

Self-administered outcome instrument in carpal tunnel syndrome

Reliability, validity and responsiveness evaluated in 102 patients

Isam Atroshi¹, Ragnar Johnsson² and Anna Sprinchorn¹

We evaluated a Swedish version of a self-administered disease-specific outcome questionnaire for carpal tunnel syndrome regarding reliability, validity and responsiveness to clinical change. It consists of multi-item scales assessing symptom severity, function, patient satisfaction and quality of life. It was given to 102 patients before and 3 months after carpal tunnel release. Test-retest reliability, studied in a subsample of 22 patients on two occasions with a 1-3-week interval, showed good agreement between the scores. Internal consistency of the scales was high (Cronbach alpha 0.80-0.95). Validity of the

scales was evaluated using the SF-36 general health questionnaire in a subgroup of 48 patients as well as items concerning patient satisfaction, showing the expected relationships between these measures. Responsiveness of the scales to clinical change, estimated by the effect size and standardized response mean, was large (0.94-1.7). We conclude that this questionnaire can provide a standardized measure of symptom severity and functional status, as well as patient satisfaction and quality of life in the carpal tunnel syndrome.

¹Hand Surgery Section, Department of Orthopedics, Håssleholm-Kristianstad Hospitals, S-291 85 Kristianstad, Sweden Tel +46 44-131260. Fax -131264. E-mail: isam.atroshi@mailbox.swipnet.se; and ²Department of Orthopedics, Lund University Hospital, Lund, Sweden
Submitted 97-04-19. Accepted 97-10-26

Standardized instruments in the form of self-administered questionnaires that measure outcomes of concern to patients including symptoms, function, satisfaction with the results of treatment and quality of life are being increasingly used in clinical practice (Katz et al. 1995). Such questionnaires can evaluate outcomes of treatment from the patient's perspective, and can facilitate comparison between different studies (Stucki et al. 1996). Validated generic (general health) instruments, such as the Medical Outcomes Study 36-item short-form health survey (SF-36), have been widely used in different populations and patient groups (Ware and Sherbourne 1992, McHorney et al. 1993, Sullivan et al. 1995). Disease-specific questionnaires have been developed to measure important clinical features in specific diseases which might not be detected by generic instruments. A self-administered questionnaire for the assessment of symptom severity and functional status in carpal tunnel syndrome (CTS) has been introduced and shown to be reliable and valid (Levine et al. 1993). We evaluated an expanded Swedish version of this CTS Instrument, with regard to reliability, validity and responsiveness to clinical change.

Patients and methods

Between May 1995 and July 1996, 111 consecutive patients with carpal tunnel syndrome underwent carpal tunnel release at Kristianstad Hospital. 1 patient sustained an injury to the operated hand during the first postoperative week and 8 patients refused to participate in the postoperative evaluation, leaving 102 patients for this study. The mean age of the patients was 52 (21-88) years and 67 were women. In each of 13 bilaterally operated patients, only the hand operated first was included. The dominant extremity was involved in 60 patients. The diagnosis of CTS was made on the basis of characteristic symptoms and signs. Two-point discrimination was greater than 6 mm in at least one radial digit in 28 patients. Thenar atrophy and/or weakness was found in 27 patients. Preoperatively, a Swedish version of the CTS Instrument (Levine et al. 1993) was mailed to all 111 patients and the Swedish version (Sullivan et al. 1995) of the SF-36 (Ware and Sherbourne 1992) was mailed to a subgroup of 53 consecutive patients.

Carpal tunnel release was performed under local anesthesia by a single surgeon using the two-portal endoscopic technique (Chow 1989, Atroshi et al. 1997).

Follow-up evaluation was performed by an independent examiner 3 months postoperatively. At the follow-up examination, an expanded postoperative version of the CTS Instrument was completed by the 102 patients included in this study. The SF-36 was completed by 48 of 53 patients.

8 patients who did not want to participate in the 3-month follow-up evaluation were contacted by telephone by an independent observer and asked about postoperative symptoms. 4 patients reported complete symptom relief, 3 reported improvement but some residual pain and/or weakness, and 1 patient reported unchanged numbness and tingling.

Outcome Instruments

Before the study, the original CTS instrument was translated from English to Swedish, taking into consideration proposed guidelines (Guillemin et al. 1993). The CTS Instrument consists of two multi-item scales (Table 1). The symptom severity scale consists of 11 items about symptoms experienced by the patient in the preceding 2 weeks. The functional status scale consists of 8 items concerning difficulties in performing 8 activities of daily living due to hand or wrist symptoms in the preceding 2 weeks. Each item on the symptom severity and functional status scales has 5 possible response options, ranging from 1 (no symptom or dysfunction) to 5 (most severe symptom or dysfunction). The symptom severity and functional status score is the mean of all answered items on each scale. Thus, a higher symptom severity or functional status score indicates worse symptoms or dysfunction.

We expanded the original CTS Instrument by adding a 2-item palmar pain scale, an 8-item satisfaction scale and four separate items regarding patient opinion, thus creating a postoperative version. The 2-item palmar pain scale refers to the severity of postoperative palmar pain/tenderness with 4 response options (none, mild, moderate, severe), and limitation of activities due to palmar pain/tenderness with 4 response options (no pain/tenderness, pain/tenderness, but no limitation, some limitation, much limitation). The 8-item satisfaction scale has 5 response options for each item (completely satisfied, very satisfied, somewhat satisfied, dissatisfied and very dissatisfied) concerning satisfaction with the results of surgery as regards relief of symptoms and dysfunction. Scale scores were calculated as the mean of the answered items on each scale. The first of the four separate items concerning patient opinion refers to satisfaction with the overall results of surgery, with 5 response options equal to those on the satisfaction scale. The second item refers to the effect of surgery on quality of life

with 7 response options (much improved, moderately improved, slightly improved, unchanged, slightly worse, moderately worse, much worse), which were summed to 5 response levels in the analysis by summing the first two and the last two options. The third and fourth items refer to willingness to have the same operation if able to make the decision again and willingness to recommend the operation to a close friend, with both items having 5 response options (definitely, probably, unsure, probably not, definitely not).

The SF-36 is a self-administered questionnaire of general health and well-being consisting of 8 multi-item scales measuring 8 health dimensions (physical functioning, role limitations because of physical health problems, bodily pain, general health perceptions, vitality, social functioning, role limitations because of emotional problems, mental health). The responses to the items are transformed into scale scores. Each of the 8 scales is scored from 0 (worst possible health) to 100 (best possible health). Thus, a higher overall score on each SF-36 scale indicates a better health state.

Analyses

The CTS Instrument scales were examined for reliability, validity and responsiveness.

Reliability includes two aspects, reproducibility and internal consistency. Reproducibility (test-retest reliability) or the degree in which the instrument yielded stable scale scores on different administrations in the absence of clinical change, was assessed in a subgroup of patients, using the reliability coefficient (Bland and Altman 1986). For evaluation of reproducibility, a randomly selected sample of 22 patients (17 women) aged 49 (22–73) years, completed the instrument preoperatively on two occasions with an interval of 14 (7–20) days. Internal consistency or the degree in which each scale in the instrument measured a single concept, was assessed for all 102 patients preoperatively and postoperatively by the Cronbach alpha coefficient which summarizes inter-item correlation in each scale (Nunnally and Bernstein 1994).

Validity or the degree in which the instrument measured what it purported to measure was studied in two ways. Construct-related validity was assessed by studying the correlations between the CTS questionnaire symptom severity and functional status scales and the SF-36 scales in the subgroup of 48 patients who completed the SF-36 preoperatively and postoperatively. We hypothesized that the CTS symptom severity and functional status scales would correlate with the SF-36 scales in a direction in which worse symptoms or dysfunctions correlate with worse

Table 1. Number of patients replying to each response option and mean score for each item in the expanded CTS questionnaire

Scale Item	Preoperative response option						Postoperative response option					
	1 n	2 n	3 n	4 n	5 n	mean score	1 n	2 n	3 n	4 n	5 n	mean score
<i>Symptom severity</i> ^a												
Severity of nocturnal pain	9	27	27	27	10	3.0	70	18	9	2	2	1.5
Frequency of nocturnal awakening due to pain	15	18	39	19	9	2.9	72	11	12	2	4	1.6
Severity of daytime pain	7	29	47	15	2	2.8	41	38	17	4	1	1.9
Frequency of daytime pain	8	22	25	21	24	3.3	42	27	14	10	5	2.1
Duration of daytime pain	8	35	24	7	25	3.1	43	30	16	9	3	2.0
Severity of numbness	1	11	34	42	12	3.5	64	17	9	7	2	1.7
Severity of weakness	18	16	35	22	7	2.8	27	35	18	16	4	2.4
Severity of tingling	4	11	50	28	7	3.2	66	26	4	5	0	1.5
Severity of nocturnal numbness/tingling	3	9	33	35	19	3.6	75	13	6	4	2	1.5
Frequency of nocturnal awakening due to numbness/tingling	7	17	45	19	12	3.1	79	9	6	3	3	1.4
Difficulty in gripping small objects	15	25	24	28	5	2.8	61	21	11	7	1	1.7
<i>Functional status</i> ^a												
Writing	32	27	23	14	0	2.2	69	17	6	6	0	1.5
Holding a book	21	27	29	17	5	2.6	59	21	9	9	0	1.7
Buttoning clothes	20	21	24	21	11	2.8	59	19	15	2	4	1.7
Gripping the telephone	29	24	21	21	2	2.4	69	12	7	4	3	1.5
Opening jars	22	21	22	21	13	2.8	42	26	13	14	4	2.1
Performing household chores	17	32	33	15	2	2.5	56	21	16	4	1	1.7
Carrying a grocery bag	17	21	24	30	6	2.9	41	25	16	12	3	2.1
Bathing and dressing	45	33	13	5	2	1.8	77	18	3	2	0	1.3
<i>Palmar pain</i> ^b												
Severity of palmar pain/tenderness							39	37	19	6	-	1.9
Limitation of activities due to palmar pain/tenderness							39	24	28	6	-	2.0
<i>Satisfaction</i> ^c												
Relief of daytime pain							63	10	11	7	2	1.7
Relief of nocturnal pain							42	19	15	10	2	2.0
Relief of tingling							63	18	6	6	2	1.6
Relief of numbness							57	21	7	8	2	1.7
Relief of sensory loss							46	17	15	8	2	1.9
Relief of weakness							26	10	23	17	5	2.6
Improvement in ability to perform activities of daily living							34	15	16	11	6	2.3
Improvement in ability to perform work							34	15	16	11	6	2.3
<i>Patient opinion</i>												
Satisfaction with the overall results ^c							51	22	19	9	1	1.9
Effect of surgery on quality of life ^d							70	9	13	2	1	1.9
Willingness to have the same operation ^e							84	3	6	1	1	1.2
Recommend the operation to a close friend ^e							77	8	6	3	1	1.4

^a 1 (no symptom or dysfunction)-5 (most severe symptom or dysfunction)

^b 1 (no pain/tenderness)-4 (severe pain/tenderness causing much limitation)

^c 1 (completely satisfied)-5 (very dissatisfied)

^d 1 (much/moderately improved)-5 (moderately/much worse)

^e 1 (definitely)-5 (definitely not)

health, and that stronger correlations would be expected with the physical than the mental SF-36 scales. The correlation between the palmar pain scale and the SF-36 bodily pain scale was also studied. The Spearman correlation coefficient was used for these analyses because of the non-normal distribution of the scores in some of the SF-36 scales.

Validity was also assessed by studying the extent to which postoperative symptom severity and functional status scores were related to patient satisfaction, using

analysis of variance. Validity of the satisfaction scale was assessed by studying its correlation with the items regarding patient opinion, using the Spearman correlation coefficient.

Responsiveness or the ability of the instrument to detect clinical change (Guyatt et al. 1987) was studied by calculating the mean scores (95% confidence intervals) for the symptom severity and functional status scales and the SF-36 scales before and after operation, and estimating the effect size (ES) (Kazis et al.

Table 2. Correlation coefficients (95% confidence intervals) between the CTS questionnaire symptom severity (SS) and functional status (FS) scales and the SF-36 scales

SF-36 scales	Preoperative		Postoperative	
	SS	FS	SS	FS
Physical functioning	0.30 (0.02–0.54)	0.62 (0.41–0.77)	0.33 (0.05–0.56)	0.56 (0.33–0.73)
Physical role	0.56 (0.33–0.73)	0.70 (0.52–0.82)	0.52 (0.28–0.70)	0.55 (0.32–0.72)
Bodily pain	0.64 (0.44–0.78)	0.65 (0.45–0.80)	0.64 (0.44–0.78)	0.67 (0.48–0.80)
General health	0.25 (-0.04–0.50)	0.45 (0.19–0.65)	0.36 (0.08–0.58)	0.49 (0.24–0.68)
Vitality	0.31 (0.03–0.55)	0.37 (0.10–0.59)	0.43 (0.17–0.64)	0.57 (0.34–0.74)
Social functioning	0.26 (-0.03–0.51)	0.36 (0.09–0.58)	0.36 (0.08–0.58)	0.46 (0.20–0.66)
Emotional role	0.23 (-0.06–0.48)	0.29 (0.01–0.53)	0.47 (0.22–0.67)	0.52 (0.28–0.70)
Mental health	0.20 (-0.09–0.46)	0.32 (0.04–0.56)	0.35 (0.07–0.58)	0.34 (0.06–0.57)

Table 3. Mean scores (95% confidence intervals) for the CTS questionnaire symptom severity and functional status scales, as related to patient satisfaction with the overall results of surgery

Satisfaction group	Patients	Symptom severity scale		Functional status scale	
		Preoperative	Postoperative	Preoperative	Postoperative
Satisfied ^a	73	3.0 (2.9–3.2)	1.4 (1.3–1.6)	2.4 (2.2–2.6)	1.4 (1.3–1.6)
Somewhat satisfied	19	3.2 (3.0–3.5)	2.2 (1.9–2.6)	2.8 (2.4–3.2)	2.2 (1.8–2.5)
Dissatisfied ^b	10	3.4 (2.8–4.0)	3.2 (2.7–3.7)	3.1 (2.5–3.7)	3.2 (2.7–3.7)

^a Includes responses: completely satisfied and very satisfied

^b Includes responses: dissatisfied and very dissatisfied

1989) and the standardized response mean (SRM) (Liang et al. 1990) for each scale.

All statistical tests were two-sided and a p-value of less than 0.05 was considered to indicate significance.

Results

The results of the individual items in the expanded CTS questionnaire studied are shown in Table 1.

Reliability

Reproducibility. The mean difference between the scores recorded in the two successive administrations of the preoperative CTS Instrument was -0.1 (SD 0.3) for the symptom severity scale and 0.08 (SD 0.35) for the functional status scale. These two mean differences did not differ significantly from zero and 95% of the differences were less than the reliability coefficient calculated as 0.64 for the symptom severity scale and 0.71 for the functional status scale.

Internal consistency. Preoperatively, the Cronbach alpha was 0.80 for the symptom severity scale (item-total correlation 0.24–0.60), and 0.87 for the functional status scale (item-total correlation 0.57–0.73). 3 months postoperatively, the Cronbach alpha was 0.92 for the symptom severity scale (item-total correlation

0.59–0.81), 0.93 for the functional status scale (item-total correlation 0.65–0.86), 0.90 for the palmar pain scale (item-total correlation 0.82–0.82) and 0.95 for the satisfaction scale (item-total correlation 0.70–0.91). The Cronbach alpha coefficient did not improve by deletion of single items on any of the scales.

Validity

The CTS symptom severity and functional status scales correlated with all SF-36 scales in the expected direction with worse symptoms or dysfunctions correlating with worse state of health (Table 2). Stronger correlations were observed with the SF-36 physical scales than mental scales. The correlation coefficient between the CTS palmar pain scale and the SF-36 bodily pain scale was 0.39 (95% CI, 0.21–0.54).

Analysis of the relationship between patient satisfaction with the overall results of surgery and the symptom severity and functional status scales showed worse postoperative scores for both scales ($p < 0.001$) in patients with lower degree of satisfaction (Table 3).

The Spearman correlation coefficient between the satisfaction scale and the items concerning patient opinion was 0.82 (95% CI, 0.74–0.88) for the item concerning satisfaction with the overall results of surgery, 0.61 (95% CI, 0.47–0.72) for the item concerning effect of surgery on self-reported quality of life,

Table 4. Mean scores (95% confidence intervals) for the CTS questionnaire and SF-36 scales and estimates of responsiveness

	Preoperative	Postoperative	ES ^a	SRM ^a
<i>CTS Instrument (n 102)</i>				
Symptom severity ^b	3.1 (3.0-3.3)	1.7 (1.6-1.9)	2.1	1.7 (1.4-2.0)
Functional status ^b	2.5 (2.4-2.7)	1.7 (1.6-1.9)	0.94	0.94 (0.72-1.16)
Palmar pain ^c	-	2.0 (1.8-2.1)	-	-
Satisfaction ^b	-	2.0 (1.8-2.2)	-	-
<i>SF-36^d (n 48)</i>				
Physical functioning	73 (67-78)	74 (68-80)	0.06	0.07 (-0.23-0.38)
Physical role	49 (37-62)	61 (48-73)	0.27	0.27 (-0.03-0.57)
Bodily pain	42 (36-48)	62 (54-70)	1.04	0.77 (0.48-1.1)
General health	68 (61-75)	71 (64-77)	0.11	0.13 (-0.21-0.47)
Vitality	55 (47-62)	70 (64-75)	0.60	0.70 (0.38-1.01)
Social functioning	84 (78-90)	87 (80-93)	0.12	0.13 (-0.18-0.44)
Emotional role	70 (58-81)	77 (66-89)	0.19	0.21 (-0.09-0.51)
Mental health	78 (72-84)	83 (78-88)	0.23	0.25 (-0.06-0.55)

^a ES effect size, SRM standardized response mean

^b Higher score (1-5) indicates worse symptoms and dysfunction or less satisfaction

^c Higher score (1-4) indicates worse palmar pain/tenderness causing limitation of activities

^d Higher score (0-100) indicates better health

0.52 (95% CI, 0.36-0.65) for the item concerning willingness to have the same operation, if able to make the decision again, and 0.43 (95% CI, 0.26-0.58) for the item concerning willingness to recommend the operation to a close friend.

Responsiveness

The estimated ES and SRM for the CTS Instrument symptom severity and functional status scales ranged from 0.94 to 2.11 indicating a large clinical change (Table 4). The 95% confidence intervals for the SRM were high for both scales with a lower limit of 1.38 for the symptom severity scale and 0.72 for the functional status scale.

The estimated ES and SRM were large for the SF-36 bodily pain scale, moderate for the vitality scale and small for the remaining scales.

Discussion

Important aspects of the evaluation of treatment outcomes are the reliability and validity of the evaluation methods. In a recent study, the discrepancy found between patients' and physicians' evaluations of outcome of total hip arthroplasty increased when patients were dissatisfied with the outcome (Lieberman et al. 1996). Orthopedic outcomes of principal concern to patients are mainly pain and function. In upper extremity disorders, pain, paresthesias and difficulties in performing activities of daily living are among the main reasons why patients seek treatment. Accordingly, evaluation of these outcomes is highly relevant.

Assessment of outcome in CTS has aroused increased interest, particularly following the introduction of endoscopic surgery (Brown et al. 1993, Katz et al. 1994). In the clinical evaluation of CTS, standardized methods of demonstrated reliability and validity have been used for the assessment of sensibility (Dellon et al. 1987) and strength (Mathiowetz et al. 1984). For symptom assessment, various methods have been used, the reliability and validity of which have not always been determined. Outcomes such as two-point discrimination and grip strength are often described as objective, while symptom outcomes are considered subjective. Until recently, functional impairment in CTS has rarely been assessed. The measurement properties of the CTS Instrument in evaluating symptoms and function in CTS have been investigated and reliability and validity comparable to those reported for sensibility and strength measures have been shown (Levine et al. 1993). Our study revealed similar findings.

Patient satisfaction, which is an important outcome variable, can be assessed more comprehensively with a multi-item scale than a single question, since satisfaction can be influenced by other factors. This is demonstrated by the varying correlations between the satisfaction scale and the four items regarding satisfaction with the overall results of surgery, effect of surgery on quality of life, willingness to have the same operation, if able to make the decision again, and to recommend the operation to a close friend. The finding that a substantial number of patients reported improved quality of life following surgical treatment demonstrates the impact of CTS on self-perceived

quality of life.

Since its introduction, the CTS Instrument has been used in a number of studies (Katz et al. 1995, Amadio et al. 1996) and the results support its sensitivity to clinical change. The CTS Instrument has proved to be more responsive to clinical change after CTS surgery than the commonly performed sensibility and strength measures (Amadio et al. 1996). Thus, reporting results of CTS treatment would be optimized if validated instruments measuring symptom and functional outcomes are used, in addition to conventional sensibility and strength measures.

Generic instruments, such as the SF-36, have been found less sensitive when used in evaluating upper extremity disorders than disease-specific instruments (Amadio et al. 1996, Beaton et al. 1996). Our findings are consistent with those reported by Amadio et al. (1996), indicating that the CTS Instrument is more responsive to clinical change in CTS than the SF-36. Moreover, the CTS instrument had greater variability in the preoperative scores, with fewer ceiling effects than the SF-36. However, the SF-36 can show the possible effect of CTS on general health and well-being. Moreover, the health dimensions measured by the SF-36 can be compared in patients with CTS and other conditions. In our study, the bodily pain scale of the SF-36 was more responsive to clinical change in CTS than the remaining scales.

Our study showed that endoscopic carpal tunnel release resulted in a large improvement in symptoms and function, as measured by the effect size and standardized response mean for the CTS Instrument. Levine et al. (1993) reported an effect size of 1.1 for the symptom severity scale and 0.7 for the functional status scale in 26 patients who completed a 3-month follow-up after open carpal tunnel release. In a series of 23 patients reported by Katz et al. (1995), the estimated effect size 3 months after carpal tunnel release, performed with open technique in the majority of the patients, was 1.3 for the symptom scale and 0.7 for the function scale. Amadio et al. (1996) reported a standardized response mean of 1.8 for the symptom scale and 1.3 for the function scale in 22 CTS patients 3 months after open carpal tunnel release. The high standardized response means shown in our study, including the high lower limits for the 95% confidence intervals, confirm the efficacy of endoscopic surgical treatment of CTS.

Comparison of CTS outcomes measured by the CTS questionnaire in different populations should be taken with caution, since populations might differ in other important aspects that could influence outcome. For such comparisons, indices of the magnitude of clinical change, such as the effect size and standard-

ized response mean, ought to be more informative than individual scores. Thus, longitudinal analyses might well give more valid results than cross-sectional analyses, if outcomes in two CTS populations are to be compared. The indices of responsiveness (effect size and standardized response mean) are valuable in showing whether statistically significant differences are also clinically important (Kazis et al. 1989, Liang et al. 1990).

Our results show that the scales of the CTS Instrument seem to be reliable, valid and responsive to clinical change. The instrument provides a consistent standardized measure of the severity of symptoms and function in CTS. Although the instrument is validated for use in CTS, it might be used in other compression neuropathies. However, its measurement properties in other compression neuropathies need to be evaluated.

Acknowledgments

The authors thank Ewald Ormstein, MD, and Christina Gummesson, MSPT, for their contributions, and Jonas Ranstam, PhD, for statistical advice. This study was supported by a grant from the Kristianstad County Council and the Swedish Medical Research Council (#17x-09509).

References

- Amadio P C, Silverstein M D, Ilstrup D M, Schleck C D, Jensen L M. Outcome assessment for carpal tunnel surgery: the relative responsiveness of generic, arthritis-specific, disease-specific and physical examination measures. *J Hand Surg (Am)* 1996; 21: 338-46.
- Atroshi I, Johnsson R, Ormstein E. Endoscopic carpal tunnel release: prospective assessment of 255 consecutive cases. *J Hand Surg (Br)* 1997; 22: 42-7.
- Beaton D E, Richards R R. Measuring function of the shoulder: a cross-sectional comparison of five questionnaires. *J Bone Joint Surg (Am)* 1996; 78: 882-90.
- Bland J M, Altman D G. Statistical methods for assessing agreement between two methods for clinical measurement. *Lancet* 1986; i: 307-10.
- Brown R A, Gelberman R H, Seiler J G et al. Carpal tunnel release: a prospective randomized assessment of open and endoscopic methods. *J Bone Joint Surg (Am)* 1993; 75: 1265-75.
- Chow J C Y. Endoscopic release of the carpal ligament: a new technique for carpal tunnel syndrome. *Arthroscopy* 1989; 5: 19-24.
- Dellon A L, Mackinnon S E, Crosby P M. Reliability of two-point discrimination measurements. *J Hand Surg (Am)* 1987; 12: 693-6.
- 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, Walter S, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chron Dis* 1987; 2: 171-8.
- Katz J N, Gelberman R H, Wright E A, Abrahamsson S-O, Lew R A. A preliminary scoring system for assessing the outcome of carpal tunnel release. *J Hand Surg (Am)* 1994; 19: 531-8.
- Katz J N, Fossel K K, Simmons B P, Swartz R A, Fossel A H, Koris M J. Symptoms, functional status and neuromuscular impairment following carpal tunnel release. *J Hand Surg (Am)* 1995; 20: 549-55.
- Kazis L E, Anderson J J, Meenan R F. Effect sizes for interpreting changes in health status. *Med Care* 1989; 27: 178-89.
- Levine D W, Simmons B P, Koris M J, Lawren H D, Gerri G H, Fossel A H, Katz J N. 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.
- Liang M H, Fossel A H, Larson M G. Comparisons of five health status instruments for orthopedic evaluation. *Med Care* 1990; 28: 634-42.
- Lieberman J R, Dorey F, Shekelle P, Schumacher L, Thomas B J, Kilgus D J, Finerman G A. Differences between patients' and physicians' evaluation of outcome after total hip arthroplasty. *J Bone Joint Surg (Am)* 1996; 78: 835-8.
- Mathiowetz V, Weber K, Volland G, Kashman N. Reliability and validity of grip and pinch strength evaluations. *J Hand Surg (Am)* 1984; 9: 222-6.
- McHorney C A, Ware J E, Raczek A E. The MOS 36-item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care* 1993; 31: 247-63.
- Nunnally J C, Bernstein I H. *Psychometric theory*. McGraw-Hill, New York 1994.
- Stucki G, Daltroy L, Liang M H, Lipson S J, Fossel A H, Katz J N. Measurement properties of a self-administered outcome measure in lumbar spinal stenosis. *Spine* 1996; 21: 796-803.
- Sullivan M, Karlsson J, Ware J E. The Swedish SF-36 health survey: I. Evaluation of data quality, scaling assumptions, reliability and construct validity across general populations in Sweden. *Soc Sci Med* 1995; 41: 1349-58.
- Ware J E, Sherbourne C D. The MOS 36-Item Short-Form Health Survey (SF-36): I. Conceptual framework and item selection. *Med Care* 1992; 30: 473-83.