Further validation of the Toronto extremity salvage score for lower extremity soft tissue sarcoma based on Finnish patients

Kask, Gilber

2021-01

Kask , G , Uimonen , M M , Barner-Rasmussen , I , Tukiainen , E J , Blomqvist , C & Repo , J P 2021 , ' Further validation of the Toronto extremity salvage score for lower extremity soft tissue sarcoma based on Finnish patients ' , Journal of Plastic, Reconstructive & Aesthetic Surgery , vol. 74 , no. 1 , pp. 71-78 . https://doi.org/10.1016/j.bjps.2020.08.007

http://hdl.handle.net/10138/341313
https://doi.org/10.1016/j.bjps.2020.08.007

Journal of Plastic, Reconstructive & Aesthetic Surgery (2020) 000, 1–8

JPRAS
An International Journal of Surgical Reconstruction
www.JPRASurg.com

Further validation of the Toronto extremity salvage score for lower extremity soft tissue sarcoma based on Finnish patients

Gilber Kask[a,b,*], Mikko M. Uimonen[c], Ian Barner-Rasmussen[a], Erkki J. Tukiainen[a], Carl Blomqvist[d], Jussi P. Repo[c]

[a] *Plastic Surgery, University of Helsinki and Helsinki University Hospital, P.O. Box 266, FI-00029, HUS, Finland*
[b] *Tampere University Hospital, Department of Orthopaedics and Traumatology, Unit of Musculoskeletal Surgery, Teiskontie 35, FI-33521, Tampere, Finland*
[c] *Department of Surgery, Central Finland Central Hospital, Keskussairaalantie 19, FI-40620, Jyväskylä, Finland*
[d] *Comprehensive Cancer Center, Helsinki University Hospital and University of Helsinki, P.O. Box 180, FI-00029, HUS, Finland*

Received 22 September 2019; accepted 1 August 2020
Available online xxx

KEYWORDS
Neoplasms;
Soft tissue;
Sarcoma;
Outcome;
Function;
Validation

Summary The most widely used patient-reported outcome (PRO) measure for soft tissue sarcoma (STS) patients is the Toronto Extremity Salvage Score (TESS). The aim of the study was to validate and test the reliability of the TESS for patients with lower extremity STS based on Finnish population data. Patients were assessed using the TESS, the QLQ-C30 Function and Quality of life (QoL) modules, the 15D and the Musculoskeletal tumour Society (MSTS) score. The TESS was completed twice with a 2- to 4-week interval. The intraclass correlation coefficient (ICC) was used for test-retest reliability. Construct validity was tested for structural validity and convergent validity. Altogether 136 patients completed the TESS. A ceiling effect was noted as 21% of the patients scored maximum points. The ICC between first and second administration of the TESS was 0.96. The results of exploratory factor analysis together with high Cronbach's alpha (0.98) supported a unidimensional structure. The TESS correlated moderately with the MSTS score (rho = 0.59, $p < 0.001$) and strongly with the mobility dimension in the 15D

Ethical review committee statement: Institutional ethical board approval.

* Corresponding author at: Plastic Surgery, University of Helsinki and Helsinki University Hospital, P.O. Box 266, FI-00029 HUS, Finland.
E-mail addresses: gilber.kask@pshp.fi (G. Kask), ian.barner-rasmussen@hus.fi (I. Barner-Rasmussen), erkki.tukiainen@hus.fi (E.J. Tukiainen), Carl.Blomqvist@hus.fi (C. Blomqvist).

https://doi.org/10.1016/j.bjps.2020.08.007

Please cite this article as: G. Kask, M.M. Uimonen and I. Barner-Rasmussen et al., Further validation of the Toronto extremity salvage score for lower extremity soft tissue sarcoma based on Finnish patients, Journal of Plastic, Reconstructive & Aesthetic Surgery, https://doi.org/10.1016/j.bjps.2020.08.007

HRQL instrument (rho = 0.76, $p < 0.001$) and the physical function in QLQ-C30 (rho = 0.83, $p < 0.001$).
The TESS instrument is a comprehensive and reliable PRO measure. The TESS may be used as a validated single index score, for lower extremity STS patients for the measurement of a functional outcome. The TESS seems to reflect patients' HRQoL well after the treatment of lower extremity soft tissue sarcomas.

Introduction

Soft tissue sarcomas (STSs) are a rare group of heterogeneous mesenchymal tumours that represents approximately 1% of all solid malignancies in adults.[1] These neoplasms can arise in any site, but the most common anatomical location is the lower extremity, accounting for more than 50% of all STS.[2]

Several methods have been described for assessing functional outcome, including patient-reported outcome (PRO), clinician-reported outcome (ClinRO) and different objective or performance-related outcome measures [3-6]. The most widely used tools for the treatment outcome assessment of lower extremity STS are the Toronto Extremity Salvage Score (TESS)[7] and the Musculoskeletal tumour Society (MSTS) score.[3-6,8,9]

Any chosen measure for the assessment of outcome should be relevant, comprehensive and comprehensible with respect to the study population [10]. The TESS is a PRO measure that can be used to assess outcome from the patient's point of view. It is extremity-specific and measures performance and physical disability in activities of daily living.[7] In addition to its proven good psychometric properties,[11-17] the minimal clinically important difference (MCID) has recently been estimated for the TESS,[18] further supporting its use. The English version of the TESS has been previously validated for lower extremity sarcoma patients [11]. TESS has been translated and validated into the Chinese, Dutch, Korean, Danish, Portuguese and Japanese languages.[12-17] Although the TESS has been previously validated and its reliability tested, no studies have focused on validating the TESS solely for lower extremity STS patients. Previous publications have used a heterogeneous population of either upper or lower extremity STS patients, the combination of soft tissue and bone sarcoma patients or patients with both benign and malignant tumours. The validation study sample should represent the population of interest,[19] in this case lower extremity soft tissue sarcoma. The TESS has been previously translated and cross-culturally adapted into Finnish and tested in a pilot study,[20] but so far it has remained psychometrically unvalidated for assessing outcomes after the treatment of lower extremity STS.

The aim of this study was to validate and assess the reliability of the lower extremity section of the TESS for patients with lower extremity STS based on Finnish population data. More specifically the purpose was: (1) to study the validity and reliability of the Finnish version of the TESS and (2) to further investigate the psychometric properties of the TESS measure in lower extremity STS patients.

Methods

Study design

The protocol of this cross-sectional study was accepted by the Ethics Committee of the Helsinki and Uusimaa Hospital District, Finland. This study report was performed using the STROBE guidelines. The results of the study were interpreted and reported according to COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) guidelines.[10] The validation of the Finnish version of the lower limb TESS was performed amongst patients treated for soft tissue tumours at Helsinki University Hospital between the years 2006 and 2015, as identified from hospital databases by ICD-10 codes as well as by the NOMESCO Classification of Surgical Procedures codes. Inclusion criteria were age of at least 18 years, surgical treatment for lower extremity STS, local disease at the time of diagnosis, minimum follow-up of 6 months, completion of the TESS twice at separate time points and written informed consent.

Demographic, clinical, surgical and oncological data were collected retrospectively, functional outcome and health-related quality of life (HRQoL) outcome data were collected prospectively. Two functional outcome measures (TESS and MSTS 1993) and two HRQoL outcome measures (QLQ-C30 and 15D) were used.

Patients were invited to participate in the study by mail. Signing the informed consent form and completion of the questionnaires confirmed patients' participation in the study. After 2–4 weeks, a second mail was sent to participants, and they completed the TESS questionnaire a second time. Within 2 weeks of receiving the signed informed consent, a specially educated sarcoma nurse interviewed the patient by telephone and assessed the MSTS score.

Tumour location in the lower extremity was defined as a tumour distal to the inguinal ligament anteriorly or iliac crest posteriorly. The classification of tumour depth was based on the relationship to the superficial fascia – tumours were defined as superficial when superficial to and not infiltrating the fascia.

Outcome measures

TESS

The TESS is a self-administered PRO questionnaire that includes 30 items regarding activity limitations in daily life, such as restrictions in body movement, mobility, self-care and performance of daily tasks and routine, in lower ex-

Please cite this article as: G. Kask, M.M. Uimonen and I. Barner-Rasmussen et al., Further validation of the Toronto extremity salvage score for lower extremity soft tissue sarcoma based on Finnish patients, Journal of Plastic, Reconstructive & Aesthetic Surgery, https://doi.org/10.1016/j.bjps.2020.08.007

tremity STS patients. The tasks are rated from 0 (not possible) to 5 (without any problem). The raw score is converted to a score ranging from 0 to 100 points, with higher scores indicating a less functional limitation.[7]

The MCID is recently reported for the TESS questionnaire.[18] MCID is the smallest difference in score, which a patient feels as beneficial and therefore could mandate a change in the patient's management.[21] MCID for the TESS in lower extremity sarcoma patients using distribution-based and anchor-based methods would be around 5 points at 6 months and range between 4 and 10 points at 12 months[18].

MSTS 1993

The MSTS 1993 score is a limb- and oncology-specific ClinRO measure, measuring functional outcome and impairment.[8] The MSTS 1993 score is a revised version of the original MSTS 1987 score[9] and places less importance on clinical parameters in favour of a functional outcome. A physician completes six items together with the patient. There are separate questionnaires for lower and upper extremities. Pain, function and emotional acceptance are measured for both lower and upper extremities. The use of walking aids, gait and walking are evaluated for the lower extremity only. All categories are rated on a scale from 0 to 5, with 5 representing normal function. Function is reported as a percentage of the maximum score.

EORTC QLQ-C30

The European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30 questionnaire is designed for assessing HRQoL in cancer patients.[22] In the current study, the Physical Function and Global Quality of life modules of the QLQ-C30 were used to measure cancer-related functional impairment and complaints as well as impairment in the perceived quality of life. A result is scored from 0 to 100, with higher scores indicating better health.

15D

The 15D is a generic and self-administered HRQoL PRO instrument.[23] The questionnaire contains 15 dimensions of health: mobility, vision, hearing, breathing, sleeping, eating, speech, excretion, usual activities, mental function, discomfort and symptoms, depression, distress, vitality and sexual activity. Each dimension has five levels, describing the patient's health state at that moment. A general HRQoL score is calculated from the results of all 15 dimensions using a formula provided by the authors of the instrument. The index score varies between 0 representing the worst imaginable HRQoL and 1 representing the best. In the current study, the mobility item of the 15D was also used separately.

Statistical analysis

The demographic and clinical data are presented as means with standard deviations (SDs) and 95% confidence intervals (CIs) or as counts with percentages. The distribution of TESS scores was assessed. Floor and ceiling effects were examined to assess scale targeting. Floor or ceiling effect was considered confirmed if more than 15% of the patients received minimum or maximum scores, respectively.[24] The 'ceiling effect' describes the situation when many subjects in the study have scores at or close to the upper limit.[25] A floor effect is the opposite situation, i.e. if over 15% of patients score minimum points. Large ceiling and floor effects may be an indication that the TESS measure lacks content validity. Furthermore, to examine the influence of clinical factors, the floor and ceiling effects were examined separately in the following subgroups: patients with reconstruction vs. no reconstruction, muscle resection vs. no muscle resection, radiation therapy vs. no radiation therapy and high vs. low tumour grade. In addition, the influence of the length of follow-up time was examined by calculating Spearman's correlation coefficient between the time of operative treatment and the first administration of the TESS.

Test-retest reliability of the TESS was examined by calculating the intraclass correlation coefficients (ICC) and 95% CIs between the scores of the first and the second administration. ICC values over 0.7 were interpreted as the sufficient stability of the scores between the administrations, whereas ICC values below 0.7 indicated unacceptable stability.[10,26] In addition, difference in the first and second administration scores was examined using paired samples *t*-test. In clinical trials, test-retest reliability is the most important type of reliability for PRO instruments.[27] It shows the stability of the instrument over time, in case no change is detected in the concept of interest.[27] The test-retest reliability is high when the score remains unchanged and when retested within a short time interval.

Construct validity of the TESS was tested regarding structural validity, convergent validity and measurement invariance. Construct validity is the extent to which the measure 'behaves' in a way that is consistent with theoretical hypotheses and represents how well scores of the instrument are indicative of the theoretical PRO construct.[28]

Structural validity is one aspect of construct validity, and it shows the degree to which the scores of a PRO instrument are an adequate reflection of the dimensionality of the construct to be measured.[29] The structural validity of the TESS was assessed with an exploratory factor analysis (EFA). EFA was conducted to test the unidimensionality of the TESS and to define the best-fit model. A factor loading value of 0.4 was used as a cut-off value in determining sufficient factor representation.[30] A parallel analysis was performed to determine the number of factors to examine in the EFA.

Internal consistency was assessed by calculating Cronbach's alpha for all 30 items of the TESS. The bootstrapping method of 1000 repetitions was used to obtain the 95% CIs.[31] An alpha value of 0.7 was used as a cut-off point, with values above it representing acceptable internal consistency.[10] Furthermore, alpha values over 0.9 indicate that there might be redundant items in the scale.

Convergent validity refers to how closely the instrument is related to other variables and other measures of the same construct and it should not correlate with dissimilar ones.[32] The convergent validity of the TESS was evaluated by examining the convergence between the TESS and the MSTS scores, the Physical Function and the Global Quality of life subscales of the QLQ-C30 and the general HRQoL 15D index score as well as its mobility dimension. The Spearman's correlation coefficients between the instruments were calculated.

Measurement invariance should be evaluated when a PRO instrument is or will be used in different patient

Please cite this article as: G. Kask, M.M. Uimonen and I. Barner-Rasmussen et al., Further validation of the Toronto extremity salvage score for lower extremity soft tissue sarcoma based on Finnish patients, Journal of Plastic, Reconstructive & Aesthetic Surgery, https://doi.org/10.1016/j.bjps.2020.08.007

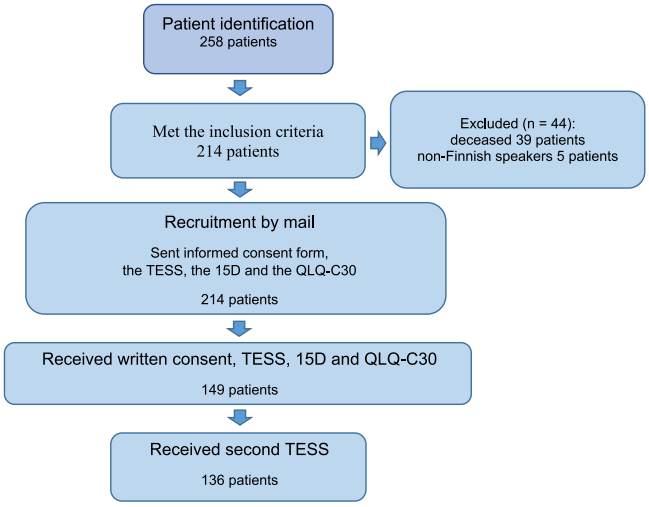

Figure 1 Patients' recruitment.

populations. Measurement invariance shows the degree to which the items of the instrument perform independently of demographic characteristics of the patients. To assess measurement invariance, the associations of the TESS and sex, age and body mass index (BMI) were examined. The influence of age and BMI on TESS was examined using Spearman's correlation coefficients. The strength of the correlations was interpreted as follows: 0.00-0.30 negligible, 0.30-0.50 weak, 0.50-0.70 moderate, 0.70-0.90 strong and 0.90-1.00 very strong.[33] The difference in the TESS between men and women was tested using independent samples *t*-test.

Statistical analyses were conducted using R 3.6.1 and SPSS 25.0 statistical software.

Results

A flow sheet of recruitment is shown in Figure 1. A total of 136 patients completed the TESS questionnaires two times in 2-4 weeks' interval. The MSTS, the 15D and the QLQ-C30 were completed by 74/136 (54%), 135/136 (99%) and 136/136 (100%) of patients, respectively. Median follow-up from surgery to first administration was 5.1 years (range: 0.5-16.4). Sociodemographic and clinical details are presented in Table 1, reconstructive methods in Table 2.

The distribution of the TESS total scores was strongly skewed towards higher scores, indicating good outcomes (Figure 2). The length of follow-up time was not associated with the TESS scores (rho = 0.11, $p = 0.21$). As none of the patients scored minimum points, there was no floor effect. A ceiling effect was present, as 29 patients (21%) obtained the maximum score. The ceiling effect was independent of reconstructive procedures performed (maximum score 16% vs. 23% for reconstruction vs. no reconstruction) or muscle resection (maximum score 17% vs. 32% for muscle resection vs. no muscle resection) as well as of tumour grade (maximum score 16% vs. 23% for high grade vs. low grade). In patients who had undergone radiation therapy, no ceiling effect was observed as the proportion of patients obtaining maximum score was 9%, whereas amongst the patients who had not undergone radiation therapy, a ceiling effect was present (maximum score 30%).

The mean scores of the first and second administration of the TESS showed no significant difference. The ICC between the first and second administration scores was 0.95 (95% CI = 0.93-0.96 and $p < 0.001$; Figure 3).

Table 1 Sociodemographic and clinical characteristics of participants.

	$N = 136$
Female, n (%)	70 (51)
Age, years, mean (SD)	65.6 (14.4)
BMI, mean (SD)	27.3 (5.4)
Tumour status (%)	
Primary	111 (81.6)
Local recurrence	25 (18.4)
Sarcoma, n (%)	
Liposarcoma	52 (38.2)
Undifferentiated pleomorphic sarcoma	27 (19.9)
Sarcoma NOS	17 (12.5)
Leiomyosarcoma	16 (11.8)
Myxofibrosarcoma	10 (7.4)
Other	14 (10.2)
Surgical procedure	
Excision	98 (72.1)
Myectomy	33 (24.3)
Amputation	4 (2.9)
Rotationplasty	1 (0.7)
Myectomy excision	
Rotationplasty	1 (0.7)
Histological margin, n (%)	
Intralesional	12 (8.8)
Marginal	81 (59.6)
Wide	43 (31.6)
Tumour size, mm and mean (SD)	80.7 (61.4)
Wound closure, n (%)	
Direct	82 (60.3)
Split-thickness skin graft	22 (16.2)
Flap reconstruction	32 (23.5)
Radiotherapy (%)	
Yes	55 (40.4)
No	81 (59.6)

SD - standard deviation, NOS - not otherwise specified and BMI - body mass index.

Table 2 Reconstructions (33 of 136 patients).

	$N = 33$
Reconstructions, n (%)	
Microvascular LD	5
Stylo cutanea directa	4
Microvascular ALT	4
Pedicular ALT	3
Stylo microvascularis gracilis	2
Transpositio musculus gastrocnemicus	2
Pedicular tensor fascia lata	2
Other	11

ALT - anterolateral free flap and LD - latissimus dorsi.

Please cite this article as: G. Kask, M.M. Uimonen and I. Barner-Rasmussen et al., Further validation of the Toronto extremity salvage score for lower extremity soft tissue sarcoma based on Finnish patients, Journal of Plastic, Reconstructive & Aesthetic Surgery, https://doi.org/10.1016/j.bjps.2020.08.007

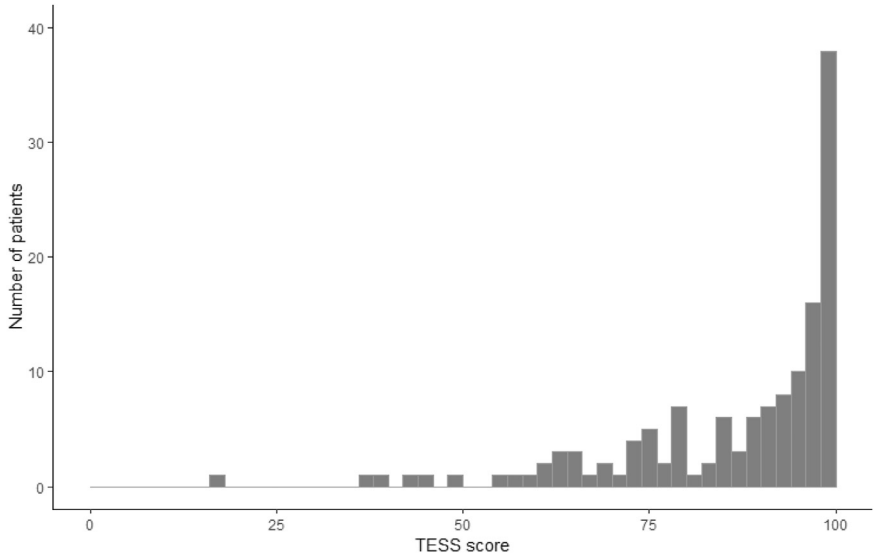

Figure 2 Distribution of TESS scores.

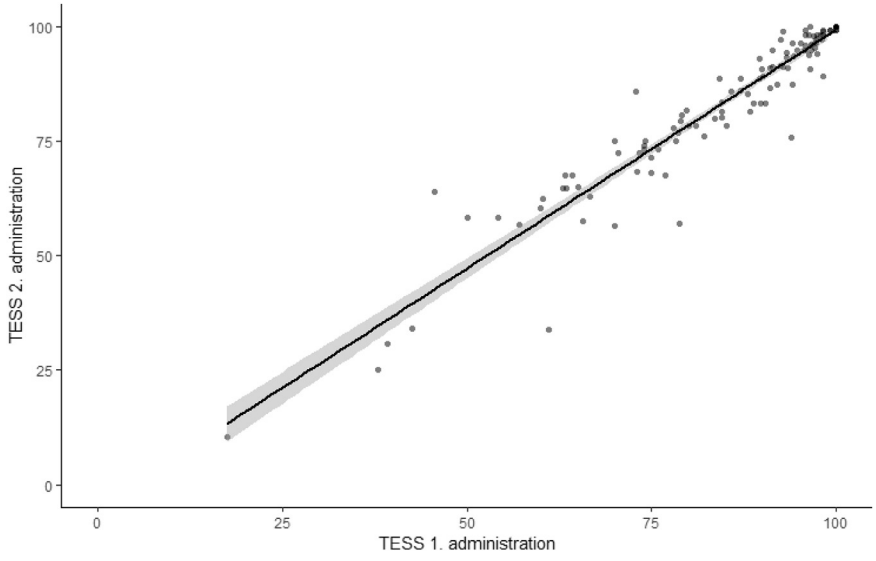

Figure 3 Linearity between first and second administration scores of the TESS.

The parallel analysis indicated the inclusion of eight factors in the EFA (Figure 4). EFA of the TESS items with eight included factors revealed four factors filling the Kaiser criteria of Eigenvalue over 1. Factor 1 (Eigenvalue = 22.3) explained 74.4% of the variance, while factors 2 (Eigenvalue = 1.9), 3 (Eigenvalue = 1.3) and 4 (Eigenvalue = 1.0) explained 6.4%, 4.2% and 3.3% of the total variance, respectively. However, all items, except item 29, loaded strongest on factor 1. Item 29 loaded strongest on factor 3 (loading value 0.557), although the loading on factor 1 was also strong (loading value 0.515). The loading values on factor 1 varied between 0.515 and 0.947 across all items. High loading of all items on factor 1 indicates unidimensional one factor structure of the TESS leaving other factors insignificant. Cronbach's alpha of the TESS items was 0.97 (95% CI = 0.97

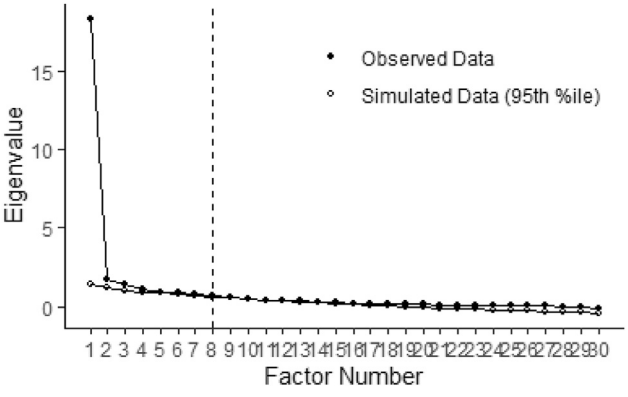

Figure 4 Scree plot for eigenvalues of the factors.

Please cite this article as: G. Kask, M.M. Uimonen and I. Barner-Rasmussen et al., Further validation of the Toronto extremity salvage score for lower extremity soft tissue sarcoma based on Finnish patients, Journal of Plastic, Reconstructive & Aesthetic Surgery, https://doi.org/10.1016/j.bjps.2020.08.007

Table 3 Predefined hypotheses and conclusions of measured psychometric properties of the TESS.

Feature	Hypothesis	Conclusion
Scale targeting		
No floor effect	Min score <15%	Confirmed
No ceiling effect	Max score <15%	Rejected
Test-retest reliability		
Correlation between baseline and repeated admission scores	ICC > 0.7	Confirmed
Construct validity		
Structural validity		
EFA	Unidimensional best-fit model	Confirmed
Internal consistency	Cronbach's alpha > 0.7	Confirmed
Convergent validity		
Convergence with MSTS score	Significant and at least low correlation with MSTS score (rho ≥ 0.3 and $p < 0.05$)	Confirmed
Convergence with 15D and Mobility dimension	Significant and at least low correlation with 15D index and 15D Mobility (rho ≥ 0.3 and $p < 0.05$)	Confirmed
Convergence with QLQ-C30 Function and Quality of life	Significant and at least low correlation with QLQ-C30 Function and Quality of life (rho ≥ 0.3 and $p < 0.05$)	Confirmed
Measurement invariance		
Independency of demographic characteristics	Non-significant or negligible associations with demographic characteristics	Confirmed/ Rejected

EFA = Exploratory Factor Analysis.
ICC = Intraclass Correlation Coefficient.

- 0.98), indicating very high internal consistency, which on the other hand, suggests item redundancy.

TESS showed strong correlations to the 15D *Mobility dimension* (rho = 0.76, 95% CI from −0.84 to −0.66 and $p < 0.001$) and the *General HRQoL* index (rho = 0.75, 95% CI = 0.66–0.82 and $p < 0.001$) as well as to the QLQ-C30 subscales of *Physical Function* (rho = 0.83, 95% CI 0.76 to 0.88 and $p < 0.001$) and *Global QoL* (rho = 0.71, 95% CI from 0.60 to 0.79 and $p < 0.001$). The correlation between TESS and MSTS was moderate (rho = 0.59, 95% CI = from 0.40 to 0.73 and $p < 0.001$).

Significant negative correlations of the TESS with age (rho = −0.23 and $p = 0.006$) and BMI (rho = −0.25 and $p = 0.006$) existed, although they were low. No significant difference between TESS scores of male and female subjects was observed ($p = 0.143$). Conclusions of measured psychometric properties of the TESS are presented in Table 3.

Discussion

The results of this study indicate the ability of the TESS to function as a single index score, as the results of EFA provided strong support for the unidimensionality of the instrument. A slight ceiling effect was noted (21%), potentially demonstrating problems in the content validity of the TESS instrument for assessing long-term outcomes after the treatment of lower extremity STS. The high test-retest reliability of the TESS is in line with previous findings, with ICC ranging between 0.87 and 0.94.[12,13]

Our results generally supported the hypothesized one-factor model of the TESS, indicating sufficient structural validity. Factor analysis is usually performed to test whether the instrument can be used as a single item score, the items should be separated into modules or the items should be presented as a profile. The EFA including eight factors suggested by parallel analysis found eventually four potential factors for the TESS. The loading values on factor 1 were, however, high across all items, suggesting the insignificance of other factors. The result can be interpreted supporting the unidimensional structure on one factor of the TESS. The studies by Xu et al. and Ogura et al. found a two-factor structure in the Chinese and Japanese versions of the TESS.[13,14] The more heterogeneous sample of lower extremity sarcoma patients could potentially be the reason for the difference. On the other hand, in previous publications investigating the internal consistency, the Cronbach's alpha has been high, ranging from 0.94 to 0.98.[11-14,16]. High Cronbach's alpha was also supporting the unidimensional structure of the TESS. Despite this discrepancy, it seems that the TESS could be used as a single index score for lower extremity STS patients. As the TESS is a single index score, the score should be presented as a sum of the TESS score and which presents one factor, the physical function. In contrast to QLQ-C30 questionnaire, where different indexes or modules can be reported separately and present different aspects of patients' HRQoL.[22]

In contrast to Kim et al. who noted a high correlation between the TESS and the MSTS score (0.77),[12] the correlation observed in the current study was slightly lower (0.59). However, Kim's study included 47% bone sarcoma patients and 50% of patients had undergone bone resection, possibly explaining the differences in correlation results. The correlation between the TESS and relevant items in QLQ-C30 and 15D global was strong. This is in accordance with

Please cite this article as: G. Kask, M.M. Uimonen and I. Barner-Rasmussen et al., Further validation of the Toronto extremity salvage score for lower extremity soft tissue sarcoma based on Finnish patients, Journal of Plastic, Reconstructive & Aesthetic Surgery, https://doi.org/10.1016/j.bjps.2020.08.007

the study by Kim et al., where the correlation was also strong (rho = 0.77 and 0.83).[12] Based on the results presented here, the TESS seems to reflect patients' HRQoL well after the treatment of lower extremity STS.

A ceiling effect was noted as 21% of our patients who obtained the maximum score. This is in line with the results of Ogura et al., where a ceiling effect was noted with 16.7% of patients who obtained the maximum score.[13] The ceiling effect was present in all subgroups, except for patients who received radiotherapy. Twenty-one percent of patients fulfilling the inclusion criteria had died, which may have caused some bias. This may have contributed to the high ceiling effect, due to the self-selection of high-functioning patients. The phenomenon is inevitable using the TESS instrument in this field. This potentially demonstrates problems in the content validity. One obvious potential reason for the ceiling effect may be that many STS patients have an excellent functional result.

The study sample should represent the population of interest,[19] but many of the previous validation studies of the TESS have been done on the heterogeneous population, including upper extremity and bone sarcoma patients. Previous studies have included 48–126 extremity tumour patients, but the amount of lower extremity STS patients ranged from zero to fifty-one or the number of patients was not clear.[11-17]

The strength of this study was its large sample size, which can be considered suitable for conducting psychometric tests.[34] All statistical analyses were based on predefined hypotheses adhering to the COSMIN guidelines.[10,34] Testing at two distinct time points gave important information about the reproducibility of the instrument. The inclusion criteria were wide to generate a representative sample of patients with lower extremity STS treated at a large tertiary academic referral centre. A further strength of the present analysis was the more homogeneous patient material in contrast to previous ones, including bone sarcoma patients (or upper extremity patients).

Conclusion

The current study was the first to validate the TESS distinctly for lower extremity STS patients using the largest sample size published thus far. Despite the ceiling effect, the TESS instrument is a comprehensive and reliable PRO measure, providing valid scores in assessing the functional outcome after the treatment of lower extremity STS. Our study shows that the instrument functions as a single index score. The TESS reflects patients' HRQoL, indicating its relevance in studying the functional outcome after the surgical treatment of lower extremity STS.

Declaration of Competing Interest

None.

Acknowledgments

Funding: This research received no specific grant or funding from any funding agency in the public, commercial, or not for profit sectors.

Data availability statement: The data that support the findings of this study are available from the corresponding author, GK, upon reasonable request. The Finnish version of the TESS and original version of the TESS[7] are added as supplementary material.

References

1. Toro JR, Travis LB, Wu HJ, et al. Incidence patterns of soft tissue sarcomas, regardless of primary site, in the surveillance, epidemiology and end results program, 1978-2001: an analysis of 26,758 cases. *Int J Cancer* 2006;**119**:2922–30 2006/10/03. doi:10.1002/ijc.22239.
2. Ogura K, Higashi T, Kawai A. Statistics of soft-tissue sarcoma in Japan: report from the bone and soft tissue tumor registry in Japan. *J Orthop Sci* 2017;**22**:755–64 2017/06/12. doi:10.1016/j.jos.2017.03.017.
3. Tang MH, Pan DJ, Castle DJ, et al. A systematic review of the recent quality of life studies in adult extremity sarcoma survivors. *Sarcoma* 2012;**2012**:171342 2012/09/12. doi:10.1155/2012/171342.
4. Furtado S, Errington L, Godfrey A, et al. Objective clinical measurement of physical functioning after treatment for lower extremity sarcoma - a systematic review. *Eur J Surg Oncol* 2017;**43**:968–93 2016/11/12. doi:10.1016/j.ejso.2016.10.002.
5. Davis AM. Functional outcome in extremity soft tissue sarcoma. *Semin Radiat Oncol* 1999;**9**:360–8 1999/10/12. doi:10.1053/SRAO00900360.
6. Kask G, Barner-Rasmussen I, Repo JP, et al. Functional outcome measurement in patients with lower-extremity soft tissue sarcoma: a systematic literature review. *Ann Surg Oncol* 2019 2019/08/14. doi:10.1245/s10434-019-07698-w.
7. Davis AM, Wright JG, Williams JI, et al. Development of a measure of physical function for patients with bone and soft tissue sarcoma. *Qual Life Res* 1996;**5**:508–16 1996/10/01.
8. Enneking WF, Dunham W, Gebhardt MC, et al. A system for the functional evaluation of reconstructive procedures after surgical treatment of tumors of the musculoskeletal system. *Clin Orthop Relat Res* 1993:241–6 1993/01/01.
9. WF E. Modification of the system for functional evaluation in the surgical management of musculoskeletal tumours. In: *Limb Salvage in Musculoskeletal Oncology*. New York, NY, USA: Churchill Livingstone; 1987. p. 626–39.
10. Prinsen CA, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Res* 2018;**27**:1147–57.
11. Davis AM, Bell RS, Badley EM, et al. Evaluating functional outcome in patients with lower extremity sarcoma. *Clin Orthop Relat Res* 1999:90–100 1999/02/12.
12. Kim HS, Yun J, Kang S, et al. Cross-cultural adaptation and validation of the Korean Toronto Extremity Salvage Score for extremity sarcoma. *J Surg Oncol* 2015;**112**:93–7 2015/06/16. doi:10.1002/jso.23947.
13. Ogura K, Uehara K, Akiyama T, et al. 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**:1098–105 2015/08/27. doi:10.1007/s00776-015-0767-8.
14. Xu L, Sun M, Sun W, et al. Cross-cultural adaptation and validation of the chinese version of toronto extremity salvage score for patients with extremity sarcoma. *Springerplus* 2016;**5**:1118 2016/08/02. doi:10.1186/s40064-016-2818-9.
15. Saraiva D, de Camargo B, Davis AM. Cultural adaptation, translation and validation of a functional outcome questionnaire (TESS) to Portuguese with application to patients with lower extremity osteosarcoma. *Pediatr Blood Cancer* 2008;**50**:1039–42 2007/11/02. doi:10.1002/pbc.21392.

Please cite this article as: G. Kask, M.M. Uimonen and I. Barner-Rasmussen et al., Further validation of the Toronto extremity salvage score for lower extremity soft tissue sarcoma based on Finnish patients, Journal of Plastic, Reconstructive & Aesthetic Surgery, https://doi.org/10.1016/j.bjps.2020.08.007

16. Saebye C, Safwat A, Kaa AK, et al. Validation of a Danish version of the Toronto Extremity Salvage Score questionnaire for patients with sarcoma in the extremities. *Dan Med J* 2014;**61**:A4734 2014/01/08.
17. Willeumier JJ, van der Wal C, van der Wal RJP, et al. Cross-cultural adaptation, translation, and validation of the Toronto extremity salvage score for extremity bone and soft tissue tumor patients in Netherlands. *Sarcoma* 2017;**2017**:6197525 2017/08/16. doi:10.1155/2017/6197525.
18. Ogura K, Uehara K, Akiyama T, et al. Minimal clinically important differences in Toronto extremity salvage score for patients with lower extremity sarcoma. *J Orthop Sci* 2019 2019/04/20. doi:10.1016/j.jos.2019.03.022.
19. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;**27**:1147–57 2018/02/13. doi:10.1007/s11136-018-1798-3.
20. Kask G, Barner-Rasmussen I, Repo J, et al. Functional outcome after lower extremity soft tissue sarcoma treatment: a pilot study based on translated and culturally adapted measures. *Scand J Surg* 2018 1457496918798210. 2018/09/07. doi:10.1177/1457496918798210.
21. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 2003;**56**:395–407 2003/06/19. doi:10.1016/s0895-4356(03)00044-1.
22. Aaronson NK, Ahmedzai S, Bergman B, et al. The European organization for research and treatment of cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;**85**:365–76 1993/03/03. doi:10.1093/jnci/85.5.365.
23. Sintonen H. The 15D instrument of health-related quality of life: properties and applications. *Ann Med* 2001;**33**:328–36 2001/08/09.
24. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Quality of Life Research* 1995;**4**:293–307.
25. . *The Cambridge Dictionary of Statistics*. 4th Ed. The Edinburgh Building, Cambridge: Cambridge University Press; 2010.
26. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;**60**:34–42.
27. Health USDo, Human Services FDACfDE, Research Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes* 2006;**4**:79 2006/10/13. doi:10.1186/1477-7525-4-79.
28. Frost MH, Reeve BB, Liepa AM, et al. What is sufficient evidence for the reliability and validity of patient-reported outcome measures? *Value Health* 2007;**10**(Suppl 2):S94–S105 2007/11/13. doi:10.1111/j.1524-4733.2007.00272.x.
29. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;**63**:737–45 2010/05/25. doi:10.1016/j.jclinepi.2010.02.006.
30. Peterson RA. A meta-analysis of variance accounted for and factor loadings in exploratory factor analysis. *Mark Lett* 2000;**11**:261–75.
31. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;**16**:297–334.
32. . *Health Measurement Scales: A practical guide to their development and use*. 5th Ed. Oxford, United Kingdom: Oxford University Press; 2015.
33. Mukaka MM. A guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 2012;**24**:69–71.
34. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Qual Life Res* 2018;**27**:1171–9 2017/12/21. doi:10.1007/s11136-017-1765-4.

Please cite this article as: G. Kask, M.M. Uimonen and I. Barner-Rasmussen et al., Further validation of the Toronto extremity salvage score for lower extremity soft tissue sarcoma based on Finnish patients, Journal of Plastic, Reconstructive & Aesthetic Surgery, https://doi.org/10.1016/j.bjps.2020.08.007