A patient-reported outcome measure for patients with Dupuytren's disease

Lærke Lannov Bendixen, Rasmus Wejnold Jørgensen & Claus Hjorth Jensen

ABSTRACT

INTRODUCTION: A Danish validated patient-reported outcome measure (PROM) specific to Dupuytren's disease (DD) does currently not exist. Such a PROM would be useful in the assessment of DD patients. The aim of this study was to translate the Southampton Dupuytren's Scoring Scheme (SDSS) into Danish and validate the translated version of the PROM.
METHODS: The SDSS was translated according to guidelines by Swaine-Verdier et al. A total of 110 patients diagnosed with DD completed the translated SDSS and evaluated the PROM. The severity of DD was assessed clinically using the Tubiana classification method. In all, 16 patients treated with collagenase injections were re-evaluated with the SDSS and the Tubiana post treatment. The reliability of the SDSS was tested in terms of internal consistency expressed as Cronbach's alpha and test-retest expressed by the intraclass correlation coefficient (ICC). Construct validation was calculated by Spearman's correlation coefficient between the SDSS and the Tubiana, and the responsiveness of the PROM was tested using point-biserial correlation and standardised response mean (SRM).
RESULTS: Cronbach's alpha was 0.76 and an ICC of 0.82 (95% confidence interval (CI): 0.53-0.93, p < 0.0001). Spearman's correlation coefficient was 0.26 (95% CI: 0.08-0.43, p = 0.007) preoperatively and 0.39 (95% CI: −0.13-0.74, p = 0.14) post treatment. The SRM was 1.96 (95% CI: 1.42-2.48), and the point-biserial correlation coefficient was 0.52 (95% CI: 0.03-0.81, p = 0.039).
CONCLUSIONS: The PROM shows good reliability, has substantial responsiveness to change and enjoys a high level of patient approval. The PROM is therefore recommended for patients with DD.
FUNDING: none.
TRIAL REGISTRATION: The study was reported to the Danish Data Protection Agency.

Dupuytren's disease (DD) is a benign fibroproliferative disease affecting the tissues of the palmar fascia leading to extension deficit and ultimately to impaired function of the affected hand [1]. The disease is most frequent among Caucasians above the age of 50 years, making it very common in Western countries including Denmark with a reported prevalence of 12% at the age 55 years, increasing with age [2, 3]. Patients with the condition may need treatment in the form of partial fasciectomy, percutaneous needle fasciotomy or collagenase injections.

Patient-reported outcome measures (PROM) are questionnaires designed to measure the patients' perception of an aspect of health or health-related quality of life, in this instance disease severity and treatment outcome [4]. Due to the fact that the patients' wellbeing and disease severity assessed by clinical measures do not always correlate, PROMs are of great significance in the overall assessment of the patients and may influence the choice of treatment. In addition, PROMs are easily performed and do not require the patient to be present at the hospital. To the authors' knowledge, a Danish validated PROM specific to DD currently does not exist.

The Southampton Dupuytren's Scoring Scheme (SDSS) is a PROM developed in 2014 designed to assess the degree of disability caused by DD [5]. The SDSS consists of five items concerning physical discomfort and impaired function of the hand in relation to various activities, which are divided into four groups; personal activities, domestic activities, hobbies, and work-related activities. The patient is asked to indicate to what degree he/she finds that the listed items are problematic due to DD on a scale consisting of five response categories. The score ranges from 0 to 20 – 0 being no discomfort or functional limitation and 20 being the worst possible discomfort and lack of hand function.

The original SDSS was validated through comparison with the Quick Disabilities of the Arm, Shoulder and Hand, also known as the QuickDASH (non-DD specific questionnaire). However, previous studies have found that the QuickDASH is inadequate for assessment of DD [6].

The Tubiana score is a clinician-based objective score of DD severity. The Tubiana method systematically evaluates extension deficit of each digital ray and sums to a score ranging from 0 to 23 – 0 being no extension deficit and 23 being severe extension deficit of all digital joints [7].

The aim of this study was to translate and validate the SDSS to provide a useful tool in the assessment of DD in Danish patients.

METHODS
Translation
The SDSS was translated in accordance with guidelines by Swaine-Verdier et al [8]. The approach taken in in the guideline is that although the forward-backward

translation method, generally considered the gold standard, provides translations that are linguistically similar to the original, there might not be conceptual equivalence between the original and the translated version. The guideline offers an alternative method, which was followed in this study.

The SDSS was initially translated by five clinicians employed at the orthopaedic department and an English professional with no healthcare-related education. The focus of the translation process was conceptual equivalence, acceptability of wording and accessibility. The translated version of the SDSS is presented in **Figure 1**.

Patients

A total of 110 patients attending the Department of Orthopaedic Surgery, Gentofte Hospital, Denmark, between February and May 2018 were included in the study. The eligibility criteria were a diagnosis of DD, either primary or recurrent, and the ability to adequately read and understand Danish. The exclusion criteria were other pathology in the ipsilateral hand. The study population is presented in **Table 1**. In relation to the preliminary consultation, patients were asked to fill out the SDSS with regards to the affected hand. At the same occasion, the extension deficit caused by DD was evaluated according to the Tubiana classification method by a medical student. A finger goniometer was used to measure the extension deficit. Furthermore, patients were asked to evaluate the questionnaire by indicating if the questions accurately and adequately reflected the problems associated with DD.

In all, 16 out of the 35 patients treated with collagenase injections were asked to attend a follow-up examination a minimum of four weeks post treatment –

once again they were asked to fill out the SDSS and the DD was evaluated again by the same medical student using the Tubiana system.

Statistics

The psychometric properties of the SDSS were evaluated in terms of internal consistency, reproducibility, responsiveness, construct validity and floor-ceiling effects [9, 10].

Reproducibility was assessed by the test-retest method [11]. A total of 20 patients re-completed the SDSS a minimum of seven days after the initial evaluation.

Since no treatment was given during this period, the assumption was that the DD remained stable between test and retest. The results were expressed by the intraclass correlation coefficient (ICC).

The internal consistency of the translated SDSS was analysed using Cronbach's alpha. Cronbach's alpha tests if all items in an instrument measure the same construct and is expressed as a test score between 0 and 1. A score between 0.7 and 0.95 is considered indicative of a good internal consistency. Test scores above 0.95 indicate that some items might measure the same parameter and therefore are superfluous.

As the true value of the SDSS construct is unknown, it is best validated through comparison with measurements that are believed to capture a similar construct. Spearman's rank correlation coefficient was calculated to investigate a possible correlation between the Tubiana and the SDSS, pre- as well as post-operatively. Spearman's rank correlation coefficient was chosen because a Shapiro-Wilk test showed that the data did not fit a normal distribution.

FIGURE 1 / The Danish version of The Southampton Dupuytren Scoring Scheme.

Angiv venligst, hvordan sygdommen påvirker dig indenfor hvert af følgende områder:

Hvor stort et problem har du med?	Intet problem (0 point)	Mindre problem (1 point)	Moderat problem (2 point)	Betydende problem (3 point)	Alvorligt problem (4 point)	Samlet score (udfyldes af personale)
Ubehag, ømhed og smerte						
Personlige aktiviteter F.eks. at vaske ansigtet, påklædning, håndvask, hårvask, at tage handsker på						
Hjemlige aktiviteter F.eks. at holde på et glas/en kop, åbne skruelåg, spise, lave mad						
Arbejds- og sociale sammenhænge F.eks. at anvende en computer, skrive i hånden, give hånd, kosmetisk udseende						
Fritid og hobby F.eks. at køre bil, cykle, ketsjersport, gør-det-selv-arbejde, at spille på musikinstrument, havearbejde						
I alt[a] (udfyldes af personalet)						

a) Mindste score = 0; højeste score = 20

TABLE 1 / Demographics of the study population (N = 110).

	n	Median (range)
Men	79	
Women	31	
Age, yrs		70 (29-96)
Primary	67	
Recurrence	43	
No treatment	24	
Collagenase injections	35	
Needle fasciotomy	10	
Partial fasciectomy	40	

In order to determine if the SDSS is able to respond to changes in the clinical condition, the standardised response mean (SRM) and a point-biserial correlation were calculated to describe internal and external responsiveness, respectively [12, 13]. Internal responsiveness is defined as the ability of a measurement to detect change over a prespecified period of time. External responsiveness is defined as the ability of a measurement to detect change corresponding to changes in a reference measure of health over a prespecified period of time. The SRM is a ratio between the mean change in test scores and the standard deviation of the change scores. SRM values of 0.2, 0.5, and 0.8 or more are considered indicative of small, moderate and large internal responsiveness, respectively. Pearson's correlation coefficient and point-biserial correlation were used to test for external responsiveness.

All of the correlations mentioned above are interpreted as follows: 0-0.10 negligible correlation, 0.10-0.39 weak, 0.40-0.69 moderate, 0.70-0.89 strong and 0.90-0.1 very strong correlation [14].

The "floor and ceiling effects" were calculated to determine if the SDSS measures the full range of the score distribution. The floor was defined as SDSS scores in the 0-1 range, and the ceiling as scores in the 19-20 range. A floor or ceiling effect is present in cases where 15% or more of the patients have scores corresponding to the floor or ceiling [15].

SPSS Statistics version 25.0 was used for statistical analysis and $p < 0.05$ was considered significant.

Trial registration: The study was reported to the Danish Data Protection Agency.

RESULTS

The overall Cronbach's alpha was 0.76. No item removal increased Cronbach's alpha. The test-retest resulted in an ICC of 0.82 (95% confidence interval (CI): 0.53-0.93, $p < 0.0001$).

The construct validation test between the Tubiana and the SDSS showed a Spearman's rank correlation coefficient of 0.26 (95% CI: 0.08-0.43, $p = 0.007$) pre-operatively and 0.39 (95% CI: –0.13-0.74, $p = 0.14$) post treatment.

A total of 15 out of 16 patients showed improvement in the SDSS post-operatively. The SRM was 1.96 with a 95% CI of 1.42-2.48.

It was not possible to detect a linear relationship between the changes in the SDSS and the Tubiana-Pearson's correlation coefficient was 0.17 (95% CI: –0.36-0.61, $p = 0.53$). However, when the data from the Tubiana were transformed into the dichotomous responses "Improvement" and "No improvement", a point-biserial correlation coefficient of 0.52 (95% CI: 0.03-0.81, $p = 0.039$) was calculated.

No floor or ceiling effect was observed preoperatively. Only 5.45% of the patients had SDDS scores in the 0-1 range and 2.72% had scores in the 19-20 range.

Post treatment no ceiling effect was observed; however, 75% of the patients had SDSS scores of 0-1.

When asked if the patients had experienced any inconveniences or problems due to DD that were not listed in the SDSS, 81 out of 110 (73.6%) patients responded "No". The patients who responded "Yes" were asked to indicate what those inconveniences might be. The results are listed in **Table 2**. When asked if any of the questions in the SDSS were irrelevant to DD, 109 out of 110 responded "No".

DISCUSSION

PROMs have grown in popularity over the past few years owing to their high compliance, low cost and their ability to provide information about the patients' perception of disease both pre- and post-operatively. The authors have chosen to translate and validate the Dupuytren-specific SDSS to provide a tool for the assessment of DD patients, and the results are promising.

With a Cronbach's alpha of 0.76 and an ICC of 0.82, the SDSS showed good reliability, meaning that all questions are relevant in measuring the same construct, and that patients give similar answers when asked to complete the questionnaire on two different occasions; provided, of course, that the disease has remained stable in the interval separating the two evaluations. The results are consistent with those of the original version [5].

The SDSS also showed considerable promise regarding responsiveness to change. An SRM of 1.96 is considered equivalent to a large internal responsiveness, and while it was not possible to detect a linear correlation between the changes in Tubiana and SDSS, the SDSS test scores decreased whenever a clinical improvement was observed and remained the same in cases where no change in the clinical condition was observed. The point-biserial correlation coefficient of 0.520 implied a moderate correlation between changes in the SDSS and changes in the Tubiana.

TABLE 2 / Individual supplementary answers.

Discomfort
Indirect tension and pain
Buzzing in the hands
Numbness
Personal activities
Put on face cream/body lotion
Hand in pockets
Toileting
Domestic activities
Easily dropping things
Reach into glove compartment
Work and social
Clapping
Carrying grandchild: infant
Interaction with others: works at a day-care centre
Hobby
Swimming, petanque, yoga
Fishing and hunting
Mixing cards
Other
Falling: trouble preventing a fall by holding onto things, pain when trying to brace the fall
Place hand flat on table
Fine motor skills
Stretch the hand
Hold onto railing/handles

73.6% of the patients did not propose any additions to the questionnaire when asked if they had experienced any inconveniences or problems due to DD not presented in the PROM. Among the patients who replied "Yes", most answers were contained in the categories already presented in the SDSS with a few exceptions such as holding onto a railing and problems or pain in relation to falling. The results imply that the SDSS does, indeed, adequately and accurately describe the patients' perception of DD.

The Spearman's rank correlation coefficient was calculated for the SDSS and the Tubiana before and after treatment. The results clearly indicate a poor, if any, correlation between the SDSS and the Tubiana. However, this does not mean that the SDSS does not have a valid place in clinical practice. It is widely accepted that objective clinical measures and subjective patient-reported measures do not always correlate due to the fact that they essentially measure different constructs [4]. A PROM with high correlation to a clinical objective measure is valuable if the goal is to have the PROM replace the objective clinical measure. If the goal is to gain additional information, then a strong correlation is undesirable.

In addition, it may be argued that the SDSS exceeds the Tubiana score when it comes to examining hand function. The Tubiana score is quickly increased if several fingers are affected, even if the extension deficit in each finger is modest. The increase in the Tubiana score is the same regardless of which joint is affected, although not all joints contribute equally to hand function. Moreover, the Tubiana does not take into account the pain and discomfort that some DD patients experience, which can also lead to functional limitation.

No floor or ceiling effect was detected preoperatively. Post treatment, a floor effect was detected. However, these calculations were based on 16 patients only. To determine whether the SDSS does have a floor effect, a larger number of DD patients should be re-evaluated post treatment. Furthermore, it is possible that the general treatment outcomes of collagenase injections are so excellent, that a floor effect cannot be avoided.

Although the results show great promise, there are certain limitations to the study. Only patients receiving collagenase injections were evaluated after treatment and only 16 patients participated in the follow-up. It is therefore not justifiable to conclude anything about the ability of the SDSS to show responsiveness to change in other treatments, i.e., percutaneous needle fasciotomy and fasciectomy. It is not possible to draw conclusions about the ability of the SDSS to detect changes in situations in which the clinical condition is exacerbated either due to disease progression or complications in relation to treatment. Secondly, the interval between test and retest varied from a week to a month and only 20 patients participated in the retest. The study has validated the SDSS using classic test theory. However, these methods have shortcomings, and more complex approaches such as confirmatory factor analyses or Item response theory would add to the validation process and should be a subject for future research [16].

The strengths of the study include the fact that the PROM has been analysed on several parameters including reliability, construct validation and responsiveness. The sample size exceeded the authors' expectations,

Clinical presentation of a patient with Dupuytren's contracture in the fourth and fifth digits.

and the evaluation of the extension deficit was performed by the same person using a finger goniometer – previous studies have shown good intraobserver agreement for classifying Tubiana [17].

Compared with other PROMs such as the URAM [18] and QuickDASH, the SDSS contains more text, which could be a matter of concern if patients are less likely to read the questions thoroughly. However, since the patient-reported evaluation of the PROM showed excellent results, there is no need to edit the SDSS.

CONCLUSIONS

The SDSS PROM shows good reliability, has substantial responsiveness to change and enjoys a high level of patient-approval and is therefore recommended for use in patients with DD. Although it is uncertain if the SDSS should influence the choice of treatment, the PROM does provide extensive information about disease severity and hand function and therefore has a place in clinical practice and research.

LITERATURE

1. Shih B, Bayat A. Scientific understanding and clinical management of Dupuytren disease. Nat Rev Rheumatol 2010;6:715-26.
2. Hindocha S, McGrouther DA, Bayat A. Epidemiological evaluation of Dupuytren's disease incidence and prevalence rates in relation to etiology. Hand (NY) 2009;4:256-69.
3. Wolfe S, Hotchkiss R, Kozin S et al. Green's Operative Hand Surgery. 7th ed. Philadelphia: Elsevier, 2017.
4. Suk M HB, Norvell DC, Helfelt DL. Musculoskeletal outcome measures and instruments. Switzerland: AO Publishing, 2005.
5. Mohan A, Vadher J, Ismail H et al. The Southampton Dupuytren's Scoring Scheme. J Plast Surg Hand Surg 2014;48:28-33.
6. Budd HR, Larson D, Chojnowski A et al. The QuickDASH score: a patient-reported outcome measure for Dupuytren's surgery. J Hand Ther 2011;24:15-20; quiz 1.
7. Tubiana R. Evaluation of deformities in Dupuytren's disease. Ann Chir Main 1986;5:5-11.
8. Swaine-Verdier A, Doward LC, Hagell P et al. Adapting quality of life instruments. Value Health 2004;7(suppl 1):S27-S30.
9. Mokkink LB, Terwee CB, Knol DL et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol 2010;10:22.
10. Mokkink LB, Terwee CB, Patrick DL et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010;19:539-49.
11. Kottner J, Audige L, Brorson S et al. Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. J Clin Epidemiol 2011;64:96-106.
12. de Yebenes Prous MJ, Rodriguez Salvanes F, Carmona OL. Responsiveness of outcome measures. Reumatol Clin 2008;4:240-7.
13. Husted JA, Cook RJ, Farewell VT et al. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000;53:459-68.
14. Schober P, Boer C, Schwarte L. Correlation coefficients: appropriate use and interpretation. Anesth Analg 2018;126:1763-8.
15. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res 1995;4:293-307.
16. Mokkink LB, Prinsen CA, Patrick DL et al. COSMIN study design checklist for patient-reported outcome measurement instruments. COSMIN, 2019. www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf (20 Dec 2019).
17. Broekstra DC, Lanting R, Werker PM et al. Intra- and inter-observer agreement on diagnosis of Dupuytren disease, measurements of severity of contracture, and disease extent. Man Ther 2015;20:580-6.
18. Beaudreuil J, Allard A, Zerkak D et al. Unite Rhumatologique des Affections de la Main (URAM) scale: development and validation of a tool to assess Dupuytren's disease-specific disability. Arthritis Care Res 2011;63:1448-55.