**Clinical Orthopaedics
and Related Research®**
A Publication of The Association of Bone and Joint Surgeons®

CrossMark

SYMPOSIUM: 2015 MEETINGS OF THE MUSCULOSKELETAL TUMOR SOCIETY AND THE
INTERNATIONAL SOCIETY OF LIMB SALVAGE

A Novel System for the Surgical Staging of Primary High-grade Osteosarcoma: The Birmingham Classification

Lee M. Jeys MSc, FRCS, Chris J. Thorne MBChB,
Michael Parry MD, FRCS, Czar Louie L. Gaston MD,
Vaiyapuri P. Sumathi MD, FRCPath, J. Robert Grimer FRCS

Abstract

Background   Chemotherapy response and surgical margins have been shown to be associated with the risk of local recurrence in patients with osteosarcoma. However, existing surgical staging systems fail to reflect the response to chemotherapy or define an appropriate safe metric distance from the tumor that will allow complete excision and closely predict the chance of disease recurrence. We therefore sought to review a group of patients with primary high-grade osteosarcoma treated with neoadjuvant chemotherapy and surgical resection and analyzed margins and chemotherapy response in terms of local recurrence.

Questions/purposes   (1) What predictor or combination of predictors available to the clinician can be assessed that more reliably predict the likelihood of local recurrence? (2) Can we determine a better predictor of local recurrence-free survival than the currently applied system of surgical margins? (3) Can we determine a better predictor of overall survival than the currently applied system of surgical margins?

Methods   This retrospective study included all patients with high-grade conventional osteosarcomas without metastasis at diagnosis treated at one center between 1997 and 2012 with preoperative chemotherapy followed by resection or amputation of the primary tumor who were younger than age 50 years with minimum 24-month followup for those still alive. A total of 389 participants matched the inclusion criteria. Univariate log-rank test and multivariate Cox analyses were undertaken to identify predictors of local recurrence-free survival (LRFS). The Birmingham classification was devised on the basis of two stems: the response to chemotherapy (good response = $\geq$ 90% necrosis; poor response = < 90% necrosis) and margins (< 2 mm or $\geq$ 2 mm). The 5-year overall survival rate was 67% (95% confidence interval [CI], 61%–71%) and 47 patients developed local recurrence (12%).

Results   Intralesional margins (hazard ratio [HR], 9.9; 95% CI, 1.2–82; p = 0.03 versus radical margin HR, 1) and a poor response to neoadjuvant chemotherapy (HR, 3.8; 95% CI, 1.7–8.4; p = 0.001 versus good response HR, 1) were independent risk factors for local recurrence (LR). The best predictor of LR, however, was a combination of

L. M. Jeys
School of Life and Health Sciences, Aston University, Birmingham, UK

C. J. Thorne
School of Clinical and Experimental Medicine, University of Birmingham, Edgbaston, UK

L. M. Jeys (✉), C. J. Thorne, M. Parry, C. L. L. Gaston, V. P. Sumathi, J. R. Grimer
The Royal Orthopaedic Hospital NHS Foundation Trust, Bristol Road South, Birmingham B31 2AP, UK
e-mail: lee.jeys@nhs.net

margins $\leq$ 2 mm and a less than 90% necrosis response to chemotherapy (Birmingham 2b HR, 19.6; 95% CI, 2.6-144; p = 0.003 versus Birmingham 1a; margin >2 mm and more than 90% necrosis HR, 1). Two-stage Cox regression model and higher Harrell's C statistic demonstrate that the Birmingham classification was superior to the Musculoskeletal Tumor Society (MSTS) margin classification for predicting LR (Harrell's C statistic Birmingham classification 0.68, MSTS criteria 0.59). A difference in overall survival was seen between groups of the Birmingham classification (log-rank test p < 0.0001), whereas the MSTS margin system was not discriminatory (log-rank test p = 0.14).

Conclusions Based on these observations, we believe that a combination of the recording of surgical margins in millimeters and the response to neoadjuvant chemotherapy can more accurately predict the risk of local recurrence than the current MSTS system. A multicenter collaboration study initiated by the International Society of Limb Salvage is recommended to test the validity of the proposed classification and if these findings are confirmed, this classification system might be considered the standard practice in oncology centers treating patients with osteosarcomas and allow more effective communication of margin status for research.

Level of Evidence Level IV, prognostic study.

Introduction

Osteosarcoma is the third most common type of neoplasia in adolescents preceded only by leukemia and lymphoma [19]. Although it can occur at any age, it is predominantly a disease that afflicts the young with a peak incidence in the second and third decades of life [9]. Numerous variables have been associated with an adverse prognosis in osteosarcoma including metastatic disease at presentation, nonosteoblastic histological subtype, tumor size, male gender, young age, tumor location, genetic variations, poor response to neoadjuvant chemotherapy, and inadequate surgical margins [4, 7, 10, 12]. However, the last two are the only factors that have been shown to independently increase the risk of local recurrence (LR) [3]. The relationship between inadequate margin and local disease recurrence has been clearly reported, although there remains substantial debate about the appropriate "thickness" of a margin for primary high-grade osteosarcoma [2, 3, 12]. From a pragmatic perspective, the interpretation of exactly what entails a marginal and wide excision as defined by Enneking et al. [11] remains inherently subjective and may vary depending on who is assessing the margin. A surgeon may classify a margin one way and the

involved pathologist might have a different designation based on histology. In both cases, the experience of the observer likely influences the margin designation.

Chemotherapy has a clear role in the management of high-grade osteosarcoma, leading to improved survival rates when combined with adequate local control by surgical resection or amputation. There is sufficient evidence to argue that a response to chemotherapy (determined by analyzing the percentage of necrotic tumor in the resected specimen after neoadjuvant chemotherapy) is an important risk factor for LR [3, 19], although surprisingly, this is not reflected in any existing staging criteria [15]. The Musculoskeletal Tumor Society (MSTS) was the first to consider the extent (margin) of the surgical excision of sarcomas and relate it into the staging matrix, on which ongoing treatment decisions are based and long-term recurrence-free prognosis is predicted [11]. However, this proposal was made in the era just before the routine use of chemotherapy for bone sarcomas and was considered to be a "surgical" staging system that could be modified by the use of adjuvants such as chemotherapy or radiation therapy. However, the existing surgical staging systems fail to adequately reflect the response to chemotherapy or define an appropriate safe metric distance from the tumor that will allow complete excision and closely predict the chance of disease recurrence.

The purpose of this study was to review a group of patients with primary high-grade osteosarcoma treated with neoadjuvant chemotherapy and surgical resection and to analyze the risk of LR based on the chemotherapy response and surgical margins achieved. We asked the following questions: (1) What predictor or combination of predictors available to the clinician can be assessed that more reliably predict the likelihood of local recurrence? (2) Can we determine a better predictor of local recurrence-free survival than the currently applied system of surgical margins? (3) Can we determine a better predictor of overall survival than the currently applied system of surgical margins?

Patients and Methods

A retrospective study was performed of our prospectively collected oncology database and all patients with a diagnosis of a primary high-grade conventional osteosarcoma treated between January 1, 1997, and December 31, 2012, at our institution were identified. Only patients who had histologically diagnosed high-grade osteosarcoma, younger than age 50 years at the time of diagnosis, and who had been managed with preoperative chemotherapy followed by surgery were included in the study. A minimum of 2 years followup was required for patients alive. From a total

of 558 eligible patients seen during the study time period, those with metastasis at diagnosis (n = 103 [18.5%]), progressed on preoperative chemotherapy and did not receive definitive surgery (n = 38 [7%]), patients diagnosed at our institution but who received final surgery at another unit (n = 7 [1%]) and followup elsewhere (n = 18 [3%]), and died from complications of chemotherapy (n = 3 [0.5%]) were excluded.

Patient demographic characteristics were collected in addition to disease-specific variables from the database and then included: sex (male or female), age at diagnosis (< 16 or > 16 years), MSTS stage (intralesional, marginal, wide, radical), nearest margin to the tumor (mm), Fletcher's classification of type of osteosarcoma [13], tumor size (< 8 cm or > 8 cm), tumor location (extremity or central), percentage chemotherapy necrosis (< 90% or ≥ 90%), vascular invasion (yes or no), complete pathological fracture (yes or no), type of operation (limb salvage or amputation), local recurrence-free survival (LRFS), and overall survival (OS). The unicortical breaks were described in the radiology reports and initially mistakenly coded as pathologic fractures when they would have been the cortical breaks seen in osteosarcomas with large soft tissue extension. They are not pathologic fractures and were included in the nonpathologic group for analysis.

Patient Characteristics

A total of 389 (70% of 558) participants matched the inclusion criteria. Mean age at diagnosis was 16.4 years (SD 0.40; range, 3–49 years), 228 were males (59%), and mean followup was 79 months (SD 3.04; range, 4–219 months). The common anatomical sites for tumor location were the femur (n = 196 [50%]) followed by the tibia (n = 99 [25%]) and the humerus (n = 51 [13%]). Osteosarcoma mean length was 104 mm (SD 2.38; range, 20-420 mm) and osteoblastic subtype was the most prevalent (51%). Sixty patients (15%) presented with a pathological fracture (56 at diagnosis and four after diagnosis) requiring cast immobilization or traction while undergoing preoperative chemotherapy. Endoprosthetic reconstructions were the most commonly carried out surgical procedures (71%) (Table 1).

Over this time period, patients received neoadjuvant chemotherapy according to the respective European Osteosarcoma Intergroup (EOI) or European and American Osteosarcoma Studies (EURAMOS) trials at the time of their diagnosis primarily receiving cisplatin, doxorubicin, and methotrexate with surgery performed after 8 weeks and 12 weeks, respectively [17, 22]. In terms of histological response to chemotherapy, 42% had ≥ 90% necrosis and 58% had < 90% necrosis. Histological examination was

performed by a histopathologist (VPS) highly experienced in bone sarcomas. The resection specimens were photographed and examined for involvement of margins. Perpendicular sections were taken from the margins before bisecting the bone along its long axis exposing most of the tumor. All sections were examined for extent of chemotherapy-induced necrosis. Necrosis was expressed as a percentage of the total tumour area [18]. Margins were measured on histologic slides in millimeters from the resection surface to the nearest tumor (Fig. 1). All relevant information was extracted from the pathology reports and where the metric information of margins was lacking, the pathology slides were reviewed for this study (n = 8).

Statistical Analysis

Descriptive statistics were used to display demographic data. Kaplan-Meier analysis was used to determine LRFS and OS with time zero defined as the date of diagnosis and censored at the date of last followup or local recurrence and death, respectively. Univariate analysis was performed comparing groups with log-rank test and significant variables underwent subsequent multivariate Cox proportional hazard analysis to identify predictors of LRFS and OS with p value < 0.05 considered significant. A new classification of surgical margins (the Birmingham classification) was devised on two stems governed by the response to chemotherapy (good response = ≥ 90% necrosis; poor response = < 90% necrosis) with subdivisions within each stem by margin (≤ 2 mm or > 2 mm). To compare the existing MSTS staging criteria with the proposed Birmingham classification for predicting LR, a two-stage Cox regression model was undertaken where after entering one variable into the model, the second variable would only enter the model if the introduction of the latter significantly improved the prediction of LR. Corresponding Harrell's C statistics were calculated for the Birmingham classification and MSTS models. This process was repeated comparing the Birmingham classification with different margin thresholds (clear versus contaminated, 1 mm, 3 mm) and response to chemotherapy. Statistical analysis was performed on Stata (Version 9.2; College Station, TX, USA).

Results

Predictors of Local Recurrence

A total of 47 patients developed LR (12%). Factors that were identified as significant predictors of LRFS were surgical margins (intralesional margin hazard ratio [HR], 9.9; 95% confidence interval [CI], 1.2–82; p = 0.03 versus

Table 1. Patient characteristics

Patient characteristic	Number	Percentage
Sex		
Male	228	59
Female	161	41
Age (years)		
< 16	218	56
≥ 16	171	44
Tumor location		
Femur	196	50
Tibia	99	25
Humerus	51	13
Fibula	23	6
Pelvis	12	3
Radius	3	< 1
Ulna	2	< 1
Calcaneum	1	< 1
Scapula	1	< 1
Talus	1	< 1
Type of operation		
Endoprosthetic reconstruction	276	71.0
Amputation	89	23
Allograft	11	3
Excision	11	3
Rotationplasty	2	< 1
Fletcher's subclassification of osteosarcoma		
Osteoblastic	199	51
Chondroblastic	34	9
Telangiectatic	26	7
Fibroblastic	19	5
Giant cell	4	1
Small cell	3	< 1
Mixed	46	12
Unclassified	58	15
Intra-/extracompartmental		
Intracompartmental (IIa)	119	31
Extracompartmental (IIb)	267	69
Pathological fracture		
Yes	60	15
No	329	85
Skip lesions		
Yes	8	2
No	381	98
Vascular invasion		
Yes	44	11
No	345	89
Chemotherapy necrosis		
≥ 90%	162	42
< 90%	227	58

Table 1. continued

Patient characteristic	Number	Percentage
Musculoskeletal Tumor Society stage		
Intralesional	26	7
Marginal	175	45
Wide	144	37
Radical	44	11
Local recurrence		
Yes	47	12
No	352	88
Total	389	

radical margin HR, 1) and response to neoadjuvant chemotherapy (response < 90% chemotherapy necrosis HR, 3.8; 95% CI, 1.7–8.4; p = 0.001 versus response ≥ 90% chemotherapy necrosis HR, 1) (Table 2; Appendix 1 [Supplemental materials are available with the online version of *CORR*®.]). With the numbers we had, we could not detect a higher risk of LR by age < 16 years or those without a pathologic fracture.

Comparison of Birmingham With MSTS for Local Recurrence

The Birmingham system offered a better prediction of LRFS than did the MSTS system in a two-stage Cox regression model, wherein introducing the Birmingham classification into a model with the MSTS variable improved the prediction of LR (MSTS HR, 1.2; 95% CI, 0.8–1.9; p = 0.3 and Birmingham HR, 1.9; 95% CI, 1.3–2.7; p < 0.0001). Introducing the MSTS variable into a model already with the Birmingham classification did not improve the prediction of LR (Birmingham HR, 2.1; 95% CI, 1.5–2.8; p < 0.0001; MSTS omitted p = 0.3). The probability of predicting LR as quantified by Harrell's C-statistic was 0.68 for the Birmingham classification and 0.59 for MSTS. This process was repeated comparing the Birmingham classification of chemotherapy response and 2-mm cutoff margin with other models using chemotherapy response and different margin cutoffs (clear/contaminated, 1 mm, 3 mm) with the Birmingham classification having the highest Harrell's C-statistic (0.66 for chemotherapy response and clear/contaminated margins, 0.65 for 1 mm, and 0.61 for 3 mm). The Kaplan-Meier curve analysis for LRFS illustrates the spread achieved between groups for the Birmingham classification (log-rank test p < 0.0001) (Fig. 2). Patients classified as Birmingham 2b (margins ≤ 2 mm, < 90% chemotherapy necrosis) were 20× more likely to develop LR than those with margins > 2 mm and ≥ 90% chemotherapy necrosis
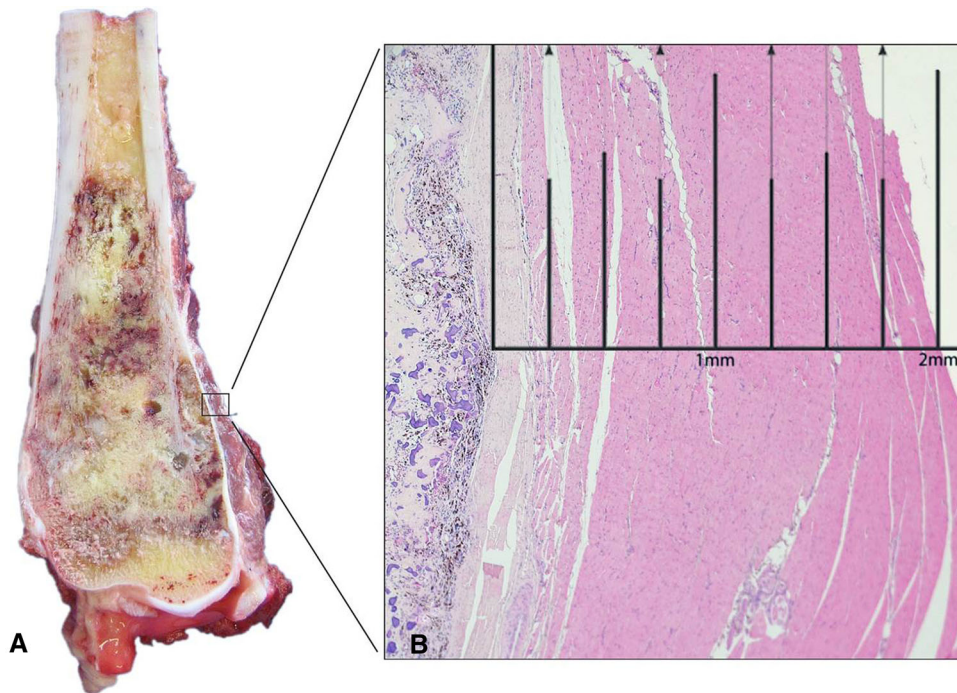
Fig. 1A–B (**A**) Cross-section of high-grade osteoblastic osteosarcoma of distal femur shows tumor extending through the cortex into the soft tissue. (**B**) On microscopy (hematoxylin and eosin; original magnification, ×1.25), the distance of viable tumor to the closest peripheral soft tissue resection margin measures 2 mm.

Table 2. Multivariate Cox proportional hazard analysis for local recurrence-free survival

Variable	Multivariate analysis		
	HR	95% CI	p value
Tumor location			
Extremity	1	0.7–4.4	0.19
Central	1.8		
Vascular invasion			
No	1	0.9–3.9	0.07
Yes	1.9		
MSTS surgical margins			
Radical	1	0.7–39	0.11
Wide	5.2	0.8–45	0.07
Marginal	6.2	1.2–82	0.03
Intralesional	9		
Chemotherapy necrosis			
≥ 90%	1	1.7–8.4	0.001
< 90%	3.8		

HR = hazard ratio; CI = confidence interval; MSTS = Musculoskeletal Tumor Society.

(Birmingham 2b HR, 19.6; 95% CI, 2.6-144; p = 0.003 versus Birmingham 1a HR, 1) (Table 3). No significant interaction between surgical margins and chemotherapy response was found on multivariate analyses. Chemotherapy response was a very strong predictor and should be part of any assessment for predicting local recurrence (not just margins). Adding necrosis to MSTS improves prediction, but the two-by-four stratification is cumbersome, spread is not well defined, and using Enneking definitions potentially affects reproducibility of margin status between centers. Using a clearly defined margin of 2 mm (or potentially clear/not clear in future studies with bigger numbers) in combination with chemotherapy response makes for an easy, reproducible classification, which can be communicated accurately between surgeons.

Comparison of Birmingham With MSTS for Overall Survival

The 5-year OS for the patient cohort was 67% (95% CI, 61%–71%). After controlling for size and tumor location, amputation (HR, 1.53; 95% CI, 1.0–2.2; p = 0.02 versus limb salvage HR, 1), vascular invasion (HR, 2.2; 95% CI, 1.4–3.3; p < 0.0001 versus no vascular invasion HR, 1), and < 90% chemotherapy necrosis (HR, 3.1; 95% CI, 2.0–4.8; p < 0.0001 versus > 90% chemotherapy necrosis HR, 1) were independent risk factors predicting OS (Table 4; Appendix 2 [Supplemental materials are available with the online version of CORR®.]). Margins stratified by MSTS criteria were not predictive of overall survival (log-rank test p = 0.14), whereas the Birmingham classification showed differences in survival between Groups 1 and 2
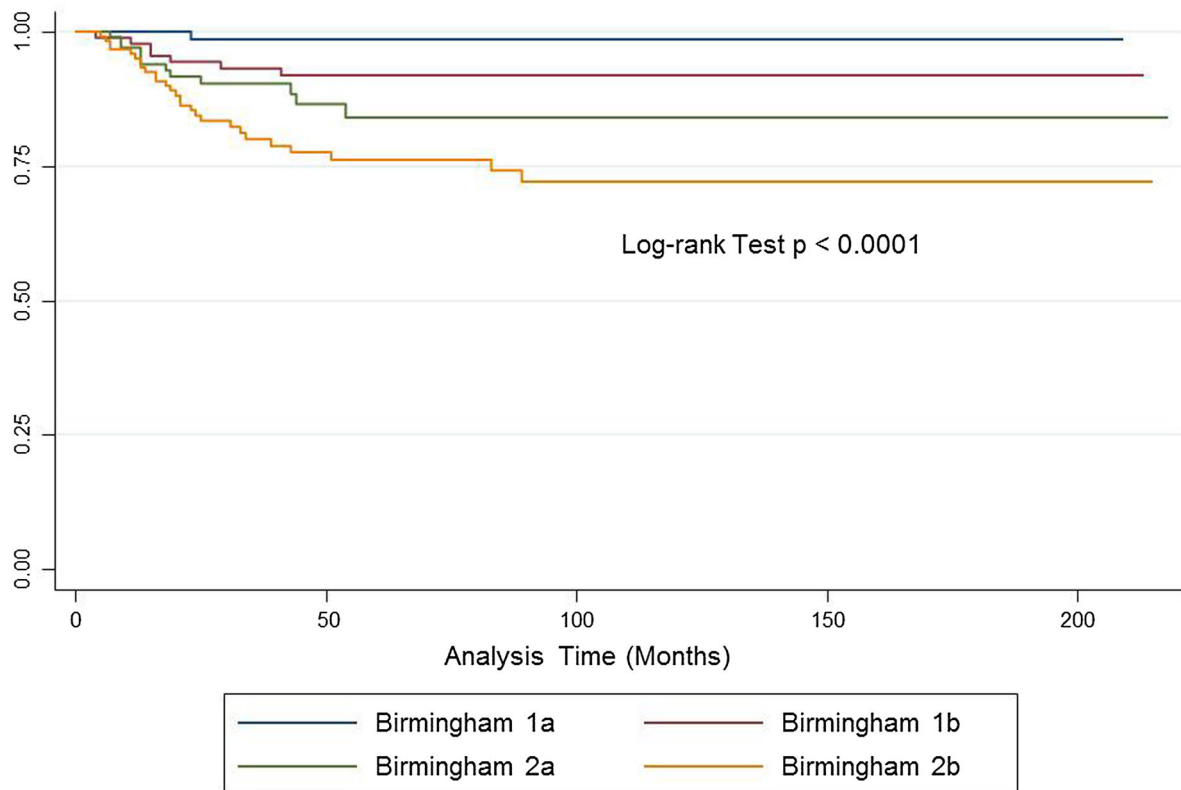
Fig. 2 Kaplan-Meier analysis shows local recurrence-free survival (LRFS) of patients with osteosarcoma using the Birmingham classification. Birmingham 1a ($\geq$ 90% chemotherapy necrosis, > 2-mm margins) 99% 5-year LRFS (95% CI, 91%–99.8%). Birmingham 1b ($\geq$ 90% chemotherapy necrosis, $\leq$ 2-mm margins) 92% 5-year LRFS (95% CI, 83%–96%). Birmingham 2a (< 90% chemotherapy necrosis, > 2-mm margins) 84% 5-year LRFS (95% CI, 73%–91%). Birmingham 2b (< 90% chemotherapy necrosis, $\leq$ 2-mm margins) 76% 5-year LRFS (95% CI, 77%–83%).

Table 3. The Birmingham classification system for margins and chemotherapy-induced necrosis and corresponding local recurrence-free survival rates

Birmingham classification	5-year LRFS	95% CI	HR	95% CI	p value
1a: Necrosis $\geq$ 90% and margins > 2 mm	98.6%	90.5%–99.8%	1		
1b: Necrosis $\geq$ 90% and margins $\leq$ 2 mm	91.7%	83.3%–95.9%	5.9	0.7–47	0.09
2a: Necrosis < 90% and margins > 2 mm	84.0%	72.5%–90.9%	11.0	1.4–84	0.02
2b: Necrosis < 90% and margins $\leq$ 2 mm	76.1%	72.5%–90.9%	19.6	2.6–144	0.003

LRFS = local recurrence-free survival; CI = confidence interval; HR = hazard ratio.

divided by chemotherapy (log-rank test p < 0.0001) but not between subdivisions by margin (Fig. 3). Adding necrosis to MSTS does improve prediction. Point of classification is that margins + response is significantly better than margins alone and that response should always be in the equation.

Discussion

Neoadjuvant chemotherapy response and surgical margins achieved after oncological resections are well-established prognostic factors for survival and LR in patients with

osteosarcoma [4, 6]. There is a clear relationship between inadequate margins and local disease recurrence, although what is defined as "marginal" and "wide" excision remains inherently subjective and may vary depending on the reporter [21]. Existing surgical staging systems fail to reflect a response to chemotherapy or define an appropriate safe metric distance from the tumor that will allow complete excision and closely predict the risk of disease recurrence. With this large, observational study, we confirmed that surgical margins and chemotherapy response were associated with LR. By incorporating surgical margins and response to chemotherapy in a new system, we

Table 4. Multivariate Cox proportional hazard analysis for overall survival

Variable	Multivariate analysis		
	HR	95% CI	p value
Tumor location			
Extremity	1	0.6–2.4	0.47
Central	1.25		
Type of operation			
Limb salvage	1	1.0–2.2	0.029
Amputation	1.53		
Size (cm)			
< 8 cm	1	0.8–1.7	0.35
> 8 cm	1.20		
Vascular invasion			
No	1	1.4–3.3	< 0.0001
Yes	2.17		
Chemotherapy necrosis			
≥ 90%	1	2.0–4.8	< 0.0001
< 90%	3.12		

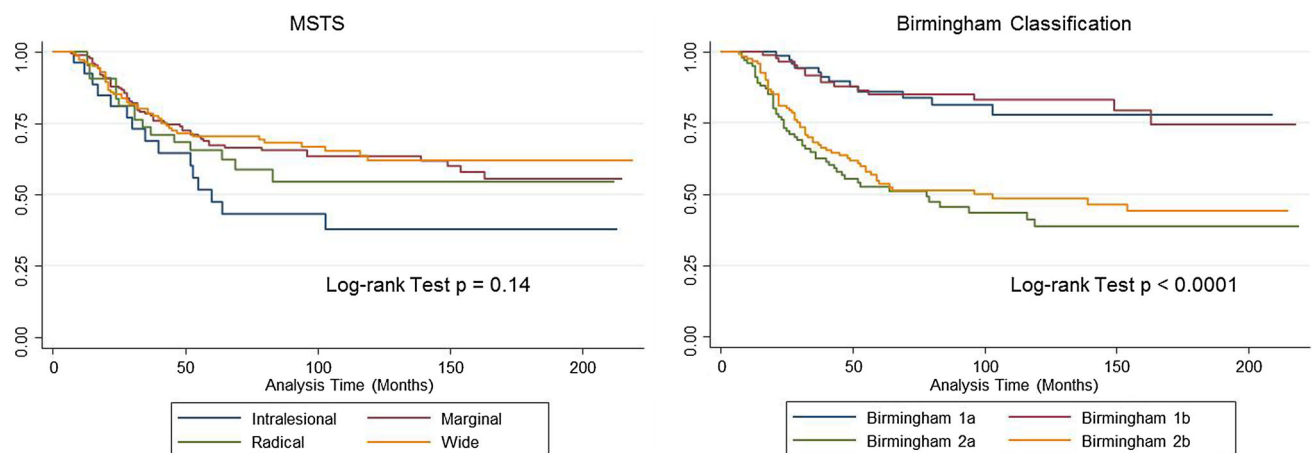HR = hazard ratio; CI = confidence interval.

Fig. 3 Kaplan-Meier analysis showing difference in overall survival (OS) between groups by the Birmingham classification compared with no differences in MSTS. Birmingham 1a (≥ 90% chemotherapy necrosis, > 2-mm margins) 86% 5-year OS (95% confidence interval [CI], 75%–92%). Birmingham 1b (≥ 90% chemotherapy necrosis, ≤ 2-mm margins) 85% 5-year OS (95% CI, 75%–91%). Birmingham 2a (< 90% chemotherapy necrosis, > 2-mm margins) 53% 5-year OS (95% CI, 42%–62%). Birmingham 2b (< 90% chemotherapy necrosis, ≤ 2-mm margins) 53% 5-year OS (95% CI, 44%–62%).

were able to predict LR and OS better than the commonly used MSTS margin criteria.

We acknowledge that there are several limitations to our study. Although this study had a relatively large number of patients, the chemotherapy regimens varied somewhat over the time of the study and both the length and intensity of preoperative chemotherapy will likely change the overall percentage of patients with good or standard responses as has been shown in other studies [17]. Also, the determination of percent necrosis is subjective and may vary even

between pathologists with extensive experience with bone sarcomas. Our cutoff of a 2-mm margin was not different from the other metric margins (Harrell's C statistics Birmingham 0.68, 1 mm 0.65, 3 mm 0.61), but we chose it because it was the best predictor of LR when taken along with chemotherapy response in our series and is an attainable target for sarcoma surgeons. With the numbers we had, there was no difference between 2-mm margins and clear/contaminated margins (Harrell's C statistic 0.66) in combination with chemotherapy, but with a larger

number of patients, this may become statistically significant. Bertrand et al. in a study of 51 patients with osteosarcomas used negative margins > 1 mm versus positive margins to predict LR [5].

The proportion of patients experiencing LR in the group of patients our study (12%) is comparable to previous studies [4, 23] and confirms that both response to chemotherapy and the surgical margin are predictive of LR (Tables 2, 4). Bacci et al. in a study of 789 nonmetastatic osteosarcoma found that inadequate surgical margins and poor response to chemotherapy were associated with local recurrence [4]. Andreou et al. attempted to define the width of surgical margin in the soft tissue periphery after resection of osteosarcomas but had significant missing data and could not correlate this with local recurrence [1].

The Birmingham classification only incorporates two predictors of LR. The authors acknowledge, and have shown through the multivariate analyses in this article, that only chemotherapy response and margins were associated with LR. An ideal classification should be simple, reproducible, and clinically relevant and the authors feel this system could be taken forward and, if adopted by a major organization such as the International Society of Limb Salvage (ISOLS), could lead for the first time to a universal, simple classification of defined surgical margins in osteosarcoma. This system has only been modeled on osteosarcoma margins and application to other tumor types (such as Ewing's sarcoma) requires further investigation, because it may not be applicable. Although the recording of the response to neoadjuvant chemotherapy is standard practice in the majority of centers treating patients with osteosarcomas of bone, the recording of surgical margins is variable. Currently, there is no consensus on how surgical margins for sarcomas are reported in different centers worldwide, although the MSTS staging remains popular with orthopaedic oncologists [16]. Our results, however, demonstrate that using both chemotherapy response and surgical margin attained in millimeters is more predictive for the development of LR when compared with the MSTS criteria. We propose that the adoption of such a simple system incorporating measurable and reproducible variables and replacing the subjectivity of "wide," "marginal," or "microscopic tumor at margin" will allow standardization of treatment and monitoring as well as improving communication about cases and aid future research. For this reason, we did not look at combining chemotherapy response with the four types of margins included in the MSTS classification to avoid using these subjective definitions of margins. Adding necrosis to the MSTS classification might have also increased the predictive value of that system.

The ability of functional imaging to predict response to chemotherapy in osteosarcomas becomes more vital with our findings, and several groups are investigating the role of PET-CT or functional MRI to preoperatively predict response to neoadjuvant chemotherapy, which may aid the preoperative planning of margins in light of this classification system [8, 14]. In particular, it may affect the decision as to the type of surgery planned if the margins are likely to be less than 2 mm and the chemotherapy response is poor because of the probable higher rate of LR after a closer margin. A more radical resection or amputation might be considered preferable in that circumstance. Whether this affects survival is still unclear because recent work has suggested that even amputation in this group of patients may not improve survival [20].

Because the classification system has been shown to predict survival and locally recurrent disease, it proves to be a useful surgical staging system; however, the authors are cognizant that the prime predictor of survival is chemotherapy response, as shown by the fact that the groups with > 90% response to chemotherapy had the best survival regardless of margin status (Fig. 3). Bielack et al. noted that incomplete surgical margins compared with macroscopically complete surgery were related to survival in an analysis of 1702 patients with osteosarcoma [6]. In their study, Bertrand et al. found that a positive margin compared with a negative margin > 1 mm was an independent predictor of local recurrence and overall survival [5]. However, as a result of low patient numbers, they were unable to show the effect of chemotherapy response on LR or indeed OS. Response to chemotherapy has otherwise been commonly associated with survival in several large osteosarcoma trials [4, 6, 23].

If our observations are confirmed by others in larger series, we propose that a combination of the recording of surgical margins in millimeters and the response to neoadjuvant chemotherapy should be the standard practice in oncology centers treating patients with osteosarcoma because we have shown that this combination can predict the risk of LR. Although intralesional versus tumor-free margins may indeed turn out to be discriminatory for LR in larger multicenter studies, the ideal of resecting the tumor with a clear margin of at least 2 mm of normal tissue in our series was a better predictor, and the authors strongly feel that a classification system that includes very close resection margins, ie, no tumor at the edge of the resection specimen, may lead to surgical errors and thus inadequate treatment, especially in those patients who are found to have a poor response to chemotherapy after surgery (25% risk of LR at 5 years). Conversely, the notion of not performing limb salvage if a margin of > 10 mm is not technically possible (a classic Enneking wide margin) has also been shown to be incorrect; a clear margin of 2 mm of normal tissue in a poor responding tumor has a 16% risk of LR at 5 years but without compromising survival. Further

studies are required to determine whether further resection, secondary amputation, or adjuvant radiotherapy is the optimal treatment for the group of patients at higher risk of LR (Birmingham 2b).

We believe the Birmingham classification may represent an improvement on the ability to predict LR and survival in patients with osteosarcoma treated with neoadjuvant chemotherapy. If our observations are confirmed by others perhaps in a large prospective multicenter validation study undertaken by the ISOLS, it could offer useful prognostic information for treating oncologists and be helpful in advising patients.

References

1. Andreou D, Bielack SS, Carrle D, Kevric M, Kotz R, Winkelmann W, Jundt G, Werner M, Fehlberg S, Kager L, Kühne T, Lang S, Dominkus M, Exner GU, Hardes J, Hillmann A, Ewerbeck V, Heise U, Reichardt P, Tunn PU. The influence of tumor- and treatment-related factors on the development of local recurrence in osteosarcoma after adequate surgery. An analysis of 1355 patients treated on neoadjuvant Cooperative Osteosarcoma Study Group protocols. *Ann Oncol.* 2011;22:1228–1235.

2. Bacci G, Ferrari S, Longhi A, Perin S, Forni C, Fabbri N, Salduca N, Versari M, Smith KV. Pattern of relapse in patients with osteosarcoma of the extremities treated with neoadjuvant chemotherapy. *Eur J Cancer.* 2001;37:32–38.

3. Bacci G, Forni C, Longhi A, Ferrari S, Mercuri M, Bertoni F, Serra M, Briccoli A, Balladelli A, Picci P. Local recurrence and local control of non-metastatic osteosarcoma of the extremities: a 27-year experience in a single institution. *J Surg Oncol.* 2007;96:118–123.

4. Bacci G, Longhi A, Versari M, Mercuri M, Briccoli A, Picci P. Prognostic factors for osteosarcoma of the extremity treated with neoadjuvant chemotherapy: 15-year experience in 789 patients treated at a single institution. *Cancer.* 2006;106:1154–1161.

5. Bertrand TE, Cruz A, Binitie O, Cheong D, Letson GD. Do surgical margins affect local recurrence and survival in extremity, nonmetastatic, high-grade osteosarcoma? *Clin Orthop Relat Res.* 2016;474:677–683.

6. Bielack SS, Kempf-Bielack B, Delling G, Exner GU, Flege S, Helmke K, Kotz R, Salzer-Kuntschik M, Werner M, Winkelmann W, Zoubek A, Jürgens H, Winkler K. Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. *J Clin Oncol.* 2002;20:776–790.

7. Bispo Júnior RZ, de Camargo OP. Prognostic factors in the survival of patients diagnosed with primary non-metastatic osteosarcoma with a poor response to neoadjuvant chemotherapy. *Clinics (Sao Paulo).* 2009;64:1177–1186.

8. Byun BH1, Kong CB, Lim I, Choi CW, Song WS, Cho WH, Jeon DG, Koh JS, Lee SY, Lim SM. Combination of 18F-FDG PET/CT and diffusion-weighted MR imaging as a predictor of histologic response to neoadjuvant chemotherapy: preliminary results in osteosarcoma. *J Nucl Med.* 2013;54:1053–1059.

9. Cotterill SJ, Wright CM, Pearce MS, Craft AW. Stature of young people with malignant bone tumors. *Pediatr Blood Cancer.* 2004;42:59–63.

10. Dahlin DC, Coventry MB. Osteogenic sarcoma. A study of six hundred cases. *J Bone Joint Surg Am.* 1967;49:101–110.

11. Enneking WF, Spanier SS, Goodman MA. A system for the surgical staging of musculoskeletal sarcoma. *Clin Orthop Relat Res.* 1980;153:106–120.

12. Ferrari S, Bacci G, Picci P, Mercuri M, Briccoli A, Pinto D, Gasbarrini A, Tienghi A, Brach del Prever A. Long-term follow-up and post-relapse survival in patients with non-metastatic osteosarcoma of the extremity treated with neoadjuvant chemotherapy. *Ann Oncol.* 1997;8:765–771.

13. Fletcher C, Unni K, Mertens F. *World Health Organisation Classification of Tumours. Pathology and Genetics of Tumours of Soft Tissue and Bone.* Lyon, France: IARC Press; 2002:262–288.

14. Gaston LL, Di Bella C, Slavin J, Hicks RJ, Choong PF. 18F-FDG PET response to neoadjuvant chemotherapy for Ewing sarcoma and osteosarcoma are different. *Skeletal Radiol.* 2011;40:1007–1015.

15. Grimer RJ, Taminiau AM, Cannon SR. Surgical outcomes in osteosarcoma. *J Bone Joint Surg Br.* 2002;84:395–400.

16. Hoang K, Gao Y, Miller BJ. The variability in surgical margin reporting in limb salvage surgery for sarcoma. *Iowa Orthop J.* 2015;35:181–186.

17. Lewis IJ, Nooij MA, Whelan J, Sydes MR, Grimer R, Hogendoorn PC, Memon MA, Weeden S, Uscinska BM, van Glabbeke M, Kirkpatrick A, Hauben EI, Craft AW, Taminiau AH; MRC BO06 and EORTC 80931 collaborators; European Osteosarcoma Intergroup. Improvement in histologic response but not survival in osteosarcoma patients treated with intensified chemotherapy: a randomized phase III trial of the European Osteosarcoma Intergroup. *J Natl Cancer Inst.* 2007;99:112–128.

18. Mangham DC, Athanasou NA. Guidelines for histopathological specimen examination and diagnostic reporting of primary bone tumours. *Clin Sarcoma Res.* 2011;1:6.

19. Provisor AJ, Ettinger LJ, Nachman JB, Krailo MD, Makley JT, Yunis EJ, Huvos AG, Betcher DL, Baum ES, Kisker CT, Miser JS. Treatment of nonmetastatic osteosarcoma of the extremity with preoperative and postoperative chemotherapy: a report from the Children's Cancer Group. *J Clin Oncol.* 1997;15:76–84.

20. Reddy KI, Wafa H, Gaston CL, Grimer RJ, Abudu AT, Jeys LM, Carter SR, Tillman RM. Does amputation offer any survival benefit over limb salvage in osteosarcoma patients with poor chemonecrosis and close margins? *Bone Joint J.* 2015;97:115–120.

21. Trovik CS, Skjeldal S, Bauer H, Rydholm A, Jebsen N. Reliability of margin assessment after surgery for extremity soft tissue sarcoma: the SSG experience. *Sarcoma.* 2012;290698.

22. Whelan JS, Bielack SS, Marina N, Smeland S, Jovic G, Hook JM, Krailo M, Anninga J, Butterfass-Bahloul T, Böhling T, Calaminus G, Capra M, Deffenbaugh C, Dhooge C, Eriksson M, Flanagan AM, Gelderblom H, Goorin A, Gorlick R, Gosheger G, Grimer RJ, Hall KS, Helmke K, Hogendoorn PC, Jundt G, Kager L, Kuehne T, Lau CC, Letson GD, Meyer J, Meyers PA, Morris C, Mottl H, Nadel H, Nagarajan R, Randall RL, Schomberg P, Schwarz R, Teot LA, Sydes MR, Bernstein M; EURAMOS collaborators. EURAMOS-1, an international randomised study for osteosarcoma: results from pre-randomisation treatment. *Ann Oncol.* 2015;26:407–414.

23. Whelan JS, Jinks RC, McTiernan A, Sydes MR, Hook JM, Trani L, Uscinska B, Bramwell V, Lewis IJ, Nooij MA, van Glabbeke M, Grimer RJ, Hogendoorn PC, Taminiau AH, Gelderblom H. Survival from high-grade localised extremity osteosarcoma: combined results and prognostic factors from three European Osteosarcoma Intergroup randomised controlled trials. *Ann Oncol.* 2012;23:1607–1616.