

**PATIENT-REPORTED  
OUTCOME  
MEASUREMENT  
GROUP, OXFORD**

**A STRUCTURED REVIEW  
OF PATIENT-REPORTED  
OUTCOME MEASURES FOR  
PROCEDURES FOR CARPAL  
TUNNEL SYNDROME**

**Report to the Department of Health  
2011**



**patient-reported**

**O**

**U**

**T**

**C**

**O**

**M**

**E**

**measures**



**A STRUCTURED REVIEW OF PATIENT-REPORTED  
OUTCOME MEASURES FOR PROCEDURES FOR CARPAL  
TUNNEL SYNDROME, 2011**

Monica Hadi  
Elizabeth Gibbons  
Ray Fitzpatrick

*Patient-reported Outcome Measurement Group  
Department of Public Health  
University of Oxford*

*Copies of this report can be obtained from:*

Elizabeth Gibbons  
Senior Research Officer  
Patient-reported Outcome Measurement Group

tel: +44 (0)1865 289402

e-mail: [elizabeth.gibbons@dph.ox.ac.uk](mailto:elizabeth.gibbons@dph.ox.ac.uk)



# **A STRUCTURED REVIEW OF PATIENT-REPORTED OUTCOME MEASURES FOR PROCEDURES FOR CARPAL TUNNEL SYNDROME**

## **CONTENTS**

<b>Executive summary</b>	<b>1</b>
<b>1. Introduction, methods, search terms and results</b>	<b>3</b>
<b>2. Generic PROMs</b>	<b>8</b>
<b>3. Condition-specific PROMs</b>	<b>11</b>
<b>4. Discussion and recommendations</b>	<b>20</b>
<b>Appendix A – Search strategy and sources</b>	<b>23</b>
<b>Appendix B – Search strategy for PubMed</b>	<b>24</b>
<b>Appendix C – Appraisal of the methodological quality of PROMs</b>	<b>25</b>
<b>Appendix D - Appraisal criteria</b>	<b>26</b>
<b>Appendix E – Summary of instruments</b>	<b>27</b>
<b>References</b>	<b>29</b>



## **EXECUTIVE SUMMARY**

### **Aims of the report**

The aims of this report are to identify Patient-reported Outcome Measures (PROMs) for procedures for Carpal Tunnel Syndrome (CTS).

The methods of the review are described and the results of the search including sources and search terms used to identify relevant published research. Details of this evidence are presented firstly for generic health status measures, preference-based measures, and condition or procedure-specific PROMs evaluated for procedures for CTS. The report concludes with discussion and recommendations.

### **Results**

Two generic instruments, which have been evaluated with patients with carpal tunnel syndrome (CTS), were identified in this review:

1. SF-36
2. Nottingham Health Profile (NHP)

Four condition-specific measures were identified in this review:

1. Boston Carpal Tunnel Questionnaire (BCTQ)
2. Disability of Hand and Shoulder Questionnaire (DASH)
3. Michigan Hand Questionnaire (MHQ)
4. Patient Evaluation Measures (PEM)

### **Recommendations**

Based on the volume of evaluations and good measurement and operational characteristics, the following are highlighted as promising PROMs for the evaluation of the quality of NHS services:

#### **Generic:**

For the measurement of comprehensive general health status, the SF-36 has more psychometric supportive evidence, is acceptable to patients, and can be considered.

#### **Condition-specific:**

Based on the volume and quality of evidence, the BCTQ is clearly the most promising condition-specific measure. There is also evidence of good psychometric performance for the DASH. However, the BCTQ has the advantage of being a CTS specific instrument and is supported by extensive research in this population.

Despite lack of evidence for preference based measures in this review, the EQ-5D is widely used in the UK context for evaluation of a range of elective procedures and can be considered further for the assessment of general health in patients undergoing carpal tunnel surgery.





## 1. INTRODUCTION

Patient-reported outcome measures (PROMs) offer enormous potential to improve the quality and results of health services. They provide validated evidence of health from the point of view of the user or patient. They may be used to assess levels of health and need in populations, and in users of services they can provide evidence of the outcomes for the purposes of audit, quality assurance and comparative performance evaluation. They may also improve the quality of interactions between health professionals and individual service users.

Lord Darzi's Interim Report on the future of the NHS recommends that patient-reported outcome measures (PROMs) should have a greater role in the NHS (Darzi 2007). The new Standard NHS Contract for Acute Services, introduced in April 2008, included a requirement to report from April 2009 on patient-reported outcome measures (PROMs) for patients undergoing Primary Unilateral Hip or Knee replacements, Groin Hernia surgery or Varicose Vein procedures. Furthermore, Lord Darzi's report 'High Quality Care for All' (2008) outlines policy regarding payments to hospitals based on quality measures as well as volume. These measures include PROMs as a reflection of patients' experiences and views. Guidance has now been issued regarding the routine collection of PROMs for the selected elective procedures (Department of Health, 2008) and since April 2009, the routine collection of PROMs for the selected elective procedures has been implemented and is ongoing. This report expands on this by reviewing the evidence of PROMs for other common elective procedures.

There are three broad categories of PROMs: generic health status, preference-based, and condition- or population-specific-measures. Generic instruments comprise items intended to be relevant to the widest range of patient conditions and the general population. Preference-based measures are also broad in content but additionally provide utilities or values regarding health (for use in, for example, cost-utility analyses of interventions). Condition-specific instruments are often more focused on a particular disease or health condition (for example, diabetes), a patient population (for example, older people), a specific problem or symptom (for example, pain), or a described function (for example, activities of daily living). For any given area of health, condition-specific instruments may have greater clinical appeal due to the inclusion of content specific to particular conditions, and the likelihood of increased responsiveness to interventions.

It has been recommended that a combination of a generic or utility measure with a specific measure be used in the assessment of patient-reported health outcomes, on the grounds that the complementary content of the two types of measure, when combined, should assess a full range of aspects of health relevant to the particular population concerned. However, consensus is often lacking as to which instrument to use for specific purposes and contexts (Garratt et al., 2002). Structured reviews of PROMs for specific health conditions or populations can provide guidance for selection. An evidence-based approach strengthens recommendations from these reviews.

Selection criteria have been defined for assessing the quality of existing PROMs (McDowell, 2006; Fitzpatrick et al., 1998; Streiner and Norman, 2003). These include

measurement issues, such as reliability, validity, responsiveness and precision, as well as practical issues, such as acceptability and feasibility.

## **Carpal Tunnel Syndrome**

Carpal tunnel syndrome (CTS) is a condition in which the median nerve is compressed at the wrist, leading to paresthesias, numbness and muscle weakness in the hand. To date, CTS is the most common compression neuropathy, estimated to occur in 5% of women and 3% of men (Atroshi et al, 1999). Most cases of CTS develop in people who are between ages of 45-64 and if left untreated, CTS may lead to permanent "nerve damage", i.e. irreversible numbness, muscle wasting and weakness.

Most cases of CTS are idiopathic as the exact reasons for the condition are unknown. However, a number of risk factors for the development of CTS have been identified including family history, age, repetitive strain to the wrist and other health conditions (such as rheumatoid arthritis, diabetes, and obesity). CTS is also common in pregnant women, however, cases that occur during pregnancy are usually resolved after birth. This may be due to the fluid retention that often occurs during pregnancy, placing additional pressure on the carpal tunnel.

There are two general types of treatment for CTS; conservative treatment and surgical release. The term conservative generally refers to treatment by medication, such as steroid injections, oral steroids, anti-inflammatory drugs, pyridoxine (Vitamin B6) and botulinum toxin B but can also be used to imply the use of alternative therapies such as exercise therapy splinting, ultrasound, yoga and laser treatment. Surgical procedures employ either endoscopic or open carpal tunnel release. It is recommended when there is static (constant, not just intermittent) numbness, muscle weakness, or atrophy which can no longer be controlled by conservative treatments. In general, milder cases of CTS usually respond well to non-surgical treatment whilst more severe cases usually require surgery to reduce the pressure on the median nerve.

The Hospital episode statistics (HES) reported that approximately 60,000 CTS procedures were carried out in England during the HES financial year 2008-2009. Of these, 65% were females, 32% were males and the mean age of patients was reported at 58 years.

### **Aim of the report**

The aim of this report is to identify Patient-reported Outcome Measures (PROMs) which have been evaluated with patients with carpal tunnel syndrome.

### **Structure of the report**

The methods of the review are described, including search strategies and search terms used to identify relevant published research regarding PROMs for people with Carpal Tunnel Syndrome. The report focuses on evidence and recommendations for PROMs evaluated in patients undergoing this procedure.

## Methods

Methods adopted were as described in previous reviews performed by the PROM group, Oxford. Comprehensive searches were conducted; articles retrieved were assessed for relevance and evidence of measurement performance and operational characteristics abstracted for each PROM identified.

### a) Search sources and terms

Several sources were searched to identify relevant articles. The primary source of evidence was the bibliographic database compiled by the PROM group in 2002 with funding from the Department of Health and hosted by the University of Oxford. In 2005, it became the property of the NHS Information Centre for Health & Social Care. The PROM database comprises over 16,000 records (available online at <http://phi.uhce.ox.ac.uk>) downloaded from several electronic databases using a comprehensive search strategy (available on request); these records have been assessed as eligible for inclusion in the bibliography and assigned keywords. The titles and abstracts of these, as well as a further 14,000 records identified as potential inclusions, were searched using the terms 'carpal tunnel' OR 'median nerve AND entrapment' OR 'median nerve AND neuropath'. The search strategy used in this review and the flowchart illustrating the search process are detailed in Appendix A and B, respectively.

Supplementary searches included scanning the reference lists of review articles and others, checking instrument websites, where found, and drawing on other bibliographic resources. Hand-searching of titles of key journals from 2007 to July 2009 was conducted. The following journals were selected:

- European Journal of Hand Surgery
- American Journal of Hand Surgery
- The Journal of Bone and Joint Surgery
- Medical Care
- Health and Quality life Outcomes
- Quality of Life Research

The National Institute for Health Research: Health Technology Assessment Programme, published research was also searched.

In addition, English-language PubMed records for the period 2007-9 (to 18 September 2009) were searched using a modified version of the Carpal Tunnel-specific terms listed above, combined with a search strategy to identify PROMs devised by the PROM Group in collaboration with the Knowledge Centre of the University of Oxford (see Appendix B).

### b) Inclusion criteria

Published articles were included if they provided evidence of measurement and/or practical properties of relevant PROMs (Fitzpatrick et al., 1998).

### *Population*

- Patients with carpal tunnel syndrome (CTS)
- English-speaking populations.

### ***Study design selection***

- studies where a principal PROM is being evaluated;
- studies evaluating several PROMs concurrently;
- trials or studies evaluating the effectiveness of interventions; where a PROM is used as an endpoint;
- prospective studies measuring patient-reported outcomes where data is available for a PROM in terms of measurement performance or operational characteristics.

### ***Specific inclusion criteria for generic, preference-based and condition-specific instruments***

- the instrument is patient-reported;
- there is published evidence of measurement reliability, validity or responsiveness following completion in the specified patient population;
- evidence is available from English-language publications, and instrument evaluations conducted in populations within the UK, North America, or Australasia;
- the instrument will ideally be multi-dimensional. It is at the reviewer's discretion to include RPOMs which are specific to a health condition but have a narrow focus, for example, a specific dimension of health, such as symptoms.

### **Exclusions**

- studies using clinician-rated instruments;
- studies evaluating the performance of non-patient reported measures of functioning or health status where a PROM is used as a comparator;
- studies with very small samples, i.e. fewer than 10 participants
- studies using incomplete versions of instruments.

### **c) Data extraction**

For all PROMs included in the review, evidence is reported for the following measurement criteria:

- reliability
- validity
- responsiveness
- precision

Operational characteristics, such as patient acceptability and feasibility of administration for staff, are also reported.

### **d) Assessment of methodological quality of PROMs**

Assessment and evaluation of the PROMs was performed by means of the criteria described in Appendix C.

## Results

Searches identified nearly 300 potentially relevant records. When assessed against the review inclusion criteria, 55 articles were included in the review (Table 1).

**Table 1: Number of articles identified by the literature review**

<i>Source</i>	<i>Results of search</i>	<i>Number of articles included in review</i>
<b>PROM bibliography: 30,350</b>	60	21
<b>PubMed 2007-September 2009</b>	221	21
<b>Hand searching</b>	17	13
<b>TOTAL</b>		<b>55</b>

## **GENERIC PROMs**

Two generic instruments were identified:

- a) SF-36
- b) NHP

### **a. SF-36**

The Medical Outcomes Study (MOS) Short Form 36-item Health Survey (SF-36) is derived from the work of the Rand Corporation during the 1970s (Ware and Sherbourne., 1992; Ware et al., 1994; Ware., 1997). It was published in 1990 after criticism that the SF-20 was too brief and insensitive. The SF-36 is intended for application in a wide range of conditions and with the general population. Ware et al., (1994; 1997) proposed that the instrument should capture both mental and physical aspects of health. International interest in this instrument is increasing, and it is by far the most widely evaluated measure of health status (Garratt et al., 2002).

Items were derived from several sources, including extensive literature reviews and existing instruments (Ware and Sherbourne., 1992; Ware and Gandek., 1998; Jenkinson and McGee., 1998). The original Rand MOS Questionnaire (245 items) was the primary source, and several items were retained from the SF-20. The 36 items assess health across eight domains (Ware., 1997), namely bodily pain (BP: two items), general health perceptions (GH: five items), mental health (MH: five items), physical functioning (PF: ten items), role limitations due to emotional health problems (RE: three items), role limitations due to physical health problems (RP: four items), social functioning (SF: two items), and vitality (VT: four items). An additional health transition item, not included in the final score, assesses change in health.

There are between two and six categorical response options for each item. Scoring uses a weighted scoring algorithm and a computer-based programme is recommended. Eight domain scores give a health profile; scores are transformed into a scale from 0 to 100, where 100 denotes the best health. Scores can be calculated when up to half of the items are omitted. Two component summary scores for physical and mental health (MPS and MCS, respectively) can also be calculated. A version of the SF-36 plus three depression questions has been developed and is variously called the Health Status Questionnaire (HSQ) or SF-36-D.

The SF-36 can be self-, interview-, or telephone-administered.

A total of 10 articles evaluating the SF-36 are included in the review. Of these, 2 were conducted in the UK (Amadio et al. (1995; Ettema et al. (2006).

### ***Validity***

Convergent construct validity has been demonstrated through observations of significant relationship between hand disability and the 8 scales of the SF-36 (Redmond et al., 2009). These ranged from a strong relationship for the PCS (-0.71) to a moderate relationship for the MCS (-0.39).

Discriminant validity was demonstrated by Turner et al. (2004) as patients with CTS and lower back pain had lower PCS and MSC scores than patients with CTS, indicating lower quality of life.

### ***Responsiveness***

In a UK investigating of outcomes pre- and post carpal tunnel release surgery, Amadio et al. (1995) found improved scores in RP, RE and Pain sub-scales following surgical release.

Other work has shown moderate to larger SRMs only in Pain and physical component summary subscales (PCS) following surgery. A study by Bessette et al. (1998) assessed 196 patients undergoing carpal tunnel release and found BP and RP scales to be the most sensitive to change. Furthermore, a more recent study by Appleby et al. (2009) showed improved scores only in the PCS; and PF, RP and BP.

Conversely, other investigations have failed to support these findings. Vaile et al. (2009) assessed improvements after corticosteroid injection and found this to be significant only in the pain subscale. Another UK study by Ettema et al. (2006) found no mean difference in the surgery group compared to the treatment group and more recently, a randomised controlled study of 101 patients undergoing either surgery or non-surgical therapy found the SF-36 to not be responsive to clinical change (Jarvik et al., 2009). Moreover, the SF-36 has been shown to be less sensitive to change when compared to condition specific measures in patients with Carpal Tunnel Syndrome (Bessette et al., 1998; Gay et al., 2003).

However, these findings do not necessarily suggest that the SF-36 is not sensitive enough to detect change after carpal tunnel release surgery, but may merely reflect the conservative treatments to be less effective.

### ***Acceptability***

Good response rates have been reported across a number of studies. A study by Bessette et al. (1998) of 196 patients undergoing surgery reported a 96% response rate at 6 months post-surgery (preoperative measures were not reported). Similarly, Jarvik et al. (2009) recorded an 87% response rate at 12 months (N=116) and Gay et al. (2003) reported a consistent 93% response rate (N=34) at 6 weeks and again at 12 weeks after surgery.

Lower response rates have also been noted with Amadio et al. (1995) reporting a 64% postal completion at 3 months post-surgery. In a study investigating the benefits of surgery versus conservative therapy, Ettema et al. (2006) reported a 69% mailed follow-up response rate with a mean follow-up of 4.8 years.

No evidence was found on appropriateness, reliability, precision, interpretability or feasibility in English-speaking population studies regarding SF-36 in relation to Carpal Tunnel Syndrome.

### **b. The Nottingham Health Profile (NHP)**

The Nottingham Health Profile (NHP) was developed in the UK during the 1970s for use in the evaluation of medical or social interventions (Hunt et al., 1980). Instrument content was derived from over 2000 statements given by 768 patients with a variety of chronic ailments and other lay people.

Part I of the instrument has 38 items across six domains: bodily pain (BP), emotional reactions (ER), energy (E), physical mobility (PM), sleep (S), and social isolation (SI). All items are statements that refer to departures from normal functioning, and relate to feelings and emotional state rather than change in behaviour. Respondents answer 'yes' or 'no' according to whether or not they feel the item applies to them in general. Positive responses are weighted and summed to give six domain scores between 0 and 100, where 100 indicates maximum limitation.

Part II of the NHP is less widely used and provides a brief indicator of handicap. The instrument may be self-, interview-, or telephone-administered.

A total of two articles evaluating the NHP are included in the review. One of these studies was conducted in the UK (Rege et al., 2001).

#### ***Responsiveness***

One UK study reported significant changes for all dimensions of the NHP (Rege et al., 2001) in a sample of 96 patients undergoing Carpal Tunnel release surgery. Furthermore, it was possible to discriminate between levels of satisfaction as dissatisfied patients were found to have higher scores pre- and post carpal tunnel surgery.

A recent study by Vaile et al. (2009) of patients undergoing corticosteroid injection partially supported these findings. The study reported significant changes for the dimensions covering BP, ER, and PM. However, no change was shown for dimensions of E, S, and SI.

#### ***Acceptability***

Rege et al. (2001) reported a response rate of 80% at baseline and respectively a postal response rate of 72% four months after Carpal tunnel release.

No evidence was found on appropriateness, reliability, validity, precision, interpretability or feasibility in English-speaking population studies regarding NHP in relation to Carpal Tunnel Syndrome.



## CONDITION-SPECIFIC PROMS

Four condition specific instruments were identified:

- a) BCTQ
- b) DASH
- c) MHQ
- d) PEM

### **a. Boston Carpal Tunnel Questionnaire (BCTQ)**

The Boston Carpal Tunnel Questionnaire (BCTQ), also referred to as the Levine Scale, Brigham and Women's Carpal Tunnel Questionnaire and Carpal Tunnel Syndrome Instrument, is the most commonly used outcome measure in the assessment of patients with carpal tunnel syndrome.

Originally developed by Levine et al. (1993), the questionnaire comprises of two scales; a Symptom Severity Scale (SSS) and a Functional Status Scale (FSS). The SSS incorporates six items for evaluation (pain, test-retest reliability of paraesthesia, numbness, weakness, nocturnal symptoms and overall functional status) whilst the FSS includes twelve functional activity items commonly affected in CTS (e.g. writing and holding a cup).

The questionnaire was developed specifically for patients with CTS and both physicians and patients were involved the item generation process. The SSS consists of 11 questions and is scored on a five-point rating scale. The FSS contains 8 items which have to be rated for degree of difficulty on a five-point scale. The overall score for both scales was calculated as the mean of the items.

The BCTQ can be self-, interview-, or telephone-administered.

A total of 36 articles evaluating the BCTQ are included in the review. Of these, 10 were conducted in the UK. Over half of the studies were recent (post 2005).

### ***Reliability***

Test-retest reliability was established by Levine et al. (1993) and Pearson's correlation coefficient was reported for SSS (0.91) and for FSS (0.93), indicating good correlation of scores on re-test. A UK study by Amirfeyz et al. (2009), assessing test-retest at 2 weeks following carpal tunnel surgery, reported a high Pearson's coefficient for the SSS (0.78) and the FSS (0.78). However, Pearson's coefficient was used for these tests rather than ICC, and the latter study only had a small sample of 43 patients.

High internal consistency has been reported for both the SSS (0.89-0.92) and the FSS (0.88-0.91) (Levine et al., 1993; Atroshi et al., 1996). Moreover, Katz et al. (1996) reported an overall Cronbach's alpha of 0.88-0.96 (varying between intervention groups), indicating high inter-item correlation within each scale.

### ***Validity***

Convergent validity has been demonstrated by moderate correlations between the SSS scores and measures of grip (0.38) and pinch (0.47) strength and between FSS scores and grip (0.50) and pinch (0.60) strength (Levine et al., 1993). All correlations were in the predicted direction; worse status is associated with worse impairment.

Convergent validity has been demonstrated by high correlation between Pre- and postoperative BCTQ scores and physical measures such as grip, key pinch and three-jaw pinch (Katz et al., 1996; Appleby et al., 2009).

Furthermore, greater satisfaction with results after surgery has been associated with greater improvement in the scores for both the SSS (0.50) and the FSS (0.32-0.54) (Levine et al., 1993; Gay et al., 2003). Pain scores (SSS domain) have also been shown to correlate with BCTQ score, indicating more functional impairment as pain is increased (Turner et al., 2004).

Discriminant validity was demonstrated by Hobby et al. (2005b) as a significant association was found between pre-operative depression and anxiety scores and the mean pre-operative scores on the BCTQ. The instrument has also been shown to discriminate between BMI scores (Bodavula et al., 2007) and age (Porter et al., 2002). Yang et al. (2009) also noted that both the SSS and the FSS dimensions of the BCTQ were able to significantly discriminate between the presence and absence of symptoms in patients with CTS.

Discriminant validity has also been supported. Jarvik et al. (2009b), in a RCT (n=116), reported that patients assigned to surgery with distal motor latency of 0.5ms or more improved more than patients assigned to non-surgical therapy at both 6 and 12 months follow-up. Similarly, Atroshi et al. (1996) classified patients as probable CTS (n=47), possible CTS (n=31) and unlikely CTS (n=36) according to the Katz-Stirrat hand diagram and found that the mean SSS score in patients classified as probable CTS was significantly higher than the mean score in patients classified as possible or unlikely. However, these findings were not replicated for the FSS domain of BCTQ.

Moreover, Jarvik et al. (2008) found the mean baseline SSS and FSS scores to be significantly higher for patients who later (1 year follow-up) had carpal tunnel decompression, indicating strong predictive validity.

### ***Responsiveness***

Responsiveness has been demonstrated for both dimension of the BCTQ after carpal tunnel release surgery (Katz et al., 1995; Guyette and Wilgis., 2004; Hobby et al., 2005; Nesbitt et al., 2006). In a large study (n=462), Bodavula et al. (2007) noted that FSS scores improved in 6 months follow-up. Same effect has been reported for the SSS dimension in a UK study (Ozyurekogulu et al., 2006).

Katz et al (1995) reported the SSS and FSS domains of the BCTQ to be 2-4 times more responsive to clinical improvement than measures of neuromuscular impairment. In a later study, Katz et al. (1996b) reported an improvement in patients who had carpal tunnel surgery for both the SSS and FSS. This effect persisted over

the 30 months follow-up whilst patients who did not have surgery showed little change in clinical status at 6, 18 and 30 months follow-up (n=340).

Bessette et al. (1998) created a weighted version of the BCTQ by asking the patients the relative importance they assigned to seven specific symptoms and functional impairments. The weighted version was compared to the original BCTQ after carpal tunnel release and was found to be slightly more responsive to change (SRMs = 1.56; 1.36 respectively). Furthermore, both the weighted and original BCTQ were found to be more sensitive to change than the SF-36. Moreover, Gay et al. (2003) found the BCTQ (ES=1.30, SRM=1.21) to be more responsive to change than the DASH (ES=0.57, SRM=0.54) at 6 weeks and again at 12 weeks after surgery. However, this was a small study with 37 patients.

Porter et al. (2002) noted improved SSS and FSS scores at 6 months follow-up from surgery. Furthermore, the improvements in both domains were found to be negatively correlated with age, indicating that patients over the age of 60 had slower recovery in symptoms and function. In a UK study, Burke et al. (2006) reported improved SSS and FSS scores at in all cases 6 months after carpal tunnel release and indicated that those in the severe and moderate SSS group had the greatest decrease (improvement) in scores.

Responsiveness to clinical change has also been demonstrated in a number of randomised controlled trials (Kaye et al., 2007; Jarvik et al., 2009; 2009b; ) and this has been reported for both domains of the BCTQ (Ettema et al., 2006; Brininger et al., 2007). In a recent UK study, Cresswell et al. (2008) compared the effectiveness of 2 types of carpal tunnel decompression (TM Indiana Tome and standard limited Palmar open incision) and reported improved SSS and FSS scores 3 months after surgery (n=115). The authors also highlighted that both domains showed a temporary deterioration immediately after surgery, but improved from the pre-operative state evident by 2 weeks. However, in their 7 years follow-up, only the SSS was found to significantly discriminate between the two surgical groups.

Postoperative responsiveness has also been demonstrated by a number of smaller studies (n<50) in north America (Levine et al., 1993; Katz et al., 1995; Bradley et al., 2003; Leit et al., 2004; Luria et al., 2008; Appleby et al., 2009; Chatterjee & Price., 2009, Gordon et al., 2009) and in the UK (Amadio et al., 1995; Badger et al., 2008; Reid et al., 2009). Responsiveness after non-surgical interventions has also been demonstrated for both domains of the BCTQ (Burke et al, 2007; Moraska et al., 2008; Amirjani et al., 2009).

Furthermore, Chatterjee & Price (2009) found the BCTQ to have an increased sensitivity to change (SRM=1.22) after surgery when compared to the MHQ (SRM=0.8). Heebner and Roddet (2008), in a randomised controlled trial of patients with CTS, reported that unlike the DASH, the BCTQ domain was able to detect a significant difference between groups at 6 months as demonstrated by improved FSS scores. However, both studies were limited by their small sample size.

### ***Acceptability***

Postal response rates have been reported between 80-93% at 6-12 weeks follow-up (Gay et al., 2003; Brininger et al., 2007; Amirfeyz et al., 2009; Reid et al., 2009) and

between 67-93% at 6-12 months follow-up after carpal tunnel release (Katz et al., 1996; Porter et al., 2002; Heebner & Roddet., 2008; Jarvik et al., 2008; 2009). Luria et al. (2008) reported that all 41 patients completed post surgery evaluation at 1 year.

High response rates have also been reported for longer follow-up periods. Katz et al (1996b) reported a 79% response rate at 30 months follow-up (n=340) whilst Kaye et al (2007) reported an 85% response rate at 2 years follow-up (n=73). Moreover, Ettema et al (2006) reported a 69% mailed follow-up response rate with a mean follow-up of 4.8 years.

Lower response rates have also been reported. A response rate of 54% was reported by Chatterjee & Price (2009) at 6-18 months postoperatively. Cresswell et al. (2008) in a 7 years follow-up period reported a 58% postal response rate (n=115).

### ***Feasibility***

Greenslade et al. (2004) established the mean time to complete BCTQ to be 5.6 minutes.

No evidence was found on appropriateness, precision or interpretability in English-speaking population studies regarding BCTQ in relation to Carpal Tunnel Syndrome.

### **b.i. The Disability of Arm, Shoulder and Hand questionnaire (DASH)**

The Disability of Arm, Shoulder and Hand questionnaire (DASH) was originally developed as a joint initiative of the American Academy of Orthopaedic Surgeon (AAOS), the council of Musculoskeletal Speciality Societies (COMSS), and the Institute for Work and Health (Toronto, Ontario, Canada). Under the review of methodologists and clinical experts, a total of 821 items were derived from 13 outcome measurement scales (Hudak et al., 1996). Thereafter, items were reduced using various strategies, including expert and field testing with patients.

The final version of the DASH is a 30-item scale and consists of two dimensions; Physical Functioning and Symptoms. The instrument is not specific to CTS and can be used to measure function in people with any of several musculoskeletal disorders of the upper limb. There are 3 scales within Physical Functioning (physical, social, and psychological) and 5 scales within Symptoms (pain, weakness, tingling and numbness, and stiffness). The items in the questionnaire emphasize upper-extremity activities and are intended to measure disability.

The response options for each item are presented as 5-point Likert scales. The overall score is scaled to a score between 0 (no disability) and 100 (maximum disability).

Both DASH questionnaires can be self-, interview-, or telephone-administered.

A total of fifteen articles evaluating the DASH are included in the review. Of these, 3 were conducted in the UK (Hobby et al., 2005; Amirfeyz et al., 2009; Appleby et al., 2009)

### ***Reliability***

Test-retest reliability was established by (Beaton et al., 2001) in a group of 86 patients and interclass correlation coefficient (ICC) was reported at 0.96. Based on a 2 weeks test-retest, Greenslade et al. (2004) reported a Pearson's correlation of 0.9 in a group of 31 patients with CTS. These findings suggest a high reproducibility for DASH. Furthermore, a recent UK study by Amirfeyz et al. (2009) reported DASH to have higher test-retest reliability than BCTQ (DASH=0.88, BCTQ=0.78). However, Pearson's coefficient was used for this test, rather than ICC, and the study only had a small sample of 43 patients.

High internal consistency was demonstrated by Hobby et al. (2005), reporting Cronbach's alpha at 0.97.

### ***Validity***

Construct validity was demonstrated through correlation of scores between DASH and BCTQ (0.73) (Beaton et al., 2001) and similarly, between DASH and PEM before (0.66) and after (0.85) carpal tunnel release surgery (Hobby et al., 2005). In addition, a recent study by Appleby et al. (2009) reported a significant correlation between DASH and the FSS (0.71) and SSS (0.51) dimensions of the BCTQ.

Further evidence of convergent validity has been demonstrated with high correlation of scores between DASH and BCTQ at 6 weeks (0.9) and 12 weeks (0.87) post-operation. Lozano et al. (2008) also showed satisfaction and DASH scores to be correlated (-.56); the lower the DASH score, the higher the satisfaction. Pre- and postoperative DASH scores have also shown to correlate with physical measures such as grip, key pinch and three-jaw pinch (Appleby et al., 2009).

Discriminant validity has also been supported as statistically lower DASH scores have been shown for patients already working with their upper limb condition and able to continue doing so (Beaton et al., 2001). Statistically different scores were also found between patients who reported to be able to do all they want to do as opposed to those who were not able to do so. Thus, change was in the anticipated direction.

Furthermore, in a recent study by Redmond et al. (2009), female patients reported more disability with significantly greater DASH scores than male patients. However, this discrimination was not hypothesised by the authors.

### ***Responsiveness***

Significantly improved scores have been reported for both domains of DASH at 3 months (Beaton et al., 2001; Hobby et al., 2005; Greenslade et al., 2004; Appleby et al., 2009;) and at 6 months post surgery follow-up (Kadzielski et al., 2008; McMillan et al., 2009). Gay et al. (2003) in a sample of 40 patients undergoing carpal tunnel release found DASH to have greater responsiveness (SMR=1.13, ES = 1.01) than SF-36 but not BCTQ. Kotsis et al. (2006) recorded moderate responsiveness (SMR = 0.7) for DASH after carpal tunnel release surgery. However, two dimensions of MHQ (Pain and Satisfaction) were found to be more responsive than the DASH.

In contrast, Heebner and Roddet. (2008), in a randomised controlled trial of patients with CTS, reported that the DASH was not able to detect any significant changes between the standard treatment group and the intervention group. However, this may

be a result of the effectiveness of the intervention, rather than the responsiveness of the DASH.

### ***Acceptability***

Postal response rates have been reported between 80-93% at 6-12 weeks after carpal tunnel release (Beaton et al., 2001; Gay et al., 2003; Greenslade et al., 2004; Amirfeyz et al., 2009). At 6 months post carpal tunnel release, Kotsis et al. (2005) reported a 98% postal response rate (n=50) whilst Kadzielski et al. (2008) reported a 66% response rate (n=49). Moreover, Heebner & Roddet. (2008) reported a 67% attendance at 6 months follow-up appointment in a randomised controlled trial.

In contrast, Lozano et al. (2008) noted a low postal response rate of 41% when they contacted 200 patients, 2 years after carpal tunnel release surgery. However, no baseline measurements had been taken from patients.

### ***Feasibility***

Greenslade et al. (2004) established the mean time to complete DASH to be 6.8 minutes.

No evidence was found on appropriateness, precision or interpretability in English-speaking population studies regarding DASH in relation to Carpal Tunnel Syndrome.

## **b.ii. The Disability of Arm, Shoulder and Hand questionnaire (DASH)**

A shorter version of DASH, the Quick DASH, consists of 19 items, 11 of which are calculated from the original DASH questionnaire. Quick DASH has been shown to highly correlate with DASH (0.79) (Niekel et al., 2009). Bialosky et al. (2009) applied the quick-DASH for evaluation of their 3 weeks randomised controlled trial in 40 patients with CTS, and found an improvement in DASH scores over time, regardless of the assigned intervention.

## **c. The Michigan Hand Outcomes Questionnaire (MHQ)**

The Michigan Hand Outcomes Questionnaire (MHQ) was developed in 1998 for use in the assessment of outcome for various hand disorders (Chung et al., 1998).

Content was guided by several established questionnaires. For instance, items from the SF-36 and the Arthritis Impact Measurement Scale were used for dimensions relating to work performance and physical function whilst items from the McGill Pain Scale and CTQ were used for the development of questions for the pain domain. Additional items were generated by a group of patients.

A total pool of 100 questions were gathered and then reduced by the use of factor analysis. The instrument has 37 items and consists of six domains: (1) overall hand function (OHF), (2) activity of daily living (ADL), (3) pain (P), (4) work performance (WP), (5) aesthetics (A) and (6) patient satisfaction with hand function (PS).

All items are in the form of questions that are ranked on a scale of 1-5. In the pain scale, high scores indicate greater pain, while in the other five scales; high scores denote better hand performance.

The raw scale score for each of the six scales is the sum of the responses of each scale item. The raw score is converted to a score ranging from 0-100. The response categories for one of the questions is reversed and re-coded.

The instrument may be self-, interview-, or telephone-administered.

A total of eight articles evaluating the MHQ are included in the review. None of the identified studies have been conducted in the UK.

### ***Reliability***

Substantial test-retest reliability was established by Chung et al. (1998), with correlation scores ranging from 0.81 for the aesthetics scale to 0.97 for the ADL scale. High internal consistency has been shown for all scales in the MHQ with Cronbach's alphas greater than 0.85 (Chung et al., 1998) and 4 of 6 scales (OHF, ADL, WP, and PS) showing Cronbach's alphas higher than 0.90.

### ***Validity***

Several aspects of validity have been demonstrated by Chung et al. (1998) in their original evaluation of the MHQ. A panel of surgeons, hand therapists and patients with hand disorders evaluated the MHQ for content validity.

Construct validity was demonstrated as Factor analysis revealed six subscales to the MHQ; OHF, ADL, P, WP, A, and PS. These findings verified the authors' hypothesis.

Convergent validity has been demonstrated by moderate correlations between similar questions in the SF-12, which ask about physical limitations due to health, and three scales in the MHQ; ADL (0.64), work performance (0.54-0.58) and pain (0.79). However, Chung et al. (1998) argue that questions in the SF-12 only inquire about health and not about hand performance, which may in turn explain why the correlations were moderate and not strong.

Divergent validity was demonstrated by different subscale scores for different afflictions of the hand.

Support has been presented for predictive validity. Chung et al. (1998) found that of the six scales in the MHQ, the Pain scale was the strongest predictor of physical function, as measured by the physical function component in the SF-12. Moreover, the Aesthetics scale was the only significant predictor of mental function component, as measured by the mental function component in the SF-12.

The aesthetics scale of the MHQ has demonstrated discriminant validity as significantly higher scores were shown for patients with carpal tunnel syndrome than patients with rheumatoid arthritis who had severe hand deformities (Chung et al., 1998).

### ***Responsiveness***

Significantly improved scores have been shown post-surgery for all dimensions (Chung et al., 1999; Weber., 2005), with pain being the most responsive dimension (Klein et al., 2002). Kotsis et al. (2005) reported the MHQ to be more responsive to change than DASH in patients undergoing carpal tunnel release surgery. Indeed, the

MHQ dimensions Pain (SRM = 0.9) and Satisfaction (SRM = 1.1) were found to be the most responsive to change, even when compared to the function/symptom dimension (SRM = 0.7) of DASH. However, these findings were based on a small sample of 50 patients.

Further support for responsiveness has been established in patients undergoing Carpal tunnel release surgery (McMillan et al., 2009). Apart from noting significantly improved scores for all domains at 3 and 6 months, the MHQ was also able to detect improvements from 3 to 6 months post-operative. Similarly, Chatterjee & Price. (2009) reported improved scores across all scales after carpal tunnel decompression at 6 months follow-up.

In contrast, a more recent study by Shauver et al. (2009) of 53 patients undergoing Carpal Tunnel release surgery only showed significant improvements in three of six domains (function, work and pain) when related to patient satisfaction.

#### ***Acceptability***

High response rates of 88% (Weber et al., 2005) and 98% (Kotis et al., 2005) have been reported for postal surveys at 6 months post Carpal Tunnel release. Klein et al. (2002) study of 20 patients undergoing release surgery, reported 100% response rate at 6 months follow-up. Patients who were unwilling or unable to attend clinic were offered phone interviews.

Lower postal response rates of 54% (Chatterjee & Price., 2009) and 49% (Chung et al., 1999) have also been reported at 6-18 months postoperatively.

Chung et al. (1998), in their original testing of the MHQ, reported a response rate of 99% amongst 200 patients at a university based clinic. Similarly, a high 100% postal response rate was also reported by Shauver et al. (2009) at 9 months post surgery.

No evidence was found on appropriateness, precision, interpretability or feasibility in English-speaking population studies regarding MHQ in relation to Carpal Tunnel Syndrome.

#### **d. Patient Evaluation Measure (PEM)**

Developed by Macey et al. (1995), Patient Evaluation Measure (PEM) consists of three components; patients' opinion on delivery of care, hand health profile and overall assessment. The instrument has a total of 18 items which are scored using a seven-point scale. In all scales, low scores indicate positive outcomes.

The authors do not report on item generation and item reduction. Similarly, there is no report on other aspects of the instrument development.

The instrument may be self-, interview-, or telephone-administered.

A total of three articles evaluating the PEM are included in the review, all of which were conducted in the UK.



### ***Reliability***

Internal consistency was demonstrated by Hobby et al. (2005), reporting Cronbach's alpha at 0.94.

### ***Validity***

Support for construct validity has been presented in a study by Forward et al. (2006) and it was confirmed that the objective measures (PEM scores) correlated with the subjective measures (e.g. increased grip strength). Hobby et al. (2005) presented a negative correlation between pre-operative overall PEM scores and grip strength (-0.54), indicating that increased grip strength is associated with lower (better) questionnaire scores.

Convergent validity has been demonstrated by strong correlations between PEM and DASH scores both preoperatively (0.66) and postoperatively (0.85) in patients with carpal tunnel syndrome (Hobby et al., 2005). In addition, a strong postoperative correlation (0.77) was shown between the change in PEM and DASH scores (Hobby et al., 2005).

Discriminant validity was demonstrated by Hobby et al. (2005b) as a significant association was found between pre-operative depression and anxiety scores and the mean pre-operative PEM. This association was also shown to be weaker at six weeks post-operation and thereafter insignificant at six months follow-up.

### ***Responsiveness***

Hobby et al (2005b) reported a significant reduction in PEM scores in a sample of 97 patients undergoing Carpal Tunnel release surgery. This was apparent at six weeks as well as six months after surgery. Similarly, Hobby et al. (2005) reported the PEM (SRM of 0.95) to show a significantly greater responsiveness to change than the DASH (SRM of 0.43).

Forward et al (2006) have supported these findings. In their study of 118 patients undergoing carpal tunnel release surgery, PEM scores significantly improved at 8-9 weeks postoperatively.

### ***Acceptability***

Hobby et al. (2005b) reported a response rate of 92% at baseline and respectively a postal response rate of 81% at six weeks and 85% at six months after Carpal tunnel release. In addition, Forward et al. (2006) reported a response rate of 94% at 8-9 weeks follow-up.

In contrast, Hobby et al. (2005), in a baseline sample of 32 patients, reported a 75% attendance at clinic for follow-up assessment at 3-6 months postoperatively (mean of 18 weeks).

Simplicity was assessed by quantifying the help needed to fill in the questionnaire and Macey et al. (1995) claim that the questionnaire is easy to complete.

No evidence was found on appropriateness, precision, interpretability or feasibility in English-speaking population studies regarding PEM in relation to Carpal Tunnel Syndrome.

## **DISCUSSION AND RECOMMENDATIONS**

### **Generic PROMs**

Table 3 summarises the evidence of measurement and operational performance of the two generic PROMs (SF-36 and NHP) identified in this review applying the adapted appraisal criteria outlined in Appendix C. There were no studies identified which applied utility measures, specifically evaluations of the EQ-5D, with patients with CTS.

Based on the volume and quality of evidence, the SF-36 is clearly the most promising generic health measure. The NHP has some indication for good performance in patients with CTS but support for the psychometric properties is limited. Whilst the evidence for SF-36 in carpal tunnel syndrome is stronger, support for responsiveness is inconsistent. Therefore, if used with CTS patients, the SF-36 should be used alongside a condition-specific instrument.

### **Condition-specific PROMs**

Four condition-specific PROMs were identified (BCTS, DASH, MHQ, PEM). Table 3 shows the appraisals of the evidence for each of the condition specific PROMs identified in this review.

The BCTQ has been evaluated mainly with patients undergoing carpal tunnel surgery as well as in a variety of interventional studies and has been shown to perform well with this population. The instrument is disease-specific and covers two important dimensions (Function and Symptoms) which are highly important to patients with CTS and have the most impact on QoL. Approximately, one third (22) of the evaluations of the BCTQ have been reported in the last five years, six of those conducted in the UK. The BCTQ has demonstrated reliability, validity and responsiveness and due to its easy administration, has proven itself to be especially appropriate within clinical settings.

The DASH has also been evaluated with patients undergoing carpal tunnel surgery as well as interventional studies. Although the volume of evaluations with this population is less than the BCTQ, the DASH questionnaire is not specific to carpal tunnel disease but developed and evaluated in a wider range of musculoskeletal disorders of the upper limb. However, the majority of evaluations with patients with CTS have been conducted in the last 5 years (3 UK studies) and the instrument has shown promising reliability, validity and responsiveness when applied to patients with CTS.

The MHQ has been utilised mainly with patients undergoing carpal tunnel surgery and has proven to be reliable, valid and responsive. However, there is a lack of evaluations with UK patients and acceptability indicators have varied across studies.

There is a limited volume of evidence for the application of PEM in patients with CTS. PEM is a UK developed outcome measure and although it appears to be promising in terms of validity and acceptability, the instrument is in need of further reliability testing and evaluations of responsiveness in patients with CTS.

### **Recommendations**

For the measurement of comprehensive general health status, the SF-36 has more psychometric supportive evidence, is acceptable to patients, and can be considered.

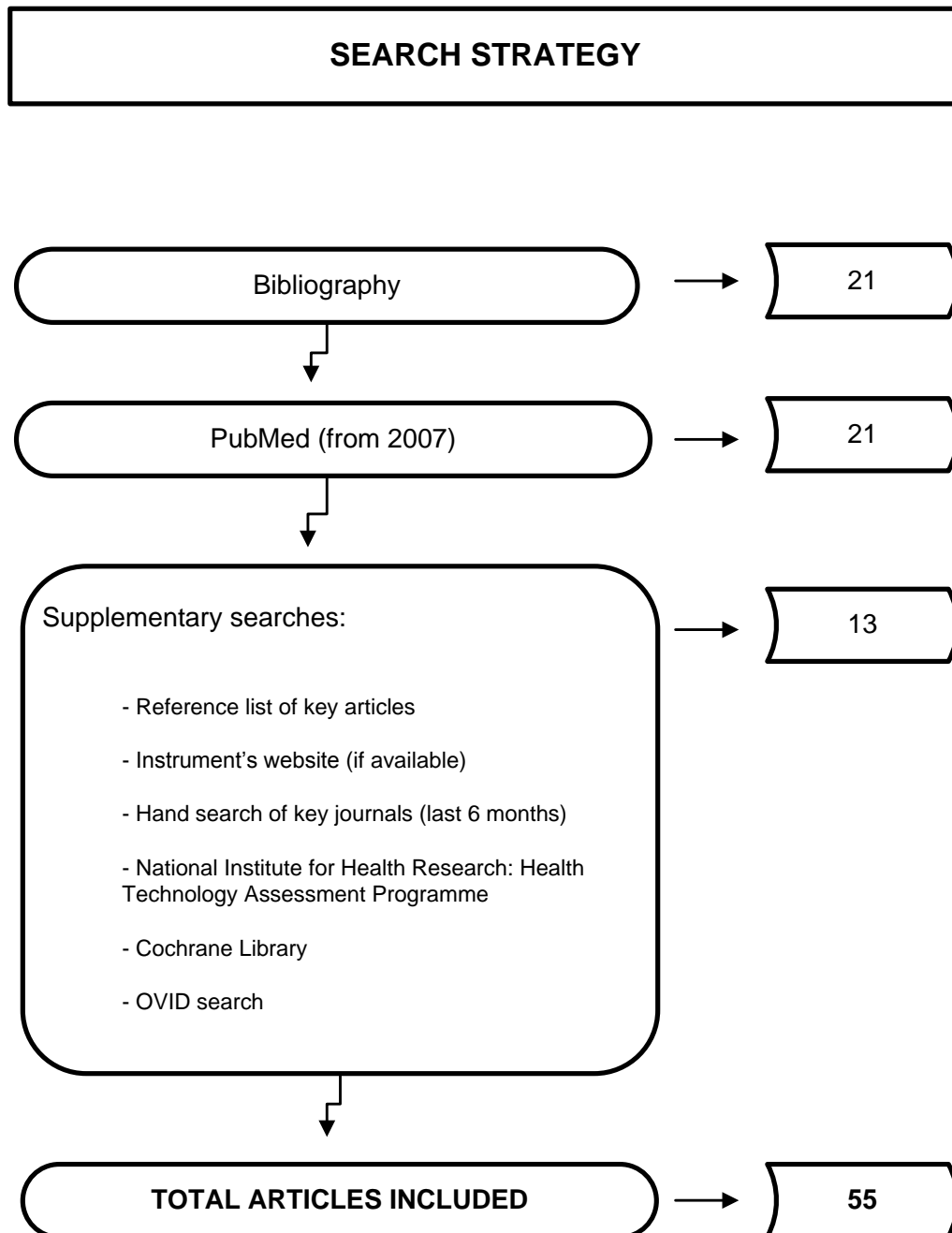
Based on the volume and quality of evidence, the BCTQ is clearly the most promising condition-specific measure. There is also evidence of good psychometric performance for the DASH, however, the BCTQ has the advantage being a CTS specific instrument and is supported by extensive research in this population.

Despite lack of evidence for preference based measures in this review, the EQ-5D is widely used in the UK context for evaluation of a range of elective procedures and can be considered further for the assessment of general health in patients undergoing carpal tunnel surgery.

Table 3: Appraisal of PROMs included in the review

PROM	Reproducibility	Internal consistency	Validity - content	Validity - construct	Responsiveness	Interpretability	Precision	Acceptability	Feasibility
<b>Generic measures</b>									
<b>SF-36</b>	0	0	0	++	+	0	0	++	0
<b>NHP</b>	0	0	0	0	+	0	0	+	0
<b>Condition-specific measures</b>									
<b>BCTQ</b>	+	++	+	+++	+++	0	0	+++	+
<b>DASH</b>	+	+	++	++	++	0	0	+++	+
<b>MHQ</b>	+	+	+	+	++	0	0	++	0
<b>PEM</b>	0	+	+	++	+	0	0	++	0

## Appendix A: Search strategy and sources



## Appendix B: Search strategy for PubMed

((((cardiovascular[tiab] OR cardiac[tiab] OR coronary[tiab] OR heart[tiab]) AND (surgery[tiab] OR revascularisation[tiab] OR revascularization[tiab] OR percutaneous[tiab] OR stent\*[tiab] OR stenting[tiab])) OR (coronary artery bypass[tw] OR CABG[tiab] OR CABS[tiab]) OR (coronary angioplasty [tiab] OR PTCA[tw] OR PCTA[tiab] OR PCI[tiab]))<sup>1</sup>

AND

(HRQL[tiab] OR HRQoL[tiab] OR QL[tiab] OR QoL[tiab] OR quality of life[tw] OR life quality[tw] OR health index\*[tiab] OR health indices[tiab] OR health profile\*[tiab] OR health status[tw] OR ((patient[tiab] OR self[tiab] OR child[tiab] OR parent[tiab] OR carer[tiab] OR proxy[tiab]) AND ((report[tiab] OR reported[tiab] OR reporting[tiab]) OR (rated[tiab] OR rating[tiab] OR ratings[tiab]) OR based[tiab] OR (assessed[tiab] OR assessment[tiab] OR assessments[tiab]))) OR ((disability[tiab] OR function[tiab] OR functional[tiab] OR functions[tiab] OR subjective[tiab] OR utility[tiab] OR utilities[tiab] OR wellbeing[tiab] OR well being[tiab]) AND (index[tiab] OR indices[tiab] OR instrument[tiab] OR instruments[tiab] OR measure[tiab] OR measures[tiab] OR questionnaire[tiab] OR questionnaires[tiab] OR profile[tiab] OR profiles[tiab] OR scale[tiab] OR scales[tiab] OR score[tiab] OR scores[tiab] OR status[tiab] OR survey[tiab] OR surveys[tiab])))

---

<sup>1</sup> Adding the terms ‘stent\*[tiab] OR stenting[tiab]’ did not capture any further records that were relevant.

### Appendix C: Appraisal of the methodological quality of PROMs

A simple rating scale (Table i) was used to rate the sum total of evidence available for each dimension or criterion against which PROMs were assessed. The dimensions or criteria are summarised in Table ii.

**Table i: Psychometric and operational criteria**

0	<i>not reported (no evaluation completed)</i>
—	<i>Evaluation evidence available indicating poor performance of instrument</i>
+	<i>Some limited evidence in favour</i>
++	<i>Good evidence in favour</i>
+++	<i>Excellent evidence in favour</i>

## Appendix D: Table ii, Appraisal criteria

Appraisal component	Definition/test	Criteria for acceptability
<b>Reliability</b>		
Test-retest reliability	The stability of a measuring instrument over time; assessed by administering the instrument to respondents on two different occasions and examining the correlation between test and re-test scores	Test re-test reliability correlations for summary scores 0.70 for group comparisons
Internal consistency	The extent to which items comprising a scale measure the same construct (e.g. homogeneity of items in a scale); assessed by Cronbach's alpha's and item-total correlations	Cronbach's alphas for summary scores $\geq 0.70$ for group comparisons Item-total correlations $\geq 0.20$
<b>Validity</b>		
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 the questionnaire development phase through pre-testing with patients. Expert opinion and literature review	Qualitative evidence from pre-testing with patients, expert opinion and literature review that items in the scale represent the construct being measured  Patients involved in the development stage and item generation
Construct validity	Evidence that the scale is correlated with other measures of the same or similar constructs in the hypothesised direction; assessed on the basis of correlations between the measure and other similar measures The ability of the scale to differentiate known-groups; assessed by comparing scores for sub-groups who are expected to differ on the construct being measured (e.g a clinical group and control group)	High correlations between the scale and relevant constructs preferably based on a priori hypothesis with predicted strength of correlation  Statistically significant differences between known groups and/or a difference of expected magnitude
Responsiveness	The ability of a scale to detect significant change over time; assessed by comparing scores before and after an intervention of known efficacy (on the basis of various methods including t-tests, effect sizes (ES), standardised response means (SRM) or responsiveness statistics)	Statistically significant changes on scores from pre to post-treatment and/or difference of expected magnitude
Floor/ceiling effects	The ability of an instrument to measure accurately across full spectrum of a construct	Floor/ceiling effects for summary scores $< 15\%$
<b>Practical properties</b>		
Acceptability	Acceptability of an instrument reflects respondents' willingness to complete it and impacts on quality of data	Low levels of incomplete data or non-response
Feasibility/burden	The time, energy, financial resources, personnel or other resources required of respondents or those administering the instrument	Reasonable time and resources to collect, process and analyse the data



**Appendix E:**

**GENERIC INSTRUMENTS**

**Table 1a:** Generic patient-reported health instruments

<i>Instrument</i>	<i>Domains (no. items)</i>	<i>Response options</i>	<i>Score</i>	<i>Administration (completion time in minutes)</i>
<b>SF-36: MOS 36-item Short Form Health Survey (36)</b>	Physical functioning (PF) (10), Role limitation-physical (RP) (4), Bodily pain (BP) (2), General health (GH) (5), Vitality (VT) (4), Social functioning (SF) (2), Role limitation-emotional (RE) (3), Mental health (MH) (5), Health transition (1)	Categorical: 2-6 options Recall: standard 4 weeks, acute 1 week	Algorithm Domain profile (0-100, 100 best health) Summary: Physical (PCS), Mental (MCS) (mean 50, sd 10)	Interview (mean values 14-15) Self (mean 12.6)
<b>Nottingham Health Profile (NHP) (38)</b>	Bodily pain (BP) (8), Emotional reactions (ER) (9), Energy (E) (3), Physical mobility (PM) (8), Sleep (S) (5), Social isolation (SI) (5)	Yes/no; positive responses weighted Recall 'general' health	Algorithm Domain profile 0-100, 100 is maximum limitation	Interview Self (10-15)

**Table 1b:** Summary of generic instruments: health status domains (*after Fitzpatrick et al., 1998*)

<i>Instrument (no. items)</i>	<i>Instrument domains</i>							
	<b>Physical function</b>	<b>Symptoms</b>	<b>Global judgement of health</b>	<b>Psychological well-being</b>	<b>Social well-being</b>	<b>Cognitive functioning</b>	<b>Role activities</b>	<b>Personal constructs</b>
<b>SF-36 (36)</b>	x	x	x	x	x		x	
<b>NHP (38)</b>	x	x		x	x			

## CONDITION SPECIFIC INSTRUMENTS

**Table 2a:** Generic patient-reported health instruments

<i>Instrument</i>	<i>Domains (no. items)</i>	<i>Response options</i>	<i>Score</i>	<i>Administration (completion time in minutes)</i>
<b>Boston Carpal Tunnel Questionnaire (BCTQ) (19)</b>	Two domains (19): Symptom Severity Scale (SSS) (11), Functional Status Scale (FSS) (8)	5-point Likert Scale  Recall: A typical twenty-four hour period during the past two weeks	Mean of the scores for the 19 individual items (1 point = best score)	Interview/Self (mean 5.6)
<b>Disability of Arm, Shoulder, and Hand questionnaire (DASH)</b>	Two domains (30): Physical functioning (Physical, Social, Psychological), Symptoms (Pain, Weakness, Stiffness, Tingling/Numbness)	5 to 7-point Likert Scale  Recall: one week	0 -100, 100 is maximum disability	Interview/Self (mean 6.8)
<b>Michigan Hand Questionnaire (MHQ)</b>	Six domains (37): Overall hand function (OHF), Activities of daily living (ADL), Pain (P), Work performance (WP), Aesthetics (A), Patient satisfaction with hand function (PS)	5-point Likert Scale	Pain Scale: high score = greater pain Other 5 scales: high scores = better performance	Interview/Self (15)
<b>Patient Evaluation Measures (PEM)</b>	Three subscales (18): Treatment (5), How my hand is now (10), Overall assessment (3)	7-point Likert Scale	Lower scores indicate better QoL	Interview/Self (5)

**Table 2b:** Summary of generic instruments: health status domains (*after Fitzpatrick et al., 1998*)

<i>Instrument (no. items)</i>	<i>Instrument domains</i>							
	<b>Physical function</b>	<b>Symptoms</b>	<b>Global judgement of health</b>	<b>Psychological well-being</b>	<b>Social well-being</b>	<b>Cognitive functioning</b>	<b>Role activities</b>	<b>Personal constructs</b>
<b>BCTQ</b>	x	x						
<b>DASH</b>	x	x		x	x			
<b>MHQ</b>	x	x					x	x
<b>PEM</b>	x	x	x					

## REFERENCES

- Alderson M, McGall D. The Alderson-McGall hand function questionnaire for patients with carpal tunnel syndrome: a pilot evaluation of a future outcome measure. *Journal of Hand Therapy* 1999;12:313-322.
- Amadio PC, Silverstein MD, Ilstrup DM, Schleck CD, Jensen LM. Outcome assessment for carpal tunnel surgery: the relative responsiveness of generic, arthritis-specific, disease-specific, and physical examination measures. *Journal of Hand Surgery* 1996;21:338-346.
- Amirjani N, Ashworth NL, Watt MJ, Gordon T, Chan KM. Corticosteroid iontophoresis to treat carpal tunnel syndrome: a double-blind randomized controlled trial. *Muscle Nerve* 2009;39:627-633.
- Appleby MA, Neville-Smith M, Parrott MW. Functional outcomes post carpal tunnel release: a modified replication of a previous study. *J Hand Ther* 2009;22:240-248.
- Atroshi I, Gummesson C, Johnsson R, et al. Prevalence of carpal tunnel syndrome in a general population. *JAMA* 1999; 282:153-158.
- Badger SA, O'Donnell ME, Sherigar JM, Connolly P, Spence RA. Open carpal tunnel release--still a safe and effective operation. *Ulster Med J* 2008;77:22-24.
- Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier CH. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. *Journal of Hand Therapy* 2001;14:128-146.
- Bessette L, Sangha O, Kuntz KM et al. Comparative responsiveness of generic versus disease-specific and weighted versus unweighted health status measures in carpal tunnel syndrome. *Medical Care* 1998;36:491-502.
- Bialosky JE, Bishop MD, Price DD, Robinson ME, Vincent KR, George SZ. A randomized sham-controlled trial of a neurodynamic technique in the treatment of carpal tunnel syndrome. *J Orthop Sports Phys Ther* 2009;39:709-723.
- Bodavula VK, Burke FD, Dubin NH, Bradley MJ, Wilgis EF. A prospective, longitudinal outcome study of patients with carpal tunnel surgery and the relationship of body mass index. *Hand (N Y)* 2007;2:27-33.
- Bradley MP, Hayes EP, Weiss AP, Akelman E. A prospective study of outcome following mini-open carpal tunnel release. *Hand Surgery* 2003;8:59-63.
- Brazier JE, Jones N, Kind P. Testing the validity of the EuroQol and comparing it with the SF-36 Health Survey questionnaire. *Quality of Life Research* 1993; 2(3):169-180.
- Brazier JE, Roberts J, Deverill M: The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics* 2002; 21(2):271-292.

Brininger TL, Rogers JC, Holm MB, Baker NA, Li ZM, Goitz RJ. Efficacy of a fabricated customized splint and tendon and nerve gliding exercises for the treatment of carpal tunnel syndrome: a randomized controlled trial. *Arch Phys Med Rehabil* 2007;88:1429-1435.

Burke FD, Wilgis EF, Dubin NH, Bradley MJ, Sinha S. Relationship between the duration and severity of symptoms and the outcome of carpal tunnel surgery. *Journal of Hand Surgery* 2006;31:1478-1482.

Burke J, Buchberger DJ, Carey-Loghmani MT, Dougherty PE, Greco DS, Dishman JD. A pilot study comparing two manual therapy interventions for carpal tunnel syndrome. *J Manipulative Physiol Ther* 2007;30:50-61.

Chan L, Turner JA, Comstock BA et al. The relationship between electrodiagnostic findings and patient symptoms and function in carpal tunnel syndrome. *Archives of Physical Medicine & Rehabilitation* 2007;88:19-24.

Chatterjee JS, Price PE. Comparative responsiveness of the Michigan Hand Outcomes Questionnaire and the Carpal Tunnel Questionnaire after carpal tunnel release. *J Hand Surg Am* 2009;34:273-280.

Chung KC, Pillsbury MS, Walters MR, Hayward RA. Reliability and validity testing of the Michigan Hand Outcomes Questionnaire. *Journal of Hand Surgery* 1998;23:575-587.

Cresswell TR, Heras-Palou C, Bradley MJ et al. Long-term outcome after carpal tunnel decompression - a prospective randomised study of the Indiana Tome and a standard limited palmar incision. *J Hand Surg Eur Vol* 2008;33:332-336.

Darzi A. Our NHS Our Future: NHS Next Stage Review Interim Report. Department of Health, London, October 2007.

Darzi A. Our NHS Our Future: High Quality Care for All. NHS Next Stage Review Final Report. Department of Health, London, June 2008.

Department of Health, London, December 2008. Guidance on the routine collection of Patient Reported Outcome Measures (PROMs) for the NHS in England 2009/10.

Department of Health, London, March 2000. National Strategic Framework for Coronary Heart Disease.

Ettema AM, Amadio PC, Cha SS, Harrington JR, Harris AM, Offord KP. Surgery versus conservative therapy in carpal tunnel syndrome in people aged 70 years and older. *Plastic & Reconstructive Surgery* 2006;118:947-958.

EuroQol Group, The. EuroQol: a new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16(3):199-208.

Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment* 1998; 2(14).

Forward DP, Singh AK, Lawrence TM, Sithole JS, Davis TR, Oni JA. Preservation of the ulnar bursa within the carpal tunnel: does it improve the outcome of carpal tunnel surgery? A randomized, controlled trial. *Journal of Bone & Joint Surgery - American Volume* 2006;88:2432-2438.

Fulton-Kehoe D, Gluck J, Wu R, Mootz R, Wickizer TM, Franklin GM. Measuring work disability: what can administrative data tell us about patient outcomes? *J Occup Environ Med* 2007;49:651-658.

Garratt AM, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. *British Medical Journal* 2002; 324(7351):1417-1421.

Gay RE, Amadio PC, Johnson JC. Comparative responsiveness of the Disabilities of the Arm, Shoulder, and Hand; the Carpal Tunnel Questionnaire; and the SF-36 to clinical change after carpal tunnel release. *Journal of Hand Surgery* 2003;28:250-254.

Gordon T, Amirjani N, Edwards DC, Chan KM. Brief post-surgical electrical stimulation accelerates axon regeneration and muscle reinnervation without affecting the functional measures in carpal tunnel syndrome patients. *Exp Neurol* 2009.

Guyette TM, Wilgis EF. Timing of improvement after carpal tunnel release. *J Surg Orthop Adv* 2004;13:206-209.

Heebner ML, Roddey TS. The effects of neural mobilization in addition to standard care in persons with carpal tunnel syndrome from a community hospital. *J Hand Ther* 2008;21:229-240.

Hobby JL, Venkatesh R, Motkur P. The effect of psychological disturbance on symptoms, self-reported disability and surgical outcome in carpal tunnel syndrome. *Journal-of-Bone-and-Joint-Surgery* 2005;87B:196-200.

Jain R, Hudak PL, Bowen CVA. Validity of health status measures in patients with ulnar wrist disorders. *Journal of Hand Therapy* 2001;14:147-153.

Jarvik JG, Comstock BA, Heagerty PJ et al. Magnetic resonance imaging compared with electrodiagnostic studies in patients with suspected carpal tunnel syndrome: predicting symptoms, function, and surgical benefit at 1 year. *J Neurosurg* 2008;108:541-550.

Jarvik JG, Comstock BA, Kliot M et al. Surgery versus non-surgical therapy for carpal tunnel syndrome: a randomised parallel-group trial. *Lancet* 2009;374:1074-1081.

Kadzielski J, Malhotra LR, Zurakowski D, Lee SG, Jupiter JB, Ring D. Evaluation of preoperative expectations and patient satisfaction after carpal tunnel release. *J Hand Surg Am* 2008;33:1783-1788.

Katz JN, Gelberman RH, Wright EA, Lew RA, Liang MH. Responsiveness of self-reported and objective measures of disease severity in carpal tunnel syndrome. *Medical Care* 1994;32:1127-1133.

- Katz JN, Punnett L, Simmons BP, Fossel AH, Mooney N, Keller RB. Workers' compensation recipients with carpal tunnel syndrome: the validity of self-reported health measures. *American Journal of Public Health* 1996;86:52-56.
- Kaye JJ, Reynolds JM. Carpal tunnel syndrome: using self-report measures of disease to predict treatment response. *Am J Orthop* 2007;36:E59-E62.
- Kotsis SV, Chung KC. Responsiveness of the Michigan Hand Outcomes Questionnaire and the Disabilities of the Arm, Shoulder and Hand questionnaire in carpal tunnel surgery. *Journal-of-Hand-Surgery* 2005;30:81-86.
- Leit ME, Weiser RW, Tomaino MM. Patient-reported outcome after carpal tunnel release for advanced disease: a prospective and longitudinal assessment in patients older than age 70. *Journal of Hand Surgery* 2004;29:379-383.
- 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. *Journal of Bone and Joint Surgery, Series A* 1993;75:1585-1592.
- Lim LLY, Valenti LA, Knapp JC, Dobson AJ, Plotnikoff R, Higginbotham N *et al.* A self-administered quality-of-life questionnaire after acute myocardial infarction. *Journal of Clinical Epidemiology* 1993; 46(11):1249-56.
- Lozano Calderon SA, Paiva A, Ring D. Patient satisfaction after open carpal tunnel release correlates with depression. *J Hand Surg Am* 2008;33:303-307.
- Luria S, Waitayawinyu T, Trumble TE. Endoscopic revision of carpal tunnel release. *Plast Reconstr Surg* 2008;121:2029-2034.
- McDowell I. *Measuring Health: a guide to rating scales and questionnaires*. 3rd ed. Oxford, UK & New York, USA: Oxford University Press, 2006.
- McMillan CR, Binhammer PA. Which Outcome Measure is the Best? Evaluating Responsiveness of the Disabilities of the Arm, Shoulder, and Hand Questionnaire, the Michigan Hand Questionnaire and the Patient-Specific Functional Scale Following Hand and Wrist Surgery. *Hand (N Y)* 2009;4:311-318.
- Moraska A, Chandler C, Edmiston-Schaetzl A, Franklin G, Calenda EL, Enebo B. Comparison of a targeted and general massage protocol on strength, function, and symptoms associated with carpal tunnel syndrome: a randomized pilot study. *J Altern Complement Med* 2008;14:259-267.
- Nesbitt KS, Innis PC, Dubin NH, Wilgis EF. Staged versus simultaneous bilateral endoscopic carpal tunnel release: an outcome study. *Plastic and Reconstructive Surgery* 2006;118:139-147.
- Nielke MC, Lindenhovius AL, Watson JB, Vranceanu AM, Ring D. Correlation of DASH and QuickDASH with measures of psychological distress. *J Hand Surg Am* 2009;34:1499-1505.

- Ozyurekoglul T, McCabe SJ, Goldsmith LJ, LaJoie AS. The Minimal Clinically Important Difference of the Carpal Tunnel Syndrome Symptom Severity Scale. *Journal of Hand Surgery* 2006;31:733-738.
- Redmond CL, Bain GI, Laslett LL, McNeil JD. Hand Syndromes Associated with Diabetes: Impairments and Obesity Predict Disability. *J Rheumatol* 2009.
- Rege AJ, Sher JL. Can the outcome of carpal tunnel release be predicted? *Journal of Hand Surgery* 2001;26:148-150.
- Reid MJ, David LA, Nicholl JE. A one-stop carpal tunnel clinic. *Ann R Coll Surg Engl* 2009;91:301-304.
- Shauver MJ, Chung KC. The minimal clinically important difference of the Michigan hand outcomes questionnaire. *J Hand Surg Am* 2009;34:509-514.
- Streiner D, Norman G. Health Measurement Scales: a practical guide to their development and use. 3<sup>rd</sup> ed. Oxford, UK & New York, USA: Oxford University Press, 2003.
- Turner JA, Franklin G, Heagerty PJ et al. The association between pain and disability. *Pain* 2004;112(3):307-314.
- Vaile JH, Mathers DM, Ramos RC, Russell AS. Generic health instruments do not comprehensively capture patient perceived improvement in patients with carpal tunnel syndrome. *Journal of Rheumatology* 1999;26:1163-1166.
- Valenti L, Lim L, Heller RF, Knapp J. An improved questionnaire for assessing quality of life after acute myocardial infarction. *Quality of Life Research* 1996; 5(1):151-61.
- Wainner RS, Fritz JM, Irrgang JJ, Delitto A, Allison S, Boninger ML. Development of a clinical prediction rule for the diagnosis of carpal tunnel syndrome. *Archives of Physical Medicine and Rehabilitation* 2005;86:609, 847-618, 848.
- Ware JE. SF-36 Health Survey. Manual and Interpretation Guide. Boston, MA: The Health Institute, New England Medical Centre, 1997. 2<sup>nd</sup> edition.
- Yang J, Boninger ML, Leath JD, Fitzgerald SG, Dyson-Hudson TA, Chang MW. Carpal Tunnel Syndrome in Manual Wheelchair Users with Spinal Cord Injury: A Cross-Sectional Multicenter Study. *Am J Phys Med Rehabil* 2009.