

# The Elbow Self-Assessment Score (ESAS): development and validation of a new patient-reported outcome measurement tool for elbow disorders

Marc Beirer<sup>1</sup> · Henrik Friese<sup>1</sup> · Andreas Lenich<sup>2</sup> · Moritz Crönlein<sup>1</sup> · Gunther H. Sandmann<sup>1</sup> · Peter Biberthaler<sup>1</sup> · Chlodwig Kirchhoff<sup>1</sup> · Sebastian Siebenlist<sup>1</sup>

Received: 21 February 2015 / Accepted: 8 May 2015 / Published online: 17 May 2015  
© European Society of Sports Traumatology, Knee Surgery, Arthroscopy (ESSKA) 2015

## Abstract

**Purpose** To develop and validate an elbow self-assessment score considering subjective as well as objective parameters.

**Methods** Each scale of the American Shoulder and Elbow Surgeons-Elbow Score, the Broberg and Morrey rating system (BMS), the Patient-Rated Elbow Evaluation (PREE) Questionnaire, the Mayo Elbow Performance Score (MEPS), the Oxford Elbow Score (OES) and the Quick Disabilities of the Arm, Shoulder and Hand (Quick-DASH) was analysed, and after matching of the general topics, the dedicated items underwent a fusion to the final ESAS's item and a score containing 22 items was created. In a prospective clinical study, validity, reliability and responsiveness in physically active patients with traumatic as well as degenerative elbow disorders were evaluated.

**Results** Validation study included 103 patients (48 women, 55 men; mean age 43 years). A high test–retest reliability was found with intraclass correlation coefficients of at least 0.71. Construct validity and responsiveness were confirmed by correlation coefficients of  $-0.80$  to  $-0.84$  and  $0.72$ – $0.84$  ( $p < 0.05$ ). Correlation coefficients of the ESAS and well-established elbow rating systems BMS, PREE, MEPS, OES and Quick-DASH were between 0.70 and 0.90 ( $p < 0.05$ ).

**Conclusions** With this novel Elbow Self-Assessment Score (ESAS), a valid and reliable instrument for a qualitative self-assessment of subjective and objective parameters (e.g. range of motion) of the elbow joint is demonstrated. Quantitative measurement of elbow function may not longer be limited to specific elbow disorders or patient groups. The ESAS seems to allow for a broad application in clinical research studying elbow patients and may facilitate the comparison of treatment results in elbow disorders. The treatment efficacy can be easily evaluated, and treatment concepts could be reviewed and revised.

**Level of evidence** Diagnostic study, Level III.

**Keywords** Elbow disorders · Self-assessment score · Outcome measurement tool · Validity · Reliability · Responsiveness

## Introduction

For elbow disorders, clinical rating systems became more and more popular in modern evaluation of treatment results [22]. The physician-based clinical examination, however, does not necessarily correlate with the patient's satisfaction [6]. Therefore, the use of self-assessment instruments as additional tools to clinically assessed parameters for a comprehensive evaluation of the elbow is increasing [11]. Self-assessment scores additionally represent easy and cost-effective tools to collect patient's relevant data in day-by-day clinical work. Long travel distances could be avoided, and even immobile patients could be reached. Despite the availability of numerous elbow-specific scores, there is no standard evaluation tool for elbow function, and we are still far from a single outcome evaluation system which is reliable, valid and sensitive to clinically relevant

✉ Marc Beirer  
marc.beirer@mri.tum.de

<sup>1</sup> Department of Trauma Surgery, Klinikum rechts der Isar, Technical University of Munich, Ismaningerstrasse 22, 81675 Munich, Germany

<sup>2</sup> Department of Orthopaedic Sports Medicine, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany



**Fig. 1** Functional abilities depicted as photographs, **a** flexion/extension, **b** pronation/supination, **c** force measurement in 90° flexion

changes [11]. A currently performed investigation to assess the quality of validation studies of elbow-specific outcome measurement tools identified the Oxford Elbow Score (OES) as high-qualitative rating system which has been validated in a heterogeneous study population [22]. Indeed, the OES focuses on subjective parameters such as pain, social psychology and disability in daily activities, but the range of motion (ROM) as an essential objective parameter in elbow disorders is rarely considered [7].

Therefore, the purpose of this prospective study was to develop and validate an all-purpose Elbow Self-Assessment Score (ESAS) for a patient-based follow-up examination considering subjective as well as objective parameters in a heterogeneous patient collective.

## Materials and methods

### Development of the scoring system

A systematic review of the literature was performed to identify valid and commonly used scoring systems regarding follow-up examination in the field of elbow disorders. *PubMed.gov* was searched for elbow-specific terms (elbow, surgery, joint and upper extremity) combined with psychometric (validity, reliability, responsiveness and follow-up) and instrument-specific terms (self-evaluation, patient-based, measurement tool, outcome measure and questionnaire). The American Shoulder and Elbow Surgeons-Elbow (ASES-E) Score [10], the Broberg and Morrey rating system (BMS) [4], the Patient-Rated Elbow Evaluation (PREE) Questionnaire [12], the Mayo Elbow Performance Score (MEPS) [5], the Oxford Elbow Score (OES) [7] and the Quick Disabilities of the Arm, Shoulder and Hand (Quick-DASH) [3] were identified as frequently used and valid assessment measurement tools in elbow disorders.

To ensure content validity of the Elbow Self-Assessment Score (ESAS), each scale of the ASES-E, the BMS, the PREE, the MEPS, the OES and the Quick-DASH was

analysed for items addressing either general topics or specific items. Subsequently, a matching of the general topics was performed, and the dedicated items underwent a fusion to the final ESAS's item. Typical functional abilities were depicted as photographs (see Fig. 1). Finally, the ESAS contains 22 items addressing three domains: pain (seven items), elbow function including range of motion (12 items) and quality of life (three items). The best and least symptomatic score for each item is set zero and the worst ten. The overall score is then converted to a scale of 100 %, whereas a value of 100 % indicates an excellent result and a value of 0 % a poor result.

### Patient collective

At our outpatient clinic, 103 consecutive patients who had suffered from soft tissue and/or osseous injuries as well as degenerative disorders of the elbow joint were included to the study. Written informed consent was obtained from each patient. The dominant side was affected in 56 cases. People with limited legal capacity, under legal supervision or suffering from psychiatric diseases, dementia or other cognitive diseases were excluded.

### Testing and evaluation of measurement qualities

#### *Floor and ceiling effects*

According to McHorney and Tarlov [13], floor and ceiling effects exist, if more than 15 % of the patients achieve the highest or lowest possible score. Similarly, we defined the presence of floor or ceiling effects, if more than 15 % of our patient collective would achieve the highest (100 points) or lowest (0 point) possible score of the ESAS.

#### *Internal consistency*

Internal consistency is defined by the degree of interrelation between the tested items [14]. The subscales are based

on a reflective model in which all items are defined by a manifestation of the same underlying construct. According to previous published studies, Cronbach's alpha was calculated per subscale and a score above 0.70 was considered as sufficient homogeneity of the subscales' items [20, 23].

#### *Test–retest reliability*

Test–retest reliability is defined as the extent to which scores of the same patients under the same conditions coincide in repeated measurements [14]. The time period between the repeated measurements should be long enough to prevent from recall of the tested items and moreover should be short enough to ensure that no change of the clinical symptoms has occurred [20]. In this study, a time period of 10–14 days after the initial examination was chosen to assess test–retest reliability. Intraclass correlation coefficients (ICCs) were calculated, and positive reliability was assumed when the ICC was at least 0.70 for all tested subscales [20].

#### *Construct validity*

Construct validity is defined as the degree to which the scores of a self-assessment instrument are consistent with a priori hypothesis, based on the assumption that the instrument validly measures the construct to be measured [14]. Construct validity was assessed by correlating the subscales of the ESAS with the subscales of the OES. In recent literature, this score was reported as a valid, reliable and responsive self-administered instrument that can be used for follow-up examinations of several types of elbow disorders and was therefore used for correlation [22]. The Pearson correlation coefficient (PCC) was calculated. Similar to previous studies, a positive construct validity was assumed when the PCC was at least 0.70 for all measured subscales [9].

#### *Responsiveness*

Responsiveness is defined as the ability of an instrument to detect changes over time of the construct to be measured [14]. Responsiveness was evaluated 4–6 months after the initial presentation of the patient. To assess responsiveness, patients completed the ESAS and a Global Perceived Effect (GPE) Score consisting of only one question per subscale on the patients' subjective opinion regarding improvement or worsening during the last months. A list of potential answers contained seven categories [much better (+3), better (+2), somewhat better (+1), no change (0), somewhat worse (–1), worse (–2), much worse (–3)] for each subscale of the ESAS. The time period of 4–6 months was chosen to be long

enough to allow for a clinical change and short enough to ensure that the patients are able to recall their health state during their initial presentation. The Spearman's correlation coefficient (SCC) was calculated. SCC between the change of the ESAS and the GPE Score of at least 0.40 was assumed to indicate positive responsiveness [23].

#### *Correlation of the ESAS with established elbow scores*

We supposed that at least a moderate correlation would be obtained between the new elbow measurement tool (ESAS) and established elbow rating systems (BMS, PREE, MEPS, OES and Quick-DASH). The PCC was calculated followed by a linear regression analysis. A positive correlation was assumed when the PCC was at least 0.70.

The study protocol was approved by the local ethics committee (Ethics Committee of the medical faculty, Technical University of Munich; study number 5536/12).

#### **Statistical analysis**

The results were compared by calculating the SCC and PCC with a linear regression analysis. A  $p$  value <0.05 determined significance. Statistics were calculated using commercially available programs (SigmaStat 3.1, SigmaPlot 8.02, Systat Software Inc., Chicago, USA).

## **Results**

### **Patients and study design**

Validity, reliability and responsiveness of the ESAS were determined in a prospective, clinical study. Between March and December 2014, 103 consecutive patients (mean age 43 years, SD 15.4 years; range 18–82 years) were asked to complete the ESAS, the BMS, the PREE, the MEPS, the OES and the Quick-DASH at initial presentation for evaluating validity. Several patients did not complete all scores correctly and had to be excluded from the study (one for the BMS, eight for the PREE, one for the MEPS, nine for the OES and 14 for the Quick-DASH). Table 1 summarises patient's diagnosis, representing a wide spectrum of traumatic and degenerative elbow disorders. Figure 2 shows the clinical study profile.

### **Floor and ceiling effects**

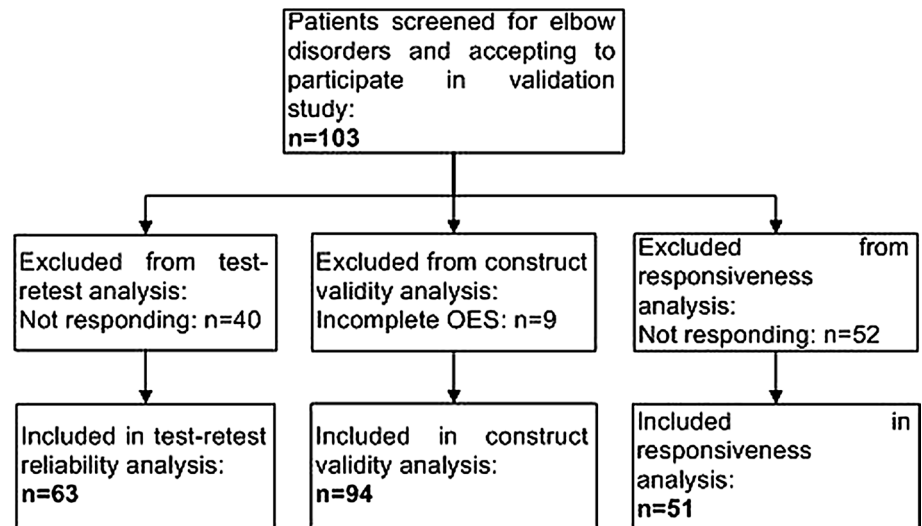
None of the patients achieved the lowest possible score, but one patient achieved the best score of the ESAS

**Table 1** Study population,  $n = 103$

Diagnosis	Total ( $n = 103$ )	Men ( $n = 55$ )	Women ( $n = 48$ )
Elbow contusion	10	9	1
Chronic olecranon bursitis	8	5	3
Radial head fracture	3	1	2
Elbow joint dislocation	18	10	8
Distal biceps tendon rupture	5	5	
State after fracture of the olecranon	6	4	2
State after fracture of the radial head	11	5	6
State after fracture of the distal humerus	5	2	3
State after fracture of the proximal forearm	1	1	
State after osteomyelitis of the distal humerus	1		1
State after elbow arthroplasty	2	1	1
State after reconstruction of the LCL	6	1	5
Lateral epicondylitis	10	5	5
Medial epicondylitis	2	1	1
Elbow osteoarthritis	4	3	1
Ulnar nerve lesion	2		2
LCL lesion	2		2
Osteonecrosis of the radial head	1		1
Ulnar tunnel syndrome	3	2	1
Radial nerve lesion	1		1
Osteochondritis dissecans	1		1
Elbow plica syndrome	1		1

LCL lateral collateral ligament

**Fig. 2** Clinical study profile; flowchart of the study process



(100 points). Thus, there were no floor or ceiling effects to be described.

**Internal consistency**

Cronbach’s alpha was calculated for each subscale of the ESAS. Values of at least 0.83 showed a high consistency for all items in one subscale (Table 2).

**Test–retest reliability**

Retest was performed at a mean of 12 days (SD 3.0 days; range 7–22 days) after the patients’ initial consultation. A total of 63 patients (61 %) returned the completed questionnaire (Fig. 2). Intraclass correlation coefficients (ICCs) were between 0.71 and 0.81 for all subscales of the ESAS (Table 2).

## Construct validity

Assessment of construct validity contained a correlation of the subscales of the ESAS with the subscales of the OES.

**Table 2** Internal consistency ( $n = 103$ ) and test–retest reliability ( $n = 63$ )

	Test mean (SD)	Retest mean (SD)	ICC	Cronbach's $\alpha$
ESAS total	50.5 (26.5)	59.7 (26.8)	0.76	
Pain	4.3 (2.6)	3.8 (2.7)	0.78	0.92
Function	4.1 (2.6)	3.4 (2.4)	0.81	0.92
QOL	5.7 (3.1)	4.5 (3.2)	0.71	0.83

SD standard deviation, ICC intraclass correlation coefficient, ESAS Elbow Self-Assessment Score, QOL quality of life

**Table 3** Pearson's correlation coefficients ( $r$ ) determined when comparing the subscales of the ESAS to the subscales of the OES,  $n = 94$

ESAS	OES (pain)	OES (function)	OES (QOL)
Pain	−0.84		
Function		−0.82	
QOL			−0.80

ESAS Elbow Self-Assessment Score, OES Oxford Elbow Score, QOL quality of life

PCC of at least  $-0.80$  was calculated for all subscales (Table 3).

## Responsiveness

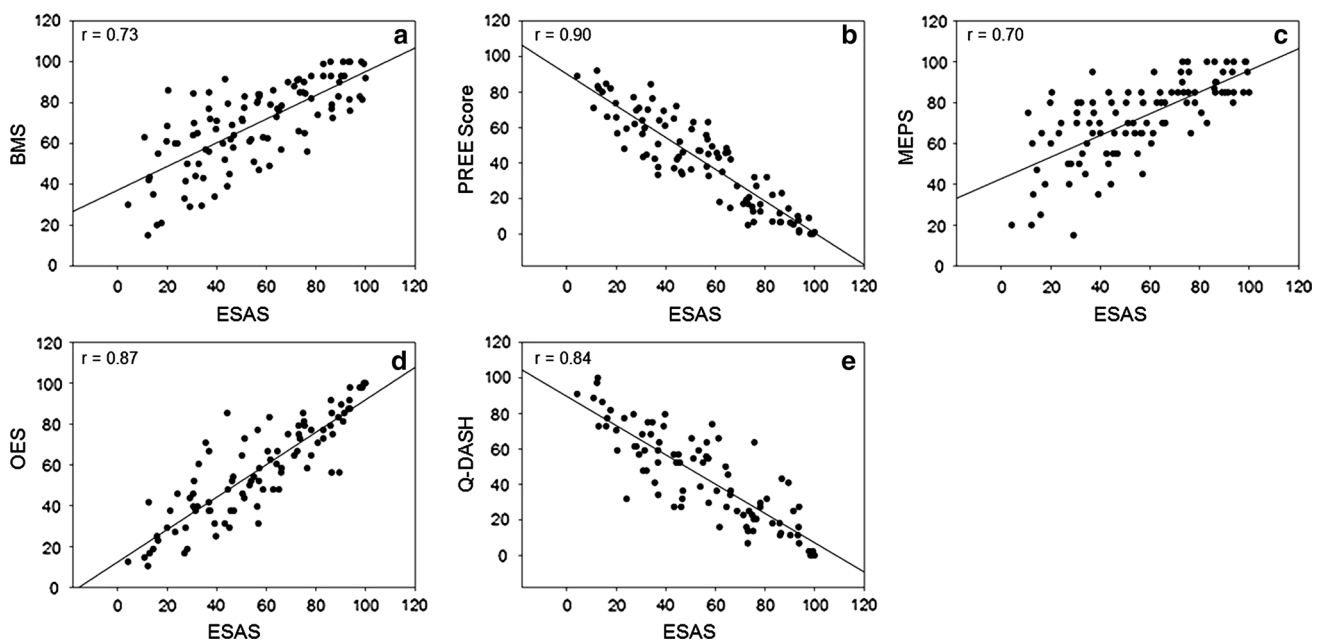
A total of 51 patients (50 %) returned the completed ESAS and GPE Score 154 days (SD 25.5 days; range 103–196 days) after the initial assessment (Fig. 2). The SCC was 0.73 for pain, 0.84 for function and 0.72 for elbow-related quality of life.

## Correlation of the ESAS with established elbow scores

Figure 3 shows the results of the correlation between the ESAS and frequently used elbow rating systems. The PCC between the ESAS and the BMS was 0.73,  $-0.90$  for the PREE, 0.70 for the MEPS, 0.87 for the OES and 0.84 for the Quick-DASH ( $p < 0.05$ ).

## Discussion

The most important finding of the present study was a positive validity, reliability and responsiveness of a novel elbow self-assessment score, the Elbow Self-Assessment Score (ESAS). Based on a single 22-item tool, this new evaluation score records subjective as well as objective parameters. With special regard to well-established elbow rating



**Fig. 3** Simple regression scatter plots of the correlation between the ESAS and the BMS (a,  $n = 102$ ), the PREE Score (b,  $n = 95$ ), the MEPS (c,  $n = 102$ ), the OES (d,  $n = 94$ ) and the Q-DASH (e,  $n = 89$ ). Solid lines represent the linear regression. Pearson's correlation coefficients ( $r$ ) are given in each panel. ESAS Elbow Self-Assess-

ment Score, BMS Broberg and Morrey Score, PREE Score Patient-Rated Elbow Evaluation Score, MEPS Mayo Elbow Performance Score, OES Oxford Elbow Score, Q-DASH Quick Disabilities of the Arm, Shoulder and Hand

systems (BMS, PREE, MEPS, OES and Quick-DASH), a high correlation was found ( $p < 0.05$ ).

In recent years, the importance and the use of self-assessment scores in outcome studies as additional measurement tools to the physician-based objective evaluation increased most likely due to their advantages in financial and logistic concerns [18] to allow for a comprehensive evaluation of the clinical outcome. Furthermore, avoiding face-to-face contact with the patients eliminates a certain observer bias in terms of the interviewer knowing the purpose of the study. On the other hand, self-assessment scores offer other possible sources of bias in terms of non- and incomplete response [15]. In the present study, a non-responding rate of 39 % in assessing test–retest reliability and 50 % in responsiveness was found. This is favourably comparable to dropout rates of other validation studies in the current literature [16, 23]. Parker and Dewey recommend reminding the participating patients by mail or telephone to increase the responding rate [15], which may be in the focus of further validation studies.

The presented study collective consisted of 103 consecutive patients with a mean age of 43 years with a male–female ratio of almost 1:1 comparable to other validation studies concerning number of patients, age and gender [7, 19, 23]. The number of different diagnoses of the presented patient collective represents the wide spectrum of elbow disorders including acute traumatic osseous and ligament injuries as well as degenerative diseases (see Table 1). Several authors prefer such a heterogenous collective of patients combining different clinical entities for validation of elbow-specific rating systems in order to allow for a universal application [7, 12, 17, 22]. Despite the limited responding rate in the presented study, the percentage of traumatic and degenerative disorders remained equal in the evaluation of test–retest reliability and responsiveness, and the broad application of the ESAS is not limited.

The statistical evaluation included the assessment of internal consistency, test–retest reliability, construct validity and responsiveness. Cronbach's  $\alpha$  of at least 0.83 resulting for all subscales stands for a high internal consistency. The different items of the same subscale (e.g. elbow pain) seem to measure the same general construct resulting in similar scores. The highest value of 0.92 found for the subscales pain and function did not exceed 0.95 that might indicate item redundancy [22]. The assessment of test–retest reliability resulted in ICCs between 0.71 and 0.81 for all subscales of the ESAS, which indicates a positive reliability. In the current literature, an exact time point for the retest assessment is missing, but in most cases, a time period of 1 or 2 weeks is considered as appropriate for determining test–retest reliability [20]. The patients evaluated in this study were instructed to complete and return the second questionnaire after 10–14 days. Nevertheless,

several patients returned the score after 7 days which may increase the risk of recall bias. A few other patients did only return the score 22 days after their initial visit increasing the possibility of a change of their clinical state. In the literature, no gold standard exists for comparison of the construct validity between elbow scores. Therefore, the decision was made to correlate the subscales of the ESAS with the subscales of a previously reported validated score [22]. For comparison, we decided for the OES—a well-established valid, reliable and responsive instrument that can be used for follow-up examination of several types of elbow injuries such as osteoarthritis, post-traumatic stiffness, epicondylitis and other conditions—as reference score. Pearson's correlation coefficients of at least  $-0.80$  resulted for all subscales of the ESAS. Compared to other validation studies, these results indicate a high construct validity in a self-reported score [2, 8]. The evaluation of responsiveness included the correlation between the GPE Score and the change in scores of the first and second ESAS. A range from 0.72 to 0.84 for the subscale pain, function and elbow-related quality of life was found, indicating high responsiveness. Since the GPE Score contains only one single question, subjective clinical change of the elbow function may have been influenced considerably by persisting symptoms although other symptoms changed considerably, thus possibly resulting in a supposed minor responsiveness, requiring a multi-item instrument [21]. In the current literature, various statistics to determine responsiveness are available; however, the method of choice remains unknown [1]. Thorborg et al. [23] showed the determination of effective size and standardised response mean in addition to the GPE Score as a considerable amendment to assess responsiveness. Convergent validity, as an expression of the relation between the ESAS and the BMS, the PREE, the MEPS, the OES and the Quick-DASH, was shown by high correlation coefficients.

This study has some weaknesses. To avoid financial and logistic burden for the participating patients, the evaluation of test–retest reliability and responsiveness was conducted at the patients' homes. This change in setting may influence the test results. Nonetheless, we consider this fact as irrelevant since the initial assessment in our clinic and the second and the third assessment at home were accomplished in self-administration. Furthermore, responsiveness was assessed by correlating a global perceived effect score with the single subscales of the ESAS. Since the GPE Score contained only one single question and the subscales of the ESAS contained between three and twelve questions, the GPE Score could be less reliable than a multi-item instrument [21], resulting in a reduced interpretability of responsiveness. In addition, the low responding rate may limit the significant responsiveness of the ESAS. Another limitation is that the ESAS has only been tested in Germany, and a

cross-cultural adaption into other languages and determination of its clinimetric properties have to be conducted before it can be used worldwide.

The universal applicability of the ESAS may result in difficulties regarding the assessment of borderline patients such as highly trained athletes or frail people being in need for care. However, due to the vast majority of patients being potentially evaluated by this tool, these drawbacks might be negligible.

To sum up, the ESAS is clinically relevant for a comprehensive elbow evaluation in daily practice. The treatment efficacy can be easily evaluated, and treatment concepts could be reviewed and revised.

## Conclusions

The Elbow Self-Assessment Score (ESAS) is a self-administrated, valid and reliable tool to assess the most important aspects of the elbow function. Based on the present data, the ESAS seems to allow for a qualitative self-assessment of subjective as well as objective parameters (e.g. ROM) of the elbow joint. The implementation of the ESAS may not be restricted to specific elbow disorders or patient groups with the aim of universal clinical applicability.

**Conflict of interest** The authors declare that they have no conflict of interest.

## References

- Angst F (2011) The new COSMIN guidelines confront traditional concepts of responsiveness. *BMC Med Res Methodol* 11:152 (**author reply 152**)
- Ashmore AM, Gozzard C, Blewitt N (2007) Use of the Liverpool Elbow Score as a postal questionnaire for the assessment of outcome after total elbow arthroplasty. *J Shoulder Elbow Surg* 16(3 Suppl):S55–S58
- Beaton DE, Wright JG, Katz JN, Upper Extremity Collaborative G (2005) Development of the QuickDASH: comparison of three item-reduction approaches. *J Bone Joint Surg Am* 87(5):1038–1046
- Broberg MA, Morrey BF (1986) Results of delayed excision of the radial head after fracture. *J Bone Joint Surg Am* 68(5):669–674
- Broberg MA, Morrey BF (1987) Results of treatment of fracture-dislocations of the elbow. *Clin Orthop Relat Res* 216:109–119
- Capuano L, Poulain S, Hardy P, Longo UG, Denaro V, Maffulli N (2011) No correlation between physicians administered elbow rating systems and patient's satisfaction. *J Sports Med Phys Fitness* 51(2):255–259
- Dawson J, Doll H, Boller I, Fitzpatrick R, Little C, Rees J, Jenkinson C, Carr AJ (2008) The development and validation of a patient-reported questionnaire to assess outcomes of elbow surgery. *J Bone Joint Surg Br* 90(4):466–473
- Dawson J, Lavis G (2012) Validity, reliability, and responsiveness of a self-reported foot and ankle score (SEFAS). *Acta Orthop* 83(6):674 (**author reply 674–675**)
- Harris KK, Dawson J, Jones LD, Beard DJ, Price AJ (2013) Extending the use of PROMs in the NHS—using the Oxford Knee Score in patients undergoing non-operative management for knee osteoarthritis: a validation study. *BMJ Open* 3(8):e003365
- King GJ, Richards RR, Zuckerman JD, Blasler R, Dillman C, Friedman RJ, Gartsman GM, Iannotti JP, Murnahan JP, Mow VC, Woo SL (1999) A standardized method for assessment of elbow function. Research Committee, American Shoulder and Elbow Surgeons. *J Shoulder Elbow Surg* 8(4):351–354
- Longo UG, Franceschi F, Loppini M, Maffulli N, Denaro V (2008) Rating systems for evaluation of the elbow. *Br Med Bull* 87:131–161
- MacDermid JC (2001) Outcome evaluation in patients with elbow pathology: issues in instrument development and evaluation. *J Hand Ther* 14(2):105–114
- McHorney CA, Tarlov AR (1995) Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 4(4):293–307
- Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, Bouter LM, de Vet HC (2006) Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. *BMC Med Res Methodol* 6:2
- Parker C, Dewey M (2000) Assessing research outcomes by postal questionnaire with telephone follow-up. TOTAL Study Group. *Trial of Occupational Therapy and Leisure. Int J Epidemiol* 29(6):1065–1069
- Pedersen CK, Danneskiold-Samsøe B, Garrow AP, Waehrens EE, Bliddal H, Christensen R, Bartels EM (2013) Development of a danish language version of the manchester foot pain and disability index: reproducibility and construct validity testing. *Pain Res Treat* 2013:284903
- Sathyamoorthy P, Kemp GJ, Rawal A, Rayner V, Frostick SP (2004) Development and validation of an elbow score. *Rheumatology (Oxford)* 43(11):1434–1440
- Siemiatycki J (1979) A comparison of mail, telephone, and home interview strategies for household health surveys. *Am J Public Health* 69(3):238–245
- Smith TO, Donell ST, Clark A, Chester R, Cross J, Kader DF, Arendt EA (2014) The development, validation and internal consistency of the Norwich Patellar Instability (NPI) score. *Knee Surg Sports Traumatol Arthrosc* 22(2):324–335
- Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC (2007) Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 60(1):34–42
- Terwee CB, Roorda LD, Dekker J, Bierma-Zeinstra SM, Peat G, Jordan KP, Croft P, de Vet HC (2010) Mind the MIC: large variation among populations and methods. *J Clin Epidemiol* 63(5):524–534
- The B, Reininga IH, El Moumni M, Eygendaal D (2013) Elbow-specific clinical rating systems: extent of established validity, reliability, and responsiveness. *J Shoulder Elbow Surg* 22(10):1380–1394
- Thorborg K, Holmich P, Christensen R, Petersen J, Roos EM (2011) The Copenhagen Hip and Groin Outcome Score (HAGOS): development and validation according to the COSMIN checklist. *Br J Sports Med* 45(6):478–491