The Spine Journal 18 (2018) 1804–1814

Clinical Study

# Predicting survival for metastatic spine disease: a comparison of nine scoring systems

A. Karim Ahmed, BS[a], C. Rory Goodwin, MD, PhD[a,b,*], Amir Heravi, BS[a], Rachel Kim, BS[a], Nancy Abu-Bonsrah, MD[a], Eric Sankey, MD[b], Daniel Kerekes, BS[a], Rafael De la Garza Ramos, MD[c], Joseph Schwab, MD[d], Daniel M. Sciubba, MD[a]

[a]*Department of Neurosurgery, The Johns Hopkins University School of Medicine, 600 North Wolfe St, Meyer 5-185, Baltimore, MD 21287, USA*
[b]*Department of Neurosurgery, Duke University Medical Center, 200 Trent Dr, Durham, NC 27710, USA*
[c]*Department of Neurological Surgery, Montefiore Medical Center/Albert Einstein College of Medicine, 3316 Rochambeau Ave, Bronx, NY 10467, USA*
[d]*Department of Orthopedic Surgery, Massachusetts General Hospital, 55 Fruit St, Boston, MA 02114, USA*

Received 28 November 2017; revised 24 February 2018; accepted 13 March 2018

**Abstract**

**BACKGROUND CONTEXT:** Despite advances in spinal oncology, research in patient-based prognostic calculators for metastatic spine disease is lacking. Much of the literature in this area investigates the general predictive accuracy of scoring systems in heterogeneous populations, with few studies considering the accuracy of scoring systems based on patient specifics such as type of primary tumor.
**PURPOSE:** The aim of the present study was to compare the ability of widespread scoring systems to estimate both overall survival at various time points and tumor-specific survival for patients undergoing surgical treatment for metastatic spine disease in order to provide surgeons with information to determine the most appropriate scoring system for a specific patient and timeline.
**STUDY DESIGN:** This is a retrospective study.
**PATIENT SAMPLE:** Patients who underwent surgical resection for metastatic spine disease at a single institution were included.
**OUTCOME MEASURES:** Areas under the receiver operating characteristic curves were generated from comparison of actual survival of patients and survival as predicted by application of prevalent scoring systems.
**METHODS:** A preoperative score for all 176 patients was retrospectively calculated utilizing the Skeletal Oncology Research Group (SORG) Classic Scoring Algorithm, SORG Nomogram, original Tokuhashi, revised Tokuhashi, Tomita, original Bauer, modified Bauer, Katagiri, and van der Linden scoring systems. Univariate and multivariate Cox proportional hazard models were constructed to assess the association of patient variables with survival. Receiver operating characteristic analysis modeling was utilized to quantify the accuracy of each test at different end points and for different primary tumor subgroups. No funds were received in support of this work. The authors have no conflicts of interest to disclose.
**RESULTS:** Among all patients surgically treated for metastatic spine disease, the SORG Nomogram demonstrated the highest accuracy at predicting 30-day (area under the curve [AUC] 0.81) and 90-day (AUC 0.70) survival after surgery. The original Tokuhashi was the most accurate at predicting 365-day survival (AUC 0.78). Multivariate analysis demonstrated multiple preoperative factors strongly associated

FDA device/drug status: Not applicable.

Author disclosures: ***AKA:*** Nothing to disclose. ***CRG:*** Grants: Burroughs Wellcome Fund (Amount not disclosed), UNCF-Merck Science Initiative (Amount not disclosed), outside the submitted work. ***AH:*** Nothing to disclose. ***RK:*** Nothing to disclose. ***NAB:*** Nothing to disclose. ***ES:*** Nothing to disclose. ***DK:*** Nothing to disclose. ***RDGR:*** Nothing to disclose. ***JS:*** Nothing to disclose. ***DMS:*** Consulting: Medtronic (Amount not disclosed), DePuy Synthes (Amount not disclosed), Stryker (Amount not disclosed), NuVasive (Amount not disclosed), K2M (Amount not disclosed), outside the submitted work.

The disclosure key can be found on the Table of Contents and at www.TheSpineJournalOnline.com.

A. Karim Ahmed is an NREF Medical Student Summer Research Fellow. C. Rory Goodwin is a UNCF Merck Postdoctoral Fellow and has received an award from the Burroughs Wellcome Fund and the Johns Hopkins Neurosurgery Pain Research Institute. Daniel M. Sciubba has consulting relationships with Medtronic, Globus, DePuy Synthes, and Orthofix. The remaining authors have nothing to disclose.

* Corresponding author. Department of Neurosurgery, Duke University Medical Center, 200 Trent Dr, Durham, NC 27710, USA. Tel.: (919) 684-7777; fax: (919) 684-7357.

*E-mail address*: rory.goodwin@duke.edu (C.R. Goodwin)

https://doi.org/10.1016/j.spinee.2018.03.011

with survival after surgery for spinal metastasis. The accuracy of each scoring system in determining survival probability relative to primary tumor etiology and time elapsed since surgery was assessed.
**CONCLUSIONS:** Among the nine scoring systems assessed, the present study determined the most accurate scoring system for short-term (30-day), intermediate (90-day), and long-term (365-day) survival, relative to primary tumor etiology. The findings of the present study may be utilized by surgeons in a personalized effort to select the most appropriate scoring system for a given patient. 

*Keywords:* Metastasis; Nomogram; Prognosis; Scoring system; Spine tumor; Surgery; Survival

## Introduction

Surgery for metastatic spine disease, the most common site of skeletal metastasis, is typically performed to address pain, mechanical instability, local tumor control, or neurologic compromise [1,2]. Despite the prevalence of spinal metastases, there is a lack of information regarding the factors that can predict both short-term and long-term prognoses for these patients [3–7]. The majority of prognostic calculators in the literature do not include multiple continuous variables in predicting a patient's survival, fail to estimate survival probability at various time points, and perform inconsistently across different primary tumor types [1,4,5,8–23].

Several input variables are common to the existing prognostic calculators and include tumor type, Karnofsky Performance Scale, presence or absence of visceral metastases, and number of vertebral segments involved. Despite these commonalities, prognostic calculators can provide inconsistent survival predictions for a given patient. The most cited predictive models in metastatic spine disease include the original Tokuhashi, revised Tokuhashi, Tomita, original Bauer, modified Bauer, Katagiri, and van der Linden scoring systems [1–35]. The more recent Classic Scoring Algorithm and Nomogram were created by the Skeletal Oncology Research Group (SORG) to determine 30-, 90-, and 365-day survival [36,37]. Today, there is still no consensus about which of these scoring systems has the greatest predictive survival accuracy relative to primary tumor type and time since surgery. The majority of the literature in this area investigates which scoring system is most superior by considering a heterogeneous population of patients, consisting of varying tumor types and prognosis [1,4,5,8–24,38,39].

The present study therefore aims to compare the most prevalent scoring systems available in their ability to determine both overall survival and tumor-specific survival for patients undergoing surgical treatment for metastatic spine disease. As such, the purpose of the present study is to arm surgeons with information to determine the most appropriate scoring system for a specific patient.

## Methods

### *Study design and subject inclusion*

After institutional review board approval (NA_00067508), a retrospective chart review was conducted that included all patients who underwent surgery for metastatic spinal tumors at a single institution between 2003 and 2016. The inclusion criteria were (1) patient age 18–100 years at the time of surgery; (2) complete and detailed electronic medical records with clinical presentation, imaging, and operative notes available; (3) patient who underwent surgical resection of a metastatic spine lesion; (4) pathologic confirmation of primary tumor etiology; (5) known survival or most recent follow-up.

### *Predictive scoring systems*

All patients were scored using the seven scoring systems most represented in the literature [1–35]. These scoring systems were calculated based on retrospective data, before the time of surgery. Scoring was performed by a study member (AKA), who did not participate in the medical or surgical management of these patients and was blinded to postoperative survival. Scoring systems included the original Tokuhashi [25], revised Tokuhashi [7], Tomita [26], original Bauer [8,9], modified Bauer [17,32], Katagiri [16], and van der Linden scoring systems (Supplementary Tables S1–S8) [29]. The recently published SORG Classic Scoring System and Nomogram have only been externally validated in one single-cohort study [36,37]. Therefore, an external validation of the SORG survival algorithm was performed with the present cohort before inclusion in the comparison.

A survival probability was calculated for each patient with respect to the given scoring system. The predictive ability of each scoring system was assessed using receiver operating characteristic (ROC) analysis at 30-, 90-, and 365-day postsurgery time points such that an area under the curve (AUC) could be calculated to quantify predictive accuracy.

### *Statistical analysis*

To assess which variables were marginally significant, a univariate Cox regression analysis ($p<.1$) was performed for all variables collected (Table 1). Categorical variables were established as age<65 years versus ≥65 years, Eastern Cooperation Oncology Group (ECOG) performance status score 0–2 versus 3–4, Charlson Comorbidity Index≤6 versus >6 (where 6 points is assigned for metastatic solid tumor) [40], primary tumor of good prognosis (lymphoma, breast, multiple myeloma, kidney, prostate, and thyroid) versus primary tumor of poor prognosis (lung, colon, rectum, bladder, esophagus, hepatocellular, melanoma, stomach, any other), 1 versus

EVIDENCE & METHODS

**Context**
There are multiple scoring systems available to guide spine surgeons regarding metastatic disease.

**Contribution**
Using cases from a single institution, the authors retrospectively applied nine of them to their cases to determine accuracy for specific time-points and tutor type. The SORG Nomogram and original Tokuhashi systems were best (at different time points) in predicting mortality.

**Implications**
The information provides some guidance as surgeons look to assess prognosis and provide rational treatments for these patients.

>1 spinal metastases, presence versus absence of metastases to lung or liver, presence versus absence of metastases to brain, history or no history of previous systemic therapy, white blood cell count<11,000/μL versus ≥11,000/μL, and hemoglobin level>10g/dL versus ≤10g/dL. The proportional hazards assumption was tested using a log-rank test and by visualizing the graph of the scaled Schoenfeld residual against time for each covariate. A multivariate Cox proportional hazards model was constructed using all the marginally significant variables (p<.1). Subsequently, covariates that were not independently associated with survival (p<.05) or that violated the proportional hazards assumption were removed in a stepwise manner. The final variables with independent and statistically significant association with patient survival were retained and reported.

The predictive abilities of prognostic scoring algorithms were tested using ROC analysis at 30-, 90-, and 365-day postsurgery time points using the calculated AUC, 95% confidence intervals (CIs), and p-values for each model. Receiver operating characteristic curves were generated using a cumulative or dynamic definition for cases and controls at each time point. Two-tailed p-values<.05 were considered statistically significant. In an effort to test the predictive accuracy of each scoring system based on primary tumor etiology, ROC curves for the most represented primary tumor types were also created. Primary tumors included in the analysis were breast, prostate, lung, and renal cell carcinoma (RCC). Two-tailed p-values<.05 were considered statistically significant. An AUC cutoff was set at 0.70 for a scoring system to be considered to have sufficient predictive accuracy [37,41].

## Results

### *Patient demographics and presentation*

Upon retrospective review of electronic medical records for patients surgically treated for spinal metastasis, 235 patients were initially identified. Of the 235 patients screened, 59 did not have complete medical records and were excluded. The remaining 176 patients met all inclusion criteria for the present study and were included. The mean age and body mass index for all patients were 60±12 years and 27±7 kg/m$^2$, respectively. The mean preoperative modified Charlson Comorbidity Index score, with age factored in, was 8±2. Of the 176 patients included, 66 (38%) had additional comorbidities before surgery. The majority of patients were male (60%) and had metastasis located in the thoracic spine (64%). Preoperatively, the most common presenting symptoms were pain (93%) and neurologic deficit (59%). American Spinal Cord Injury Association scores were not available, retrospectively. A vertebrectomy or corpectomy with stabilization was performed in 122 (69%) cases (Table 1). The decision to operate was made by the attending surgeon at this single institution and followed standard treatment guidelines, from widely

Table 1
Patient demographics and baseline characteristics

| Characteristic | Value |
|---|---|
| Age* (y) | 60±12 |
| BMI*,† (kg/m$^2$) | 27±7 |
| Modified Charlson Comorbidity Index score* (points) | 8±2 |
| Additional comorbidities | 66 (38%) |
| Male sex | 105 (60%) |
| Surgical and clinical characteristics | |
| Pathologic fracture | 129 (73%) |
| Pain | 164 (93%) |
| Time from start of neurologic symptoms and surgery | |
| No neurologic symptoms | 72 (41%) |
| <14 d | 53 (30%) |
| ≥14 d | 51 (29%) |
| Preoperative ASIA impairment scale | |
| Neurologic deficit (A, B, C, or D) | 104 (59%) |
| No neurologic deficit (E) | 72 (41%) |
| ECOG performance status | |
| Score 0–2 (≤50% of waking hours bed or chair bound) | 109 (62%) |
| Score 3–4 (>50% of waking hours bed or chair bound) | 66 (38%) |
| Time between primary cancer diagnosis and surgery | |
| <30 d | 38 (22%) |
| ≥30 d | 138 (78%) |
| Number of mobile spine levels undergoing operation | |
| 1 level | 17 (10%) |
| 2 levels | 11 (6%) |
| ≥3 levels | 67 (38%) |
| Surgery | |
| Vertebrectomy or corpectomy with stabilization | 122 (69%) |
| Decompression and stabilization | 40 (23%) |
| Decompression alone | 11 (6%) |
| Stabilization alone | 3 (2%) |
| Location | |
| Cervical | 36 (21%) |
| Thoracic | 112 (64%) |
| Lumbar | 51 (29%) |
| Combined | 30 (17%) |

ASIA, American Spinal Cord Injury Association; ECOG, Eastern Cooperation Oncology Group; BMI, body mass index.

Note: One hundred seventy-six patients with metastatic spine tumors were included from a single institution.

* Values depicted as mean±standard deviation.

† BMI was available for 125 of the 176 patients.

accepted literature in spinal oncology, at the time of surgery [2,42,43]. Surgery for metastatic spine disease was palliative in nature, to address neurologic dysfunction, instability, or pain. Because this was a retrospective study, it was not possible to characterize the exact surgical indication for each case, as many patients presented with multiple possible indications (ie, pain and myelopathy). All patients were managed, consulted on, and followed up by a multidisciplinary team of radiologists, medical oncologists, and radiation oncologists. As such, the decision to pursue surgery was made based on the best interest of the patient and clinical judgment by the treating team. As evidenced by the mean age at surgery and modified Charlson Comorbidity Index, patients undergoing surgery were relatively healthy and believed to benefit from surgery. All patients had oncologic staging before surgery, and those with significant comorbidities or systemic burden were not offered surgery.

### *Survival by primary tumor etiology*

Primary tumor etiology included hepatocellular, lung, breast, diffuse large B-cell lymphoma, melanoma, Merkel cell, multiple myeloma, pancreatic adenocarcinoma, plasmacytoma, prostate, RCC, sarcoma, squamous cell, thyroid, and bladder cancer. The greatest tumor-specific mean overall survival from the time of surgery was for patients with multiple myeloma (1,658 days), and the poorest prognosis was for patients with bladder cancer (55 days) (Table 2). The overall survival for all 176 patients was 282 days (95% CI 207–374) from the time of surgery (Fig. 1).

Table 2
Primary tumor type and survival

| Primary tumor | Number of patients* | Pathologic fracture† | Survival from surgery (d) |
|---|---|---|---|
| Adrenocortical | 2 (1%) | 1 (50%) | 150 |
| Basal cell | 1 (1%) | 0 (0%) | 121 |
| Breast | 37 (21%) | 31 (84%) | 695 |
| CLL | 1 (1%) | 1 (100%) | 978 |
| Colorectal | 3 (2%) | 2 (67%) | 278 |
| Hepatocellular | 2 (1%) | 1 (50%) | 306 |
| Lung | 34 (19%) | 22 (65%) | 158 |
| DLBCL | 1 (1%) | 0 (0%) | 139 |
| Melanoma | 5 (3%) | 5 (0%) | 491 |
| Merkel cell | 1 (1%) | 0 (0%) | 82 |
| Multiple myeloma | 16 (9%) | 15 (94%) | 1,658 |
| Pancreatic | 1 (1) | 1 (100%) | 374 |
| Plasmacytoma | 3 (2%) | 3 (100%) | 1,607 |
| Prostate | 27 (15%) | 10 (37%) | 372 |
| Renal cell | 36 (20%) | 33 (92%) | 730 |
| Sarcoma | 1 (1%) | 0 (0%) | 200 |
| Squamous cell | 2 (1%) | 1 (50%) | 87 |
| Thyroid | 2 (1%) | 2 (100%) | 623 |
| Bladder | 1 (1%) | 1 (100%) | 55 |

CLL, chronic lymphocytic leukemia; DLBCL, diffuse large B-cell lymphoma.

* The percentage of patients of a given primary tumor etiology is represented out of 176 patients included in the study.

† Percent pathologic fracture is represented as the proportion of patients of a respective primary tumor etiology.

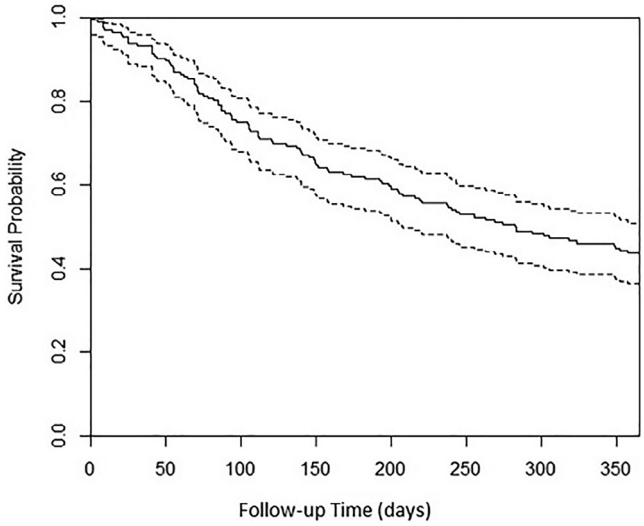


Fig. 1. Kaplan-Meier curve for patient survival within 1 year after surgery. After surgery, the median age of survival for the 176 patients included in the present study was 282 days (95% confidence interval [CI] 207–374).

### *External validation of the SORG survival algorithm*

The SORG Classic Scoring Algorithm was utilized to calculate a predictive score for each patient (Supplementary Table S1, Fig. S1) [36,37]. Patients with scores of 0–2, 3–4, and 5–12 were determined to be in the good, intermediate, and poor prognosis groups, respectively. The mean respective survival times for the good, intermediate, and poor

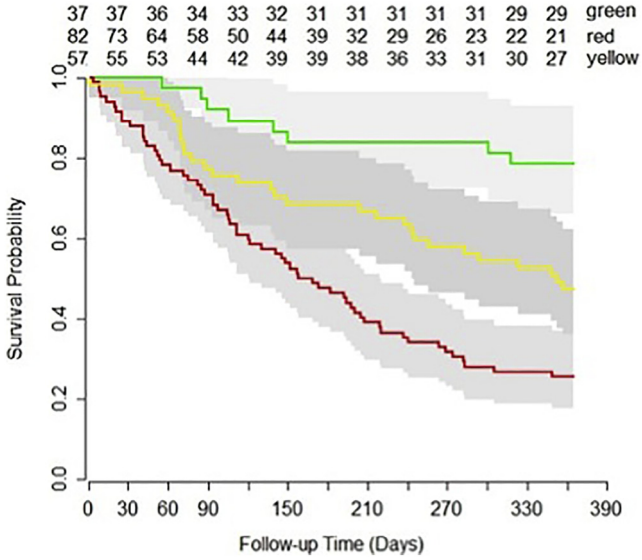


Fig. 2. Kaplan-Meier survival curves for SORG Classic Scoring Algorithm. The good prognosis group (0–2), intermediate prognosis group (3–4), and poor prognosis (5–12) group had median survival of 880 days (95% confidence interval [CI] 516–1227), 352 days (95% CI 238–484), and 163 days (95% CI 111–207) from the time of surgery, respectively. Stratified patient survival within 1 year of follow-up with the corresponding 95% CI for each line. The number of at-risk patients in each group is specified above the graph at 30-day checkpoints. Green: Good prognosis group (score 0–2). Yellow: Intermediate prognosis group (score 3–4). Red: Poor prognosis group (score 5–12).

prognosis groups were 1261±98, 552±51, and 338±39 days, respectively. Fifteen patients were still alive at the most recent follow-up (mean 1,994 days).

Patients were stratified based on prognosis group, and a Kaplan-Meier survival analysis was performed (Fig. 2). The median survival times for the good, intermediate, and poor prognosis groups were 880 days (95% CI 516–1227), 352 days (95% CI 238–484), and 163 days (95% CI 111–207) from the time of surgery, respectively. A log-rank test revealed statistical significance (p<.015) between each defined prognostic group based on the preoperative SORG Classic Scoring Algorithm. Satisfactory receiver operating characteristic curves, as presented in Tables 5–7, demonstrate external validation of the Nomogram.

### *Factors associated with survival*

A univariate Cox regression analysis was performed. Older age (≥65 years, hazard ratio [HR] 1.02 [95% CI 1.01–1.04], p=.003), additional comorbidities (Charlson Comorbidity Index>6, where 6 points is assigned for metastatic solid tumor; HR 1.57 [95% CI 1.14–2.16], p=.006), presence of pathologic fractures at affected spinal level (HR 0.58 [95% CI 0.41–0.82], p=.002), acute neurologic deficit (<14 days, HR 1.61 [95% CI 1.11–2.35], p=.012), poor ECOG performance status (ECOG 3 or 4; HR 2.10 [95% CI 1.53–2.89], p<.001), primary cancer type with poor prognosis (lung, colon, rectum, bladder, esophagus, hepatocellular, melanoma, stomach, and any other not included in good prognosis group—consisting of lymphoma, breast, multiple myeloma, kidney, prostate, and thyroid; HR 2.83 [95% CI 1.98–4.04], p<.001), metastasis to the lungs or the liver (HR 1.94 [95% CI 1.35–2.78], p<.001), brain metastasis (HR 2.53 [95% CI 1.53–4.16], p<.001), and hemoglobin levels (≤10 g/dL, HR 0.89 [95% CI 0.81–0.98], p=.012) were significantly associated with survival (Table 3).

Upon multivariate analysis, the following covariates had an independent and statistically significant association with decreased survival: older age (HR 1.02 [95% CI 1.01–1.03], p=.024), poor ECOG performance status (HR 2.23 [95% CI 1.59–3.12], p<.001), primary cancer type with poor prognosis (HR 2.61 [95% CI 1.78–3.81], p<.001), metastasis to the lungs or the liver (HR 1.61 [95% CI 1.12–2.31], p=.010), and brain metastasis (HR 2.38 [95% CI 1.41–4.05], p=.001) (Table 3).

Upon multivariate analysis, the following covariates had an independent and statistically significant association with decreased survival: older age (HR 1.02 [95% CI 1.01–1.03], p=.024), poor ECOG performance status (HR 2.23 [95% CI 1.59–3.12], p<.001), primary cancer type with poor prognosis (HR 2.61 [95% CI 1.78–3.81], p<.001), metastasis to the lungs or the liver (HR 1.61 [95% CI 1.12–2.31], p=.010), and brain metastasis (HR 2.38 [95% CI 1.41–4.05], p=.001) (Table 4).

### *Overall predictive accuracy of survival by nine scoring systems*

Among all patients surgically treated for metastatic spine disease, the SORG Nomogram demonstrated the highest accuracy at predicting 30-day (AUC 0.81) and 90-day (AUC 0.70) survival after surgery (Tables 5 and 6). The original Tokuhashi was the most accurate at predicting 365-day survival (AUC 0.78) (Table 7).

The SORG Nomogram (AUC 0.81), Classic Scoring Algorithm (AUC 0.77), original Tokuhashi (AUC 0.74), van der Linden (AUC 0.74), Katagiri (AUC 0.73), and revised Tokuhashi (AUC 0.71) were the only scoring systems to achieve sufficient accuracy at predicting 30-day survival (Table 5). The SORG Nomogram and Katagiri were the only scoring systems considered sufficiently accurate at predicting 90-day survival after surgery (AUC 0.70 for both) (Table 6). All systems but the original Bauer system were found to be sufficient at predicting 365-day survival (AUC 0.69) (Table 7). Therefore, the SORG Nomogram and Katagiri were the only scoring systems to achieve sufficient accuracy at predicting 30-, 90-, and 365-day survival for all patients.

### *Predictive accuracy of survival: excluding hematologic malignancies*

Receiver operating characteristic AUC curves were generated for all patients excluding those with multiple myeloma, lymphoma, and plasmacytoma (n=155). Of this group, the SORG Nomogram had the highest accuracy at predicting 30-day survival (AUC 0.79), and the original Tokuhashi had the highest accuracy at predicting 365-day survival (Table 5–7). No scoring system achieved sufficient accuracy at predicting 90-day survival for this subgroup, although the original Tokuhashi and Linden systems came closest (both AUC 0.69) (Table 6).

Five systems accurately predicted both 30- and 365-day survival in this cohort: the SORG Classic Scoring (30-day AUC 0.75, 365-day AUC 0.75), original Tokuhashi (0.75, 0.81), revised Tokuhashi (0.72, 0.81), Katagiri (0.70, 0.76), and Linden (0.76, 0.77) systems (Tables 5 and 7). The Tomita accurately predicted 365-day survival only (AUC 0.70).

### *Predictive accuracy of survival: breast cancer*

For patients with metastatic breast cancer of the spine, the SORG Nomogram was the most accurate (AUC 0.99) at predicting 30-day survival after surgery—significantly outperforming all other scoring systems at this time point (p<.05 for all). The 30-day survival estimate was also sufficiently accurate for the Linden (AUC 0.89), original Tokuhashi (AUC 0.88), and SORG Classic Scoring (AUC 0.81) systems. The Linden was the most accurate predictive scoring system at 90 days (AUC 0.88) and at 365 days (AUC 0.89) (Tables 6 and 7). The original Tokuhashi was the only system to achieve sufficient accuracy at predicting 30-day (AUC 0.88), 90-day (AUC 0.71), and 365-day (AUC 0.82) survival after surgery for patients with breast cancer (Table 5–7).

### *Predictive accuracy of survival: lung cancer*

For patients with metastatic lung cancer of the spine, the Linden scoring system most accurately predicted 30-day

Table 3
Univariate Cox regression analysis

| | β Regression coefficient* | HR† | p-Value |
|---|---|---|---|
| Demographic characteristics | | | |
| Age | 0.020±0.007 | 1.02 (1.01–1.04) | .003‡ |
| Sex | 0.004±0.162 | 1.00 (0.73–1.38) | .978 |
| BMI | | | |
| 18.5–30 kg/m$^2$ | | | |
| <18.5 kg/m$^2$ | 0.200±0.464 | 1.22 (0.49–3.03) | .666 |
| >30 kg/m$^2$ | –0.309±0.217 | 0.73 (0.48–1.12) | .155 |
| Comorbidities | 0.449±0.164 | 1.57 (1.14–2.16) | .006‡ |
| Clinical and surgical characteristics | | | |
| Pathologic fractures | –0.548±0.177 | 0.58 (0.41–0.82) | .002‡ |
| Pain | –0.421±0.301 | 0.66 (0.36–1.19) | .162 |
| Time between start of neurologic symptoms and surgery | | | |
| No neurologic symptoms | Reference | Reference | Reference |
| <14 d | 0.479±0.191 | 1.61 (1.11–2.35) | .012‡ |
| ≥14 d | 0.051±0.193 | 1.05 (0.72–1.53) | .792 |
| ASIA impairment scale (preoperative) | | | |
| No neurologic deficit (E) | Reference | Reference | Reference |
| Neurologic deficit (A, B, C, D) | 0.278±0.162 | 1.32 (0.96–1.81) | .087 |
| ECOG performance status | | | |
| Score 0–2 (≤50% waking hours bed or chair bound) | Reference | Reference | Reference |
| Score 3–4 (>50% waking hours bed or chair bound) | 0.743±0.163 | 2.10 (1.53–2.89) | <.001‡ |
| Oncologic status | | | |
| Primary cancer type§ | | | |
| Good prognosis group | Reference | Reference | Reference |
| Poor prognosis group | 1.040±0.182 | 2.83 (1.98–4.04) | <.001‡ |
| >1 mobile spine metastases (mobile spine defines as C3–C6, L2–L4) | –0.105±0.164 | 0.90 (0.65–1.24) | .52 |
| Other bone metastases | 0.159±0.164 | 1.17 (0.85–1.62) | .332 |
| Visceral metastases | | | |
| None | Reference | Reference | Reference |
| Lung or liver | 0.660±0.184 | 1.94 (1.35–2.78) | <.001‡ |
| Brain | 0.927±0.254 | 2.53 (1.53–4.16) | <.001‡ |
| Prior local radiation therapy | –0.205±0.160 | 0.81 (0.60–1.12) | .201 |
| Prior systemic therapy | 0.169±0.162 | 1.18 (0.86–1.63) | .295 |
| Laboratory values | | | |
| White blood cell count (1,000/μL) | –0.009±0.011 | 0.99 (0.97–1.01) | .379 |
| Hemoglobin (g/dL) | –0.116±0.046 | 0.89 (0.81–0.98) | .012‡ |
| Calcium (mg/dL) | –0.131±0.092 | 0.88 (0.73–1.05) | .157 |
| Creatinine (mg/dL) | 0.087±0.306 | 1.09 (0.60–1.99) | .775 |
| Platelet (1,000/mL) | –0.000±0.000 | 1.00 (1.00-1.00) | .719 |

ASIA, American Spinal Cord Injury Association; ECOG, Eastern Cooperation Oncology Group; BMI, body mass index; HR, hazard ratio.

* The values are given as the b coefficient and the standard error.

† The values are given as the HR, with the 95% confidence interval (CI) in parentheses.

‡ These p-values were significant and had a two-tailed p-value of <.05.

§ The good prognosis group included lymphoma, breast cancer, multiple myeloma, kidney cancer, prostate cancer, and thyroid cancer. The poor prognosis group included lung cancer, colon cancer, rectal cancer, bladder cancer, esophageal cancer, liver cancer, melanoma, gastric cancer, and other cancers.

survival after surgery (AUC 0.72). None of the scoring systems were sufficiently accurate at predicting 90-day survival after surgery for this group (Table 6). The SORG Nomogram was the most accurate for 365-day (AUC 0.85) survival. The revised Tokuhashi also achieved sufficient accuracy at 30 days (AUC 0.71). Both SORG predictive systems were significantly more accurate than the Tomita, original Bauer, modified Bauer, and Katagiri at 365 days for patients with metastatic lung cancer (p<.05 for all) (Table 7).

*Predictive accuracy of survival: prostate cancer*

For patients with metastatic prostate cancer of the spine, the SORG Classic Scoring, SORG Nomogram, original Tokuhashi, revised Tokuhashi, Katagiri, and Linden were equally accurate at 30 days (AUC 1 for all). All of these scoring systems, with the exception of the SORG Nomogram, were also sufficiently accurate at predicting 365-day survival (Table 7). However, the most accurate scoring system at 365 days was the Linden (AU 0.76), and no scoring systems were sufficiently accurate at predicting 90-day survival for prostate cancer (Table 6).

*Predictive accuracy of survival: renal cell carcinoma*

For patients with RCC, the SORG Classic Scoring system most accurately predicted 30-day survival (AUC 0.82) after

Table 4
Multivariate Cox regression analysis

| | β Regression coefficient* | HR† | p-Value |
|---|---|---|---|
| Age | 0.017±0.008 | 1.02 (1.01–1.03) | .024‡ |
| ECOG performance status | | | |
| Score 0–2 (≤50% waking hours bed or chair bound) | Reference | Reference | Reference |
| Score 3–4 (>50% waking hours bed or chair bound) | 0.800±0.172 | 2.23 (1.59–3.12) | <.001‡ |
| Primary cancer type§ | | | |
| Good prognosis group | Reference | Reference | Reference |
| Poor prognosis group | 0.958±0.194 | 2.61 (1.78–3.84) | <.001‡ |
| Visceral metastases | | | |
| None | Reference | Reference | Reference |
| Lung or liver | 0.476±0.185 | 1.61 (1.12–2.31) | .010‡ |
| Brain | 0.869±0.270 | 2.38 (1.41–4.05) | .001‡ |

ECOG, Eastern Cooperation Oncology Group.

Notes: Explanatory variables were dropped by stepwise backward elimination in the following order: (1) preoperative ASIA score, (2) time between the start of neurologic symptoms and the surgical procedure, (3) presence of a pathologic fracture due to violation of proportional hazard assumption, (4) presence of any additional comorbidity, and (5) hemoglobin levels.

* The values are given as the b coefficient and the standard error.

† The values are given as the hazard ratio (HR), with the 95% confidence interval (CI) in parentheses.

‡ These p-values were significant and had a two-tailed p-value of <.05.

§ The good prognosis group included lymphoma, breast cancer, multiple myeloma, kidney cancer, prostate cancer, and thyroid cancer. The poor prognosis group included lung cancer, colon cancer, rectal cancer, bladder cancer, esophageal cancer, liver cancer, melanoma, gastric cancer, and other cancers.

Table 5
Area under the curve (AUC) from receiver operating characteristic (ROC) analysis for all patients at 30 days after surgical treatment for metastatic spine disease

| Scoring method | All patients | All patients excluding hematologic malignancies† | Breast | Lung | Prostate | Renal cell carcinoma |
|---|---|---|---|---|---|---|
| Classic | **0.77*** (0.66–0.8773) | **0.75** (0.63–0.87) | **0.81** (0.69–0.92) | 0.55 | **1.00**‡ (1.00–1.00) | **0.82** (0.68–0.97) |
| SORG Nomogram | **0.81** (0.69–0.93) | **0.79** (0.66–0.92) | **0.99** (0.96–1.00) | 0.61 | **1.00**‡ (1.00–1.00) | **0.81** (0.64–0.99) |
| Original Tokuhashi | **0.74** (0.64–0.85) | **0.75** (0.64–0.85) | **0.88** (0.78–0.97) | 0.64 | **1.00**‡ (1.00–1.00) | 0.68 |
| Revised Tokuhashi | **0.71** (0.57–0.85) | **0.72** (0.58–0.86) | 0.68 (0.54–0.82) | **0.71** (0.46–0.97) | **1.00**‡ (1.00–1.00) | 0.63 |
| Tomita | 0.59 | 0.57 | 0.24 | 0.56 | 0.42 | 0.55 |
| Original Bauer | 0.63 | 0.61 | 0.31 | 0.67 | **0.73** (0.61–0.86) | 0.59 |
| Modified Bauer | 0.59 | 0.56 | 0.24 | 0.57 | 0.44 | 0.55 |
| Katagiri | **0.73** (0.57–0.89) | **0.70** (0.54–0.87) | 0.25 | 0.64 | **0.94** (0.88–1.00) | 0.65 |
| Linden | **0.74** (0.60–0.88) | **0.76** (0.62–0.89) | **0.89** (0.80–0.98) | **0.72** (0.47–0.98) | **0.73** (0.61–0.86) | **0.76** (0.63–0.89) |

SORG, Skeletal Oncology Research Group.

Notes: The most accurate scoring system for a group of patients is highlighted in red. Scoring systems that were sufficiently accurate (AUC>0.70) appear in bold.

* Values are given as the AUC (95% confidence interval [CI]) for values achieving sufficient accuracy.

† Hematologic malignancies include patients with multiple myeloma, plasmacytoma, or lymphoma. It was not possible to generate an AUC for 30-day survival because there were no deaths during this period for this group.

‡ One patient with metastatic prostate cancer passed away within 30 days after surgery.

Table 6
Area under the curve (AUC) from receiver operating characteristic (ROC) analysis for all patients at 90 days after surgical treatment for metastatic spine disease

| Scoring method | All patients | All patients excluding hematologic malignancies* | Breast | Lung | Prostate | Renal cell carcinoma | Hematologic malignancies |
|---|---|---|---|---|---|---|---|
| Classic | 0.67 | 0.65 | 0.57 | 0.63 | 0.58 | **0.71** (0.52–0.89) | **0.74** (0.55–0.93) |
| SORG Nomogram | **0.70** (0.61–0.79) | 0.67 | 0.59 | 0.67 | 0.52 | **0.70** (0.51–0.89) | **0.87** (0.72–1.00) |
| Original Tokuhashi | 0.69 | 0.69 | **0.71** (0.42–0.99) | 0.62 | 0.65 | 0.65 | **0.89** (0.80–0.99) |
| Revised Tokuhashi | 0.67 | 0.67 | 0.68 | 0.63 | 0.63 | 0.62 | **0.97** (0.92–1.00) |
| Tomita | 0.60 | 0.59 | 0.53 | 0.53 | 0.42 | 0.53 | 0.61 |
| Original Bauer | 0.63 | 0.61 | 0.52 | 0.55 | 0.46 | 0.57 | 0.68 |
| Modified Bauer | 0.62 | 0.59 | 0.54 | 0.51 | 0.44 | 0.53 | 0.63 |
| Katagiri | **0.70** (0.60–0.79) | 0.67 | 0.53 | 0.61 | 0.60 | 0.63 | **0.92** (0.84–1.00) |
| Linden | 0.66 | 0.69 | **0.88** (0.74–1.00) | 0.65 | 0.62 | **0.74** (0.58–0.89) | **0.84** (0.73–0.95) |

SORG, Skeletal Oncology Research Group.

Notes: The most accurate scoring system for a group of patients is highlighted in red. Scoring systems that were sufficiently accurate (AUC >0.70) appear in bold.

Table 7
Area under the curve (AUC) from receiver operating characteristic (ROC) analysis for all patients at 365 days after surgical treatment for metastatic spine disease

| Scoring method | All patients | All patients excluding hematologic malignancies* | Breast | Lung | Prostate | Renal cell carcinoma | Hematologic malignancies |
|---|---|---|---|---|---|---|---|
| Classic | **0.77** (0.70–0.84) | **0.75** (0.67–0.82) | **0.71** (0.52–0.90) | **0.84** (0.66–1.00) | **0.73** (0.53–0.92) | **0.72** (0.53–0.90) | **0.81** (0.59–1.00) |
| SORG Nomogram | **0.78** (0.71–0.85) | **0.75** (0.68–0.83) | **0. 72** (0.51–0.93) | **0.85** (0.69–1.00) | 0.64 | **0.71** (0.52–0.89) | **0.81** (0.62–1.00) |
| Original Tokuhashi | **0.78** (0.71–0.85) | **0.81** (0.74–0.88) | **0. 82** (0.66–0.97) | 0.69 | **0.74** (0.55–0.94) | **0.78** (0.61–0.95) | **0.73** (0.53–0.94) |
| Revised Tokuhashi | **0.77** (0.70–0.84) | **0.81** (0.74–0.88) | **0.77** (0.61–0.93) | **0.79** (0.56–1.00) | **0.76** (0.57–0.94) | **0.75** (0.57–0.93) | **0.75** (0.53–0.97) |
| Tomita | **0.70** (0.62–0.78) | **0.70** (0.62–0.79) | 0.56 | 0.50 | 0.61 | 0.69 | 0.48 |
| Original Bauer | 0.69 | 0.68 | 0.59 | 0.48 | 0.59 | 0.67 | 0.41 |
| Modified Bauer | **0.71** (0.64–0.79) | 0.69 | 0.55 | 0.50 | 0.64 | 0.67 | 0.50 |
| Katagiri | **0.78** (0.72–0.85) | **0.76** (0.69–0.84) | 0.61 | 0.62 | **0.76** (0.57–0.94) | **0.73** (0.56–0.91) | **0.77** (0.58–0.95) |
| Linden | **0.71** (0.63–0.78) | **0.77** (0.69–0.84) | **0.89** (0.81–0.98) | **0.77** (0.59–0.94) | **0.76** (0.59–0.94) | **0.84** (0.69–0.99) | **0.70** (0.49–0.92) |

SORG, Skeletal Oncology Research Group.

Notes: The most accurate scoring system for a group of patients is highlighted in red. Scoring systems that were sufficiently accurate (AUC>0.70) appear in bold.

surgery for metastatic spine disease. The Linden system most accurately predicted both 90-day (AUC 0.74) and 365-day survival (AUC 0.84) (Tables 5–7). The SORG Classic Scoring (AUC 0.82, 0.71, 0.72), the SORG Nomogram (AUC 0.81, 0.70, 0.71), and the Linden system (AUC 0.76, 0.74, 0.84) were sufficiently accurate at predicting survival at 30, 90, and 365 days after surgery. However, the SORG Nomogram outperformed Tomita and modified Bauer at 30 days, and Linden at 365 days ($p<.05$ for all). The original Tokuhashi (AUC 0.78), the revised Tokuhashi AUC 0.75), and the Katagiri (AUC 0.73) were accurate for this subgroup at 365 days only (Table 7).

*Predictive accuracy of survival: lymphoma, plasmacytoma, and multiple myeloma*

Of the patients with lymphoma, multiple myeloma, or plasmacytoma of the spine, none passed away within 30 days of surgery for spine disease. The 90-day survival was most accurately predicted by the revised Tokuhashi (AUC 0.97) and the 365-day survival by the SORG Nomogram (AUC 0.81) (Tables 6 and 7). The SORG Classic Scoring (AUC 0.74, 0.81), SORG Nomogram (AUC 0.87, 0.81), original Tokuhashi (AUC 0.89, 0.73), revised Tokuhashi (AUC 0.97, 0.75), Katagiri (AUC 0.92, 0.77), and Linden (AUC 0.84, 0.70) were sufficiently accurate at predicting survival 90 and 365 days after surgery. The SORG Nomogram was significantly more accurate than the SORG Classic Scoring system at 90 days ($p<.05$) and more accurate than the Tomita, original Bauer, and modified Bauer at both 90 and 365 days ($p<.05$ for all) (Tables 6 and 7).

## Discussion

Traditional scoring systems in metastatic spine disease are beneficial in distinguishing long-term survival (>12 months) from short-term survival (<3 months) for patients with good or poor prognosis, respectively. However, given the mean overall survival of patients diagnosed with a metastatic spine lesion in the literature is ~10 months [1,2,23,42] (9.4 months for our cohort), there is a necessity for more accurate predictive methods in the short- to intermediate-term survival interval. The present study was concerned with the accuracy of predictive scoring systems with respect to tumor-specific prognosis in metastatic spine disease.

In the present study, univariate and multivariate analyses demonstrated that several previously investigated factors were significantly associated with decreased survival (Tables 3 and 4). Notably, presence of preoperative pathologic fracture was associated with better survival in our independent cohort. This may be due to confounding, whereby the majority of pathologic fractures presented in patients with tumors of good prognosis. Pathologic fracture was seen in 84% of patients with breast cancer, 94% of patients with multiple myeloma, and 92% of patients with RCC, which were three of the largest cohorts characterized by three of the longest mean survival times (Table 2).

In the current study, the SORG Nomogram was the most accurate scoring system at predicting 30- and 90-day survival, and the original Tokuhashi was the most accurate at predicting 365-day survival. A previous study comparing the SORG Nomogram with the Tokuhashi and SORG classic scores also found the Nomogram to be superior at predicting 3-month survival, but further found that superiority to persist at the 12-month range [37]. The inclusion of multiple myeloma, lymphoma, and plasmacytoma is controversial in scoring systems of metastatic spine disease because they are hematological malignancies [17,23,36]. Several scoring systems, such as the Tokuhashi [23] and Tomita systems [26], exclude these patients in their original texts. To accurately and directly compare the validity of each scoring system, we elected to assess predictive accuracy by both including and excluding patients with hematologic malignancies of the spine. Interestingly, the original Tokuhashi score, which did not include such malignancies in the original work, was the most accurate at predicting survival at 365 days after surgery.

Excluding these patients, the SORG Nomogram and Original Tokuhashi were still the most accurate at predicting 30- and 365-day survival, respectively (none of the nine scoring systems reached significant accuracy at 90 days for this cohort).

A study by Aoude and Amiot demonstrated that both the revised Tokuhashi score and the Tomita score were accurate in predicting poor prognosis (survival less than 6 months), but that the Tomita score was limited in distinguishing between moderate (6–12 months) and good (>12 months) prognosis for patients undergoing surgery for spine metastases [24]. In the present study, the revised Tokuhashi was sufficiently accurate at 30 and 365 days after surgery, whereas, in contrast to this previous study, the Tomita score was only accurate at predicting 365-day survival. Breast cancer, lung cancer, prostate cancer, RCC, and hematologic malignancies represent some of the most common spinal malignancies [1,2,44]. As such, it is valuable to compare the accuracy of predictive scoring systems in estimating tumor-specific survival for these etiologies, and these results are presented earlier (Tables 5–7). Interestingly, the original and modified Bauer were the only scoring systems to include lymphoma and multiple myeloma in their original classification [8,9,17,18], but were not found to be accurate at predicting survival for these patients in the present study.

The revised Tokuhashi score is often cited as the most accurate predictive scoring system for spinal metastases [7,11–13,19,22,26–35]. Ulmar et al. and Wang et al. demonstrated that, unlike the original Tokuhashi score, only the revised Tokuhashi score accurately predicted survival for patients undergoing surgery for breast cancer metastases to the spine [7,25,27,28,31]. This is in contrast to the present study, where, although the original Tokuhashi did not have the greatest accuracy at any individual time point for patients with metastatic breast cancer, it was the only scoring system to achieve sufficient accuracy at all time points (30, 90, and 365 days) after surgery (Tables 5–7).

In a study of the revised Tokuhashi score in metastatic lung cancer by Hessler et al. [14], the predicted survival was only similar to actual survival in 67.1% of cases—likely attributed to recent medical advances in lung cancer treatment [14,19,21,23,35]. However, the revised Tokuhashi was demonstrated here to have sufficient accuracy at 30 and 365 days for patients with metastatic lung cancer. Similar to previous studies [27,28,31], the Tomita score failed to accurately predict survival in patients with RCC.

The present study successfully validated the SORG Classic Scoring Algorithm and Nomogram in an independent cohort at a single institution and compared the survival accuracy of nine predictive scoring systems in metastatic spine disease. This was accomplished for all patients, all patients excluding hematologic malignancies, and by primary tumor type. To date, this is one of the largest studies comparing the accuracy of multiple scoring systems, the only study to include nine scoring systems, and the only study to compare primary tumor-specific prognostic accuracy for this many scoring systems and time points. In doing so, clinicians can utilize the tables outlined here (Tables 5–7) to choose the most appropriate scoring system for patients with metastatic spine disease relative to both primary tumor etiology and time after surgery.

Despite the advantages of the present study, there are limitations. The present sample size used in the external validation of the SORG Nomogram is underpowered compared with the original cohort of 649 patients. For tumor-specific prognosis, not all primary tumor etiologies are equally represented, introducing possible sampling bias. In addition, Karnofsky Performance Status, which was found to be significant in other studies [5–7,22,23,25], is not included in this analysis. The present study also carries with it the limitations inherently associated with retrospective studies, including selection and information bias. There are additional limitations regarding single institution studies. As such, further prospective multi-institutional studies are appropriate to replicate and confirm the findings presented herein.

## Conclusions

Multivariate analysis demonstrated preoperative factors strongly associated with survival after surgery for spinal metastasis. For all patients with various primary tumor types, SORG Nomogram demonstrated the highest accuracy at predicting 30- and 90-day survival after surgery for metastatic spine disease. The original Tokuhashi had the highest 365-day predictive accuracy after surgery. Similarly, the most accurate scoring system to predict postoperative survival probability was assessed relative to primary tumor etiology for short-term (30-day), intermediate (90-day), and long-term (365-day) survival after surgery. To date, this is one of the largest studies comparing the accuracy of multiple scoring systems, the only comprehensive study to include nine scoring systems, and the only study to compare primary tumor-specific prognostic accuracy for this many scoring systems. Clinicians may use the tables outlined here (Tables 5–7) to determine the most accurate scoring system for a specific patient, based on primary tumor etiology and prognosis. The findings of the present study can improve decision-making in the care for patients with metastatic spine disease, and offer a more individualized tool for informing patients.

## Supplementary material

Supplementary material related to this article can be found at https://doi.org/10.1016/j.spinee.2018.03.011.

## References

[1] Coleman RE. Clinical features of metastatic bone disease and risk of skeletal morbidity. Clin Cancer Res 2006;12:6243s–6249s. doi:10.1158/1078-0432.CCR-06-0931.

[2] Sciubba DM, Petteys RJ, Dekutoski MB, Fisher CG, Fehlings MG, Ondra SL, et al. Diagnosis and management of metastatic spine disease. A review. J Neurosurg Spine 2010;13:94–108. doi:10.3171/2010.3.SPINE09202.

[3] Ghori AK, Leonard DA, Schoenfeld AJ, Saadat E, Scott N, Ferrone ML, et al. Modeling 1-year survival after surgery on the metastatic spine. Spine J 2015;15:2345–50. doi:10.1016/j.spinee.2015.06.061.
[4] Sioutos PJ, Arbit E, Meshulam CF, Galicich JH. Spinal metastases from solid tumors. Analysis of factors affecting survival. Cancer 1995;76:1453–9.
[5] Tatsui H, Onomura T, Morishita S, Oketa M, Inoue T. Survival rates of patients with metastatic spinal cancer after scintigraphic detection of abnormal radioactive accumulation. Spine 1996;21:2143–8.
[6] Tokuhashi Y, Ajiro Y, Umezawa N. Outcome of treatment for spinal metastases using scoring system for preoperative evaluation of prognosis. Spine 2009;34:69–73. doi:10.1097/BRS.0b013e3181913f19.
[7] Tokuhashi Y, Matsuzaki H, Oda H, Oshima M, Ryu J. A revised scoring system for preoperative evaluation of metastatic spine tumor prognosis. Spine 2005;30:2186–91.
[8] Bauer HC, Wedin R. Survival after surgery for spinal and extremity metastases. Prognostication in 241 patients. Acta Orthop Scand 1995;66:143–6.
[9] Bauer H, Tomita K, Kawahara N, Abdel-Wanis ME, Murakami H. Surgical strategy for spinal metastases. Spine 2002;27:1124–6.
[10] Chen H, Xiao J, Yang X, Zhang F, Yuan W. Preoperative scoring systems and prognostic factors for patients with spinal metastases from hepatocellular carcinoma. Spine 2010;35:E1339–46. doi:10.1097/BRS.0b013e3181e574f5.
[11] Eap C, Tardieux E, Goasgen O, Bennis S, Mireau E, Delalande B, et al. Tokuhashi score and other prognostic factors in 260 patients with surgery for vertebral metastases. Orthop Traumatol Surg Res 2015;101:483–8. doi:10.1016/j.otsr.2015.03.007.
[12] Enkaoua EA, Doursounian L, Chatellier G, Mabesoone F, Aimard T, Saillant G. Vertebral metastases: a critical appreciation of the preoperative prognostic Tokuhashi score in a series of 71 cases. Spine 1997;22:2293–8.
[13] Hernandez-Fernandez A, Vélez R, Lersundi-Artamendi A, Pellisé F. External validity of the Tokuhashi score in patients with vertebral metastasis. J Cancer Res Clin Oncol 2012;138:1493–500. doi:10.1007/s00432-012-1222-2.
[14] Hessler C, Vettorazzi E, Madert J, Bokemeyer C, Panse J. Actual and predicted survival time of patients with spinal metastases of lung cancer: evaluation of the robustness of the Tokuhashi score. Spine 2011;36:983–9. doi:10.1097/BRS.0b013e3181e8f7f8.
[15] Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol 2008;26:1364–70. doi:10.1200/JCO.2007.12.9791.
[16] Katagiri H, Takahashi M, Wakai K, Sugiura H, Kataoka T, Nakanishi K. Prognostic factors and a scoring system for patients with skeletal metastasis. J Bone Joint Surg Br 2005;87:698–703. doi:10.1302/0301-620X.87B5.15185.
[17] Leithner A, Radl R, Gruber G, Hochegger M, Leithner K, Welkerling H, et al. Predictive value of seven preoperative prognostic scoring systems for spinal metastases. Eur Spine J 2008;17:1488–95. doi:10.1007/s00586-008-0763-1.
[18] Majeed H, Kumar S, Bommireddy R, Klezl Z, Calthorpe D. Accuracy of prognostic scores in decision making and predicting outcomes in metastatic spine disease. Ann R Coll Surg Engl 2012;94:28–33. doi:10.1308/003588412X13171221498424.
[19] Quraishi NA, Manoharan SR, Arealis G, Khurana A, Elsayed S, Edwards KL, et al. Accuracy of the revised Tokuhashi score in predicting survival in patients with metastatic spinal cord compression (MSCC). Eur Spine J 2013;22(Suppl. 1):S21–6. doi:10.1007/s00586-012-2649-5.
[20] Rades D, Dunst J, Schild SE. The first score predicting overall survival in patients with metastatic spinal cord compression. Cancer 2008;112:157–61. doi:10.1002/cncr.23150.
[21] Tabouret E, Cauvin C, Fuentes S, Esterni B, Adetchessi T, Salem N, et al. Reassessment of scoring systems and prognostic factors for metastatic spinal cord compression. Spine J 2015;15:944–50. doi:10.1016/j.spinee.2013.06.036.
[22] Tang V, Harvey D, Park Dorsay J, Jiang S, Rathbone MP. Prognostic indicators in metastatic spinal cord compression: using functional independence measure and Tokuhashi scale to optimize rehabilitation planning. Spinal Cord 2007;45:671–7. doi:10.1038/sj.sc.3102024.
[23] Tokuhashi Y, Uei H, Oshima M, Ajiro Y. Scoring system for prediction of metastatic spine tumor prognosis. World J Orthop 2014;5:262–71. doi:10.5312/wjo.v5.i3.262.
[24] Aoude A, Amiot L-P. A comparison of the modified Tokuhashi and Tomita scores in determining prognosis for patients afflicted with spinal metastasis. Can J Surg 2014;57:188–93. doi:10.1503/cjs.012013.
[25] Tokuhashi Y, Matsuzaki H, Toriyama S, Kawano H, Ohsaka S. Scoring system for the preoperative evaluation of metastatic spine tumor prognosis. Spine 1990;15:1110–13.
[26] Tomita K, Kawahara N, Kobayashi T, Yoshida A, Murakami H, Akamaru T. Surgical strategy for spinal metastases. Spine 2001;26:298–306.
[27] Ulmar B, Naumann U, Catalkaya S, Muche R, Cakir B, Schmidt R, et al. Prognosis scores of Tokuhashi and Tomita for patients with spinal metastases of renal cancer. Ann Surg Oncol 2007;14:998–1004. doi:10.1245/s10434-006-9000-5.
[28] Ulmar B, Richter M, Cakir B, Muche R, Puhl W, Huch K. The Tokuhashi score: significant predictive value for the life expectancy of patients with breast cancer with spinal metastases. Spine 2005;30:2222–6.
[29] van der Linden YM, Dijkstra SPDS, Vonk EJA, Marijnen CAM, Leer JWH. Dutch Bone Metastasis Study Group. Prediction of survival in patients with metastases in the spinal column: results based on a randomized trial of radiotherapy. Cancer 2005;103:320–8. doi:10.1002/cncr.20756.
[30] Walker MP, Yaszemski MJ, Kim CW, Talac R, Currier BL. Metastatic disease of the spine: evaluation and treatment. Clin Orthop 2003;(415 Suppl.):S165–75. doi:10.1097/01.blo.0000092977.12414.f9.
[31] Wang M, Bünger CE, Li H, Wu C, Høy K, Niedermann B, et al. Predictive value of Tokuhashi scoring systems in spinal metastases, focusing on various primary tumor groups: evaluation of 448 patients in the Aarhus spinal metastases database. Spine 2012;37:573–82. doi:10.1097/BRS.0b013e31822bd6b0.
[32] Wibmer C, Leithner A, Hofmann G, Clar H, Kapitan M, Berghold A, et al. Survival analysis of 254 patients after manifestation of spinal metastases: evaluation of seven preoperative scoring systems. Spine 2011;36:1977–86. doi:10.1097/BRS.0b013e3182011f84.
[33] Yamashita T, Siemionow KB, Mroz TE, Podichetty V, Lieberman IH. A prospective analysis of prognostic factors in patients with spinal metastases: use of the revised Tokuhashi score. Spine 2011;36:910–17. doi:10.1097/BRS.0b013e3181e56ec1.
[34] Zhang D, Xu W, Liu T, Yin H, Yang X, Wu Z, et al. Surgery and prognostic factors of patients with epidural spinal cord compression caused by hepatocellular carcinoma metastases: retrospective study of 36 patients in a single center. Spine 2013;38:E1090–5. doi:10.1097/BRS.0b013e3182983bf8.
[35] Zoccali C, Skoch J, Walter CM, Torabi M, Borgstrom M, Baaj AA. The Tokuhashi score: effectiveness and pitfalls. Eur Spine J 2016;25:673–8. doi:10.1007/s00586-015-4339-6.
[36] Paulino Pereira NR, Janssen SJ, van Dijk E, Harris MB, Hornicek FJ, Ferrone ML, et al. Development of a prognostic survival algorithm for patients with metastatic spine disease. J Bone Joint Surg Am 2016;98:1767–76. doi:10.2106/JBJS.15.00975.
[37] Paulino Pereira NR, McLaughlin L, Janssen SJ, van Dijk CN, Bramer JAM, Laufer I, et al. The SORG nomogram accurately predicts 3- and 12-months survival for operable spine metastatic disease: external validation. J Surg Oncol 2017;115:1019–27. doi:10.1002/jso.24620.
[38] Dardic M, Wibmer C, Berghold A, Stadlmueller L, Froehlich EV, Leithner A. Evaluation of prognostic scoring systems for spinal metastases in 196 patients treated during 2005–2010. Eur Spine J 2015;24:2133–41. doi:10.1007/s00586-014-3482-9.
[39] Bollen L, Wibmer C, Van der Linden YM, Pondaag W, Fiocco M, Peul WC, et al. Predictive value of six prognostic scoring systems for spinal

bone metastases: an analysis based on 1379 patients. Spine 2016;41:E155–62. doi:10.1097/BRS.0000000000001192.

[40] Whitmore RG, Stephen JH, Vernick C, Campbell PG, Yadla S, Ghobrial GM, et al. ASA grade and Charlson Comorbidity Index of spinal surgery patients: correlation with complications and societal costs. Spine J 2014;14:31–8. doi:10.1016/j.spinee.2013.03.011.

[41] Forsberg JA, Wedin R, Bauer HC, Hansen BH, Laitinen M, Trovik CS, et al. External validation of the Bayesian Estimated Tools for Survival (BETS) models in patients with surgically treated skeletal metastases. BMC Cancer 2012;12:493. doi:10.1186/1471-2407-12-493.

[42] Fisher CG, Dipaola CP, Ryken TC, Bilsky MH, Shaffrey CI, Berven SH, et al. A novel classification system for spinal instability in neoplastic disease: an evidence-based approach and expert consensus from the Spine Oncology Study Group. Spine 2010;15:E1221–9.

[43] Laufer I, Rubin DG, Lis E, Cox BW, Stubblefield MD, Yamada Y, et al. The NOMS framework: approach to the treatment of spinal metastatic tumors. Oncologist 2013;18:744–51.

[44] American Cancer Society. Cancer facts & figs. 2016. Cancer.org 2017. Available at https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2016.html Accessed February 10, 2017.