THE PATIENT OUTCOMES OF SURGERY-HAND/ARM (POS-HAND/ARM): A NEW PATIENT-BASED OUTCOME MEASURE

S. J. CANO, J. P. BROWNE, D. L. LAMPING, A. H. N. ROBERTS, D. A. MCGROUTHER and N. A. BLACK

From the Health Services Research Unit, London School of Hygiene & Tropical Medicine, London, UK

The purpose of this study was to develop and validate a new patient-based outcome measure for hand/arm disorders for use in audit, clinical trials and effectiveness studies. There were three stages. First, we carried out interviews with 40 patients with hand/arm disorders to develop and pilot questionnaire content. Second, in a postal survey with 165 pre- and 181 post-surgery patients, we reduced the number of items and identified scales. Third, in a postal survey with 132 pre- and 204 post-surgery patients we evaluated the psychometric properties of the measure.

Findings confirmed the acceptability, reliability, validity and responsiveness to clinical change of the questionnaire. The Patient Outcomes of Surgery-Hand/Arm (POS-Hand/Arm) is a new surgical outcome measure that can be used before and after surgery (29 and 33 items, respectively) to evaluate and compare new techniques, surgical teams and units.

Journal of Hand Surgery (British and European Volume, 2004) 29B: 5: 477–485

Keywords: questionnaire, outcomes, hand surgery, arm surgery, symptoms, quality of life

INTRODUCTION

Hand/arm disorders that require surgery represent a significant public health burden due to high volume (Department of Health, 2003), negative impact on patients (Higgs et al., 1995), and high disability costs (Katz et al., 1996). Rigorous evaluation of the outcomes of both the disorder and its treatment is an essential aspect of evidence-based health care, but has not been routinely done in this and other types of related surgery (Goldacre et al., 1996). In hand/arm surgery, clinical measures such as range of motion have been used to evaluate improvement (Macey et al., 1995), but patient-based outcomes such as pain and activities of daily living have not been well investigated (Amadio, 1997).

Although psychometrically sound measures have been used to assess patient-based outcomes in carpal tunnel syndrome (Levine et al., 1993) and the psychological impact of rheumatoid arthritis on hands (Vamos et al., 1990), they have limited applicability across the range of conditions requiring hand/arm surgery. Three main site-specific measures have been used to evaluate outcomes in hand/arm disorders. The Patient Evaluation Measure (PEM; Macey et al., 1995) assesses patient satisfaction with treatment, general hand functioning and activities of daily living. The Disabilities of Arm Shoulder and Hand (DASH; Hudak et al., 1996) measures symptoms and functional status in patients with disorders of the upper limb. The Michigan Hand Questionnaire (MHQ; Chung et al., 1998) assesses hand function, activities of daily living, work performance, pain, aesthetics and satisfaction. Of these, the PEM has not been formally psychometrically evaluated and the DASH and MHQ were both developed for use in North America.

We describe the development and psychometric evaluation of a new patient-based measure for evaluating outcomes in surgery for hand/arm disorders. Our objective was to develop a measure designed specifically to evaluate the outcomes of surgery, and include all clinically relevant domains as agreed by hand surgery experts representing the British Association of Plastic Surgeons (Cano et al., 2003). We established that the new measure should be site-specific to the hand/arm but applicable across a range of disorders requiring surgery, that it should target elective non-malignant hand/arm surgery and be suitable for use before and after surgery for use in audit, clinical trials and effectiveness studies, and that it should meet rigorous measurement criteria (e.g. acceptable, reliable, valid, responsive to clinical change).

PATIENTS AND METHODS

We followed international guidelines (Scientific Advisory Committee of the Medical Outcome Trust, 2002) for the development and validation of health outcome measures. This rigorous, three-stage, gold standard methodology has not previously been used in developing outcome measures in hand/arm surgery (see Table 1). In Stage 1, we generated a pool of items from patient interviews, expert opinion and a review of the literature. In Stage 2, we field tested this large pool of items by postal survey to select the questions that showed the best scientific performance and we identified scales. In Stage 3, we evaluated the measurement properties of the new measure by a further postal survey of an independent group of patients. Ethics approval was obtained for each recruitment site before patients were invited to take part in this study.

Table 1—Questionnaire development and validation: comparison of the POS-Hand/Arm and existing measures

Method/Evaluation		*Measure*		
		POS-Hand/Arm	*DASH*	*MHQ*
Item Generation				
	Patient interviews	♦		♦
	Literature	♦	♦	♦
	Expert opinion	♦	♦	♦
	Develop conceptual model	♦		
Item Reduction				
	Expert opinion	♦	♦	
	Item redundancy	♦		
	Endorsement frequencies	♦	♦	
	Missing data	♦		
	Factor analysis	♦	♦	♦
	Tests of scaling assumptions	♦		
Psychometric Analyses				
	Acceptability	♦		
	Internal consistency reliability	♦	♦	♦
	Item total correlations	♦		
	Test–retest reliability	♦		♦
	Validity: within scale	♦		♦
	Validity: comparison with other measures	♦		♦
	Responsiveness	♦		♦

Stage 1—Item Generation

An initial pool of 75 questions about the health impact of hand/arm conditions and surgery was generated from three sources: a comprehensive literature review, multidisciplinary expert opinion, and semi-structured interviews with 17 patients recruited randomly from computerized lists in a large plastic surgery department. All English-speaking adult patients, 18 years and over who had undergone elective surgery for non-malignant hand/arm disorders were eligible to participate in this stage of the study. Terminally ill patients were excluded on ethical grounds. The item pool was then piloted (also known as pre-testing) with 23 patients in order to clarify ambiguities in the wording of items.

Stage 2—Item Reduction

The 75-item questionnaire was administered by postal survey to 236 pre-surgery patients and 262 post-surgery patients (including all pre-surgery patients and those who were missed at the pre-surgery assessment due to administrative omissions; $n = 26$). These patients were recruited consecutively from computerized lists at two large plastic surgery departments. All English-speaking adult patients who had been given a surgery date were eligible to participate in this stage of the study, except for terminally ill patients who were excluded on ethical grounds. Patients were sent a questionnaire on average 2 weeks before surgery and 3 months after surgery. All patients who returned a pre-surgery questionnaire (165 of the 236) were sent the 3-month post-surgery questionnaire to evaluate responsiveness. A second subsample of patients ($n = 88$) was randomly selected to receive a second post-surgery questionnaire, 2 weeks after returning the first post-surgery questionnaire, to evaluate test-retest reproducibility. Standard techniques (Dillman, 1978) were used to ensure a high response rate including personalized letters, standardized instructions and follow-up reminder letters.

We used standard psychometric methods (Scientific Advisory Committee of the Medical Outcome Trust, 2002) to evaluate the psychometric performance of individual items and a strategy for item reduction (Table 2, row 1) developed in our previous work (Lamping et al., 2002, 2003).

Stage 3—Psychometric Evaluation

We evaluated the psychometric properties of the POS-Hand/Arm in 196 pre-surgery and 316 post-surgery patients (this included all pre-surgery patients plus a sub-sample of patients who were sent a post-surgery questionnaire only; $n = 120$). These were recruited consecutively from three large plastic surgery departments through computerized patient lists. This was done by postal survey using the same methods as in Stage 2. We also recruited a sub-sample of patients who were assessed only at 3 months post-surgery, for the post-surgery analysis.

We compared the POS-Hand/Arm with three existing measures, including a generic health status measure – the Medical Outcomes Study Short-Form Health Survey (SF-36; Ware and Sherbourne, 1992) – and the two existing hand/arm specific measures – the DASH (Hudak et al., 1996) and the MHQ (Chung et al., 1998). We then used "gold standard" psychometric tests and criteria (Nunnally and Bernstein, 1994; Streiner and

Table 2—Psychometric tests and criteria*

Psychometric property	*Definition/test*	*Criteria for acceptability*
1. Item reduction	Identify items for possible elimination due to weak psychometric properties; assessed on the basis of item analysis, factor analysis and tests of scaling assumptions.	Applied to all items: • missing data $<5\%$ • item redundancy (inter-item correlations <0.70) • maximum endorsement frequencies $<80\%$ (ie the proportion of respondents who endorse each response category) • aggregate adjacent endorsement frequencies $>10\%$ • item–total correlations >0.30 • items in the physical activities scale rated as 'not relevant' $<35\%$. In principal components factor analysis, all items load on the first unrotated factor >0.30. Scaling success (i.e. higher item-own scale correlations than item-other scale correlations).
2. Acceptability	Data quality; assessed by completeness of data and score distributions.	Applied to all items: • missing data $<5\%$ • maximum endorsement frequencies $<80\%$
3. Reliability		
3.1. Internal consistency	The extent to which items in a scale measure the same construct; assessed by Cronbach's α (Cronbach, 1951) and item-total correlations.	Cronbach's α coefficients for summary scores >0.70; item-total correlations >0.30.
3.2. Test–retest reproducibility	The stability of a scale; assessed on the basis of correlations between repeat administrations of the scale on two occasions.	Intraclass correlations (ICCs) >0.70 between test and retest scores.
4. Validity		
4.1. Content validity	The extent to which the content of a scale is representative of the conceptual domain it is intended to cover; assessed qualitatively during questionnaire development.	Qualitative evidence from pre-testing, expert opinion and literature review that items are representative of all important domains relevant to outcomes in surgery for hand/arm disorders.
4.2. Construct validity		
4.2.1. Within-scale analyses	Evidence that the scale measures a single construct, that items can be combined to form a summary score, and that sub-scales measure distinct but related constructs; assessed on the basis of internal consistency, item–total correlations, intercorrelations between scales, and tests of scaling assumptions.	Cronbach's α coefficients for summary scores >0.70; item-total correlations >0.30; intercorrelations between scales $r = 0.30$ to 0.70; scaling successes
4.2.2. Analyses against external criteria		
4.2.2.1. Convergent validity	Evidence that the scale is correlated with other measures of the same or similar constructs; assessed on the basis of correlations between the new scale and measures of similar constructs.	Hypotheses based on the degree of conceptual similarity between measures: (i) higher correlations between the POS-Hand/Arm Symptom scale and other measures of symptoms (e.g. DASH Function/Symptom score, MHQ Pain scale) than with measures of psychological functioning (e.g. SF-36 Mental Health, Role Limitations-Emotional); (ii) moderate correlations ($r = 0.30-0.70$) between the POS-Hand/Arm Psychological Functioning/Cosmetic Appearance scale and other psychological functioning or cosmetic appearance scales (e.g. SF-36 Mental Health and MHQ Aesthetics scales).
4.2.2.2. Discriminant validity	Evidence that the scale is not correlated with measures of different constructs; assessed on the basis of correlations with age, sex and social class.	Low correlations with age, sex and social class ($r<0.30$).
4.2.2.3. Known group differences/hypothesis testing	The ability of the scale to differentiate known groups; assessed by comparing scores between groups whose scores on the scale are expected to differ.	POS-Hand/Arm scores should be significantly higher for those who report improvement after surgery than for those who report no improvement.
5. Responsiveness	The ability of scale to detect clinically significant change following a treatment of known efficacy; assessed by examining scores before and after surgery and calculating an effect size statistic (mean change score divided by standard deviation of pre-surgery scores).	Moderate to large (0.50–0.80) effect sizes (Kazis et al., 1989).

*Adapted from Lamping et al., (2002, 2003).

Table 3—Respondent Characteristics

			Item reduction		*Psychometric evaluation*	
	Semi-structured interviews	*Pre-test*	*Pre-surgery*	*Post-surgery*	*Pre-surgery*	*Post-surgery*
N	17	23	165	181	132	204
Characteristics	Descriptive statistics					
Age						
Mean (SD)	51 (12)	56 (19)	55 (15)	56 (16)	57 (15)	57 (15)
Range	20–68	19–83	18–87	21–91	20–90	20–87
Characteristics	Percentages					
Gender						
Male	41	48	57	43	43	50
Female	59	52	43	57	57	50
Ethnicity						
White	100	100	98	96	96	96
Employment status						
Employed	82	—	43	41	44	41
Retired	6	—	32	35	34	38
Unable to work	6	—	10	7	11	9
Other	6	—	15	17	11	12
Surgery						
Carpal tunnel syndrome	24	17	18	31	31	35
Dupuytren's fasciectomy	18	—	15	13	23	21
Joint surgery (e.g. arthroplasty)	24	35	24	25	26	23
Tendon surgery (e.g. trigger finger)	12	9	16	12	7	6
Mass (e.g. ganglion, nail) excision			12	9	8	5
Other	22	39	15	10	5	10

Norman, 1995) to evaluate acceptability, reliability, validity, and responsiveness (Table 2, rows 2–4).

RESULTS

Stage 1—Item Generation

Statements generated from interviews resulted in 75 items. The items were grouped into six clinically relevant domains (symptoms, physical activities, psychological functioning, cosmetic appearance, complications and satisfaction). The review of the literature and expert opinion confirmed these domains. Questions asked patients about the health impact of hand/arm disorders during the previous 4 weeks. This time period was judged to be clinically sensible based on disease and surgical variables. Item response categories were chosen so as not to have more than five based on standard guidelines (Streiner and Norman, 1995). Pre-testing resulted in minor changes to the content of the questionnaire (results not presented). Respondent characteristics of patients who participated in interviews and pre-testing are shown in Table 3 (columns 1 and 2).

Stage 2—Item Reduction

Of the 236 patients who received the POS-Hand/Arm pre-surgery, 165 (70%) returned completed questionnaires. A total of 181 (69%) post-surgery patients returned completed questionnaires at 3 months. Respondents represented a wide range of age, employment status and disease groups (Table 3, columns 3 and 4). All 165 patients who returned questionnaires in the pre-surgery sample were included in the responsiveness sub-sample. Of these, 121 (73%) returned post-surgery questionnaires. Finally, 61 (69%) patients in the test–retest sub-sample returned completed questionnaires.

Just under half of the items in the original item pool were eliminated as a result of item reduction analyses. Factor analysis and tests of scaling assumptions (Ware et al., 1997) resulted in a 29-item pre-surgery version of the POS-Hand/Arm with three scales (physical activities, symptoms, and psychological functioning/cosmetic appearance) and a 33-item post-surgery version which includes the three pre-surgery scales plus a post-surgery satisfaction scale (Appendices I and II).

Stage 3—Psychometric Evaluation

Of the 196 patients who received the POS-Hand/Arm pre-surgery, 132 (67%) returned completed questionnaires. A total of 204 (65%) post-surgery patients returned completed questionnaires at 3 months. Respondents represented a wide range of age, employment and disease groups (Table 3, columns 5 and 6). Of the 115 patients who returned pre-surgery questionnaires and who were included in the responsiveness

Table 4—Acceptability, Reliability and Tests of Scaling Assumptions of the POS-Hand/Arm

Psychometric property	*POS-Hand/Arm Pre-surgery (29 items)*	*POS-Hand/Arm Post-surgery (33 items)*
Acceptability (%; $N = 156-203$)		
Missing data for items	0.0–6.1	0.5–4.9
Cronbach's α		
Reliability ($N = 124-203$)		
Physical Activities	0.91	0.94
Symptoms	0.88	0.93
Psychological Functioning/Cosmetic Appearance	0.89	0.92
Satisfaction	NA	0.76
Item–total correlations (mean, range; $N = 156-203$)		
Physical Activities	0.58 (0.38–0.77)	0.72 (0.62–0.78)
Symptoms	0.65 (0.43–0.76)	0.71 (0.56–0.81)
Psychological Functioning/Cosmetic Appearance	0.73 (0.67–0.79)	0.80 (0.74–0.87)
Satisfaction	NA	0.57 (0.40–0.67)
Test–retest reliability (ICC; $N = 37$)		
Physical Activities	NA	0.73
Symptoms	NA	0.87
Psychological Functioning/Cosmetic Appearance	NA	0.94
Satisfaction	NA	0.93
Tests of scaling assumptions ($N = 156-203$)		
Scaling successes (%)	96.5	93.9

sub-sample, 77 (67%) returned post-surgery questionnaires. Finally, 37 (65%) patients in the test-retest sub sample returned completed questionnaires.

Acceptability (Table 4): In general, the proportion of missing data was low and responses were well distributed across response categories.

Reliability (Table 4): All four POS-Hand/Arm scales met the Cronbach's alpha criterion for internal consistency reliability. Item–total correlations ranged from 0.38 to 0.87. Test–retest intraclass correlations all exceed 0.72.

Construct validity (within-scale analyses; Table 5): Within-scale analyses supported the construct validity of the POS-Hand/Arm (intercorrelations between scales range from 0.38 to 0.73). Tests of scaling assumptions revealed that the majority of items were classified as scaling successes (Table 4).

Construct validity (comparison with other measures; Table 5): The results supported hypotheses about correlations between the POS-Hand/Arm and other measures (SF-36 $r = 0.13-0.62$, DASH $r = 0.63-0.88$ and MHQ $r = 0.19-0.77$) and correlations with age, sex, and social class were low. In particular, the Physical Activities and Symptoms scales of the POS-Hand/Arm were highly correlated with DASH Function/Symptom Score ($r \geqslant 0.80$) and MHQ Work and Activities of Daily Living (ADL) scales ($r \geqslant 0.72$). Also, high correlations were found between the POS-Hand/Arm Symptoms scale and MHQ Pain scale ($r = 0.77$). However, moderate correlations were found first between the POS-Hand/Arm Physical Activities scale and MHQ Pain, Function and Aesthetics scales ($r \geqslant 0.30$) and second between the POS-Hand/Arm Symptoms scale and MHQ Pain and Aesthetics scales ($r \geqslant 0.31$). All correlations between the POS-Hand/Arm Psychological Functioning/Cosmetic Appearance and DASH and MHQ scales were low to moderate ($r = 0.19-0.63$).

Responsiveness (Table 6). All scales of the POS-Hand/Arm showed moderate to large effect sizes in patients who improved (range 0.47–1.08) and differentiated between improvers and non-improvers ($P < 0.0001$), indicating good responsiveness.

DISCUSSION

Evidence-based health policy emphasizes the importance of using scientifically rigorous patient-based outcome measures to evaluate the impact of disease and treatment (Commission for Health Improvement, 2003). The POS-Hand/Arm is a psychometrically sound surgical outcome measure that can be used before and after surgery to evaluate and compare new techniques, surgical teams and units. In addition, it is acceptable to patients and has a simple checklist format that patients can complete easily and quickly. The POS-Hand/Arm assesses outcomes that are both surgically relevant (e.g. power, mobility, sensation and pain) (British Association of Plastic Surgeons, 2000) and important to patients (e.g. daily activities, psychological functioning and cosmetic appearance).

Although there is overlap in some of the domains measured by the POS-Hand/Arm and existing measures such as the DASH and MHQ, the POS-Hand/Arm is the only measure that evaluates outcomes of surgery, which has been developed for use in the UK with direct patient input (Fitzpatrick et al., 1999). The overlap supports the validity of all three as measures of the health impact of hand/arm disorders, although this study provides evidence that the POS-Hand/Arm clearly assesses different areas of psychological/cosmetic functioning and patient satisfaction.

The POS-Hand/Arm, like other patient-based measures of outcome in hand surgery, evaluates different

Table 5—Validity – Intercorrelations in the scales of the POS-Hand/Arm and correlations between the POS-Hand/Arm and other measures

Instrument	*Scale/dimension/variable*	*Physical Activities*	*POS-Hand/Arm Scale Symptoms*	*Psychological Functioning/ Cosmetic Appearance*
POS-Hand/Arm (pre-surgery)	Physical activities	—	—	—
	Symptoms	0.64	—	—
	Psychological Functioning/Cosmetic Appearance	0.52	0.38	—
POS-Hand/Arm (post-surgery)	Physical activities	—	—	—
	Symptoms	0.73	—	—
	Psychological Functioning/Cosmetic Appearance	0.69	0.60	—
	Satisfaction	0.45	0.55	0.44
SF-36	Role limitations–emotional	0.28	0.21	0.46
	Role limitations-physical	0.45	0.34	0.40
	Bodily pain	0.49	0.49	0.37
	Vitality	0.30	0.25	0.35
	General health perceptions	0.29	0.13	0.45
	Social functioning	0.38	0.23	0.43
	Physical functioning	0.62	0.23	0.48
	Mental health	0.18	0.31	0.50
	Physical component summary score (PCS)	0.58	0.35	0.39
	Mental component summary score (MCS)	0.08	0.21	0.42
DASH	Function/symptom score	0.88	0.80	0.63
MHQ	ADL	0.77	0.73	0.26
	Work	0.73	0.72	0.19
	Pain	0.56	0.77	0.28
	Function	0.46	0.52	0.22
	Aesthetics	0.30	0.31	0.42
Age		0.04	0.16	0.10
Sex		−0.28	−0.20	−0.14
Social class		−0.12	−0.22	−0.31

Table 6—Responsiveness of the POS-Hand/Arm

	Physical Activities	*Symptoms*	*Psychological Functioning/ Cosmetic Appearance*	*Satisfaction*
Improved				
Change score: mean (SD)	12.7	20.0	9.9	—
P	<0.001	<0.001	<0.001	—
Effect size	0.66	1.08	0.47	—
No improvement				
Change score: mean (SD)	−2.9	4.4	−7.6	—
P	0.47	0.23	0.79	—
Effect size	−0.17	0.30	−0.44	—

aspects of outcome than clinical outcome measures such as grip strength and range of motion. This has been demonstrated in previous studies of hand/arm disorders (Chung et al., 1999; Katz et al., 1996) and in several other areas of health care evaluation (Cleary, 1997; Lamping et al., 2001). Direct comparisons between clinical and patient-based measures are not appropriate as they capture different aspects of outcome.

The POS-Hand/Arm has a number of uses. First it is ideal for comparative audit of surgical teams and units. The pre-surgery version can be used to adjust for case-mix so that post-surgical comparisons between groups can be performed on a fair basis. Second, routinely collected POS-Hand/Arm outcome data can be used to identify subgroups of patients who benefit most from different types of surgery. Third, the new measure

provides a rigorous method for evaluating both the benefits and adverse events of treatment in comparative effectiveness studies regarding new surgical techniques. As a site-specific measure of outcome in surgery for hand/arm disorders, the POS-Hand/Arm can be used alone, or in conjunction with a generic measure such as the SF-36. It is recommended that comprehensive assessment of outcome should include a combination of generic and specific measures (Patrick and Deyo, 1989).

One limitation of this study is that the psychometric properties of the POS-Hand/Arm were evaluated in a sample of hand/arm plastic surgery patients from three hospitals in the south-east region of England, and it needs to be examined in other patient samples. Some properties of a measure are inherently difficult to evaluate, and in particular, the responsiveness of the POS-Hand/Arm requires further testing in clinical trials. However, to some extent scale development is an open and never-ending process and, as more data accumulates, psychometric estimates may need to be revised and slight modifications may be required. The usefulness of the new POS-Hand/Arm as a surgical outcome tool can only be demonstrated through multiple applications in different studies.

Acknowledgements

We wish to thank the people with hand/arm conditions who participated in this study and the many people who helped with patient recruitment: Mr Tony Heywood, Mr Peter Budny, Mr Mark Scott, Mrs Elizabeth Neal and staff in the Department of Plastic and Reconstructive Surgery at Stoke Mandeville Hospital; Mr Peter Mahaffey, Mr Miles Dickson, Mr Nick James and staff in the Department of Plastic Surgery at Lister Hospital; and Mr Adi Grobelaar and Ms Sue Forbes and staff in the Department of Plastic Surgery at Mount Vernon Hospital. We would also like to acknowledge advice provided by Dr Jeremy Hobart and Dr Sara Schroter during item reduction and psychometric analysis. The study was funded by the British Association of Plastic Surgeons and Stoke Mandeville Burns and Reconstructive Surgery Research Trust in the first year, the NHS Executive Anglia and Oxford Region R&D Directorate (Years 2–3), and the NHS Executive North Thames Region Health Services Research Fellowship to Dr Cano (Years 2–4).

References

Amadio PC (1997). Outcomes assessment in hand surgery: What's new? Clinics in Plastic Surgery, 24: 191–194.

British Association of Plastic Surgeons (2000). Plastic Surgery. www.baps.co.uk.

Cano SJ, Browne JP, Lamping DL (2003). Patient-based measures of outcome in plastic surgery: current approaches and future directions. British Journal of Plastic Surgery, 57: 1–11.

Chung K, Hamill J, Walters M, Hayward R (1999). The Michigan hand outcomes questionnaire (MHQ): assessment of responsiveness to clinical change. Annals of Plastic Surgery, 42: 619–622.

Chung KC, Pillsbury MS, Walters MR, Hayward RA (1998). Reliability and validity testing of the Michigan hand outcomes questionnaire. Journal of Hand Surgery , 23A: 575–587.

Cleary P (1997). Subjective and objective measures of health: which is better when? Journal of Health Services Research and Policy, 2: 3–4.

Commission for Health Improvement (2003). Clinical audit programme for England and Wales. www.chi.nhs.uk.

Cronbach LJ (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16: 297–334.

Department of Health (2003). Hospital Episode Statistics. www.doh.gov.uk/hes.

Dillman D. *Mail and telephone surveys: the total design method*. New York, Wiley, 1978.

Fitzpatrick R, Jenkinson C, Klassen A, Goodacre T (1999). Methods of assessing health-related quality of life and outcome for plastic surgery. British Journal of Plastic Surgery, 52: 251–255.

Goldacre MJ, Ferguson JA, Goodacre TEE (1996). Demographic and epidemiological aspects of plastic surgery: profile of a supradistrict speciality. Health Services Management Research, 9: 55–60.

Higgs PE, Edwards D, Martin DS, Weeks PM (1995). Carpal tunnel surgery outcomes in workers: effect of workers' compensation status. Journal of Hand Surgery, 20A: 354–360.

Hudak PL, Amadio PC, Bombardier C and the Upper Extremity Collaborative Group (UECG) (1996). Development of an upper extremity outcome measure: the DASH (Disabilities of the Arm, Shoulder and Hand). American Journal of Industrial Medicine, 29: 602–608.

Katz JN, Punnett L, Simmons BP, Fossel AH, Mooney N, Keller RB (1996). Workers' compensation recipients with carpal tunnel syndrome: the validity of self-reported health measures. American Journal of Public Health, 86: 52–56.

Kazis LE, Anderson JJ, Meenan RF (1989). Effect sizes for interpreting changes in health status. Medical Care, 27: S178–S189.

Lamping DL, Browne JP, Smith SC (2001). Patient-based measures to evaluate surgical outcomes: myths and realities. Annals of the Royal College of Surgeons of England (Suppl), 83: 78–81.

Lamping DL, Schroter S, Kurz X, Kahn SR, Abenhaim L. (2003). Evaluation of outcomes in chronic venous disorders of the leg: development of a scientifically rigorous, patient-reported measure of symptoms and quality of life. Journal of Vascular Surgery, 37: 410–419.

Lamping DL, Schroter S, Marquis P, Marrel A, Duprat-Lomon I, Sagnier PP (2002). The community-acquired pneumonia symptom questionnaire: a new, patient-based outcome measure to evaluate symptoms in patients with community acquired pneumonia. Chest, 122: 920–929.

Levine DW, Simmons BP, Koris MJ, Daltroy LH, Hohl GG, Fossel AH, et al (1993). A self-administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome. Journal of Bone and Joint Surgery (American) 75: 1585–1592.

Macey AC, Burke FD, Abbott K, Barton NJ, Bradbury E, Bradley A et al. (1995). Outcomes of hand surgery. Journal of Hand Surgery (British), 20: 841–855.

Nunnally JC, Bernstein IH *Psychometric theory*. 3 edn. New York, McGraw-Hill, 1994.

Patrick D, Deyo R. (1989) Generic and disease-specific measures in assessing health status and quality of life. Medical Care, 27: S217–S232.

Scientific Advisory Committee of the Medical Outcome Trust (2002). Assessing health status and quality-of-life instruments: Attributes and review criteria. Quality of Life Research, 11: 193–205.

Streiner DL, Norman GR *Health measurement scales: a practical guide to their development and use*. 2nd edn. Oxford, Oxford University Press, 1995.

Vamos M, White G, Caughey D (1990). Body image in rheumatoid arthritis: the relevance of hand appearance to desire for surgery. British Journal of Medical Psychology, 63: 267–277.

Ware J, Sherbourne C (1992). The MOS 36-item Short-Form Health Survey (SF-36): I. Conceptual Framework and item selection. Medical Care, 30: 473–483.

Ware JE, Harris WJ, Gandek B, Rogers BW, Reese PR (1997). MAP-R for windows: multitrait/multi-item analysis program-revised user's guide version 1. Boston, MA: Health Assessment Lab.

Appendix I

The final version of the pre-surgery POS-Hand/Arm includes 29 evaluative items forming three scales: physical activities (12 items), symptoms (12 items), psychological functioning and cosmetic appearance (5 items). Three summary scale scores are generated by summing items and then transformed to a 0-100 scale. High scores indicate better health.

POS-Hand/Arm (Pre-Surgery Version)

Instructions: Patients like you who need an operation on the hand/arm are bothered by different problems. The following questions ask about problems you may have been bothered by during the past 4 weeks.

1. During the past 4 weeks, how much were you bothered by each of the following problems in your hand/arm? (Please circle the number in the box that best describes your situation.)

	Not at all	*A little*	*Moderately*	*Quite a bit*	*Extremely*
a. Pain in your hand/arm whilst performing your usual daily activities	1	2	3	4	5
b. Cramp in your hand/arm	1	2	3	4	5
c. Stiffness in your hand/arm	1	2	3	4	5
d. Joints locking in your hand/arm	1	2	3	4	5
e. Tightness in the skin of your hand/arm	1	2	3	4	5
f. Problems with grip	1	2	3	4	5
g. Numbness in your hand/arm (eg loss of sensation, hand/arm felt dead)	1	2	3	4	5
h. Pins and needles or tingling sensations in your hand/arm	1	2	3	4	5
i. Swelling in your hand/arm	1	2	3	4	5
j. Weakness/loss of strength in your hand/arm	1	2	3	4	5
k. Restricted movement of your hand/arm (ie having difficulty with your normal range of movement)	1	2	3	4	5

2. During the past 4 weeks, has your hand/arm problem limited you in your usual daily activities? Please indicate whether your hand/arm problem limits you a lot, limits you a little, or does not limit you at all in these activities by circling the appropriate number. If you do not usually do a particular activity listed below, please circle the number in the last column (''Don't usually do'').

	Limited a lot	*Limited a little*	*Not limited at all*	*Don't usually do*
a. Picking up coins	1	2	3	4
b. Folding paper	1	2	3	4
c. Pouring from a teapot	1	2	3	4
d. Using the television/video remote control	1	2	3	4
e. Dressing yourself	1	2	3	4
f. Putting on gloves	1	2	3	4
g. Shaving or putting on make-up	1	2	3	4
h. Writing	1	2	3	4
g. Washing up dishes	1	2	3	4
h. Holding the telephone receiver	1	2	3	4
i. Turning a key in a lock	1	2	3	4
j. Going to the toilet	1	2	3	4

3. During the past 4 weeks, how often did you:

	All of the time	*Most of the time*	*Some of the time*	*A little of the time*	*None of the time*
a. Not get the amount of sleep that you needed because of your hand/arm problem?	1	2	3	4	5

4. During the past 4 weeks, how often has your hand/arm problem caused you to feel:

	All of the time	*Most of the time*	*Some of the time*	*A little of the time*	*None of the time*
a. Unattractive?	1	2	3	4	5
b. Self-conscious about your hand/arm?	1	2	3	4	5
c. A lack of confidence?	1	2	3	4	5

5. During the past 4 weeks, how bothered were you by:

	Not at all	*A little*	*Moderately*	*Quite a bit*	*Extremely*
a. The appearance of your hand/arm when performing your daily activities?	1	2	3	4	5
b. People's reactions to your hand/arm?	1	2	3	4	5

Appendix II

The post-surgery version includes 33 evaluative items forming four scales, including the same three scales as the 29-item version, plus a satisfaction scale (4 items). Four summary scale scores are generated by summing items and then transformed to a 0–100 scale. High scores indicate better health.

POS-Hand/Arm (Post-Surgery Version)

[AS APPENDIX I—ITEMS 1–4] PLUS THE FOLLOWING:

5. During the past 4 weeks, how much were you bothered by the following problem in your hand/arm? (Please circle the number in the box that best describes your situation.)

	Not at all	*A little*	*Moderate-ly*	*Quite a bit*	*Extremely*
a. Lumpiness in the operation scar?	1	2	3	4	5

6. Are the results of the operation on your hand/arm:
1 Better than you expected?
2 About what you expected?
3 Worse than you expected?

7. Has your recovery from operation on your hand/arm so far been:
1 Faster than you expected?
2 About the same as you expected?
3 Slower than you expected?
4 You did not know how long it would take?

8. If a friend had similar hand/arm problems that you had before your operation, would you recommend the same operation you had?
1 Definitely would recommend it
2 Probably would recommend it
3 Not sure
4 Probably would not recommend it
5 Definitely would not recommend it

Received: 13 November 2003
Accepted after revision: 2 June 2004
Dr. Stefan Cano, Senior Research Fellow, Neurological Outcomes Measures Unit, National Hospital for Neurology and Neurosurgery, 4th Floor Queen Mary Wing, Queen Square, London, WC1N 3BG, UK. Tel.: + 44 207 837 3611 × 3436; fax: + 44 207 813 6505; E-mail: scano@rfc.ucl.ac.uk

doi:10.1016/j.jhsb.2004.06.002 available online at http://www.sciencedirect.com