Original Research

The AJCC 8th Edition Staging System for Soft Tissue Sarcoma of the Extremities or Trunk: A Cohort Study of the SEER Database

Justin M. M. Cates, MD, PhD

Abstract

Background: The AJCC recently published the 8th edition of its cancer staging system. Significant changes were made to the staging algorithm for soft tissue sarcoma (STS) of the extremities or trunk, including the addition of 2 additional T (size) classifications in lieu of tumor depth and grouping lymph node metastasis (LNM) with distant metastasis as stage IV disease. Whether these changes improve staging system performance is questionable. **Patients and Methods:** This retrospective cohort analysis of 21,396 adult patients with STS of the extremity or trunk in the SEER database compares the AJCC 8th edition staging system with the 7th edition and a newly proposed staging algorithm using a variety of statistical techniques. The effect of tumor size on disease-specific survival was assessed by flexible, nonlinear Cox proportional hazard regression using restricted cubic splines and fractional polynomials. **Results:** The slope of covariate-adjusted log hazards for sarcoma-specific survival decreases for tumors >8 cm in greatest dimension, limiting prognostic information contributed by the new T4 classification in the AJCC 8th edition. Anatomic depth independently provides significant prognostic information. LNM is not equivalent to distant, non-nodal metastasis. Based on these findings, an alternative staging system is proposed and demonstrated to outperform both AJCC staging schemes. The analyses presented also disclose no evidence of improved clinical performance of the 8th edition compared with the previous edition. **Conclusions:** The AJCC 8th edition staging system for STS is no better than the previous 7th edition. Instead, a proposed staging system based on histologic grade, tumor size, and anatomic depth shows significantly higher predictive accuracy, with higher model concordance than either AJCC staging system. Changes to existing staging systems should improve the performance of prognostic models. Until such improvements are documented, AJCC committees should refrain from modifying established staging schemes.

J Natl Compr Canc Netw 2018;16(2):144–152
doi: 10.6004/jnccn.2017.7042

Considerable changes in the clinical and pathologic staging of soft tissue sarcoma (STS) are presented in the 8th edition of the *AJCC Cancer Staging Manual*.[1] There are some obvious improvements, such as the creation of separate staging schemes for different anatomic locations. It is well-known that STSs arising within the extremities or trunk, retroperitoneum or abdominopelvic cavities, or the head/neck region show inherent differences in their biological behavior and clinical courses.[2–13] However, other changes in the 8th edition, such as creating 2 additional T (size) classifications in lieu of tumor depth and grouping lymph node metastasis (LNM) with distant metastasis as stage IV disease, are not substantiated by a similar extent of clinical experience or published evidence.

This study was performed using data extracted from the SEER database to evaluate the changes implemented in the 8th edition of the *AJCC Cancer Staging Manual* for staging STS and to compare its clinical performance with that of the 7th edition and a proposed revised staging algorithm. The AJCC 8th edition offers no improvement in prognostication compared with the 7th edition, and furthermore is inferior to a revised staging scheme incorporating tumor depth. Until pro-

From the Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, Tennessee.
Submitted June 20, 2017; accepted for publication September 27, 2017.
The authors have disclosed that they have no financial interests, arrangements, affiliations, or commercial interests with the manufacturers of any products discussed in this article or their competitors.
Correspondence: Justin M. M. Cates, MD, PhD, Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Medical Center North, C-3322, 1161 21st Avenue South, Nashville, TN 37232. E-mail: justin.m.cates@vanderbilt.edu

posed modifications to a staging system demonstrate improved performance of prognostic models, AJCC committees should refrain from implementing alterations of staging schemes.

Patients and Methods

The SEER program of the NCI collects and publishes cancer incidence and survival data from 20 geographically diverse, population-based cancer registries representing approximately 28% of the US population. The SEER database was queried for malignant soft tissue tumors arising at all anatomic sites diagnosed from 1973 to 2013 to ensure all possible cases were extracted (see eAppendix 1 and supplemental eTable 1, available with this article at JNCCN.org, for SEER*Stat version used, database queried, search criteria, histologic and anatomic site categorizations, and details regarding exclusion of cases). Exclusion criteria included histologic diagnosis not recommended for AJCC staging, anatomic sites other than soft tissue of the extremity or trunk, cases diagnosed before 1990, cases without confirmation of diagnosis by histopathologic or cytopathologic examination, patients <18 years of age, and cases in which primary curative surgery was either not performed or could not be confirmed. Cases with tumor size <0.5 cm (n=65) were excluded based on sparsity of data in this size range. Cases with tumor size listed as >40 cm (n=75) were excluded for the same reason in addition to apparent rounding to the nearest 5 cm. The final cohort consisted of 21,396 cases. This population-based study of a publicly available deidentified patient database was exempt from Institutional Research Board approval.

Variables downloaded included information on patient age, sex, year of diagnosis, anatomic site, site-specific disease extent, lymph node and/or distant metastasis at diagnosis, AJCC TNM stage (Table 1),[14,15] histologic subtype of sarcoma, histologic grade, tumor size, primary treatment, disease-free survival status, and follow-up interval. Cases with follow-up intervals recorded as 0 months were recoded as 0.5 months if incomplete dates were available but survival of at least 1 day was documented (n=232); other cases with missing follow-up data (n=21), complete dates available with 0 days of survival (likely autopsy entries; n=44), or incomplete

Table 1. Staging Systems for Soft Tissue Sarcoma of the Extremity or Trunk

Staging System	Description
AJCC 7th edition[a]	
T1a	Tumor ≤5 cm in greatest dimension, superficial
T1b	Tumor ≤5 cm in greatest dimension, deep
T2a	Tumor >5 cm in greatest dimension, superficial
T2b	Tumor >5 cm in greatest dimension, deep
N0	No regional lymph node metastasis
N1	Regional lymph node metastasis
M0	No distant metastasis
M1	Distant metastasis
Stage groups	
Stage IA	T1a/b; N0; M0; G1
Stage IB	T2a/b; N0; M0; G1
Stage IIA	T1a/b; N0; M0; G2/3
Stage IIB	T2a/b; N0; M0; G2
Stage III	T2a/b; N0; M0; G3
	Any T; N1; M0; any G
Stage IV	Any T; Any N; M1; any G
AJCC 8th edition[b]	
T1	Tumor ≤5 cm in greatest dimension
T2	Tumor >5 cm and ≤10 cm in greatest dimension
T3	Tumor >10 cm and ≤15 cm in greatest dimension
T4	Tumor >15 cm in greatest dimension
N0	No regional lymph node metastasis or unknown lymph node status
N1	Regional lymph node metastasis
M0	No distant metastasis
M1	Distant metastasis
Stage groups	
Stage IA	T1; N0; M0; G1
Stage IB	T2, T3, T4; N0; M0; G1
Stage II	T1; N0; M0; G2/3
Stage IIIA	T2; N0; M0; G2/3
Stage IIIB	T3, T4; N0; M0; G2/3
Stage IV	Any T; N1; M0; any G
	Any T; any N; M1; any G
Vanderbilt staging system	
T1a	Tumor ≤5 cm in greatest dimension, superficial
T1b	Tumor ≤5 cm in greatest dimension, deep
T2a	Tumor >5 cm and ≤10 cm in greatest dimension, superficial
T2b	Tumor >5 cm and ≤10 cm in greatest dimension, deep
T3a	Tumor >10 cm in greatest dimension, superficial
T3b	Tumor >10 cm in greatest dimension, deep
N0	No regional lymph node metastasis or unknown lymph node status
N1	Regional lymph node metastasis
M0	No distant metastasis
M1	Distant metastasis
Stage groups	
Stage I	Any T; N0; M0; G1
	T1a; N0; M0; G2
Stage II	T1b; N0; M0; G2
	T2a/b; N0; M0; G2
	T3a/b; N0; M0; G2
	T1a/b; N0; M0; G3
	T2a; N0; M0; G3
Stage IIIA	T2b; N0; M0; G3
	T3a; N0; M0; G3
Stage IIIB	T3b; N0; M0; G3
	Any T; N1; M0; any G
	Any T; any N; M1; G1
Stage IV	Any T; any N; M1; G2/3

[a]Used with the permission of the American Joint Committee on Cancer (AJCC), Chicago, IL. The original source for this material is the AJCC Cancer Staging Manual, Seventh Edition (2010) published by Springer Science and Business Media LLC, www.springer.com.
[b]Used with the permission of the American Joint Committee on Cancer (AJCC), Chicago, IL. The original source for this material is the AJCC Cancer Staging Manual, Eighth Edition (2017) published by Springer Science and Business Media LLC, www.springer.com.

dates available with 0 days of follow-up possible (n=5) were excluded.

Statistical Analyses

The effect of sarcoma size on disease-specific survival (DSS) was assessed by flexible Cox proportional hazard regression using restricted cubic splines[16] and fractional polynomials.[17–20] For spline models, 5 knots were set at default locations.[21] Powers searched in 1- and 2-degree fractional polynomials ranged from –2 to 3, including square roots and the natural logarithm of the independent variable. The relative prognostic impact of predictive factors was assessed using a nomogram.[22]

Several different statistical methods were applied to compare different staging schemes. Kaplan-Meier curves were plotted and compared to assess the degree of discrimination between tumor stage categories. Subsequent pairwise comparisons of adjacent staging categories were performed using the Sidak method for multiple comparisons. The predictive accuracy of each staging system for determining 5-year DSS was evaluated by comparing the areas under receiver operating characteristic (ROC) curves generated from logistic regression. Three different concordance indices were also used for model comparisons: Harrell's *c*,[23,24] Somers' *D*,[25] and Gönen and Heller's *K*.[26] Because the first 2 indices neglect censored outcomes that occur before events,[27] this potential source of bias was minimized by excluding 5,089 patients censored before 5 years of clinical follow-up and comparing the differences of the estimated indices between nested regression models. These statistics were computed using bootstrap methods (500 replications) on separate training and validation sets semirandomly created after sorting on each staging system to ensure relatively similar percentages of each tumor stage in the bootstrapped samples. Gönen and Heller's *K* calculates the probability that a patient with a higher hazard ratio (HR) dies earlier than one with a lower HR. This index is based on model parameters and observed distributions of covariates and is independent of censoring bias. Therefore, all study cases were used to compute Gönen and Heller's *K* via separate bootstrapping methods. Finally, the amount of variation in observed outcomes explained by the various regression models was assessed using O'Quigley's ρ^2_k and Royston's modification thereof (R^2)[28,29]; standard errors of these estimates were obtained using bootstrap techniques (500 replications). The Bayesian information criterion based on the number of deaths secondary to sarcoma, and not simply overall sample size, was also calculated for each staging system as a measure of model fit.[30–33] All results are from 2-sided hypothesis tests using α=0.05. All analyses were performed using Stata v13.1 (Stata Corp., College Station, TX).

Results

The SEER Cohort of Sarcomas Arising in the Extremities and Trunk

Clinicopathologic characteristics of the entire cohort (N=21,396) are provided in supplemental eTable 2. Mean age was 59 years (SD, 18 years; median, 59 years; interquartile range [IQR], 45–73 years). Mean tumor size was 8.7 cm (SD, 6.5 cm; median, 7 cm; IQR, 4–12 cm). Median follow-up for censored patients was 63 months (range, 0.5–287 months). A total of 4,050 patients (19%) died of sarcoma a median of 22 months after surgical resection (range, 0.5–258 months).

Survival Analysis

All variables tested were statistically significant in univariate sarcoma-specific survival analysis, with profound effects seen for histologic grade and distant metastasis, as expected (supplemental eTable 3). Because HRs for radiation therapy before, during, or after surgical resection were not significantly different from each other (data not shown), these groups were combined for multivariable analysis. The only variable that lost statistical significance on multivariable regression was neoadjuvant chemotherapy, and therefore this variable was excluded. Unless otherwise stated, all subsequent multivariable regression analyses included the covariates listed in supplemental eTable 3.

Assessing Size as a Continuous Predictor Variable for DSS

Because tumor size is such a critical parameter by which STSs are staged in each edition of the AJCC staging system, the relationship between the log hazard and tumor size was investigated. Coefficients for HRs from Cox proportional hazard models of DSS were calculated using multivariable restricted cubic

splines and multivariable fractional polynomial regression adjusted for all covariates listed in supplemental eTable 3. Predicted log HRs for tumor size from the linear and nonlinear models are plotted in Figure 1, which demonstrates a biphasic effect of tumor size on adjusted HRs for DSS. The effect of increasing size is greatest for tumors up to 8 cm, with a log-linear effect of smaller magnitude for larger tumors. The AJCC 8th edition T categories also reflect these findings, showing decreasing contrasts between T categories with increasing tumor size.

Tumor Depth as an Independent Prognostic Factor

Another difference among previous AJCC staging systems is consideration of tumor depth (superficial or deep to superficial fascia). Comparison of superficial and deep tumors in multivariable regression demonstrated that deep tumors confer a 55% increased risk for sarcoma-specific death after accounting for confounding covariates (95% CI, 35%–78%; *P*<.0005; supplemental eTable 3). Predicted HRs for sarcoma-specific survival by tumor depth and histologic grade after adjusting for interactions between tumor depth and size and the presence or absence of metastasis show that for intermediate- and high-grade tumors, anatomic depth contributes significant prognostic information (Figure 2).

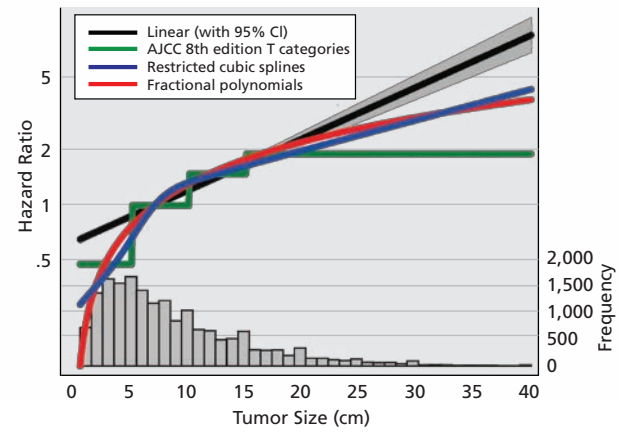

Figure 1. Effect of tumor size as a continuous variable on hazard ratio (HR) for sarcoma-specific survival. Predicted log HRs for tumor size were modeled assuming a linear effect (95% confidence interval [CI] also shown) and compared with those predicted by 2 different techniques for flexible, nonlinear regression (restricted cubic splines and fractional polynomials). Predicted HRs for AJCC 8th edition T classifications are also plotted for comparison. HRs for tumor size in each model were adjusted for year of diagnosis, patient age, sex, anatomic site (extremity vs trunk), anatomic depth (deep vs superficial), histologic grade, neoadjuvant or adjuvant radiation therapy, and presence of nodal or distant metastasis. All plots are centered on median tumor size (7 cm). A histogram of tumor size is also provided for reference.

Prognostic Impact of LNM

Categorization of LNM is inconsistent across different AJCC staging editions, probably because nodal metastasis is relatively uncommon for STS (supplemental eTable 2). Regardless, evidence from the SEER database suggests that nodal and distant metastatic disease should not be combined as stage IV disease. After controlling for covariates, pairwise comparison of HRs showed significant differences between distant and nodal metastasis (HR, 1.60; 95% CI, 1.18–2.16; *P*=.002), as well as between nodal metastasis and localized sarcomas (HR, 2.88; 95% CI, 2.19–3.80; *P*<.0005; Figure 3).

Derivation of a Revised Staging Scheme

Considering these potential issues with the AJCC 8th edition staging system, a revised staging system was derived based on assessment of the relative impact of predictive factors. Using nomographic techniques, histologic grade, tumor size, and distant metastasis were confirmed to have profound effects on DSS (supplemental eFigure 1). Given the relationship between tumor size and its adjusted log HR (Figure 1), the AJCC 8th edition T3 (>10 cm and ≤15 cm) and T4 (>15 cm) categories were combined. Lymph node status was considered separately from distant metastasis as well (Figure 3). Other potential anatomic pre-

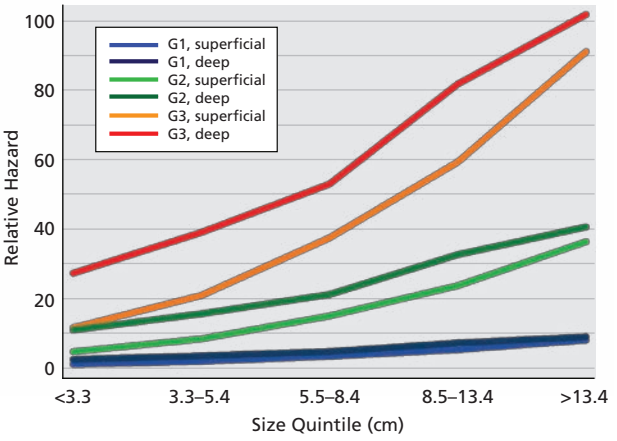

Figure 2. Predicted relative hazard ratios for sarcoma-specific survival for cases without nodal or distant metastasis by tumor depth and histologic grade over quintiles of tumor size. Plots are relative to superficial, low-grade tumors <3.3 cm in greatest dimension and are adjusted for interactions between tumor depth and size. Abbreviations: G1, low-grade (grade 1); G2, intermediate-grade (grade 2); G3, high-grade (grade 3).

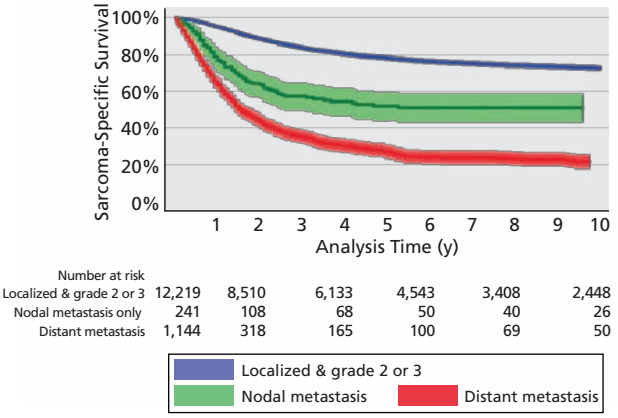

Figure 3. Kaplan-Meier sarcoma-specific survival curves with 95% confidence bands for patients with intermediate- or high-grade sarcomas localized at presentation, lymph node (but not distant) metastasis at presentation, or distant, non-nodal metastasis at presentation.

dictive factors evaluated included tumor depth and location. Anatomic depth showed a slightly greater predictive potential than anatomic location in multivariable Cox regression models (supplemental eTable 3 and eFigure 1), and therefore this variable was incorporated into the proposed staging algorithm.

The Vanderbilt staging system for STS of the extremity and trunk (Table 1) was then developed from empirical observations of sarcoma-specific survival of cases classified by histologic grade, modified AJCC 8th edition T classification, anatomic depth, and nodal or distant metastasis (supplemental eFigure 2 and eTable 4). In contrast to the AJCC 7th and 8th edition staging systems, the Vanderbilt staging system for STS separates patients into 5 relatively evenly distributed prognostic groups (Table 2).

Comparison of Vanderbilt Staging System Versus AJCC Editions

The 7th and 8th editions of the AJCC staging system for STSs were then compared with the Vanderbilt staging system to determine which, if any, is most predictive of patient outcome. Visual inspection of Kaplan-Meier plots for DSS showed poor separation for stages IA and IB in the AJCC 8th edition staging system (Figure 4). Likewise, stages IA and IB and stages IIA and IIB in the AJCC 7th edition staging system show poor discriminative ability. In contrast, the Vanderbilt staging system shows significant contrast between each adjacent stage group (Figure 4). Pairwise comparison of HR coefficients confirm these impressions (Table 3).

Table 2. Number of Patients in Each Stage Group

Staging System	N	%
AJCC 8th edition		
Stage IA	870	5.6
Stage IB	2,126	13.8
Stage IIA	4,122	26.7
Stage IIB	3,849	25.0
Stage III	3,071	19.9
Stage IV	1,385	9.0
Missing	5,973	—
AJCC 7th edition		
Stage IA	870	5.6
Stage IB	2,126	13.7
Stage IIA	4,122	26.6
Stage IIB	1,551	10.0
Stage III	5,659	36.6
Stage IV	1,144	7.4
Missing	5,924	—
Vanderbilt staging system		
Stage I	2,246	21.7
Stage II	3,475	33.6
Stage IIIA	1,842	17.8
Stage IIIB	1,896	18.3
Stage IV	872	8.4
Missing	11,065	—

Predictive Ability for 5-Year DSS

Next, the capability of each staging system to predict 5-year survival was assessed by logistic regression using the subset of cases not missing values for any of the examined staging systems (supplemental eTable 5). In this subset, 3,365 patients survived at least 5 years after diagnosis and 1,914 (36%) died within 5 years of initial diagnosis. ROC curves generated from logistic regression models (Figure 5) showed good predictive ability for each AJCC staging system (AJCC 8th edition, 78.9% [SE, 0.6%]; AJCC 7th edition, 78.6% [SE, 0.6%]) compared with a simple predictive model containing information only on the presence of nodal or distant metastasis at diagnosis (63.0% [SE, 0.6%]). However, the Vanderbilt staging system showed significantly higher accuracy (80.4% [SE, 0.6%]) than both AJCC staging systems ($P<.00005$).

Comparison of Concordance Indices and Measures of Model Fit

The Vanderbilt staging system showed significantly higher concordance with clinical outcomes than the AJCC staging systems for 2 of the 3 indices calculated (Table 4); concordances of the AJCC staging systems were not significantly different. Two measures of degree of variation explained by the regression models were also compared. Values for both ρ^2_k and

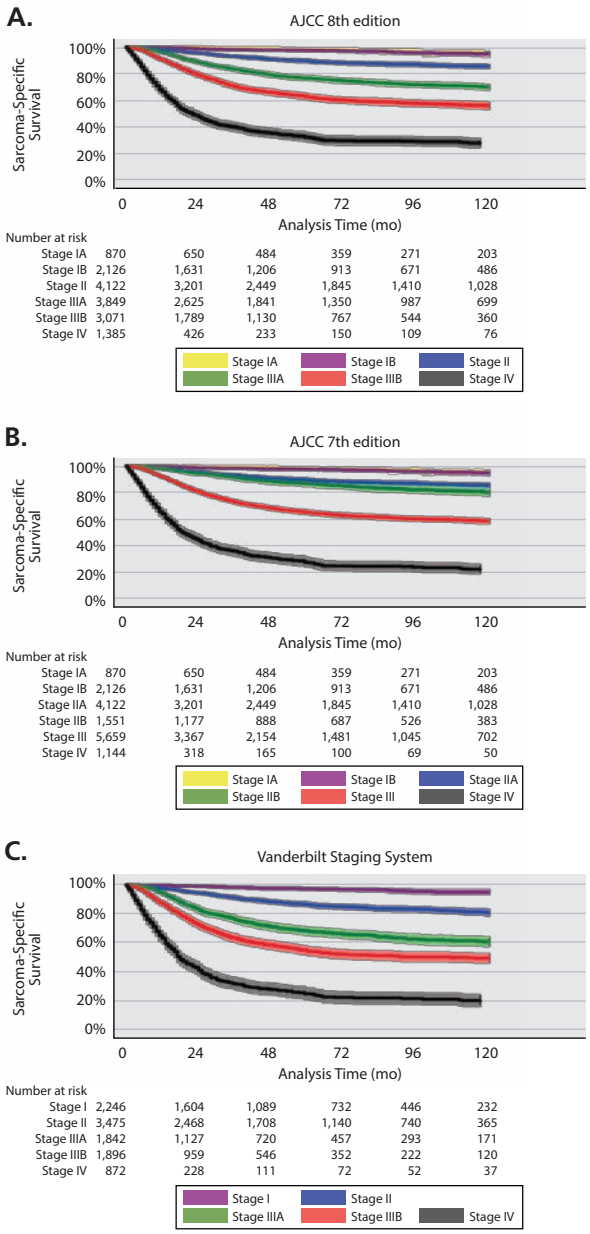

Figure 4. Kaplan-Meier sarcoma-specific survival curves with 95% confidence bands by (A) AJCC 8th edition, (B) AJCC 7th edition and (C) Vanderbilt staging algorithms for soft tissue sarcoma of the extremity or trunk.

R^2 were highest with the Vanderbilt staging system. This proposed staging system was also associated with the lowest Bayesian information criterion, suggesting that of these 3 regression models, the Vanderbilt staging system generates the best-fitting model.

Sensitivity Analysis

Because some of the statistical analyses were performed using only cases without missing data for staging variables, sensitivity analysis was performed on a subset of cases diagnosed between 2000 and 2013 (n=17,096). No marked differences were obtained for any of the results. The number of patients with discordant stage assignments and the corresponding sarcoma-specific survival rates are presented in supplemental eTables 6 and 7.

Table 3. Pairwise Comparisons of Hazard Ratios for Adjacent Tumor Stages (N=10,317)

	Hazard Ratio (95% CI)	*P* Value
AJCC 8th edition		
Stage IB vs IA	1.89 (0.66–5.38)	.47
Stage II vs IB	2.99 (1.94–4.60)	<.0005
Stage IIIA vs II	2.48 (2.02–3.04)	<.0005
Stage IIIB vs IIIA	1.72 (1.49–1.99)	<.0005
Stage IV vs IIIB	2.77 (2.40–3.20)	<.0005
AJCC 7th edition		
Stage IB vs IA	1.89 (0.66–5.38)	.47
Stage IIA vs IB	2.99 (1.94–4.60)	<.0005
Stage IIB vs IIA	1.43 (1.07–1.90)	.007
Stage III vs IIB	2.74 (2.15–3.49)	<.0005
Stage IV vs III	3.44 (3.00–3.94)	<.0005
Vanderbilt staging system		
Stage II vs I	4.57 (3.21–6.50)	<.0005
Stage IIIA vs II	2.51 (2.12–2.97)	<.0005
Stage IIIB vs IIIA	1.59 (1.37–1.85)	<.0005
Stage IV vs IIIB	2.71 (2.34–3.15)	<.0005

Discussion

Prognostic stratification of STS using the AJCC TNM staging system is difficult because of the wide distribution of anatomic sites affected by sarcoma, the diverse biological behaviors of different histologic subtypes of sarcoma, the relative infrequency of LNM, and the perceived necessity of dichotomizing or otherwise categorizing a continuous variable (size) as a primary staging criterion.[21,34] The first of these problems is now rectified in the 8th edition of the *AJCC Cancer Staging Manual*,[1] which provides separate staging algorithms for sarcomas of the extremity and trunk, retroperitoneum, or head/neck regions. However, whether the latest AJCC staging system for sarcomas of the extremity and trunk is an improvement compared with previous editions concerning the other aforementioned problems is unclear. The results of this study suggest that it is not.

Multiple statistical analyses failed to disclose any suggestion of improved clinical performance associated with the 8th edition of the AJCC staging system.

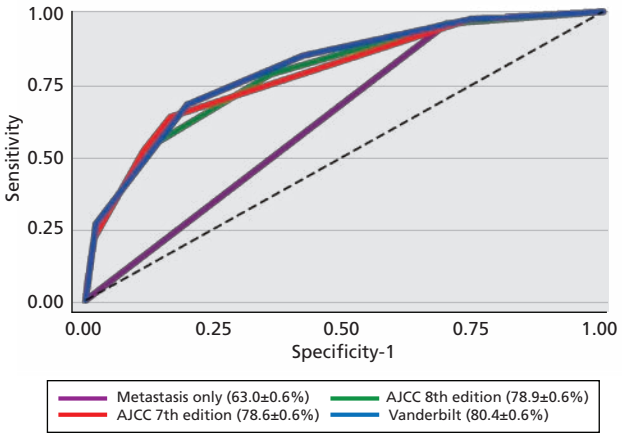

Figure 5. Receiver operating characteristic curves for 5-year sarcoma-specific survival by AJCC 7th and 8th edition (ed) and Vanderbilt staging systems (N=5,284). A model consisting of only lymph node or distant metastasis is included for comparison. Calculated predictive accuracy (with standard errors) for each staging system is also provided.

This is rather disappointing, particularly considering the loss of data recorded in the new system (tumor depth) as well as the cost in time and effort incurred by physicians and tumor registrars in converting their reporting practices to the new AJCC system. In contrast, a proposed staging system incorporating anatomic depth, slightly modifying AJCC 8th edition T categories, and separating cases with LNM from those with distant metastasis (as in the AJCC 7th edition) does show significant improvement in predicting 5-year survival and model concordance.

A major problem with all AJCC STS staging systems is reliance on dichotomization or categorization of tumor size.[2,13,35–40] Categorizing continuous data generates regression coefficients weighted by the distribution of data within each category, which almost always fails to capture the true nonlinear relationship between a continuous variable and its log hazard.[21] This caveat aside, although T categorization in the AJCC 8th edition staging system seems to recapitulate the nonlinear effect of tumor size, subcategorization of tumors >10 cm in dimension does not increase prognostic utility. Until an adequate surrogate for primary tumor growth and/or extension is implemented in the sarcoma staging system, this issue will remain one of the foremost challenges in prognostication of STS. One possible alternative is Enneking et al's concept of compartmental status, in which a sarcoma is considered intracompartmental if "within a well-delineated anatomic compartment" defined by "major fascial septae and the tendinous origins and insertions of muscles in soft tissues."[13] Conversely, extracompartmental tumors "diffusely [infiltrate] poorly demarcated adventitial planes and spaces" or arise within "ill-defined interfascial spaces and planes…limited only by loose areolar tissues that favor occult microextension."[13] Analysis of compartmental status could not be performed in this study because data regarding local tumor extension are poorly documented for STS in the SEER database. At this point, the Enneking staging system has not been tested thoroughly in staging STS.

Another problem with the AJCC 7th and 8th editions is the disregard for independent prognostic information provided by tumor depth. Deep tumors are associated with worse outcomes than superficial ones, even after controlling for tumor size and histologic grade.[2,9,36,41–43] Finally, there is no rationale for staging lymph nodes as equivalent to distant, nonnodal metastasis. In the current study, sarcomas associated with nodal metastasis showed significantly lower risk of disease-specific death than distant, nonnodal metastases. Other investigators have reported similar findings.[2,44,45] Regardless, risk stratification of LNM is not likely a major factor affecting the performance of prognostic models because <5% of patients with sarcoma present with nodal metastasis.

The current AJCC staging scheme may also be improved through stratification by histologic risk group, because histologic subtype remains an inde-

Table 4. Concordance Indices, Measures of Explained Variation, and Bayesian Information Criterion[a]

Index	AJCC 8th Edition	AJCC 7th Edition	Vanderbilt System	*P* Value
Harrell's *c* (SE)	0.746 (0.009)[b]	0.744 (0.009)[b]	0.760 (0.009)[c]	.0001
Somers' *D* (SE)	0.538 (0.022)[b]	0.534 (0.021)[b]	0.576 (0.022)[c]	<.0005
Gönen & Heller's *K* (SE)	0.731 (0.024)	0.730 (0.023)	0.739 (0.022)	.68
O'Quigley's ρ^2_k (SE)	0.572 (0.029)[b]	0.588 (0.022)[b]	0.613 (0.022)[c]	<.0005
Royston's R^2 (SE)	0.446 (0.025)[b]	0.462 (0.023)[b]	0.486 (0.023)[c]	.0001
BIC	34,866.37	34,800.0	34,638.8	—

Abbreviations: BIC, Bayesian information criterion; SE, standard error.
[a]Cells sharing footnotes (b,c) were not significantly different in post hoc tests.

pendent prognostic factor, even after controlling for tumor grade.[2,9,35,36,38,46,47] That a single staging system can or should be used for risk stratification of dozens of histologic subtypes of sarcoma, ranging from relatively indolent tumors such as atypical lipomatous tumor to highly aggressive tumors like primitive neuroectodermal tumor, is unrealistic.[38] In the future, the AJCC might also consider incorporating pertinent prognostic genomic/molecular data as it becomes available.[48]

The strengths of the SEER database are its size, established quality assurance program, and internal and external validity.[49,50] It is therefore a useful tool by which to validate or compare different staging algorithms.[50] There are some drawbacks to using the SEER database, however.[51] It does not allow for evaluation of disease-free survival because time to local or distant disease recurrence is not available. Extracting relevant data from historical variables regarding extent of disease is difficult.[52] Continued use of the Collaborative Stage schema (https://cancerstaging.org/cstage) for recording local extent of disease should allow for a detailed assessment of this predictive factor and application of the staging principles Enneking et al advocated in future studies. Data regarding neoadjuvant and adjuvant therapy, which is important to consider in studies of patient outcome, are less robust than other recorded variables in the SEER database. The accuracy of SEER data on histologic diagnosis and Fédération Nationale des Centres de Lutte Contre le Cancer (FNCLCC) grade is also debatable given the lack of central review by expert soft tissue pathologists. Because sarcoma subtype diagnosis may be unreliable in the SEER database, this variable was not considered in this analysis. In addition, many database records lack data for one or more variables. For example, tumor stage in the Vanderbilt staging system could not be derived for more than half of SEER cases because of missing data. Therefore, sensitivity analysis was performed on the subset of cases diagnosed between 2000 and 2013. Marked differences in results were not observed, suggesting that missing data did not cause significant analytical bias.

Conclusions

Site-specific staging systems in the 8th edition of the *AJCC Cancer Staging Manual* resolve a major problem heretofore unaddressed in staging STSs. Although the data presented here show that the new site-specific staging system for extremity and trunk sarcomas performs no better than the previous non–site-specific version, this does not necessarily invalidate the merit or utility of anatomic site-specific staging systems. Separate staging systems for different anatomic sites may be an important first step in improving the staging of STS in general. Clearly, the current staging system can be improved by including additional important predictor variables, as demonstrated with the Vanderbilt staging system. Alternatively, a staging system incorporating histologic risk group might also result in better risk stratification. At the very least, the AJCC should not issue broad mandatory changes in staging algorithms without evidence that such changes improve clinical and pathologic staging.

References

1. Amin MB, Edge S, Greene F, et al, eds. AJCC Cancer Staging Manual. 8th ed. Cham, Switzterland: Springer International Publishing; 2017.
2. Maki RG, Moraco N, Antonescu CR, et al. Toward better soft tissue sarcoma staging: building on American Joint Committee on Cancer staging systems versions 6 and 7. Ann Surg Oncol 2013;20:3377–3383.
3. Nathan H, Raut CP, Thornton K, et al. Predictors of survival after resection of retroperitoneal sarcoma. Ann Surg 2009;250:970–976.
4. Kotilingam D, Lev D, Lazar A, Pollock R. Staging soft tissue sarcoma: evolution and change. CA Cancer J Clin 2006;56:282–291.
5. Zagars GK, Ballo MT, Pisters PWT, et al. Prognostic factors for patients with localized soft tissue sarcoma treated with conservation surgery: an analysis of 1225 patients. Cancer 2003;97:2530–2543.
6. Finn HA, Simon MA. Staging systems for musculoskeletal neoplasms. Orthopedics 1989;12:1365–1371.
7. Dudhat SB, Mistry RC, Varughese T, et al. Prognostic factors in head and neck soft tissue sarcomas. Cancer 2000;89:868–872.
8. Eeles RA, Fisher C, A'Hern RP, et al. Head and neck sarcomas: prognostic factors and implications for treatment. Br J Cancer 1993;68:201–207.
9. Kattan MW, Leung DH, Brennan MF. Postoperative nomogram for 12-year sarcoma-specific death. J Clin Oncol 2002;20:791–796.
10. Gutierrez JC, Perez EA, Franceschi D, et al. Outcomes for soft-tissue sarcoma in 8,249 cases from a large state cancer registry. J Surg Res 2007;141:105–114.
11. Lewis JJ, Leung D, Woodruff JM, Brennan MF. Retroperitoneal soft-tissue sarcoma: analysis of 500 patients treated and followed at a single institution. Ann Surg 1998;228:355–365.
12. LeVay J, O'Sullivan B, Catton C, et al. Outcome and prognostic factors in soft tissue sarcoma in the adult. Int J Radiat Oncol Biol Phys 1993;27:1091–1099.
13. Enneking WF, Spanier SS, Goodman MA. A system for the surgical staging of musculoskeletal sarcoma. Clin Orthop Relat Res 1980;153:106–120.
14. Edge SB, Byrd DR, Compton CC, et al, eds. AJCC Cancer Staging Manual. 7th ed. New York, NY: Springer; 2010.
15. Greene FL, Compton CC, Fritz AG, et al, eds. AJCC Cancer Staging Atlas. 6th ed. New York, NY; Springer; 2006.
16. Durrleman S, Simon R. Flexible regression models with cubic splines. Stat Med 1989;8:551–561.
17. Royston P, Sauerbrei W. Building multivariable regression models with continuous covariates in clinical epidemiology—with an emphasis on fractional polynomials. Methods Inf Med 2005;44:561–571.

18. Royston P, Altman DG. Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling. J R Stat Soc Ser C (Applied Stat) 1994;43:429–467.
19. Sauerbrei W, Meier-Hirmer C, Benner A, Royston P. Multivariable regression model building by using fractional polynomials: description of SAS, STATA and R programs. Comput Stat Data Anal 2006;50:3464–3485.
20. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. Int J Epidemiol 1999;28:964–974.
21. Harrell FE Jr. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. 2nd ed. Cham, Switzerland: Springer International Publishing; 2015.
22. Zlotnik A, Abraira V. A general-purpose nomogram generator for predictive logistic regression models. Stata J 2015;15:537–546.
23. Harrell FE Jr, Lee KL, Pollock BE. Regression models in clinical studies: determining relationships between predictors and response. J Natl Cancer Inst 1988;80:1198–1202.
24. Harrell FE, Califf RM, Pryor DB, et al. Evaluating the yield of medical tests. JAMA 1982;247:2543–2546.
25. Somers RH. A new asymmetric measure of association for ordinal variables. Am Sociol Rev 1962;27:799–811.
26. Gönen M, Heller G. Concordance probability and discriminatory power in proportional hazards regression. Biometrika 2005;92:965–970.
27. Koziol JA, Jia Z. The concordance index C and the Mann-Whitney parameter Pr(X>Y) with randomly censored data. Biometrical J 2009;51:467–474.
28. O'Quigley J, Xu R, Stare J. Explained randomness in proportional hazards models. Stat Med 2005;24:479–489.
29. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol 2013;13:33.
30. Neath AA, Cavanaugh JE. The Bayesian information criterion: background, derivation, and applications. WIREs Comp Stat 2012;4:199–203.
31. Volinsky CT, Raftery AE. Bayesian information criterion for censored survival models. Biometrics 2000;56:256–262.
32. Posada D, Buckley TR. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst Biol 2004;53:793–808.
33. Burnham KP, Anderson DR. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. 2nd ed. New York, NY: Springer-Verlag; 2002.
34. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med 2006;25:127–141.
35. Pisters PW, Leung DH, Woodruff J, et al. Analysis of prognostic factors in 1,041 patients with localized soft tissue sarcomas of the extremities. J Clin Oncol 1996;14:1679–1689.
36. Brennan MF, Antonescu CR, Moraco N, Singer S. Lessons learned from the study of 10,000 patients with soft tissue sarcoma. Ann Surg 2014;260:416–422.
37. Ramanathan RC, A'Hern R, Fisher C, Thomas JM. Modified staging system for extremity soft tissue sarcomas. Ann Surg Oncol 1999;6:57–69.
38. Callegaro D, Miceli R, Bonvalot S, et al. Development and external validation of two nomograms to predict overall survival and occurrence of distant metastases in adults after surgical resection of localised soft-tissue sarcomas of the extremities: a retrospective analysis. Lancet Oncol 2016;17:671–680.
39. Raut CP, Miceli R, Strauss DC, et al. External validation of a multi-institutional retroperitoneal sarcoma nomogram. Cancer 2016;122:1417–1424.
40. Bagaria SP, Wagie AE, Gray RJ, et al. Validation of a soft tissue sarcoma nomogram using a national cancer registry. Ann Surg Oncol 2015;22:398–403.
41. Coindre JM, Terrier P, Bui NB, et al. Prognostic factors in adult patients with locally controlled soft tissue sarcoma: a study of 546 patients from the French Federation of Cancer Centers Sarcoma Group. J Clin Oncol 1996;14:869–877.
42. Gaynor JJ, Tan CC, Casper ES, et al. Refinement of clinicopathologic staging for localized soft tissue sarcoma of the extremity: a study of 423 adults. J Clin Oncol 1992;10:1317–1329.
43. Wunder JS, Healey JH, Davis AM, Brennan MF. A comparison of staging systems for localized extremity soft tissue sarcoma. Cancer 2000;88:2721–2730.
44. Riad S, Griffin AM, Liberman B, et al. Lymph node metastasis in soft tissue sarcoma in an extremity. Clin Orthop Relat Res 2004;426:129–134.
45. Behranwala KA, A'Hern R, Omar AM, Thomas JM. Prognosis of lymph node metastasis in soft tissue sarcoma. Ann Surg Oncol 2004;11:714–719.
46. Zagars GK, Ballo MT, Pisters PW, et al. Prognostic factors for patients with localized soft-tissue sarcoma treated with conservation surgery and radiation therapy. Cancer 2003;97:2530–2543.
47. Lahat G, Tuvin D, Wei C, et al. New perspectives for staging and prognosis in soft tissue sarcoma. Ann Surg Oncol 2008;15:2739–2748.
48. Chibon F, Lagarde P, Salas S, et al. Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity. Nat Med 2010;16:781–787.
49. Park HS, Lloyd S, Decker RH, et al. Overview of the Surveillance, Epidemiology, and End Results database: evolution, data variables, and quality assurance. Curr Probl Cancer 2012;36:183–190.
50. Lloyd S, Park HS, Decker RH, et al. Using the Surveillance, Epidemiology, and End Results database to investigate rare cancers, second malignancies, and trends in epidemiology, treatment, and outcomes. Curr Probl Cancer 2012;36:191–199.
51. Park HS, Lloyd S, Decker RH, et al. Limitations and biases of the Surveillance, Epidemiology, and End Results database. Curr Probl Cancer 2012;36:216–224.
52. Nathan H, Pawlik TM. Limitations of claims and registry data in surgical oncology research. Ann Surg Oncol 2008;15:415–423.

See JNCCN.org for supplemental online content.