

Reliability and Validity of the Musculoskeletal Tumor Society Scoring System for the Upper Extremity in Japanese Patients

Kosuke Uehara MD, Koichi Ogura MD, Toru Akiyama MD,
Yusuke Shinoda MD, Shintaro Iwata MD, Eisuke Kobayashi MD,
Yoshikazu Tanzawa MD, Tsukasa Yonemoto MD, Hirotaka Kawano MD,
Akira Kawai MD

Received: 8 December 2016 / Accepted: 19 May 2017 / Published online: 30 May 2017
© The Association of Bone and Joint Surgeons® 2017

Abstract

Background The Musculoskeletal Tumor Society (MSTS) scoring system developed in 1993 is a widely used disease-specific evaluation tool for assessment of physical function in patients with musculoskeletal tumors; however, only a few studies have confirmed its reliability and validity.

The institution of one or more of the authors (KO, SI, TY, AK) has received, during the study period, funding from the National Cancer Center Research and Development Fund (Grant 26-A-4). All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request. *Clinical Orthopaedics and Related Research*® neither advocates nor endorses the use of any treatment, drug, or device. Readers are encouraged to always seek additional information, including FDA-approval status, of any drug or device prior to clinical use. Each author certifies that his institution approved or waived approval for the human protocol for this investigation and that all investigations were conducted in conformity with ethical principles of research.

This work was performed at The University of Tokyo Hospital, Tokyo, Japan.

K. Uehara, K. Ogura (✉), Y. Shinoda, H. Kawano
Department of Orthopaedic Surgery, The University of Tokyo
Hospital, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8655, Japan
e-mail: ogura-ky@umin.ac.jp

K. Ogura, E. Kobayashi, Y. Tanzawa, A. Kawai
Department of Musculoskeletal Oncology, National Cancer
Center Hospital, Tokyo, Japan

T. Akiyama
Department of Orthopaedic Surgery, Saitama Medical Center,
Jichi Medical University, Saitama, Japan

S. Iwata, T. Yonemoto
Division of Orthopedic Surgery, Chiba Cancer Center, Chiba,
Japan

Questions/purposes The aim of this study was to validate the MSTS scoring system for the upper extremity (MSTS-UE) in Japanese patients with musculoskeletal tumors for use by others in research. Does the MSTS-UE have: (1) sufficient reliability and internal consistency; (2) adequate construct validity; and (3) reasonable criterion validity in comparison to the Toronto Extremity Salvage Score (TESS) or SF-36?

Methods Reliability was performed using test-retest analysis, and internal consistency was evaluated with Cronbach's alpha coefficient. Construct validity was evaluated using a scree plot to confirm the construct number and the Akaike information criterion network. Criterion validity was evaluated by comparing the MSTS-UE with the TESS and SF-36.

Results The test-retest reliability with intraclass correlation coefficient (0.95; 95% CI, 0.91–0.97) was excellent, and internal consistency with Cronbach's α (0.7; 95% CI, 0.53–0.81) was acceptable. There were no ceiling and floor effects. The Akaike Information Criterion network showed that lifting ability, pain, and dexterity played central roles among the components. The MSTS-UE showed substantial correlation with the TESS scoring scale ($r = 0.75$; $p < 0.001$) and fair correlation with the SF-36 physical component summary ($r = 0.37$; $p = 0.007$). Although the MSTS-UE showed slight correlation with the SF-36 mental component summary, the emotional acceptance component of the MSTS-UE showed fair correlation ($r = 0.29$; $p = 0.039$).

Conclusions We can conclude that the MSTS is not an adequate measure of general health-related quality of life; however, this system was designed mainly to be a simple measure of function in a single extremity. To evaluate the mental state of patients with musculoskeletal tumors in the upper extremity, further study is needed.

Introduction

The humerus is the third-most-common site for primary bone tumors such as osteosarcoma or Ewing's sarcoma [25], and soft tissue sarcomas such as liposarcoma or angiosarcoma also arise in the upper extremity, especially the upper arm. Wide resection is the mainstay of local treatment for primary malignant bone and soft tissue tumors, but this often results in various degrees of functional impairment. Surgical procedures for the upper extremity should take into account different factors than do procedures for the lower extremity; for example, in the upper extremity, retention of precise movement may be more important than massive muscle power, which is more important for lower extremity tumor resections. To evaluate functional outcome after surgery for musculoskeletal tumors of the upper extremity, some means of assessing manual dexterity and elbow and shoulder function for reaching or lifting activity are required.

The Musculoskeletal Tumor Society (MSTS) scoring system developed in 1993, which is completed by a member of the treatment team (rather than the patient, and so is not considered a patient-reported outcomes tool), was designed to measure functional outcome and quality of life after treatment for musculoskeletal tumors. It was developed in 1985 [7] and revised in 1993 [8]. Since then, this system has been used in many studies, and it has become a commonly used functional assessment tool [13]. Although the MSTS scoring system for the lower extremity has been translated, culturally adapted, and validated for use in Japanese patients [12], to our knowledge, the MSTS scoring system for the upper extremity (MSTS-UE) has not been validated for Japanese patients. The rationale for our study was to validate the MSTS-UE for use by others in research. However, to our knowledge, only one study has confirmed the reliability and validity [26]. In addition, several questionnaire tools such as the Toronto Extremity Salvage Score (TESS) [6] and SF-36 allow patients to self-rate their physical function or health-related quality of life [4]. Unlike the MSTS, the TESS is a patient-completed evaluation system developed for individuals who have undergone limb-preservation surgery for tumors of the extremities. The SF-36 score has been validated in patients with musculoskeletal disorders and is widely used for measuring health outcomes. However, it is a generic questionnaire and has the potential disadvantage of being less sensitive to clinical change in patients with disorders specific to an anatomic region or disease [17].

The aim of our study was to perform a validation analysis of the MSTS-UE in Japanese patients with musculoskeletal tumors of the upper extremity from the viewpoint of psychometric characteristics; specifically, we

sought to evaluate whether the MSTS-UE has (1) sufficient reliability and internal consistency; (2) adequate construct validity; and (3) reasonable criterion validity in comparison to the TESS or SF-36.

Patients and Methods

This study was cross-sectional in design, and approval was obtained from the institutional review boards of the participating institutions. Patients who met the following criteria were included: (1) those with a diagnosis of intermediate or malignant bone or soft tissue tumors of the upper extremity or shoulder girdle based on the 2013 WHO classification [9]; (2) age between 12 and 85 years; (3) having a minimum interval of 6 months after definitive surgery; and (4) without local recurrence or distant metastasis after surgery. Patients were recruited between August and December 2014, and 53 agreed to participate. The clinical and demographic characteristics of the participants are detailed (Table 1). In this study, six patients with atypical lipomatous tumors and four with giant cell tumors of bone, for which relatively less-invasive surgery usually is done, were included.

The MSTS-UE is based on analysis of factors pertinent to the patient as a whole and those specific to the affected upper limb [7]. It contains six categories: pain, function, emotional acceptance, hand positioning, manual dexterity, and lifting ability. Each of these categories is assigned a value of 0 to 5 points, and the total summed score is divided by the maximum possible score (30 points) and then multiplied by 100 to obtain the final score.

The Japanese version of the MSTS scoring system was approved by the Japanese Orthopaedic Association Musculoskeletal Tumor Committee [26]. In brief, to develop this version, a translation to Japanese was prepared along with a crosscultural adaptation of the MSTS scoring system using approaches devised by others [3, 11]. The English version of the MSTS scoring system was translated separately by five native Japanese musculoskeletal oncology surgeons bilingual in Japanese and English. Because each sentence was relatively simple, no professional medical translator was used. Subsequently, all independent translations were compared and combined in one document. Although backtranslation was not performed, the final version was approved by all translators. All of the translators reached a consensus that no modification was necessary from the viewpoint of crosscultural adaptation, because each item description in the original version fit well with the modern Japanese lifestyle and appeared appropriate. Since its release, the Japanese version of the MSTS has been used commonly in clinical settings for Japanese studies [10, 14, 15].

Table 1. Descriptive characteristics of the study population

Characteristic	Number of patients (%)
Overall	53 (100)
Age (years); mean [SD]	53 (range, 16–82) [16.5]
Gender	
Male	30 (57)
Female	23 (43)
Time from surgery (months); mean [SD]	56 (range, 6–288) [55.5]
Tumor location	
Soft tissue	37 (70)
Shoulder	15
Upper arm	8
Forearm	9
Hand	5
Bone	16 (30)
Scapula	4
Humerus	7
Radius	3
Other	2
Histologic diagnosis	
Soft tissue	
Myxoid liposarcoma	10
Well-differentiated liposarcoma	6
Undifferentiated pleomorphic sarcoma	4
Synovial sarcoma	4
Epithelioid sarcoma	3
Dermatofibrosarcoma protuberans	3
Other	7
Bone	
Osteosarcoma	6
Chondrosarcoma	5
Giant cell tumor of bone	4
Undifferentiated high-grade pleomorphic sarcoma of bone	1
Type of surgery	
Amputation	6 (11)
Resection only	16 (30)
Resection + reconstruction without prosthesis	24 (45)
Resection + reconstruction with prosthesis	4 (8)
Curettage	3 (6)

Psychometric Characteristics of the MSTS-UE

To validate the MSTS-UE, psychometric analysis of 53 patients was conducted. Reliability was evaluated by test-retest analysis. Second testing using the MSTS-UE was done for the same patients at 2 to 5 weeks after the first test,

with confirmation from patients, based on their response on the MSTS-UE questionnaire, that their condition did not change. The reproducibility (test-retest reliability) of the MSTS-UE was assessed by calculating the intraclass correlation coefficient (ICC) using a two-way random effects model, absolute agreement between the responses in the first and second tests for each category, and the total score. In addition, floor and ceiling effects were analyzed for each category and the total score. Such effects were considered to be present if greater than 15% of the respondents achieved the lowest (floor effect) or highest (ceiling effect) point scores [22]. The scale we used for the ICC was: 0.00–0.20, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; and 0.81–1.00, nearly perfect.

Internal consistency was established by calculating Cronbach's α coefficient, which reflects the strength of relationships among the six categories in the MSTS-UE. The scale for Cronbach's α was: excellent $0.9 \leq \alpha$; good $0.8 \leq \alpha < 0.9$; and acceptable $0.7 \leq \alpha < 0.8$. Construct validity also was evaluated to examine whether the MSTS-UE does indeed measure what it seeks to measure. A scree plot was analyzed to determine the best number of constructs. The degrees of correlation among the items in the MSTS-UE were evaluated using the Akaike Information Criterion (AIC) network to examine the latent structure of the MSTS-UE construct validity; this is a graphic modeling method that assesses relationships among items [1, 2, 5, 12, 18, 19]. The Categorical Data Analysis Program (The Japanese Institute of Statistical Mathematics, Tachikawa, Japan) was used to conduct crosstable analyses involving all combinations of questionnaire items and searched for the best subset and categorization of explanatory items simultaneously and then matching combinations were automatically indicated using the AIC [20]. To validate the construct validity, the AIC network can provide a robust result because it can be obtained not only from a linear data set, but also a nonlinear one. This model also can clarify the item relationship visually. The MSTS-UE has just six items and also six domains; one item seems almost to represent one domain. In such a situation, the questionnaire often has just a single latent structure if each item is not similar. Thus, we assumed the MSTS-UE had just a single latent structure when performing the AIC network analysis.

Criterion validity, which evaluates how well one measure predicts an outcome for another measure, was evaluated by comparing the MSTS-UE with the TESS and SF-36, which were validated previously for sufficient reliability and reasonable validity in the Japanese population [2, 19]. Correlations among these measures were assessed using Spearman's correlation coefficients normal distribution.

Statistical Analysis

All statistical analyses were performed using SPSS Version 18.0 (SPSS Inc, Chicago, IL, USA). The scores were reported as mean values \pm SD. The threshold for significance was a probability less than 0.05. Power analysis was performed in advance. There were six items for the MSTS-UE, assuming a coefficient α of 0.90 and CI of 0.1, therefore approximately 50 patients were enough for this study [21].

The total MSTS-UE scores for the 53 patients ranged from 50% to 100%, and the mean score was 85% (SD, 12.9). There were three missing data items for hand positioning, manual dexterity, and lifting ability for three patients.

Results

Reliability and Floor and Ceiling Effects

The ICC between the test and retest total scores of the MSTS-UE was 0.95 (95% CI, 0.91–0.97), confirming the high reproducibility of the MSTS scoring system (Table 2). Lifting ability showed the highest ICC (0.93; 95% CI, 0.87–0.96) and dexterity the lowest ICC (0.74; 95% CI, 0.58–0.85). Patients with intermediate bone or soft tissue tumors such as atypical lipomatous tumors or giant cell tumors of bone usually need less-invasive surgery. In this study, six patients with atypical lipomatous tumors were included. However, only two of these patients had the maximum score. None of the four patients with giant cell tumors of bone had the maximum score.

No patients were assigned the lowest possible total score of 0, indicating that there were no floor effects in this small survey population, and similarly, five patients were assigned the highest possible total score of 100 (9.6%), indicating that there were no ceiling effects in this small survey population.

Table 2. Reliability of the total MSTS-UE score and of each component

Reference score	Mean	SD	ICC	95% CI	p Value
Total (%)	85	12.9	0.95*	0.88–0.96	< 0.001
Pain	5	0.6	0.88*	0.79–0.93	< 0.001
Function	4	1.2	0.89*	0.81–0.93	< 0.001
Emotional acceptance	4	1.4	0.83*	0.72–0.90	< 0.001
Hand positioning	4	1.0	0.87*	0.78–0.93	< 0.001
Manual dexterity	5	0.5	0.74*	0.58–0.85	< 0.001
Lifting ability	4	1.2	0.93*	0.87–0.96	< 0.001

* $p < 0.001$; MSTS-UE = Musculoskeletal Tumor Society upper extremity; ICC = intraclass correlation coefficient.

Internal Consistency

The overall Cronbach's α coefficient was 0.7, suggesting an acceptable level of internal consistency.

Construct Validity

The scree plot showed the MSTS-UE had a single construct. The AIC calculation of the MSTS scoring system yielded 35 ($= {}_6C_2 + {}_6C_3$) minimal distance assortments, that is, degree of independence, for the two- or three-item groupings. Based on the spatial association of the calculation for each item (AIC network), we were able to show that "pain," and "dexterity," were related to three, and "lifting ability" was related to four other items, indicating these three items had a central role among the six factors of the system (Fig. 1).

Criterion Validity

The criterion validity was evaluated by correlating the total MSTS-UE score with the total TESS score and each component of the SF-36 (Table 3). The total MSTS-UE score significantly correlated with the total TESS score ($r = 0.75$; $p < 0.001$), but correlation was only fair with the SF-36 physical component summary ($r = 0.37$; $p = 0.007$); it showed slight correlation with the mental component summary. The MSTS-UE emotional acceptance score showed fair correlation with the SF-36 mental component summary ($r = 0.29$; $p = 0.039$).

Discussion

The MSTS scoring system is a commonly used functional evaluation system for patients who undergo surgery for musculoskeletal tumors [13]. We attempted to clarify the validity of the MSTS-UE using psychometric analysis. We found that the MSTS-UE has sufficient reliability with a

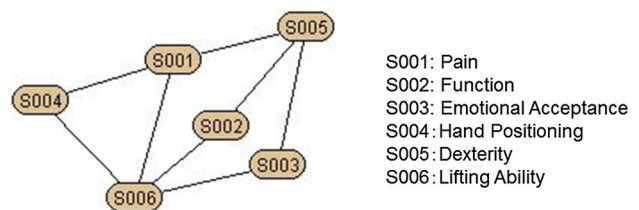


Fig. 1 The AIC network of the MSTS scoring system showed that "pain," and "dexterity," were related to three, and "lifting ability" was related to four other items, indicating these three items had a central role among the six factors of the system.

Table 3. Criterion validity: correlation between total MSTS-UE score and TESS or SF-36 components

Reference score	Mean	SD	Spearman's correlation coefficients	95% CI	p Value
TESS	87	16.5	0.75 [‡]	0.60–0.85	< 0.001
SF-36 (physical component summary)	45	16.5	0.37 [†]	0.11–0.58	0.007
SF-36 (mental component summary)	53	9.1	0.05	–0.22 to 0.32	0.719
SF-36 (physical functioning)	49	19.2	0.60 [‡]	0.39–0.75	< 0.001
SF-36 (role physical)	40	13.4	0.47 [‡]	0.23–0.66	< 0.001
SF-36 (bodily pain)	50	9.1	0.04	–0.24 to 0.30	0.796
SF-36 (general health)	51	9.4	0.32 [*]	0.51–0.54	0.022
SF-36 (vitality)	51	8.9	< 0.01	–0.27 to 0.27	0.981
SF-36 (social functioning)	48	12.2	0.34 [*]	0.71–0.56	0.016
SF-36 (role emotional)	47	12.4	0.13	–0.15 to 0.39	0.362
SF-36 (mental health)	52	9	0.24	–0.03 to 0.48	0.089

* $0.05 > p \geq 0.01$; [†] $0.01 > p \geq 0.001$; [‡] $p < 0.001$; MSTS-UE = Musculoskeletal Tumor Society – Upper Extremity; TESS = Toronto Extremity Salvage Score.

high ICC by test-retest analysis, acceptable internal consistency with a moderate Cronbach's α coefficient, adequate construct validity indicated by the AIC network, and reasonable criterion validity in comparison to the TESS or SF-36.

Our study was limited in that it was performed with Japanese patients using a Japanese version of the MSTS-UE developed by the Japanese Orthopaedic Association committee on tumors [24] on the basis of proposed guidelines for crosscultural adaptation [3, 11]. Intercultural differences therefore might have affected the outcomes, and it may not apply to other patient populations. The Japanese versions of MSTS scoring system and TESS are available for any orthopaedic doctors who want to use them and are available from the Japanese Orthopaedic Association website (https://www.joa.or.jp/member/committee/diagnosis/pdf/tess_ue.pdf) free of charge [23]. The second limitation was the heterogeneity of the tumors in the studied patients, who had intermediate bone or soft tissue tumors, such as atypical lipomatous tumor or giant cell tumor of bone, for which relatively less-invasive surgery usually is performed. This variance could have affected the internal consistency (that indicates the correlations between different items on the same test) of the results. The third limitation was the small sample size; the results of the construct validity analysis in particular may not replicate in a larger representative sample nor for other language versions of the MSTS rating system.

From our results, we conclude that the MSTS-UE has sufficient reliability and reasonable validity for clinical use. Although some studies have tried to confirm the validity of the MSTS-UE, our study is the first of which we are aware to verify the reliability of the MSTS-UE using

psychometric analysis. In 2016, Iwata et al. [12] reported that the MSTS scoring system for the lower extremity had sufficient reliability and internal consistency, adequate construct validity, and reasonable criterion validity. Wada et al. [26] reported reasonable criterion validity of the MSTS-UE with the TESS and SF-36, but they did not verify its reliability, internal consistency, or construct validity. Lee et al. [16] validated the MSTS for the upper and lower extremities, but their study included only eight patients with upper extremity tumors. Obviously their sample size was insufficient to declare reasonable reliability and validity for the MSTS-UE independently [16]. Our results showed satisfactory reliability with superior ICC and no ceiling or floor effect.

Psychometric analysis confirmed adequate validity of the items included in the MSTS-UE. To evaluate the validity of a measuring tool such as a questionnaire or functional analysis, a psychometric approach usually is used. In the current study, construct and criterion validity analyses were performed to reveal the latent structure of the MSTS scoring system using the AIC network and comparison with the SF-36 and TESS. The AIC network is one of the graph theory models and codes between nodes mean not only those two items have a strong relationship, but also the strength of the relationship (the shorter this cord is, the stronger the relationship those items have). The result of the AIC network showed that “pain,” and “dexterity,” were related to three, and “lifting ability” was related to four other items, indicating these three items had a central role among the six factors of the system. Criterion validity showed a substantial correlation between the total MSTS-UE score and the TESS and a slight correlation with the SF-36 physical component summary.

However, there was slight correlation between the total MSTS-UE score and the mental component summary of the SF-36. The SF-36 was designed and is widely used as a general health and health-related quality of life assessment tool, therefore this finding is not surprising. However, the correlation between the emotional acceptance component of the MSTS and the mental component summary of the SF-36 had a fair coefficient value. The correlation coefficients of 0.60, 0.47, and 0.34 for the MSTS-UE with the physical functioning, role physical, and social functioning scores of the SF-36 were consistent with those in a previous validation study of criterion validity (0.45, 0.60, and 0.43, respectively) [17]. The SF-36 score has been validated in patients with musculoskeletal disorders and is widely used for measuring health outcomes. However, it is a generic questionnaire and has the potential disadvantage of being less sensitive to clinical change in patients with disorders specific to an anatomic region or disease process [26].

Our study showed that the MSTS-UE is a reliable and valid instrument for assessment of physical function in patients with upper extremity sarcoma, although the correlation with mental outcome is not strong. We can conclude that the MSTS-UE is not an adequate measure of general health-related quality of life, however, this system was designed mainly to be a simple measure of function in a single extremity. To evaluate mental state of the patients with musculoskeletal tumors in the upper extremity, further study is needed.

Acknowledgments We thank Kazuo Saita MD (Department of Orthopaedic Surgery, Saitama Medical Center, Jichi Medical University, Saitama, Japan), Hirokazu Chuman MD (Department of Musculoskeletal Oncology, National Cancer Center Hospital, Tokyo, Japan), and Hirotaka Kawano MD (Department of Orthopaedic Surgery, The University of Tokyo Hospital, Tokyo, Japan) for clinical assistance; and Misuzu Mori, Yoko Kato (Department of Musculoskeletal Oncology, National Cancer Center Hospital, Tokyo, Japan), and Keiko Hishiki (Department of Orthopaedic Surgery, Saitama Medical Center, Jichi Medical University, Saitama, Japan) for administrative support with data collection.

References

1. Akaike H. Information Theory and an Extension of the Maximum Likelihood Principle. In Petrov BN, Csaki F, eds. *2nd International Symposium on Information Theory*. Budapest, Hungary: Akademia Kiado; 1973.
2. Akiyama T, Uehara K, Ogura K, Shinoda Y, Iwata S, Saita K, Tanzawa Y, Nakatani F, Yonemoto T, Kawano H, Davis AM, Kawai A. Cross-cultural adaptation and validation of the Japanese version of the Toronto Extremity Salvage Score (TESS) for patients with malignant musculoskeletal tumors in the upper extremities. *J Orthop Sci*. 2017;22:127–132.
3. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)*. 2000;25:3186–3191.
4. Brazier JE, Harper R, Jones NM, O’Cathain A, Thomas KJ, Usherwood T, Westlake L. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ*. 1992;305:160–164.
5. Carta MG, Massidda D, Moro MF, Aguglia E, Balestrieri M, Caraci F, Dell’Osso L, Di Sciascio G, Drago F, Faravelli C, Hardoy MC, Calo S, Pollice R, Fortezzo A, Akiskal H. Comparing factor structure of the Mood Disorder Questionnaire (MDQ): in Italy sexual behavior is euphoric but in Asia mysterious and forbidden. *J Affect Disord*. 2014;155:96–103.
6. Davis AM, Wright JG, Williams JI, Bombardier C, Griffin A, Bell RS. Development of a measure of physical function for patients with bone and soft tissue sarcoma. *Qual Life Res*. 1996;5:508–516.
7. Enneking W. *Modification of the System for Functional Evaluation of Surgical Management of Musculoskeletal Tumors*. New York, NY: Churchill-Livingstone; 1987.
8. Enneking WF, Dunham W, Gebhardt MC, Malawar M, Pritchard DJ. A system for the functional evaluation of reconstructive procedures after surgical treatment of tumors of the musculoskeletal system. *Clin Orthop Relat Res*. 1993;286:241–246.
9. Fletcher CD, Bridge JA, Hogendoorn P, Mertens F, eds. *World Health Organization Classification of Tumours of Soft Tissue and Bone*. 4th ed. Lyon, France; 2013.
10. Fujiwara T, Ogura K, Kobayashi E, Tanzawa Y, Nakatani F, Chuman H, Kawai A. Clinical outcomes of surgical treatments for primary malignant bone tumors arising in the acetabulum. *Sarcoma*. 2015;2015:430576.
11. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993;46:1417–1432.
12. Iwata S, Uehara K, Ogura K, Akiyama T, Shinoda Y, Yonemoto T, Kawai A. Reliability and validity of a Japanese-language and culturally adapted version of the Musculoskeletal Tumor Society Scoring System for the lower extremity. *Clin Orthop Relat Res*. 2016;474:2044–2052.
13. Janssen SJ, van Rein EA, Paulino Pereira NR, Raskin KA, Ferrone ML, Hornicek FJ, Lozano-Calderon SA, Schwab JH. The discrepancy between patient and clinician reported function in extremity Bbne metastases. *Sarcoma*. 2016;2016:1014248.
14. Kawashima H, Ogose A, Hotta T, Ariizumi T, Yamagishi T, Endo N. Extensor reconstruction of the knee using the fibular transposition technique after proximal tibial resection. *Knee*. 2016 Dec 1. [Epub ahead of print]
15. Kito M, Yoshimura Y, Isobe K, Aoki K, Momose T, Suzuki S, Tanaka A, Sano K, Akahane T, Kato H. Clinical outcome of deep-seated atypical lipomatous tumor of the extremities with median-term follow-up study. *Eur J Surg Oncol*. 2015;41:400–406.
16. Lee SH, Kim DJ, Oh JH, Han HS, Yoo KH, Kim HS. Validation of a functional evaluation system in patients with musculoskeletal tumors. *Clin Orthop Relat Res*. 2003;411:217–226.
17. Lingard EA, Katz JN, Wright RJ, Wright EA, Sledge CB, Kinemax Outcomes Group. Validity and responsiveness of the Knee Society Clinical Rating System in comparison with the SF-36 and WOMAC. *J Bone Joint Surg Am*. 2001;83:1856–1864.
18. Nagao M, Doi T, Saita Y, Kobayashi Y, Kubota M, Kaneko H, Takazawa Y, Ishijima M, Kurosawa H, Kaneko K, Nozawa M, Ikeda H, Kim SG. A novel patient-reported outcome measure for anterior cruciate ligament injury: evaluating the reliability, validity, and responsiveness of Japanese anterior cruciate ligament questionnaire 25. *Knee Surg Sports Traumatol Arthrosc*. 2016;24:2973–2982.
19. Ogura K, Uehara K, Akiyama T, Iwata S, Shinoda Y, Kobayashi E, Saita K, Yonemoto T, Kawano H, Chuman H, Davis AM, Kawai A. Cross-cultural adaptation and validation of the

- Japanese version of the Toronto Extremity Salvage Score (TESS) for patients with malignant musculoskeletal tumors in the lower extremities. *J Orthop Sci.* 2015;20:1098–1105.
20. Seichi A, Hoshino Y, Doi T, Akai M, Tobimatsu Y, Iwaya T. Development of a screening tool for risk of locomotive syndrome in the elderly: the 25-question Geriatric Locomotive Function Scale. *J Orthop Sci.* 2012;17:163–172.
 21. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use.* 4th ed. New York, NY: Oxford University Press; 2008.
 22. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34–42.
 23. The Japanese Orthopaedic Association. The Japanese versions of MSTS scoring system and TESS. Available at: https://www.joa.or.jp/member/committee/diagnosis/pdf/tess_ue.pdf. Accessed May 10, 2017.
 24. The Japanese Orthopaedic Association Committee on Tumors. *[General Rules of Clinical and Pathological Studies on Malignant Bone Tumours] [in Japanese]* 3rd ed. Tokyo, Japan: Kanehara Shuppan Co; 2000.
 25. Unni KK. *Dahlin's Bone Tumors: General Aspects and Data on 11,087 Cases.* 5th ed. Philadelphia, PA: Lippincott-Raven; 1996.
 26. Wada T, Kawai A, Ihara K, Sasaki M, Sonoda T, Imaeda T, Yamashita T. Construct validity of the Enneking score for measuring function in patients with malignant or aggressive benign tumours of the upper limb. *J Bone Joint Surg Br.* 2007;89:659–663.