

ORIGINAL ARTICLE

The Southampton Dupuytren's Scoring Scheme

Arvind Mohan¹, Jane Vadher², Hiba Ismail³ & David Warwick⁴

¹Orthopaedic Registrar, Department of Orthopaedics, Southampton University Hospitals, Southampton, UK, ²Hand Therapist, Department of Orthopaedics, Southampton University Hospitals, Southampton, UK, ³Medical Statistician, Department of Orthopaedics, Southampton University Hospitals, Southampton, UK and ⁴Reader in Orthopaedics, Consultant Hand Surgeon, Department of Orthopaedics, Southampton University Hospitals, Southampton, UK

Abstract

The aim of this study was to construct and validate a simple patient-related outcome score to quantify the disability caused by Dupuytren's disease (DD), thus enabling prioritisation of treatment, to allow reliable audit of surgical outcome and to support future research. The Southampton Dupuytren's Scoring System (SDSS) was developed in a staged fashion according to the recommendations of The Derby Outcomes Conference. (1) Item generation; (2) Item reduction; (3) Internal consistency; (4) Test–re-test; (5) Field management; (6) Sensitivity to change standardised response mean; and (7) Criterion validity: ability of the SDSS to measure what it is supposed to measure. Internal consistency measured with Cronbach's alpha indicated acceptable reliability. The test–re-test correlation coefficient showed high reliability with SDSS. Field-testing showed SDSS ratings to be higher than the QuickDASH (Disability of the arm, shoulder and hand) ratings evaluated by the patients who answered both questionnaires. Standardised response mean was more sensitive for SDSS compared with QuickDASH showing sensitivity to change. Criterion validity was used to assess if the SDSS was measuring what it is supposed to measure comparing the SDSS with QuickDASH. A highly significant correlation was found between the two scoring systems. SDSS is a disease-specific patient-related outcome measure with a good internal consistency and performs better than QuickDASH in terms of test–re-test reliability and sensitivity to change. SDSS shows better field-testing attributes suggesting that it is a relatively more patient and practitioner friendly scoring system. This study proposes to the SDSS is a useful patient-related outcome measure for DD.

Key Words: Dupuytren's contracture, Southampton Dupuytren's scoring scheme, quick DASH

Introduction

Surgeons must show the outcome of their procedures for many reasons, to include improving patient care, facilitating clinical research, and fulfilling the surgeon's professional responsibilities. In addition, patient-related outcome measures may influence both a patient's choice of surgeon and the allocation of healthcare resources. Various scoring schemes have been developed to evaluate the results of hand surgery procedures. Most of these are limb-specific and only a few are specific to certain hand conditions. DASH (Disability of the arm, shoulder and hand) [1], QuickDASH [2], Michigan Hand Score [3], and Patient Evaluation Measure [4] are all limb-specific; the Boston Carpal Tunnel Questionnaire [5] and the Mayo Wrist Scores [6] are related to individual conditions or joints.

Over the past decade, patient-based questionnaires have been generally accepted as an important addition to traditional physician-based measurements because the data gleaned are *subjective* and thereby relevant to the patient's perception of success or otherwise—this ultimately being the purpose of treatment. Furthermore, patient-compiled questionnaires are easier to administer than objective schemes, as they can be completed by post or telephone, thus increasing follow-up rates.

However, most surgeons would also expect a measurable *objective* outcome such as range of movement or radiological features to judge their technical success and to allow comparison with other methods of treatment.

Dupuytren's disease (DD) is a condition commonly leading to operation. However, there is no specific scoring system for assessing either disability or the outcome of treatment in these patients. Scoring systems like DASH and QuickDASH, whilst extensively used, are not specific for patients with DD. Because many of the sections within these schemes (e.g. pain, tingling, sleep disturbance) are not reflected in patients with DD, the scheme is likely to produce a low score in someone nevertheless greatly troubled by their condition. This low score can only be improved by a small proportion, even if the patient feels that the operation has been a great success for their individual problem: the scheme has poor sensitivity to change. If the outcome measure has such poor sensitivity to change, then the purpose of outcome measurement is not fulfilled.

There is no subjective scoring system which is specific for DD. The aim of this study was to construct a simple self-administered scheme to quantify the disability caused by the disease which is sensitive to change, thus to enable prioritisation of treatment, allow reliable audit of surgical outcome, and support future research.

Patient and methods

The Southampton Dupuytren's Scoring System (SDSS) (Figure 1) was developed in a staged fashion according to the recommendations of Derby Outcomes Conference [7]. Approval for this study was obtained from the Ethics Committee in our institution.

Southampton dupuytren's scoring scheme					
Name					
Address.....					
Hospital no					
Hand dominance					
Which fingers are affected					
<i>Please indicate how the condition affects you in each of the following areas:</i>					
How much trouble do you have with:	No problem	Minor inconvenience	Modest inconvenience	Definitely troublesome	Severe problem
Discomfort					
Personal activities, eg: washing face, dressing, washing hands, washing hair, putting on gloves.					
Domestic activities, eg; holding a glass/cup, opening jars, eating, cooking.					
Work / Social interaction, eg: using the computer, writing, shaking hands, cosmetic appearance.					
Hobbies, eg, driving/cycling, racket sports, DIY, playing musical instruments, gardening.					
Score (staff to complete)					
Total					

Figure 1. The Southampton scoring scheme. No problem = 0; Minor inconvenience = 1; Modest inconvenience = 2; Definitely troublesome = 3; Severe problem = 4; Minimum score = 0; Maximum score = 20.

Stage 1: Item generation

Twenty patients with an established diagnosis of DD attending the Dupuytren’s clinic (before surgical intervention) were invited to participate and fill in a questionnaire, listing any problems, and scoring them on a 5-point scale (no problem, mild inconvenience, modest inconvenience, definitely troublesome, severe problem). Patients with a history of trauma, infection, stroke, or any other operation on the finger other than for Dupuytren’s contracture, were excluded from this study.

Stage 2: Item reduction

We reduced the number of questions by taking the most frequent problems from the 20 respondents and then grouping them into five different categories. Each question was scored from 0–4, with a score of 0 for no problem and 4 for severe problem. The SDSS was thereby generated with a total score from 0–20.

Stage 3: Pre-testing

Internal consistency of the scoring system was measured using Cronbach’s alpha. Cronbach’s alpha provides an indication of the average correlation between the items that make up the scale.

The *test–re-test* was used to test the reliability of the SDSS. Patients with an established diagnosis of DD attending the Dupuytren’s clinic (before surgical intervention) were included in this study and patients with a history of trauma, infection, stroke, or any other operation on the finger other than for Dupuytren’s contracture were excluded from this study.

Sixty-one patients scored the SDSS in clinic and then 3 weeks later by telephone for an intra-observer variation analysis. A QuickDASH was also completed to examine the differences between a standard upper limb score and the SDSS disease-specific score. A 3-week period is preferable to 2 weeks as there is a smaller memory effect than at 2 weeks and there is no likelihood of the underlying condition changing within 3 weeks,

Table I. Field-testing results of SDSS and QuickDASH scoring systems.

	Mean			95% CI of the difference		Sig. (2-tailed) <i>p</i> -value
	SDSS	QuickDASH	Paired differences	Lower	Upper	
SDSS vs QuickDASH						
Clarity/easy to understand	9.5	7.6	1.9	1.0	2.8	< 0.05
Applicable questions	8.6	6	2.6	1.7	3.5	< 0.05
Completed satisfactorily	9.6	7.5	2.1	1.4	2.8	< 0.05

as recommended by Fitzpatrick et al. [8]. The first questionnaire was filled in by the patient in the clinic and the second was filled in through a telephone interview.

Stage 4: Field management

This stage involved a field trial, whereby 10 patients filled in a questionnaire. These were all pre-surgical cases. The patients were interviewed after completion to determine the clarity and applicability of questions as well as patient performance in completing the questionnaire. These variables were scored on a scale from 0–10.

Stage 5: Testing attributes

Sensitivity to change was assessed by the standardised response mean (SRM). SRM was calculated by dividing the mean difference of the two readings by the standard deviation of the difference. This creates a standardised score so that two instruments, originally on different scales, can be compared, as suggested by Liang et al. [9]. All patients scored the postoperative SDSS and QuickDASH score over the telephone ~ 6 months after the operation.

In order to assess whether the SDSS is measuring what it is supposed to measure (criterion validity), we compared the SDSS with another measure of function, the QuickDASH. We chose the QuickDASH, as it is much easier to complete than DASH, yet validity is (at least in general, if not for DD) unchanged, as shown by Beaton et al. [2]. QuickDASH generates a total score of 100.

Clinical measurement

The deformity was measured with a goniometer and a total extension loss (summation of fixed flexion at the MCPJ, PIPJ, and DIPJ) was calculated. If more than one digit was operated on, the worst deformity was included for this study. The total extension loss was also measured after the operation and used as the basis for the change in deformity.

Statistical methods

Internal consistency was assessed using Cronbach's alpha [10]. The *reliability* of the SSDS was tested using the test–re-test. The paired sample *t*-test was used for the *field-testing*. *Sensitivity to change* was assessed by the standardised response mean (SRM). A negative SRM expresses an improvement in the patient's symptoms, and the greater (more negative) the SRM, the greater the improvement. *Criterion validity* was used to assess if the SDSS was measuring what it was supposed to be measuring. *Pearson's correlation* was used to assess any correlation between the questionnaire score and the degree of preoperative and postoperative deformity.

Results

Internal consistency

The overall Cronbach's alpha was 0.8739 (Cronbach's alpha of 0.8–0.9 indicates good reliability); this score also shows that the questions are not too similar, since a Cronbach's alpha of 0.95 and higher shows that some items are redundant.

Test–re-test reliability

The test–re-test correlation coefficient was 0.789 between the scores at a 3-week interval, indicating that a high degree of reliability was found between SDSS measurements.

Field-testing

Table I shows the mean of each variable for the SDSS and for the QuickDASH. The 0–10 scale is an arbitrary range where 0 indicates a poorly designed and understandable questionnaire and 10 reflects one that is most easily understandable and clear for the patients. The SDSS ratings were found to be higher than the QuickDASH ratings evaluated by the patients who answered both questionnaires (paired *t*-test, $p < 0.05$).

Sensitivity to change

Table II shows standardised response mean (SRM) 6 months after the operation was significantly more sensitive for SDSS than QuickDASH.

Criterion validity

In order to assess if the SDSS is measuring what it is supposed to measure, we compared the SDSS with another measure of a similar contrast in the same study. In our case the SDSS was correlated with QuickDASH. A highly significant correlation (Pearson correlation = 0.598) was found between the two scoring systems, which indicates the validity of the SDSS, measuring what it is supposed to measure as shown by Lü and Fang [11].

Score and relation to degree of deformity

The SDSS score and QuickDASH score before the operation (Figures 2a and 3a) correlated poorly with the degree of

Table II. Sensitivity to change.

	SDSS	QuickDASH
Mean score change	–4.7	–9.5
Standard deviation	2.7	7.9
SRM	–1.8	–1.2
<i>p</i> -value	< 0.05	< 0.05
95% Confidence interval	–2.1 to –1.3	–1.8 to 0.4

The *p*-value indicates the level of significance for changes between 6 months follow-up vs baseline (BL) data using a paired *t*-test.

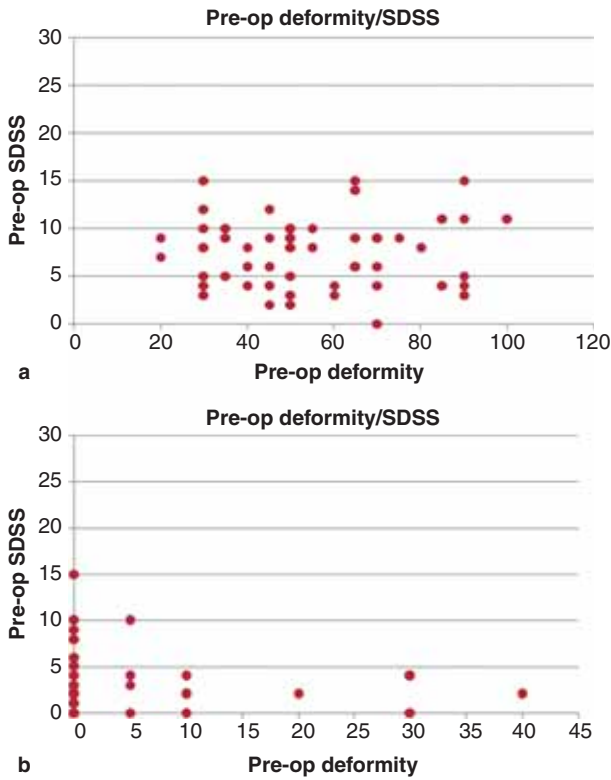


Figure 2. (a, b) Correlation between preoperative and postoperative deformity with the respective SDSS Score. Deformity = Total extension deficit. x-axis: Preoperative and postoperative deformities of the finger; y-axis: Preoperative and postoperative SDSS score.

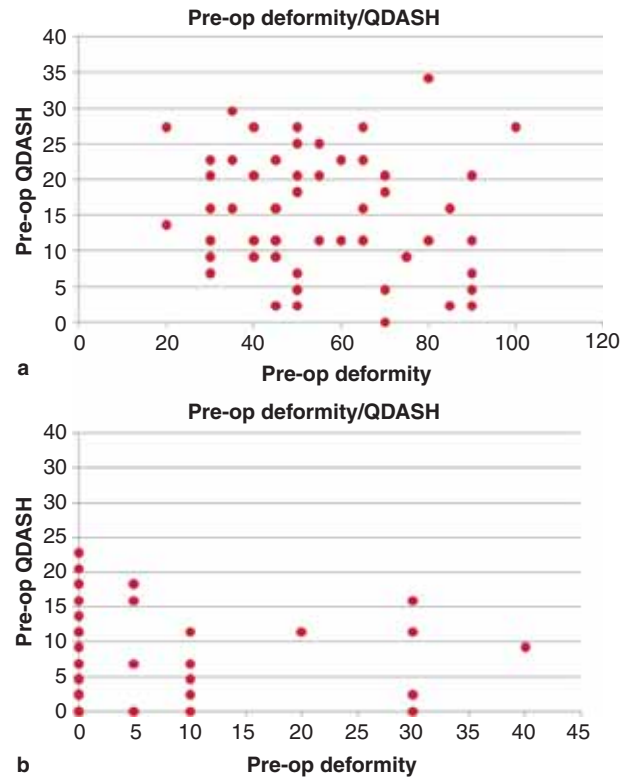


Figure 3. (a, b) Correlation between preoperative and postoperative deformity with the respective QuickDASH Score. Deformity = Total extension deficit. QDASH = QuickDASH. x-axis: Preoperative and postoperative deformities of the finger; y-axis: Preoperative and postoperative QuickDASH score.

deformity. As demonstrated by (Figure 2a) (SDSS and deformity) and (Figure 3a) (QuickDASH and deformity) there is no correlation between the questionnaire’s score and the degree of preoperative deformity. For example, a small degree of deformity can affect the patient significantly such that he/she will score highly in either questionnaire. Conversely, some patients had a high degree of deformity but had a low score, indicating that the deformity did not greatly impact on their activities of daily living. After the operation, both SDSS and

QuickDASH scores improved considerably with better sensitivity to change for the SDSS (Figure 4), but again there was poor correlation between residual deformity and function (Figures 2b and 3b).

Discussion

Various scoring systems for specific hand conditions have been developed in the past. These include subjective, objective, and mixed scoring systems. The URAM scale is a disease-

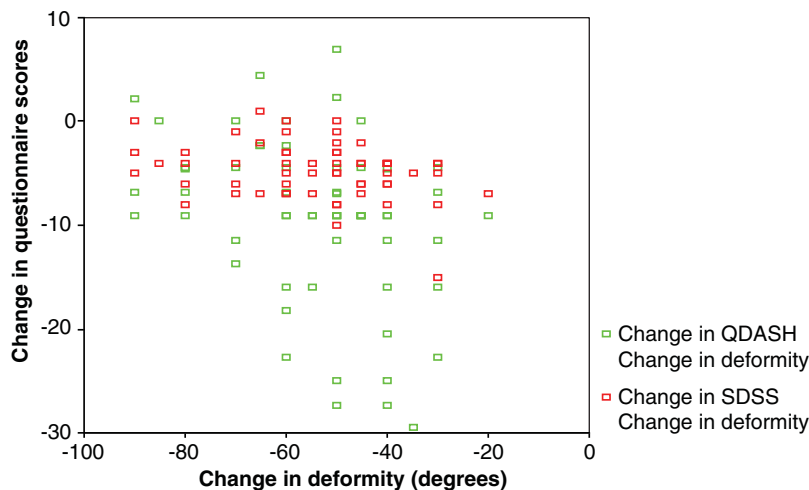


Figure 4. Correlation between change in deformity and change in SDSS and QuickDASH. Correlation coefficient between the change in SDSS and the change in deformity is -0.28 . Correlation coefficient between the change in QuickDASH and the change in deformity is -0.36 . x-axis: Change in deformity (degrees); y-axis: Change in questionnaire scores (SDSS and Quick DASH).

specific scoring system for DD [12]. This has recently been reported and involves medical experts involved in question generation. Traditionally, global upper limb schemes like the QuickDASH scoring system have been used, the ability of which to assess individual hand problems may be diluted. The ICF (International Classification of Functioning, Disability and Health) Core Set development project for hand conditions was a cooperative effort between the German Social Accident Insurance (DGUV), the Institution for Statutory Accident Insurance and Prevention in the Health and Welfare Services (BGW) (Germany), and the ICF Research Branch. It is not a disease-specific score and consists of Comprehensive and Brief ICF core set for hand conditions, with 117 and 23 ICF categories, respectively [13].

SDSS is proposed as a disease-specific scoring system for DD. In this study we have analysed the SDSS and compared it with the QuickDASH. Both these scoring systems are fully subjective and take into consideration the patient's perception of the disease disability.

Our results suggest that, in a population (rather than an individual), the initial measured DD deformity does not correlate with disability as shown by QuickDASH or SDSS. This is an intriguing finding. Our findings are consistent with other authors [11,14,15], who similarly found no correlation between deformity and function using DASH, but found some improvement after the operation. Probably objectively-measured deformity does not capture the multi-factorial nature of disability. The result might to some extent be explained because we used just the worst-affect digit for calculation—in some patients there may have been other affected digits, themselves contributing to disability. However, studies have found that the number of fingers involved also had no effect on function [15,16]. Other scoring schemes have found that preoperative deformity correlated with the Sollerman score, as shown by Draviraj and Chakrabarti [17]. Some patients with no residual deformity after the operation still continued to have disability. The explanation is not clear; perhaps residual loss of flexion, problems elsewhere in the hand, dissatisfaction with the treatment process, or other general health issues could account for this. There has been an interest in analysing results using an anchor-based approach to determine a clinically important difference (CID) and range of motion and then link it to patient satisfaction [17].

However, we also showed that, once the deformity has been corrected, disability is improved, as shown on both SDSS and QuickDASH but with better sensitivity to change for QuickDASH.

Objective physician-based measurement of deformity, and change thereof, have the advantage of comparison without the potential bias intrinsic in patient-related outcome measures, wherein other factors such as the hospital environment, expectation management by the surgeon, the patient's health, and demeanour may influence the score.

There are certain limitations of this study, like the technique of item reduction was based on our observation of the disability caused by DD; an alternative method would have been to use open-ended questions to patients with DD and then reduce from that bank. Nevertheless, Cronbach's alpha which measures the internal consistency was 0.87; this represents good consistency, suggesting that the questions were not too similar. The SDSS

does not take into account hand dominance. We have grouped dissimilar concepts together within sub-headings (e.g. work, social, etc.) for ease of completion; alternative groupings may affect the discrimination of the SDSS. We cannot be sure that the different test-re-test environment (first in clinic and second by telephone) did not affect the result.

Since we have shown a poor correlation between the SDSS and deformity, a future study should examine this further. The effect of multiple digital involvement on hand function would be clarified—a much larger sample size is needed as multiple digit DD is not so common and there would be considerable variability depending on the digit involved and degree of deformity in the second or even third involved digit. This present study was underpowered to make a valid analysis.

Subjective measurements, whilst having the advantage of patient-orientated focus and easy data acquisition, nevertheless may be subject to other influences. Correlation with patient satisfaction scores and general health measures (e.g. SF-36) would provide further insights.

In summary, SDSS is a suitable disease-specific scoring system for DD. It has good internal consistency and performs better than QuickDASH in terms of test-re-test reliability and sensitivity to change. SDSS shows better field-testing attributes suggesting that it is a relatively more patient- and practitioner-friendly scoring system.

We propose the SDSS is a useful alternative scoring system for patients with DD. We also note that the poor correlation between deformity and function is intriguing and requires both further study and also consideration of the relative merits of objective and subjective outcome measures. When evaluating different forms of treatment, the surgeon will still have to rely on objectively-acquired range of movement to avoid the confounding factors in subjective measures; however, when examining one's own results, the patient's perspective—acquired with subjective patient-related outcome measurement—must also be considered.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

- [1] Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: The DASH (Disabilities of the Arm, Shoulder and Hand). *Am J Industrial Med* 1996;29: 602–8.
- [2] Beaton DE, Katz JN, Wright JG. Development of the Quick-DASH: comparison of three item reduction approaches. *J Bone Joint Surg Am* 2005;87:1038–46.
- [3] Chung KC, Pillsbury MS, Walters MR, Hayward RA. Reliability and validity testing of Michigan Hand Outcomes Questionnaire. *J Hand Surg* 1998;23A:575–87.
- [4] Macey AC, Burke FD. Outcomes of hand surgery. *J Hand Surg* 1995;20B:841–55.
- [5] Levine DW, Simmons BP, Koris MJ, et al. A self-administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome. *J Bone Joint Surg Am* 1993;75:1585–92.
- [6] Amadio PC, Berquist TH, Smith DK, et al. Scaphoid malunion. *J Hand Surg* 1989;14A:679–87.

- [7] Bindra RR, Dias JJ, Heras-Palau C, et al. Assessing outcome after hand surgery: the current state. *J Hand Surg* 2003;28B: 289–94.
- [8] Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;2:i–iv; 1–74.
- [9] Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care* 1990;28: 632–42.
- [10] Cronbach L. Coefficient Alpha and the internal structure of tests. *Psychometrika* 1951;16:297–333.
- [11] Lü Y, Fang J. *Advanced medical statistics*. World Scientific Publishing 2003;1st Ed;217–18.
- [12] 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 (Hoboken)* 2011;63:1448–55.
- [13] <http://www.icf-research-branch.org/icf-core-sets-projects/other-health-conditions/development-of-icf-core-sets-for-hand-conditions.html>. accessed 3 November 2012.
- [14] Degreef I, Vererfve PB, De Smet L. Effect of severity of Dupuytren contracture on disability. *Scand J Plast Reconstr Surg Hand Surg* 2009;43:41–2.
- [15] Engstrand C, Borén L, Liedberg GM. Evaluation of activity limitation and digital extension in Dupuytren’s contracture three months after fasciectomy and hand therapy interventions. *J Hand Ther* 2009;22:21–7.
- [16] Zyluk A, Jagielski W. The effect of the severity of the Dupuytren’s contracture on the function of the hand before and after surgery. *J Hand Surg Eur* 2007; 32:326–9.
- [17] Draviraj KP, Chakrabarti I. Functional outcome after surgery for Dupuytren’s contracture: a prospective study. *J Hand Surg* 2004;29A:804–8.