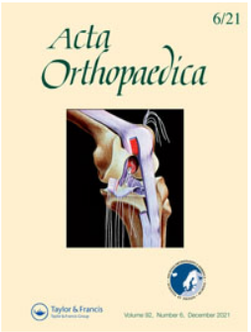

Acta Orthopaedica Scandinavica

ISSN: 0001-6470 (Print) (Online) Journal homepage: informahealthcare.com/journals/iort19

The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: Reliability and validity of the Swedish version evaluated in 176 patients

Isam Atroshi, Christina Gummesson, Bodil Andersson, Elsa Dahlgren & Anita Johansson

To cite this article: Isam Atroshi, Christina Gummesson, Bodil Andersson, Elsa Dahlgren & Anita Johansson (2000) The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: Reliability and validity of the Swedish version evaluated in 176 patients, Acta Orthopaedica Scandinavica, 71:6, 613-618, DOI: 10.1080/000164700317362262

To link to this article: https://doi.org/10.1080/000164700317362262

Published online: 08 Jul 2009.

The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire

Reliability and validity of the Swedish version evaluated in 176 patients

Isam Atroshi[1], Christina Gummesson[2], Bodil Andersson[3], Elsa Dahlgren[4] and Anita Johansson[4]

[1]Departments of Orthopedics, Hässleholm-Kristianstad Hospitals, SE-291 85 Kristianstad, Sweden. Tel +46 44131260. E-mail: isam.atroshi@mailbox.swipnet.se; and [2]Physical Therapy, Lund University, Lund, Sweden, [3]Departments of Orthopedics, Hässleholm-Kristianstad Hospitals, Kristianstad, Sweden, and [4]Primary Care, Eslöv, Sweden
Submitted 00-02-21. Accepted 00-07-30

ABSTRACT – The disabilities of the arm, shoulder and hand (DASH) questionnaire is a self-administered region-specific outcome instrument developed to measure upper-extremity disability and symptoms. The DASH consists mainly of a 30-item disability/symptom scale. We performed cross-cultural adaptation of the DASH to Swedish, using a process that included double forward and backward translations, expert and lay review, as well as field-testing to achieve linguistic and conceptual equivalence. The Swedish version's reliability and validity were then evaluated in 176 patients with upper-extremity conditions. The patients completed the DASH and SF-12 generic health questionnaire before elective surgery or physical therapy. Internal consistency of the DASH was high (Cronbach alpha 0.96). Test-retest reliability, evaluated in a subgroup of 67 patients who completed the DASH on two occasions, with a median interval of 7 days, was excellent (intraclass correlation coefficient 0.92). Construct validity was shown by a positive correlation of DASH scores with the SF-12 scores (worse upper-extremity disability correlating with worse general health), stronger correlation with the SF-12 physical than with the mental health component, correlation of worse DASH scores with worse self-rated global health, and ability to discriminate among conditions known to differ in severity. The Swedish version of the DASH is a reliable and valid instrument that can provide a standardized measure of patient-centered outcomes in upper-extremity musculoskeletal conditions.

■

Standardized validated questionnaires that measure treatment outcomes and health-related quality of life have become increasingly important in clinical research (Guyatt et al. 1993). These outcome instruments (usually self-administered questionnaires) consist of multi-item scales that measure specific health dimensions such as physical function or pain. Generic instruments such as the Short Form (SF)-36 (Ware and Sherbourne 1992, Sullivan et al. 1995) and its shorter version, the SF-12 (Ware et al. 1996), have been used as health-related quality of life measures in various musculoskeletal disorders. Because generic instruments may not be able to detect small but important changes related to specific disorders, disease- or joint-specific instruments have been created. However, development and use of disease-specific instrument for every musculoskeletal condition may not be practical (Davis et al. 1999, Swiontkowski et al. 1999). Region-specific outcome instruments have therefore been introduced.

The disabilities of the arm, shoulder and hand (DASH) outcome measure was developed by the American Academy of Orthopedic Surgeons as a region-specific instrument for measuring upper-extremity disability and symptoms (Hudak et al. 1996). The DASH has been shown to be reliable and valid in a patient population with elbow disorders (Turchin et al. 1998) and another with various upper-extremity disorders (McConnel et al. 1999).

Cross-cultural adaptation of validated outcome instruments has been advocated in order to facilitate their use in international multicenter clinical trials (Ware et al. 1995). This would also reduce the need for developing new instruments that have the same purpose (Deyo et al. 1994). To maintain the validity of the original instrument while taking into consideration important cultural differences, a specific methodology has been developed for the adaptation process (Guillemin et al. 1993, Ware et al. 1995, Lohr et al. 1996).

We have performed cross-cultural adaptation of the DASH to Swedish and evaluated the Swedish version's reliability and validity in patients with upper-extremity conditions.

Material and methods

The DASH questionnaire

The main part of the DASH is a 30-item disability/symptom scale concerning the patient's health status during the preceding week (McConnel et al. 1999). The items (Table 1) ask about the degree of difficulty in performing various physical activities because of an arm, shoulder or hand problem (21 items), the severity of each of the symptoms of pain, activity-related pain, tingling, weakness and stiffness (5 items), as well as the problem's effect on social activities, work, and sleep and its psychological impact (4 items). The procedure by which these particular items were selected followed established methods (Streiner and Norman 1995, McConnel et al. 1999). The DASH also contains two optional 4-item scales concerning the ability to perform sports and/or to play a musical instrument (sport/music scale), and the ability to work (work scale). Each item has 5 response choices, ranging from "no difficulty or no symptom" to "unable to perform activity or very severe symptom", and is scored on a 1- to 5-point scale. The scores for all items are then used to calculate a scale score ranging from 0 (no disability) to 100 (severest disability). The score for the disability/symptom scale is called the DASH score.

The adaptation process

The American version of the DASH was translated to Swedish by 2 bilingual translators whose first language was Swedish and with 1 having a medical background in both the United States and Sweden and the other having no medical background. These two "forward" translations were reviewed and discussed by the 2 translators and a synthesis of them was formed (differences were resolved by consensus). This version was translated back to English by 2 other bilingual translators whose first language was English. Both were blinded to the concepts being investigated and had no medical background. The translations were reviewed by members of a committee comprising one of the forward translators, an outcome methodologist, and two health professionals. The other translators were contacted when necessary. Discrepancies were resolved by consensus to achieve conceptual equivalence. A prefinal version was

Table 1. The items in the disabilities of the arm, shoulder and hand (DASH) disability/symptom scale

1	Opening a tight or new jar
2	Writing
3	Turning a key
4	Preparing a meal
5	Pushing open a heavy door
6	Placing an object on a shelf above the head
7	Doing heavy household chores [a]
8	Gardening or doing yard work
9	Making a bed
10	Carrying a shopping bag or briefcase
11	Carrying a heavy object (over 5 kg)
12	Changing a light bulb overhead
13	Washing or blowing dry the hair
14	Washing the back
15	Putting on a pullover sweater
16	Using a knife to cut food
17	Recreational activities that require little effort [a]
18	Recreational activities that require taking some force or impact through the arm, shoulder or hand [a]
19	Recreational activities that require moving the arm freely [a]
20	Managing transportation needs (getting from one place to another)
21	Sexual activities [b]
22	Social activities
23	Work and other daily activities
24	Pain
25	Pain when performing activities
26	Tingling
27	Weakness
28	Stiffness
29	Difficulty in sleeping
30	Impact on self-image

[a] Specific activities are given as examples
[b] Item unanswered by 10% of the patients (compared to 0–4% for the other items)

Table 2. Pretreatment scores in the disabilities of the arm, shoulder and hand (DASH) disability/symptom scale in different diagnostic groups

		DASH score [a]	
Group	n	mean (SD)	median
Shoulder disorder			
Surgical [b]	30	43 (15)	46
Nonsurgical	31	35 (20)	34
Tennis elbow	12 [c]	39 (14)	36
Carpal tunnel syndrome	25 [d]	40 (19)	34
Trapeziometacarpal arthrosis [e]	7	48 (16)	46
Tenosynovitis [e]	11	36 (15)	33
Wrist/hand ganglion [e]	12	11 (18)	5
Duputryn's disease [e]	16	21 (23)	9
Other	32	32 (23)	26

[a] Higher score (0-100) indicates greater disability
[b] Arthroscopic acromioplasty (except for one open acromioplasty)
[c] 4 surgical
[d] 22 surgical
[e] All surgical

created and subjected to field-testing on 26 patients (19 women), having a mean age of 55 (25–72) years, with different upper extremity conditions.

The final Swedish version of the DASH was then evaluated with regard to reliability and validity, using psychometric tests (Nunnally and Bernstein 1994, Lohr et al. 1996).

Patients

Patients with upper-extremity musculoskeletal conditions planned for surgical treatment at an orthopedic department, or for physical therapy at a primary care unit, were considered for inclusion in this study. Exclusion criteria were: age below 18 years, symptom duration of less than 2 months, or inability to complete questionnaires due to cognitive impairment or language difficulties. The DASH and the SF-12 generic health instrument were administered to 186 consecutive eligible patients. 10 patients were excluded because of having 4 or more unanswered items in the DASH disability/symptom scale. 176 patients (57% women), with a mean age of 52 (18–85) years, completed the questionnaires before elective surgery (n 125) or physical therapy (n 51) for upper-extremity conditions (Table 2).

Analyses

Internal consistency of each DASH scale was assessed with the Cronbach alpha. Test-retest reliability was analyzed in 67 of the patients (55% women), with a mean age of 51 (19–74) years, who completed the DASH on two occasions (prior to treatment) with a median interval of 7 (3–17) days. The one-sample t-test and intraclass correlation coefficient (ICC) were used for this analysis. The ICC analysis was then performed separately in the surgical (n 30) and nonsurgical patients, as well as in the patients who completed the questionnaires with a test-retest interval of less than 8 days (n 34) and those with an interval of 8 days or longer.

Validity of the DASH was assessed at different stages. Face and content validity (items' relevance and adequacy for the intended use) were judged by the experts and health professionals involved in the study and the patients who participated in the field-testing. These aspects of validity were also assessed in the clinical study by examining the completeness of item responses, the distribution of the scores, and the magnitude of ceiling and floor effects (i.e., proportion of best and worst possible scores, respectively). Criterion-related validation was not possible because of absence of a criterion standard for upper-extremity disability. Construct validity was assessed by testing several predefined hypotheses concerning the expected relationships between the DASH and other measures. First, the correlations between the DASH scores and the two (physical and mental health) component scores of the SF-12 were examined. It was hypothesized that the DASH scores would correlate positively with the SF-12 scores and that the correlations would be stronger with the physical health than with the mental health component scores. The Pearson correlation coefficient (r) was used for this analysis, which comprised 150 patients because 26 had at least one unanswered SF-12 item. Construct validity was also assessed by analyzing the relationship between the DASH scores and the responses to the SF-12 item regarding self-rated global health, which has 5 response choices (excellent, very good, good, fair, poor). It was hypothesized that the DASH would discriminate among different states of health and that worse DASH scores would correlate with worse

self-rated health. Analysis of variance was used for statistical testing. Finally, construct validity was assessed by analyzing the DASH scores according to a number of specific diagnoses. It was hypothesized that the DASH would discriminate among patient groups with diseases known to differ in severity.

The number of patients recruited was considered to be adequate in providing reliability and correlation coefficients with good precision.

Results

Adaptation

The questionnaire's instructions to responders as well as most of the items and response choices could be translated with little, if any difficulty. In a few items and response choices, the forward and backward translations had important discrepancies reflecting language-specific/cultural differences in expression. These could be resolved satisfactorily by using the item or response choice wording that gave the best conceptual equivalence. In 2 items regarding recreational activities, some of the activities provided as examples were judged to be infrequently performed in Sweden and were therefore replaced by other activities. In these 2 items and in 2 others (recreational activities and household chores), more examples of commonly performed activities were added. The prefinal version was judged to possess face and content validity as a measure of upper-extremity disability and symptoms.

The prefinal version performed well in field-testing. The patients stated that the items were clear and that most of them were relevant to their upper extremity condition. The average time taken by the patients to answer all items was about 10 minutes. The final version was endorsed by the American Academy of Orthopedic Surgeons as the official Swedish translation of the DASH.

Reliability

Internal consistency of the disability/symptom scale was high (Cronbach alpha 0.96). The item-total correlations were substantial for all items (range 0.42–0.82). Cronbach alpha was 0.94 for both the work and sport/music scales. Test-retest reliability analysis showed a mean score difference of 0.7 (95% CI -1.2–2.6, $p = 0.5$) and an ICC of 0.92 (95% CI 0.88–0.95) for the disability/symptom scale indicating excellent agreement between the scores. The ICC in the surgical patients was 0.93 (95% CI 0.86–0.97) and in the nonsurgical patients 0.91 (95% CI 0.84–0.96). The ICC was 0.90 (95% CI 0.81–0.95) in the patients who completed the questionnaires with an interval of less than 8 days and 0.94 (95% CI 0.88–0.97) in those with a test-retest interval of 8 days or longer.

Validity

Examination of the disability/symptom scale showed that completeness of item responses was good, with all items (except one) being answered by more than 95% of the patients (Table 1). Item responses had good distribution, with mean item scores ranging from 1.4 to 3.3. The mean DASH score for all the patients was 34 (SD 20), and the median score was 33 (range 0–78). Two patients had a best possible score, and 26 (15%) had scores below 10.

The DASH scores correlated positively with the SF-12 scores (worse upper-extremity disability correlating with worse general health), showing stronger correlation with the SF-12 physical health component scores ($r = 0.74$, 95% CI 0.66–0.81) than with the mental health component scores ($r = 0.51$, 95% CI 0.38–0.62). Analysis of the DASH scores, grouped according to the responses to the SF-12 item concerning self-rated global health, showed significant differences among the 5 groups, with worse DASH scores found in patients reporting worse health ($p < 0.001$). The DASH scores were worse in patients with shoulder disorders, tennis elbow, carpal tunnel syndrome, or trapeziometacarpal arthrosis, than in patients with wrist/hand ganglia or Dupuytren's disease (Table 2).

Discussion

In this study, Swedish adaptation of the DASH was performed following a systematic standardized approach. Similar approaches used in the translation of other health instruments, such as the SF-36, achieved good linguistic and conceptual

equivalence (Ware et al. 1995). The adaptation process can disclose important country-specific differences that, if not addressed, might influence the performance of a questionnaire and the interpretability of its results. Differences in questionnaire scores between populations in 2 countries should reflect true differences in health status rather than differences caused by the translation.

Our findings show that the Swedish version of the DASH is a reliable and valid region-specific outcome measure. It should be a valuable tool in clinical research of upper-extremity musculoskeletal disorders and for clinicians managing patients with such disorders. Upper-extremity disorders are associated with considerable health care and work disability costs (Feuerstein et al. 1998). It is therefore important to have standardized reliable and valid measures of their health burden and of the effectiveness of current and future treatment methods in improving patient-centered outcomes.

Although disease-specific instruments can be very sensitive in detecting health changes related to specific upper-extremity diseases (Atroshi et al. 1999), developing and administering a disease-specific instrument for every condition is neither practical nor necessary. There might be a place for disease-specific instruments in diseases that are very common (Swiontkowski et al. 1999). The DASH region-specific instrument can be used for a large number of conditions (Davis et al. 1999) and has shown higher responsiveness (sensitivity to health change) than generic instruments, when used for an upper-extremity disease (Kirkley et al. 1998). Because one of the main uses of the DASH is in the assessment of treatment benefits, the responsiveness of the Swedish version needs to be evaluated. This, as well as obtaining population norms for the DASH, will enhance the interpretability of the scores.

We found excellent reliability in this self-administered questionnaire. Many commonly used conventional physical measures often described as "objective", such as range-of-motion measurement, may not have this high level of reliability (Triffitt et al. 1999). The interval chosen to assess test-retest reliability of health questionnaires has varied in different studies. Some authors have recommended a 1- to 2-week test-retest interval to minimize the patient's recall of the previous answers (Deyo et al. 1991). Others have used the questionnaire on two successive days (Daltroy et al. 1996, Dawson et al. 1996); such a short interval would reduce the possibility of change in health status influencing the results. Our study showed that, in patients with upper-extremity conditions, shorter and longer test-retest intervals gave similar levels of reliability for the DASH. In fact, the result for the intraclass correlation coefficient was identical to that reported for the original DASH, which was assessed in 28 patients using a 2- to 3-day test-retest interval (Turchin et al. 1998).

The original DASH was administered to 368 patients with different upper-extremity conditions, and the results showed a mean DASH score of 38 (SD 22) and a median score of 35 (McConnel et al. 1999). The finding of similar scores in the present study supports the linguistic/conceptual equivalence of the Swedish version and the comparability of interpretations across countries, which is of crucial importance (Ware et al. 1995). Similarly, the findings reported for the original DASH (McConnel et al. 1999) as well as those of the present study showed that only 1% of patients' scores were at the ceiling level and no scores were at the floor level. This means that the DASH can detect improvement or worsening of health status in most patients.

The use of a region-specific outcome measure in assessing treatment benefits has the advantage of requiring a smaller sample size than a generic measure because of its higher responsiveness (Bessette et al. 1998). However, for the purpose of comparing the burden of two diseases on health-related quality of life, generic measures would be needed. Although combining the DASH with a generic outcome instrument would expand the health domains measured, it would also increase the respondent and administrative burden.

Various scoring systems have been developed and used to assess treatment results in patients with shoulder, elbow or wrist disorders. These systems have usually included different combinations of physical measures, such as range of motion and strength, as well as some functional measures. Combining measurements of different dimensions (i.e., body dimension and activity dimension) into a single score might cause problems

in interpretation and comparison. These scoring systems have been shown to produce widely differing results in the same patients and to correlate weakly with patients' perception of outcome (Turchin et al. 1998). Conventional clinician-measured outcomes may still be important in assessing the results of treatment in terms of improvement in joint motion or strength. However, these can not capture the different dimensions of health-related quality of life, which is central to patients.

This study was supported by research grants from the Kristianstad County Council, the Skåne County Council and the Swedish Foundation for Health Care Sciences and Allergy Research.

We thank Manfred Hofer for his assistance with data collection, Associate Professor Ragnar Johnsson and Professor Charlotte Ekdahl for their helpful comments, and Gert-Uno Larsson, Ewald Ornstein and Jan Rading for their support.

We acknowledge the work of the American Academy of Orthopedic Surgeons and the Institute for Work & Health, Toronto, Canada, in developing the DASH.

Atroshi I, Gummesson C, Johnsson R, Sprinchorn A. Symptoms, disability, and quality of life in patients with carpal tunnel syndrome. J Hand Surg (Am) 1999; 24: 398-404.

Bessette L, Sangha O, Kuntz K M, Keller R B, Lew R A, Fossel A H, Katz J N. Comparative responsiveness of generic versus disease-specific and weighted versus unweighted health status measures in carpal tunnel syndrome. Med Care 1998;: 36: 491-502.

Daltroy L H, Cats-Baril W L, Katz J N, Fossel A H, Liang M H. The North American Spine Society lumbar spine outcome assessment instrument: reliability and validity tests. Spine 1996; 21: 741-9.

Davis A M, Beaton D E, Hudak P, Amadio P, Bombardier C, Cole D, Hawker G, Katz J N, Makela M, Marx R G, Punnett L, Wright J G. Measuring disability of the upper extremity: a rationale supporting the use of a regional outcome measure. J Hand Ther 1999; 12: 269-74.

Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. J Bone Joint Surg (Br) 1996; 78: 593-600.

Deyo R A, Diehr P, Patrick D L. Reproducibility and responsiveness of health status measures: statistics and strategies for evaluation. Control Clin Trials 1991; 12: 142S-58S.

Deyo R A, Andersson G, Bombardier C, Cherkin D C, Keller R B, Lee C K, Liang M H, Lipscomb B, Shekelle P, Spratt K F. Outcome measures for studying patients with low back pain. Spine 1994; 19: 2032S-6S.

Feuerstein M, Miller V L, Burrell L M, Berger R. Occupational upper extremity disorders in the federal workforce: prevalence, health care expenditures, and patterns of work disability. J Occup Environ Med 1998; 40: 546-55.

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.

Guyatt G H, Feeny D H, Patrick D L. Measuring health-related quality of life. Ann Intern Med 1993; 118: 622-9.

Hudak P L, Amadio P C, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand). Am J Ind Med 1996; 29: 602-8.

Kirkley A, Griffin S, McLintock H, Ng L. The development and evaluation of a disease-specific quality of life measurement tool for shoulder instability: the Western Ontario shoulder instability index (WOSI). Am J Sports Med 1998; 26: 764-72.

Lohr K N, Aaronson N K, Alonso J, Burnam M A, Patrick D L, Perrin E B, Roberts J S. Evaluating quality-of-life and health status instruments: development of scientific review criteria. Clin Ther 1996; 18: 979-92.

McConnel S, Beaton D E, Bombardier C. Disabilities of the arm, shoulder and hand: the DASH outcome measure user's manual. First Edition. Institute for Work & Health, Toronto, Ontario 1999.

Nunnally J C, Bernstein I. Psychometric theory. Third Edition. McGraw-Hill, New York 1994.

Streiner D L, Norman G R. Health measurement scales: a practical guide to their development and use. Second Edition. Oxford University Press, Oxford 1995.

Sullivan M, Karlsson J, Ware J E J. The Swedish SF-36 health survey: evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. Soc Sci Med 1995; 41: 1349-58.

Swiontkowski M F, Buckwalter J A, Keller R B, Haralson R. The outcomes movement in orthopaedic surgery: where we are and where we should go. J Bone Joint Surg (Am) 1999; 81: 732-40.

Triffitt P D, Wildin C, Hajioff D. The reproducibility of measurement of shoulder movement. Acta Orthop Scand 1999; 70: 322-4.

Turchin D C, Beaton D E, Richards R R. Validity of observer-based aggregate scoring systems as descriptors of elbow pain, function, and disability. J Bone Joint Surg (Am) 1998; 80: 154-62.

Ware J E J, Sherbourne C D. The MOS 36-item short-form health survey (SF-36): conceptual framework and item selection. Med Care 1992; 30: 473-83.

Ware J E J, Keller S D, Gandek B, Brazier J E, Sullivan M. Evaluating translations of health status questionnaires: methods from the IQOLA project. Int J Technol Assess Health Care 1995; 11: 525-51.

Ware J J, Kosinski M, Keller S D. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996; 34: 220-33.