

Validation of a Danish version of the Toronto Extremity Salvage Score questionnaire for patients with sarcoma in the extremities

Casper Sæbye¹, Akmal Safwat², Anne Kathrine Kaa³, Ninna Aggerholm Pedersen¹ & Johnny Keller³

ABSTRACT

INTRODUCTION: The Toronto Extremity Salvage Score (TESS) questionnaire is a self-administered questionnaire designed to assess physical disability in patients having undergone surgery of the extremities. The aim of this study was to validate a Danish translation of the TESS.

MATERIAL AND METHODS: The TESS was translated according to international guidelines. A total of 22 consecutive patients attending the regular outpatient control programme were recruited for the study. To test their understanding of the questionnaires, they were asked to describe the meaning of five randomly selected questions from the TESS. The psychometric properties of the Danish version of TESS were tested for validity and reliability. To assess the test-retest reliability, the patients filled in an extra TESS questionnaire one week after they had completed the first one.

RESULTS: Patients showed good understanding of the questionnaire. There was a good internal consistency for both the upper and lower questionnaire measured by Cronbach's alpha. A Bland-Altman plot showed acceptable limits of agreement for both questionnaires in the test-retest. There was also good intraclass correlation coefficients for both questionnaires. The validity expressed as Spearman's rank correlation coefficient comparing the TESS with the QLQ-C30 was 0.89 and 0.90 for the questionnaire on upper and lower extremities, respectively.

CONCLUSION: The psychometric properties of the Danish TESS showed good validity and reliability.

FUNDING: not relevant.

TRIAL REGISTRATION: not relevant.

Sarcoma is a rare type of cancer, representing approximately 1% of all newly diagnosed cancers. The preferred treatment for soft tissue sarcoma in the extremities is limb sparing surgery, while it previously was amputation [1-5]. Measuring functional outcome after surgical treatment is of paramount importance. However, the patient population has a significant heterogeneity with respect to age, anatomic site of disease, tumour type, reconstructive techniques and the extent and type of tissue excised during surgery. Therefore, it is important to have

a measure that takes heterogeneity into account when measuring functional outcome.

There are several ways to measure outcome after cancer treatments. Quality of life questionnaires such as the QLQ-C30 and Short Form (SF)-36 [6, 7] are often used. However, the functional outcome seems to be a more important criterion in surgical treatment of extremity sarcomas. There are several ways to measure the functional outcome, one is clinical measurement of muscle strength and range of motion [8]; another option is to estimate outcome based on a combination of symptoms and mobility as in the Musculoskeletal Tumor Society (MSTS) score [4, 9]. However, to take into account the rarity and heterogeneity of sarcomas, the patient-completed questionnaire, the Toronto Extremity Salvage Score (TESS) was developed [10, 11].

The TESS questionnaire is a disease-specific, self-administered questionnaire for patients with sarcoma in the extremities. The intent of the TESS is to evaluate a single domain – physical disability – based on patients' reports of their function, with a view to monitoring and evaluating the physical function of individuals and groups of patients over time and to measuring change in function following different therapeutic interventions [10].

We have translated the TESS questionnaire according to the guidelines presented by Guillemin et al [12] in order to account for the linguistic and cross-cultural differences that can affect the outcome. The aim of this study was to validate the Danish translation and to re-test the psychometric properties of the TESS in a Danish population.

MATERIAL AND METHODS

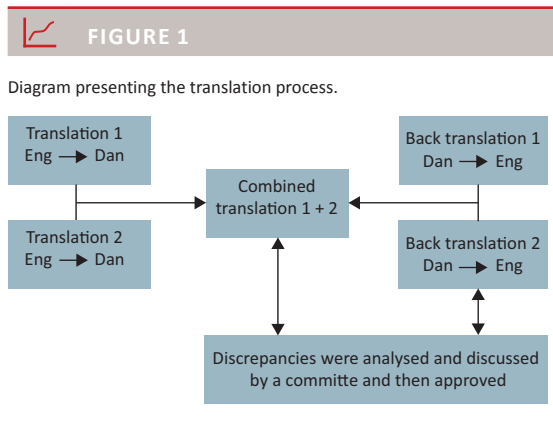
The Toronto Extremity Salvage Score questionnaire

The TESS was designed specifically to assess the physical function in musculoskeletal tumour patients aged 12-85 years. The described physical function is subsumed by the International Classification of Impairments, Disabilities and Handicaps definition of disability. There is an upper and a lower extremity version consisting of 29 and 30 questions, respectively, regarding everyday activities such as working, dressing, mobility and leisure. Each question

ORIGINAL ARTICLE

- 1) Department of Experimental Clinical Oncology, Sarcoma Centre, Aarhus University Hospital
- 2) Department of Oncology, Sarcoma Centre, Aarhus University Hospital
- 3) Department of Orthopaedics, Sarcoma Centre, Aarhus University Hospital

Dan Med J
2014;61(1):A4734



is rated on a scale from one to five, with five being the value considered for normal performance.

The total score is calculated as a percentage of the maximum score [10]. The Danish TESS questionnaires are available for download on the Danish Sarcoma Group's homepage [13].

The QLQ-C30 questionnaire

The QLQ-C30 is a self-administered 30-item questionnaire developed to assess the health-related quality of life. It is composed of several scales: five functional scales (Physical, Role, Emotional, Cognitive and Social Functioning); nine symptom scales (Fatigue, Nausea/Vomiting, Pain, Dyspnoea, Insomnia, Appetite Loss, Constipation, Diarrhoea and Financial Difficulties); and a Global Health Status/Quality of Life (QOL) score. A high score on the Global Health Status/QOL scale and functional scales represents a high level of quality of life and a high level of functionality [14]. The QLQ-C30 has been validated for the Danish population [15].

Translation process

The translation process included two independent translations made by two bilingual persons. Afterwards, the two translators analysed and compared the two translations and combined them into one. During a conference in which all the authors of this article participated both the two separate translations as well as the combined one were discussed and a consensus was reached. The consensus version was then translated back into the original language (English) by two independent translators.

The two forward translators and one of the back translators all had a clinical background and experience. The second back translator had no clinical experience or any relation to healthcare. In a second conference counting all those involved, all translated versions were evaluated and consensus was achieved on the final Danish version (Figure 1).

Validation process

The Danish TESS questionnaires were tested in 24 consecutive follow-up patients who had undergone limb-sparing surgery in the extremities for sarcoma and aggressive benign tumours at the Sarcoma Centre in Aarhus University Hospital. The patients were then asked to explain how they understood a set of five randomly selected questions from the upper or the lower TESS questionnaire. The time used to complete the TESS questionnaire was recorded for all test persons. Furthermore, the patients were given one more copies of the TESS for completion one week later. The patients were also asked to complete the QLQ-C30 questionnaire. The patients were informed and the data were collected by the first author, who has not otherwise been involved in the patients' diagnosis, treatment or follow-up.

Patients' characteristics

One was excluded due to recurrence of the disease and one due to an uncompleted first questionnaire. The patient group consisted of 13 females and nine males with a mean age of 41 years (range 19-66 years). There were six patients with disease in the upper extremities and 16 with disease in the lower extremities.

Statistics

To evaluate the internal consistency of the TESS, we chose Cronbach's alpha [16] as it provides a measurement of the strength of the relationship among the questions of the questionnaire. The stronger the relationship – i.e. the more questions contribute to the same concept – the higher the Cronbach's alpha value will be. A priori, Cronbach's alpha ≥ 0.80 is considered acceptable [17].

Test-retest reliability is the ability to obtain the same results repeatedly and independently of time. To show the absolute differences and limits of agreement (LOA) between the test- and the retest data, we used the Bland-Altman plot. Furthermore, we calculated the intraclass correlation coefficients (ICC) to show the proportion of variance of an observation due to between-subject variability in the true scores.

The validity is an index showing how well the questionnaire measures what it is supposed to measure. We measured validity using the Spearman's rank correlation coefficient to assess the correlation between the two functional outcome measures in the TESS and QLQ-C30.

Trial registration: not relevant.

RESULTS

The translators encountered no major difficulty in relation to the terms used and, therefore, very few changes were made in the synthesis stage. Likewise, very few

terms were corrected in the back translation and expert committee review phases.

The answers from the five randomly picked questions showed that they were understood as intended. The mean time used to complete the TESS questionnaire was 5 min. and 34 sec. (range 1:16-15:26).

The test for internal consistency resulted in a total Cronbach's alpha of $R = 0.90$ for the upper extremities questionnaire and $R = 0.94$ for the lower extremities questionnaire.

A total of 21 (95%) patients returned the second questionnaire. A Bland-Altman plot for the upper extremity questionnaire showed that the mean difference was -0.83 , while lower and upper LOAs were -3.97 and 2.31 , respectively. The lower extremity questionnaires' Bland-Altman plot showed that the mean difference was 1.27 , while lower and upper LOAs were -12.38 and 14.92 , respectively.

We found that the ICC for the upper and lower extremity questionnaires were 0.96 (95% confidence interval (CI): $0.755-0.994$) and 0.88 (95% CI: $0.697-0.956$), respectively.

The mean score for the TESS questionnaire was 84 (range 60-100), and for the QLQ-C30 it was 74 (range 27-100). The overall Spearman's rank correlation coefficient between the TESS and the QLQ-C30 was 0.94 . The Spearman's rank correlation coefficient for the upper extremity questionnaire was 0.89 , and for the lower extremity questionnaire it was 0.90 .

DISCUSSION

The development of the original TESS questionnaire included testing of its psychometric properties, which showed that the questionnaire was a valid, reliable and sensitive tool for assessment of the functional outcome for sarcoma patients in Canada [10]. In our translation of the TESS into Danish, we used the guidelines presented by Guillemin et al. which take linguistic and cross-cultural bias into account [12]. For example, the phrase "feeling downhearted and blue" is an untranslatable idiom, and therefore a suitable Danish equivalent must be found. When Italian scientists translated the Oxford Shoulder Score they encountered a question concerning a task that was not a common practice in the Italian culture and therefore had to modify that item for the Italian population [18]. In our translation of the TESS questionnaire, we encountered no major linguistic or cultural discrepancies. The short completion time indicates that the questionnaire was well understood and easy to fill out. Our completion time was similar, although lower, compared with results from other studies [11, 17]. The answers to the five randomly selected questions from the questionnaires also confirmed that they were well understood and easy to fill out.

The good internal consistency achieved in the present study agreed with current literature [17]. The Bland-Altman plot showed an upper and lower LOA for the TESS questionnaires that meant that there was a very low error of measurement over time for a patient. The ICCs of 0.96 (95% CI: $0.755-0.994$) and 0.88 (95% CI: $0.697-0.956$) for the upper and lower extremity questionnaires, respectively, demonstrated a high level of test-retest reliability, which corresponded well with previously published data [10].

A comparison was made between the TESS and the functional part of the QLQ-C30. We chose the QLQ-C30 because it is a validated questionnaire designed for cancer patients [14, 15]. Our results showed that there was a good correlation between the TESS and the QLQ-C30. The Spearman's rank correlation coefficient of 0.89 and 0.90 for the upper and lower extremity questionnaires, respectively, indicated that the TESS and the QLQ-C30 did not yield identical results. However, this was to be expected since the TESS is a 29/30-item questionnaire, while the QLQ-C30 function score is a five-item questionnaire.

The present study did not test the Danish TESS's responsiveness, i.e. the ability of an instrument to detect accurately change when it has occurred. This would have added strength and is therefore recommended for future studies.

Recently, a study was conducted to determine if gender and/or age-specific as well as cultural differences in a normal population would influence the score and appropriateness of the questions in the TESS questionnaire. The study was conducted in Australia and Britain. The results showed that though the TESS could be used as an international scoring system, it is age-sensitive

To test the respondents' understanding of the questionnaires, they were asked to describe the meaning of five randomly selected questions from the Toronto Extremity Salvage Score.



[19]. It is important, therefore, to consider age as a possible confounder when using the questionnaire.

A study which compared the TESS with the SF-36 and the 1987 and 1993 versions of the Musculoskeletal Tumor Society Score (MSTS) concluded that when choosing a functional outcome measurement for extremity sarcoma patients, the TESS has superior measurement properties compared to the other questionnaires [11]. However, others have argued that the TESS measures the patient's ability to cope, whereas the MSTS estimates functional impairment and therefore measures different aspects of the functional outcome [3]. Given that the TESS is superior at measuring physical disability and that the MSTS measures other aspects such as pain and acceptance of the treatment, and is partially patient-reported and physician-reported, we find that using the TESS does not exclude the use of MSTS.

A patient-administered questionnaire such as TESS may have a bias with regard to functional outcome, and therefore the value of self-reported outcome may be limited in insurance cases and when allocation of public services is involved. Further studies in this area are needed before the TESS can be considered an expedient tool for such assessments.

CONCLUSION

We have translated the TESS questionnaire into Danish and validated the Danish questionnaire. We found that it was easy to understand and the time patients used to complete the questionnaire was acceptable. The questionnaire showed good internal consistency. The test-retest result was also good and demonstrated that the questionnaire was reliable. The questionnaire showed good validity when compared to the QLQ-C30 functional score. Based on these results, we conclude that the Danish translation is valid and can be used in the Danish population. Further studies to investigate TESS responsiveness and other topics are recommended.

CORRESPONDENCE: Casper Sæbye, Afdeling for Eksperimentel Klinisk Onkologi, Aarhus Universitetshospital, Nørrebrogade 44, Bygning 5, 8000 Aarhus C, Denmark. E-mail: caspersaebye@studmed.au.dk

ACCEPTED: 19 September 2013

CONFLICTS OF INTEREST: none. Disclosure forms provided by the authors are available with the full text of this article at www.danmedj.dk

ACKNOWLEDGEMENT: Krishna Maria Olsen has contributed to the project as a translator on one of the back-translations.

LITERATURE

- Nielsen OS, Keller JO, Dombernowsky P. Sarcomas. *Ugeskr Læger* 2002;164:3036-9.
- Skubitz KM, D'Adamo DR. Sarcoma. *Mayo Clin Proc* 2007;82:1409-32.
- Aksnes LH, Bauer HC, Jebsen NL et al. Limb-sparing surgery preserves more function than amputation: a Scandinavian sarcoma group study of 118 patients. *J Bone Joint Surg Br* 2008;90:786-94.
- Johansen R, Nielsen OS, Keller J. Functional outcome in sarcomas treated with limb-salvage surgery or amputation. *Sarcoma* 1998;2:19-23.
- Davis AM, Devlin M, Griffin AM et al. Functional outcome in amputation versus limb sparing of patients with lower extremity sarcoma: a matched case-control study. *Arch Phys Med Rehabil* 1999;80:615-8.

- Barrera M, Teall T, Barr R et al. Health related quality of life in adolescent and young adult survivors of lower extremity bone tumors. *Pediatr Blood Cancer* 2012;58:265-73.
- Malek F, Somerson JS, Mitchel S et al. Does limb-salvage surgery offer patients better quality of life and functional capacity than amputation? *Clin Orthop Relat Res* 2012;470:2000-6.
- Robinson MH, Spruce L, Eeles R et al. Limb function following conservation treatment of adult soft tissue sarcoma. *Eur J Cancer* 1991;27:1567-74.
- 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;(286):241-6.
- Davis AM, Wright JG, Williams JJ et al. Development of a measure of physical function for patients with bone and soft tissue sarcoma. *Qual Life Res* 1996;5:508-16.
- Davis AM, Bell RS, Badley EM et al. Evaluating functional outcome in patients with lower extremity sarcoma. *Clin Orthop Relat Res* 1999;(358):90-100.
- 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;46:1417-32.
- www.sarkom.dk.
- 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.
- Koller M, Aaronson NK, Blazeby J et al. Translation procedures for standardised quality of life questionnaires: The European Organisation for Research and Treatment of Cancer (EORTC) approach. *Eur J Cancer* 2007;43:1810-20.
- Cronbach L. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297-334.
- 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.
- Murena L, Vulcano E, D'Angelo F et al. Italian cross-cultural adaptation and validation of the Oxford Shoulder Score. *J Shoulder Elbow Surg* 2010;19:335-41.
- Clayer M, Doyle S, Sangha N et al. The Toronto Extremity Salvage Score in unoperated controls: an age, gender, and country comparison. *Sarcoma* 2012;2012:717213.