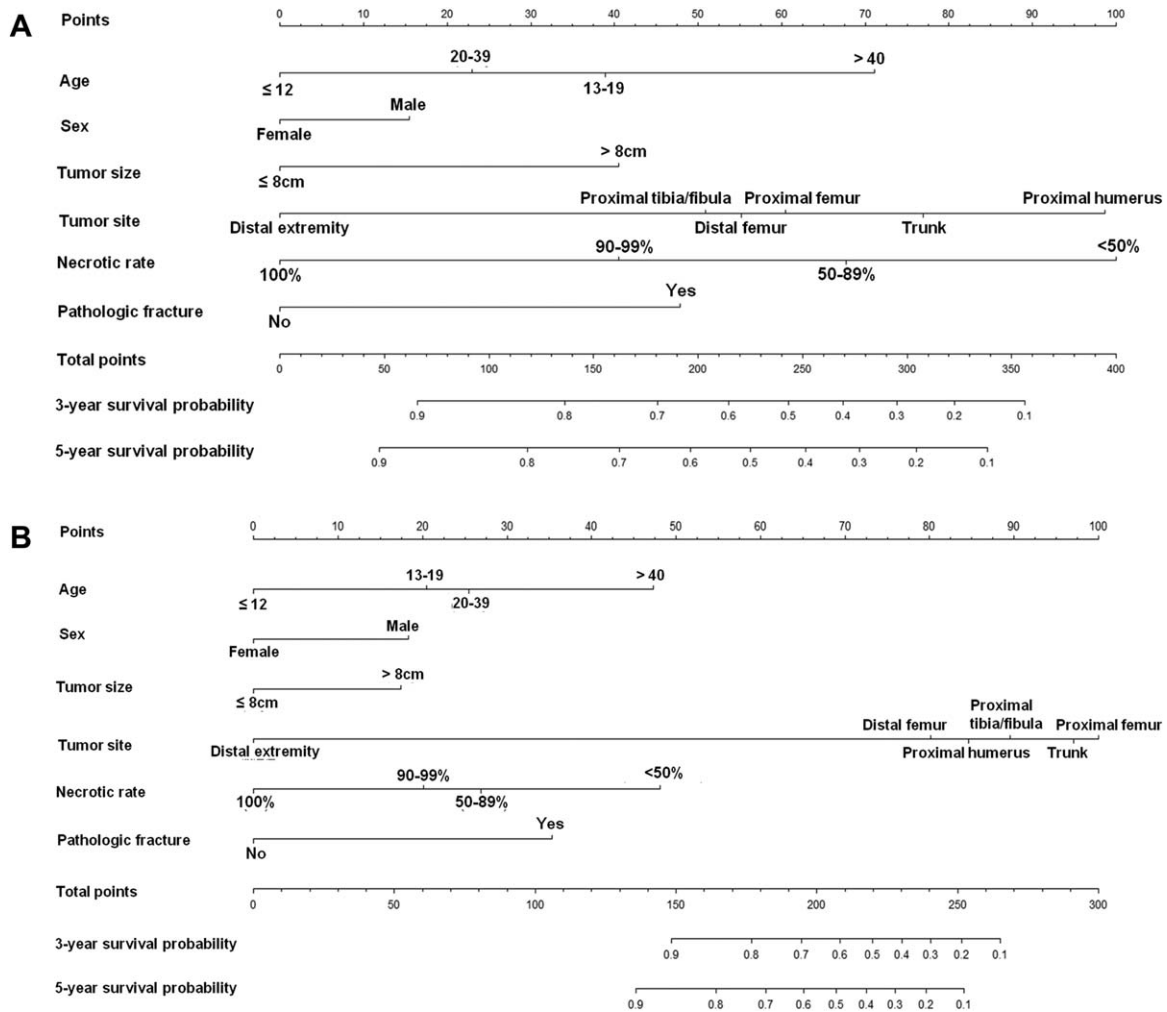
## DISCUSSION

Recently, nomograms have become widely accepted as models for risk prediction in cancer because they can provide an appreciation of the magnitude of the impact of individual factors on outcome probability, in that they are more predictive than the conventionally used American Joint Committee on Cancer stage alone.<sup>23</sup> However, few data that could be applied to a prognostic nomogram for sarcoma have been available,<sup>24</sup> especially with respect to

bone sarcoma, because of its rarity and the difficulty in collecting a sufficient number of cases.<sup>16,19</sup> In the current study, we developed and validated multivariate nomograms internally and externally that incorporated several commonly evaluable factors simultaneously to allow individual estimations of 3- and 5-year MFS and OAS for patients with nonmetastatic osteosarcoma treated by surgery and chemotherapy with curative intent. This model uses many of the discrete factors that have been shown to affect risk, and it combines them into a user-friendly calculation tool.

It is not unusual for musculoskeletal oncologists to encounter patients who develop distant metastasis or suffer tumor-related death even though they receive identical definitive treatment, and the ability to identify such higher risk patients would undoubtedly be valuable. To date, however, quantitative data that could be useful for this purpose have been lacking, and currently, there is ongoing controversy regarding the appropriate timing of follow-up because of the dilemma of avoiding unnecessary frequent imaging tests versus any resulting delay in the detection of distant metastasis, which may ultimately lead



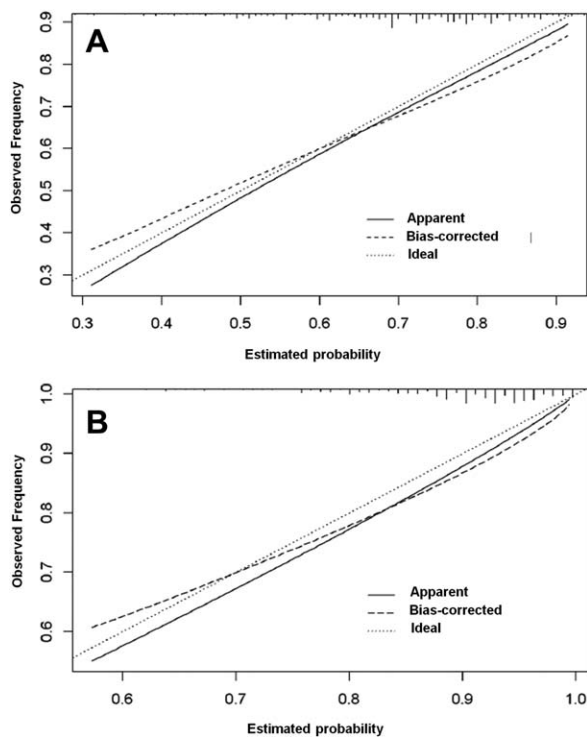


**Figure 2.** Nomogram predicting the probability of (A) distant metastasis and (B) overall survival at 3 and 5 years. The patient's value for each variable is plotted on the appropriate scale, and vertical lines are drawn to the line of points to obtain the corresponding scores. All scores should be summed to obtain the total point score. The total point score on the total point line is plotted, and a vertical line is drawn down to the bottom line. The corresponding value shows the predicted probability of distant metastasis or death.

to tumor-related death. However, it cannot be denied that the appropriate timing of follow-up and the planning of imaging tests should be stratified according to the degree of risk in individual patients. Our nomograms make it easier to derive individualized predictions of patient outcome; they allow physicians to recognize the risk of distant metastasis or tumor-related death after treatment and provide a more informative explanation for their patients.

To date, only 2 prognostic nomograms for osteosarcoma have been available.<sup>16,19</sup> The first nomogram was designed to predict the 5-year probability of metastasis after neoadjuvant chemotherapy and surgery; it was based on 365 patients with nonmetastatic osteosarcoma treated at a single center. It stratified the patients into 108 risk

groups on the basis of 4 variables: age at diagnosis (3 categories), primary tumor site (3 categories), tumor size (3 categories), and histologic response to chemotherapy (4 categories).<sup>16</sup> The second nomogram was designed to predict the 5-year probability of metastasis after neoadjuvant chemotherapy and limb salvage surgery; it was based on 91 patients with stage IIB osteosarcoma treated at a single center. It stratified the patients into 16 risk groups on the basis of 4 variables: serum alkaline phosphatase (ALP) level at diagnosis (2 categories), primary tumor site (2 categories), presence or absence of joint invasion of the tumor (2 categories), and histologic response to chemotherapy (2 categories).<sup>19</sup> Although they are brief instruments and easy to use, a nomogram that has the potential



**Figure 3.** Calibration plot for 3-year probabilities of (A) distant metastasis and (B) death for external validation.

to predict the outcome in more detail on the basis of a larger number of patients would be more ideal.

The current nomograms appear to have several differences and advantages resulting from their higher statistical power in comparison with the previously reported nomograms.<sup>16,19</sup> First, they enable more detailed prediction by incorporating a greater number of important factors and dividing patients into 768 risk groups based on 6 variables: age at diagnosis (4 categories), sex (2 categories), primary tumor site (6 categories), tumor size (2 categories), histologic response to chemotherapy (4 categories), and presence or absence of pathologic fracture (2 categories). Second, the higher statistical power in our series enabled us to classify the primary tumor site in a more detailed manner. Because the location of a tumor, such as a proximal extremity versus a distal extremity, is known to be a significant factor associated with prognosis, this characteristic should be dichotomized, even for the same bone. Third, we were able to incorporate the presence of pathologic fracture, one of the potentially important factors affecting prognosis, into the nomograms. Although to date the impact of pathologic fracture has been controversial in previous large series,<sup>25-27</sup> we have shown that pathologic fracture is significantly associated with both

MFS and OAS. Fourth, prognosis stratified by age was treated differently. In our series, we confirmed that OAS and MFS were poorest in patients who were 40 years old or older, and this was consistent with the widely accepted concept that elderly patients with osteosarcoma have a poorer outcome than younger patients.<sup>15,28</sup> This trend may have resulted from the intensity of the chemotherapy employed. Fifth, although it is commonly recognized that tumor size is a strong prognostic indicator, it had a weaker impact than expected on OAS and MFS in comparison with other variables. These results were concordant with the previously reported nomogram.<sup>16</sup> Among the possible reasons for this, the tumor size at diagnosis may have a weaker impact than the tumor size at the time of surgery because osteosarcoma is sensitive to chemotherapy. Another reason may be that extremely large tumors (>15 cm) were rare in our series; the median size was 9.0 cm (interquartile range, 7.0-12.0 cm). In addition, these results should be interpreted carefully because of the relatively wide 95% CI for OAS for age and tumor site, which may have biased the results. Sixth, although female sex hormones have been reported to be associated with a better prognosis *in vitro*,<sup>29</sup> no previous reports have demonstrated this in an actual large patient cohort. In our series, males had statistically significantly worse survival than females. Finally, there have been no externally validated prognostic nomograms for osteosarcoma, and our nomograms are the first to be externally validated with an independent patient cohort.

There were several limitations to the current study. First, it was performed with retrospective data, and we were unable to determine the decision process for treatment because we extracted data only for patients with musculoskeletal sarcoma treated surgically. Therefore, the indications for surgery may have differed among the participating hospitals, and this could have resulted in a selection bias. We sought to overcome this potential problem by collecting the largest number of cases possible. Second, although our nomograms were developed with data from the largest cohort accumulated so far, it may be necessary to collect many more patients for a nomogram predicting OAS with fewer endpoints versus a nomogram for MFS. This led to a relatively wide 95% CI for OAS. Third, our nomograms used information that became available only after surgical treatment. Although the nomograms may be predictively accurate, they cannot be used for patients requiring preoperative prognosis-based decisions regarding tumor resectability or nonresectability when severe functional deterioration is likely, as is frequently the case for osteosarcoma of the spine or pelvis. This may be the

reason that patients with truncal tumors had better outcomes than expected. Fourth, we did not include certain major prognostic factors, such as the level of ALP or lactate dehydrogenase (LDH).<sup>30</sup> The first reason for this was that ALP or LDH data were not available for some patients, so the inclusion of these markers would have led to lower statistical power. The second reason was that the levels of ALP and LDH were sometimes measured with different methods, and the standard values used among the hospitals differed. Therefore, considering the nature of this multi-institutional study, we felt it necessary to omit the data for ALP and LDH. However, it is likely that the inclusion of these serological markers or known molecular markers such as P-glycoprotein,<sup>31</sup> CXCR4,<sup>32</sup> and ezrin<sup>33</sup> would improve the predictive ability of future nomograms. Fifth, the trunk cases had better outcomes than expected (better than proximal humerus cases) with respect to MFS. The first possible reason is that the trunk category is too all-inclusive (being composed of the pelvic bone, spine, rib, clavicle, sternum, and scapula), and this heterogeneity of the category may be associated with the unexpected result. Another reason for this unexpected result is advanced pelvic or spine cases, which were not able to be indicated radical surgery, may have been excluded from the current cohort, and such a selection bias might be responsible for the unexpected result. Finally, although our nomograms were reliable with relatively good concordance indices for external validation, further external validation using a cross-racial patient cohort or prospectively collected data in the future will also increase model reliability.

In conclusion, on the basis of the largest patient cohort reported so far, we have created useful nomograms for the prediction of MFS and OAS for patients with nonmetastatic osteosarcoma, and we have externally validated them with an independent patient cohort. These nomograms use only common and readily available variables for the accurate prediction of survival, and they offer an improvement over current osteosarcoma staging, as demonstrated by adequate nomogram discrimination and calibration, and more dynamic and robust risk stratification. This tool can be used for patient counseling and for establishing appropriate surveillance strategies after surgery.

## FUNDING SUPPORT

This study was funded by a Grant-in-Aid for Scientific Research from the Ministry of Education and Science (B, no. 22390296) and by the National Cancer Center Research and Development Fund (23-A-10).

## CONFLICT OF INTEREST DISCLOSURES

The authors made no disclosures.

## REFERENCES

- Rosen G, Marcove RC, Huvos AG, et al. Primary osteogenic sarcoma: eight-year experience with adjuvant chemotherapy. *J Cancer Res Clin Oncol.* 1983;106(suppl):55-67.
- Eilber F, Giuliano A, Eckardt J, Patterson K, Moseley S, Goodnight J. Adjuvant chemotherapy for osteosarcoma: a randomized prospective trial. *J Clin Oncol.* 1987;5:21-26.
- Link MP, Goorin AM, Miser AW, et al. The effect of adjuvant chemotherapy on relapse-free survival in patients with osteosarcoma of the extremity. *N Engl J Med.* 1986;314:1600-1606.
- Link MP, Goorin AM, Horowitz M, et al. Adjuvant chemotherapy of high-grade osteosarcoma of the extremity. Updated results of the Multi-Institutional Osteosarcoma Study. *Clin Orthop Relat Res.* 1991;8-14.
- Bielack SS, Kempf-Bielack B, Delling G, 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;20:776-790.
- Meyers PA, Schwartz CL, Krailo M, et al. Osteosarcoma: a randomized, prospective trial of the addition of ifosfamide and/or muramyl tripeptide to cisplatin, doxorubicin, and high-dose methotrexate. *J Clin Oncol.* 2005;23:2004-2011.
- Goorin AM, Schwartzentruber DJ, Devidas M, et al. Presurgical chemotherapy compared with immediate surgery and adjuvant chemotherapy for nonmetastatic osteosarcoma: Pediatric Oncology Group Study POG-8651. *J Clin Oncol.* 2003;21:1574-1580.
- Bacci G, Bertoni F, Longhi A, et al. Neoadjuvant chemotherapy for high-grade central osteosarcoma of the extremity. Histologic response to preoperative chemotherapy correlates with histologic subtype of the tumor. *Cancer.* 2003;97:3068-3075.
- Patel SJ, Lynch JW Jr, Johnson T, et al. Dose-intense ifosfamide/doxorubicin/cisplatin based chemotherapy for osteosarcoma in adults. *Am J Clin Oncol.* 2002;25:489-495.
- Wilkins RM, Cullen JW, Odom L, et al. Superior survival in treatment of primary nonmetastatic pediatric osteosarcoma of the extremity. *Ann Surg Oncol.* 2003;10:498-507.
- Crews KR, Liu T, Rodriguez-Galindo C, et al. High-dose methotrexate pharmacokinetics and outcome of children and young adults with osteosarcoma. *Cancer.* 2004;100:1724-1733.
- Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti A. *AJCC Cancer Staging Manual.* 7th ed. New York, NY: Springer; 2010:281-290.
- Heck RK Jr, Stacy GS, Flaherty MJ, Montag AG, Peabody TD, Simon MA. A comparison study of staging systems for bone sarcomas. *Clin Orthop Relat Res.* 2003;64-71.
- Bacci G, Longhi A, Versari M, Mercuri M, Briccoli A, Picci P. Prognostic factors for osteosarcoma of the extremity treated with neoadjuvant chemotherapy: 15-year experience in 789 patients treated at a single institution. *Cancer.* 2006;106:1154-1161.
- Carsi B, Rock MG. Primary osteosarcoma in adults older than 40 years. *Clin Orthop Relat Res.* 2002;53-61.
- Kim MS, Lee SY, Lee TR, et al. Prognostic nomogram for predicting the 5-year probability of developing metastasis after neo-adjuvant chemotherapy and definitive surgery for AJCC stage II extremity osteosarcoma. *Ann Oncol.* 2009;20:955-960.
- Bramer JA, Abudu AA, Grimer RJ, Carter SR, Tillman RM. Do pathological fractures influence survival and local recurrence rate in bony sarcomas? *Eur J Cancer.* 2007;43:1944-1951.
- Davis AM, Bell RS, Goodwin PJ. Prognostic factors in osteosarcoma: a critical review. *J Clin Oncol.* 1994;12:423-431.
- Kim SH, Shin KH, Kim HY, et al. Postoperative nomogram to predict the probability of metastasis in Enneking stage IIB extremity osteosarcoma. *BMC Cancer.* 2014;14:666.
- Iwamoto Y, Tanaka K, Isu K, et al. Multiinstitutional phase II study of neoadjuvant chemotherapy for osteosarcoma (NECO study) in Japan: NECO-93J and NECO-95J. *J Orthop Sci.* 2009;14:397-404.



21. Iwamoto Y, Tanaka K. The activity of the Bone and Soft Tissue Tumor Study Group of the Japan Clinical Oncology Group. *Jpn J Clin Oncol.* 2012;42:467-470.
22. Kim MS, Cho WH, Song WS, Lee SY, Jeon DG. Time dependency of prognostic factors in patients with stage II osteosarcomas. *Clin Orthop Relat Res.* 2007;463:157-165.
23. Bianco FJ Jr. Nomograms and medicine. *Eur Urol.* 2006;50:884-886.
24. Kattan MW, Leung DH, Brennan MF. Postoperative nomogram for 12-year sarcoma-specific death. *J Clin Oncol.* 2002;20:791-796.
25. Jaffe N, Spears R, Eftekhari F, et al. Pathologic fracture in osteosarcoma. Impact of chemotherapy on primary tumor and survival. *Cancer.* 1987;59:701-709.
26. Abudu A, Sferopoulos NK, Tillman RM, Carter SR, Grimer RJ. The surgical treatment and outcome of pathological fractures in localised osteosarcoma. *J Bone Joint Surg Br.* 1996;78:694-698.
27. 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:49-57.
28. Aksnes LH, Hall KS, Folleraas G, et al. Management of high-grade bone sarcomas over two decades: the Norwegian Radium Hospital experience. *Acta Oncol.* 2006;45:38-46.
29. Ouyang ZX, Li XA. Inhibitory effects of tamoxifen and doxorubicin, alone and in combination, on the proliferation of the MG63 human osteosarcoma cell line. *Oncol Lett.* 2013;6:970-976.
30. Bacci G, Longhi A, Ferrari S, et al. Prognostic significance of serum alkaline phosphatase in osteosarcoma of the extremity treated with neoadjuvant chemotherapy: recent experience at Rizzoli Institute. *Oncol Rep.* 2002;9:171-175.
31. Wunder JS, Bull SB, Aneliunas V, et al. MDR1 gene expression and outcome in osteosarcoma: a prospective, multicenter study. *J Clin Oncol.* 2000;18:2685-2694.
32. Laverdiere C, Hoang BH, Yang R, et al. Messenger RNA expression levels of CXCR4 correlate with metastatic behavior and outcome in patients with osteosarcoma. *Clin Cancer Res.* 2005;11:2561-2567.
33. Khanna C, Wan X, Bose S, et al. The membrane-cytoskeleton linker ezrin is necessary for osteosarcoma metastasis. *Nat Med.* 2004;10:182-186.