


Construct validity, floor and ceiling effects, data completeness and magnitude of change for the eight-item HAKIR questionnaire: a patient-reported outcome in the Swedish National Healthcare Quality Registry for hand surgery

Hand Therapy
0(0) 1–14
© The Author(s) 2020
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1758998320966018
journals.sagepub.com/home/hth


Ingela K Carlsson^{1,2} , Elisabeth Ekstrand^{1,3}, Mikael Åström⁴, Kerstin Stihl⁵ and Marianne Arner^{5,6}

Abstract

Introduction: The aim of this study was to evaluate the construct validity, floor and ceiling effects, data completeness and magnitude of change over time for the eight-item patient questionnaire (HQ-8) in the Swedish Healthcare Quality Registry for hand surgery (HAKIR).

Methods: Construct validity was investigated through predefined hypotheses and correlation statistics between the single items in HQ-8 (pain on load, pain on motion without load, pain at rest, stiffness, weakness, numbness, cold sensitivity and ability to perform daily activities) and QuickDASH. Floor and ceiling effects and data completeness were analysed at preoperative ($n = 13,197$), three months ($n = 10,702$) and one year ($n = 9,986$) responses from hand surgery patients. Effect sizes were calculated for pre- and postoperative change scores in elective conditions and postoperative scores for acute conditions.

Results: Correlation coefficients at pre, 3 and 12 months ranged from 0.44 to 0.79 in the total group. No ceiling effect occurred, but a floor effect for the total group was noted for all items at all follow-ups. Missing responses were < 2.6% except for cold sensitivity. The effect sizes varied from small to large for individual items in elective diagnoses. For acute injuries, small effect sizes were found.

Discussion: This study provides evidence of construct validity of HQ-8, lack of ceiling effect, expected floor effect, good data completeness and an ability to detect changes over time. The results indicate that HQ-8 measures unique aspects of disability. The HQ-8 could complement the Quick-DASH in describing patient-reported outcomes after hand surgery.

Keywords

Patient-reported outcome measures, hand, quality of health care, surveys and questionnaires

Date received: 23 March 2020; accepted: 22 September 2020

Introduction

The Swedish National Healthcare Quality Registries (NQRs) have brought about considerable improvements in different fields of medicine.^{1,2} The first NQRs in the 1970s were orthopaedic arthroplasty registries, registering all joint prostheses that had been

¹Department of Hand Surgery, Skåne University Hospital, Malmö, Sweden

²Department of Translational Medicine – Hand Surgery, Lund University, Malmö, Sweden

³Department of Health Sciences, Lund University, Lund, Sweden

⁴Data Analytics and Register Centrum, Region Skåne, Sweden

⁵Department of Hand Surgery, Södersjukhuset, Stockholm, Sweden

⁶Karolinska Institutet, Department of Clinical Science and Education, Södersjukhuset, Stockholm, Sweden

Corresponding author:

Ingela K Carlsson, Department of Hand Surgery, Skane University Hospital, Jan Waldenströms gata 5, 205 02 Malmö, Sweden.
Email: ingela.carlsson@med.lu.se

implanted and removed. As a result, Sweden now has one of the lowest revision rates for knee and hip implants in the world.^{3,4}

The first NQR specifically designed for hand surgery, called HAKIR (short form for hand surgery in Swedish) began collecting data in 2010 and includes all operations performed at the seven specialist departments for hand surgery in Sweden.⁵ The registry is web-based with the option of paper surveys, and has the broad ambition of collecting patient-reported outcomes for all operated patients.⁶ The registry design and method of collecting patient-reported outcomes have previously been described in a review article.⁵

Evaluating outcomes after hand surgery is complex since the distinction between a good or a bad result cannot be made simply by measuring joint motion or grip strength or following the rate of revision of joint implants. Indications for hand surgical procedures are often perceived symptoms like pain, weakness, stiffness and paraesthesia, commonly experienced by patients.⁷⁻¹⁹ Patient-reported outcome measures (PROMs) are therefore important to include in a NQR for hand surgery to give a more complete picture of treatment results. The short version of the Disabilities of the Arm, Shoulder and Hand (Quick-DASH)²⁰ as well as a the eight item questionnaire (HQ-8) were included in HAKIR since its inception (Figure 1). The HQ-8 includes seven questions concerning perceived symptoms in the affected/injured hand and one question about the ability to perform activities of daily living (Figure 1).

A prerequisite for PROMs in a quality registry such as HAKIR is that the systematic collection of data is based on psychometrically sound outcome instruments. This includes measurement properties such as reliability and validity.²¹ The Quick-DASH has retained equal measurement properties as the original DASH with strong evidence for reliability and validity.²² These measurement properties, however, have not been evaluated for the HQ-8.

Important measurement properties in PROMs are content and construct validity. Content validity is defined as the degree to which the content of the instrument is an adequate reflection of the construct to be measured.^{23,24} It refers to the relevance, comprehensiveness and comprehensibility of the included items for the construct, target population and context of interest and is therefore essential to evaluate.²⁴ Construct validity refers to the degree in which scores in an instrument relate to other measures that are consistent with predefined hypotheses concerning the concept being measured.²⁵

An outcome measure should be responsive enough to identify true and clinically meaningful changes in function, not just a change due to random error.²⁶

Floor and ceiling effects occur when a considerable proportion of respondents endorse the best or worst score.²⁷ The measure is then unable to discriminate between respondents at either extreme of the scale. This may indicate a lack of content validity and responsiveness as well as reduced reliability because changes cannot be measured.²⁵ Data completeness (item-response) is another feature important for the content validity of questionnaires.

Content validity was evaluated as part of earlier work; this is briefly outlined in the methods section below. The aim of the present study was to investigate construct validity as well as floor and ceiling effect, data completeness and the ability to detect magnitude of change for the single-item questions (HQ-8) included in HAKIR.

Methods

All operations performed at each participating department are included in the basic registration in HAKIR. Surgical codes and reoperations are registered by hospital staff. Patients are asked to complete the HQ-8 and QuickDASH before, as well as 3 and 12 months after their surgery.⁵ Exclusion criteria for questionnaires are: age below 16 years, cognitive problems, reoperation within 1 year, lack of a Swedish social security number and patients who have declined participation in the registry. Routine follow-ups have shown that more than 80% of all surgeries performed at the participating units are registered, to date more than 120,000 operations. Response rates for the questionnaires have varied since the start in 2010 from around 30% to about 60% at present. The response rates are similar between web and paper forms.

The data collection for the present study took place between 1 February 2010 and 31 December 2015, including 13,197 preoperative questionnaires, 10,702 at 3 months and 9,986 at 12 months after surgery. The response rate was 45% for 3 months and 47% for 12 months questionnaires. The three time-points include different numbers of questionnaires as not all patients returned the questionnaires at all follow-ups. The response rates are calculated on the number of answered questionnaires at each time point divided by the number of questionnaires sent out at the same time.

Twelve common diagnostic subgroups were selected based on ICD-10 coding and surgical codes (KKÅ97) and analysed separately.^{28,29} The elective subgroups were: thumb osteoarthritis, Dupuytren's contracture, de Quervain's, ganglion, trigger finger, carpal tunnel syndrome and ulnar nerve entrapment. The acute subgroups were: digital nerve injuries, extensor and flexor



Date of birth (social security no) (yyyymmdd-nnnn):

Patient questionnaire HQ-8 (arm/hand)

Date (yyyy-mm-dd) - -

I am (please indicate your writing hand): Left handed Right handed Ambidextrous

Arm/hand that was operated on: Left Right

This questionnaire reports on problems that you have had this past week in the hand/arm that was operated on. Please tick the alternative that best corresponds to any of your problems.

1. Pain on load

No problems 0 10 20 30 40 50 60 70 80 90 100 Worst problems imaginable

2. Pain on motion without load

No problems 0 10 20 30 40 50 60 70 80 90 100 Worst problems imaginable

3. Pain at rest

No problems 0 10 20 30 40 50 60 70 80 90 100 Worst problems imaginable

4. Stiffness

No problems 0 10 20 30 40 50 60 70 80 90 100 Worst problems imaginable

5. Weakness

No problems 0 10 20 30 40 50 60 70 80 90 100 Worst problems imaginable

6. Numbness / tingling in fingers

No problems 0 10 20 30 40 50 60 70 80 90 100 Worst problems imaginable

7. Cold Sensitivity (discomfort on exposure to cold)

No problems 0 10 20 30 40 50 60 70 80 90 100 Worst problems imaginable

8. Ability to perform daily activities

No problems 0 10 20 30 40 50 60 70 80 90 100 Worst problems imaginable

Figure 1. The eight-item HAKIR Questionnaire (HQ-8).

tendon injuries, thumb ligament injuries and finger fractures (see Online Supplementary Tables S1–S20).

The study was approved by the Regional Ethics Committee in Stockholm (Dnr: 2017/2023–31).

The HQ-8 questionnaire

Development of the HQ-8. The HAKIR registry was developed between 2008 and 2009 and included a

process in six stages for examining content validity of a suitable PROM instrument. This process took place prior to the present study and is briefly outlined here.

Screening of available patient-reported outcome measures for hand surgery. A literature survey and interdisciplinary group discussions concerning available PROMs were performed by professionals at the hand surgery department in Stockholm. Some pre-existing

PROMs were too general, capturing health-related quality of life or disability in the whole upper extremity rather than describing specific hand-related symptoms. Some instruments were focussed mainly on patients' opinions on delivery of care, others were diagnosis specific, not available in Swedish, or not validated in a Swedish context.^{20,30–36} The consensus of the group was that a questionnaire including physical symptoms relevant for a broad spectrum of hand-related diagnoses would be the best choice in the context of an NQR and could serve as a complement to a region-specific PROM such as the commonly used Quick-DASH.^{20,30}

Item development. Two senior hand surgeons (including the registry holder and co-author, MA) suggested seven common and important symptoms for clinical decision-making and evaluation. The selection of symptoms was based on the following reasons: items 1 to 3 on pain; on load, on motion without load and at rest were selected to reflect variability in pain experience; item 4 on reduced range of motion since it is a prominent symptom in Dupuytren's contracture and common after hand trauma; item 5 on reduced strength which can be caused by pain or following nerve injuries or entrapments; item 6 on reduced sensibility or numbness which are common symptoms in nerve entrapments and injuries; item 7 on hand function in daily activities was added as a general measure of perceived function in the operated hand. A visual analogue scale (VAS) was chosen to reflect perceived problems during the last week in the hand/arm relevant for surgery.

Field testing procedure. Between 2008 and 2009, two hand surgery departments distributed an anonymously answered paper version of the new questionnaire to 471 patients before surgery and 177 patients postoperatively. The 177 paired responses before and after surgery displayed a wide spread of responses for the seven questions and depending on diagnosis, showing reduced, as well as increased symptoms two months after surgery. Patients were also asked to grade how well the questionnaire had captured the symptoms that they regarded as most important on a VAS (0 – not at all to 100 – completely). The mean VAS response was 82 and free text comments were generally positive regarding the relevance of included items. Erroneous markings outside of the line or double responses were noted in 15% of the surveys. This created an incentive to develop a web questionnaire, which was introduced during 2010.

Cognitive interviews. Content validity was further evaluated as part of a Master's thesis in which seven patients participated in a cognitive interview (think aloud)^{37–39} while they responded to the questionnaire.

The interviews took place in a quiet room at the hand department and were performed in 2009 by an occupational therapist with long clinical experience in specialised hand rehabilitation. The included patients (three men and four women, age 25–68 years) were randomly selected at the clinic. The diagnoses were: rheumatoid arthritis, Dupuytren's contracture, thumb fracture and ligament injury of the wrist. Patients' views on the relevance of items and ambiguities in formulations were noted. An oral summary with each patient confirmed that comments were correctly understood. The field notes were then transcribed and analysed with content analyses by the interviewer (fourth author, KS).⁴⁰ Some deficiencies in the questionnaire with room for misunderstanding of instructions were found and question number 6 was considered difficult to answer since it included two questions in one (reduced sensibility and numbness). Responses outside the VAS line were also noted.

Verbal consent was obtained prior to the interview and written and verbal information emphasising the voluntary nature of participation was given by the interviewer. All patients were informed about how data would be analysed and were assured of confidentiality. Ethical approval was not required according to guidelines for Masters' theses at the Karolinska Institute.

Review by a national interdisciplinary expert group. A national group representing all seven hand surgery departments in Sweden and consisting of surgeons, physiotherapists, occupational therapists, nurses and a statistician was formed to further review the questionnaire, all having long scientific and/or clinical experience. The group had four half-day meetings during 2009. The *relevance* of the included symptoms was discussed and was based on clinical experiences in the group and a literature review focusing on patients' own experiences of symptoms.^{8–10,12,14–16} An 8th item concerning discomfort/problems on cold exposure was added to increase *completeness* in breadth of symptoms. Verbal anchors were changed to (0 representing no problem and 100 worst problem imaginable) and words in items 4 to 7 were changed to; stiffness, weakness, numbness/tingling in fingers and ability to perform daily activities. All changes were made to increase *comprehensibility*. The final version of the eight-item HQ-8 was reached through consensus and unanimous decision. In order to reduce respondent burden, the number of included items was restricted, excluding questions, e.g., concerning fine-motor skills, grip function and aesthetics (Figure 1).

Web-survey to hand departments. As a final step before registry start 1 February 2010 in Stockholm,

the questionnaire was also sent out to all department heads at the seven hand departments asking for comments. No further changes were suggested.

HQ-8 – Scale design. During 2010 to 2013, the design used was as a horizontal visual analogue scale (0 = no problem, 100 = worst problem imaginable). This was changed in 2013 to a numeric 11 point box scale (NRS-11)⁴¹ with numerical descriptors (0, 10, 20 ... 100) upon the box. The reason for the change was the transition to web questionnaires for smartphones, where VAS-lines could be of variable length depending on screen size. The verbal anchors and all items remained unchanged (Figure 1). Patients were asked to respond to their experienced problems during the last week in the hand/arm relevant for surgery. In case of surgery for an acute injury, patients were asked to estimate perceived problems prior to the injury.

The Quick-DASH

Patients in HAKIR were also asked to respond to the Swedish version of the 11-item QuickDASH.²⁰ The Quick-DASH is a region-specific self-report outcome instrument quantifying physical function and symptoms in persons with musculoskeletal conditions of the upper limb.^{22,30}

Data analysis

Construct validity was investigated through predefined hypotheses and correlation statistics (Spearman's rank correlation coefficient, r_s). The strength of the correlations was interpreted as: $r_s < 0.5$ low; 0.5 to < 0.7 moderate; > 0.7 high.⁴² The analyses were made for the total group and in 12 subgroups of hand surgery diagnoses.

It was hypothesised that:

1) scores from item 8 in HQ-8 (ability to perform daily activities) would correlate positively with the total Quick-DASH score for the total group and in a selection of hand surgery diagnoses. A point estimate of the correlation of 0.70 or greater is considered a high correlation²¹;

2) scores from items 1 to 3 in HQ-8 (pain on load, pain on motion without load, pain at rest) would correlate positively with the total Quick-DASH score for the total group and that scores from pain on load would show the largest correlation coefficient with a point estimate of the correlation of 0.70 or greater;

3) scores from item 5 in HQ-8 (weakness) would show a point estimate of the correlation of 0.70 or greater and a stronger positive correlation with the total Quick-DASH score compared to items 4, 6 and

7 (stiffness, numbness and cold sensitivity) for the total group.

Analyses were based on completed questionnaires for each follow-up. A high correlation between the VAS and NRS has previously been established.⁴³ We therefore aggregated the data from the VAS and NRS responses in the present study.

Floor (score < 5) and ceiling effects (score > 95) were calculated for each of the HQ-8 questions at pre-surgery, 3 and 12 months follow-up. A threshold of $> 15\%$ was defined as a floor or ceiling effect.²⁷

Data completeness (item-response) was calculated as the number of missing item responses for each HQ-8 question in relation to responded questionnaires. A threshold of $> 15\%$ was defined as unacceptable.⁴⁴

The magnitude of change was quantified by effect size calculations. For elective hand-related diagnoses, the mean paired change between pre- and 3 months follow-up as well as pre- and 12 months follow-up was divided by the dispersion measure (SD) of the pre-operative (baseline) score in each HQ-8 question. For acute hand-related diagnoses, the mean paired change between 3 and 12 months follow-up was divided by the dispersion measure (SD) of the 3-months score.^{45,46} According to Cohen's criteria, an effect size of 0.20 is considered small, 0.50 is medium and 0.80 is large.⁴⁷ Data were analysed using Microsoft Excel 2013 (Microsoft Cooperation, Redmond, Washington, United States) and the SPSS software package version 23 (IBM Cooperation, Amonk, New York, United States).

Results

Sample characteristics

In total, data from 33,885 questionnaires were analysed; 13,197 before operation, 10,702 at 3 months and 9,986 at 12 months after hand surgery. Mean age was 52 (14–99) years, with 50% men and 50% women. Thirty-six percent of the questionnaires were completed in a web format and 64% in a pencil and paper format.

Construct validity

The correlation coefficients between the eight single-item questions in HQ-8 and the total score in Quick-DASH for the total group at pre, 3 and 12 months follow-up ranged from 0.44 to 0.79. The corresponding r_s values for the 12 subgroups showed a range from 0.29 to 0.85. The strongest correlation ($r_s = 0.79$) in the total group was noted for item 8 (ability to perform daily activities) and the total Quick-DASH score at the 3 months follow-up. The weakest correlation ($r_s = 0.44$) in the total group was noted for item 7 (cold sensitivity)

HQ-8 questions	Preoperative	3 months	12 months
1. Pain on load	0.69	0.76	0.74
2. Pain on motion without load	0.68	0.67	0.65
3. Pain at rest	0.65	0.58	0.58
4. Stiffness	0.5	0.58	0.58
5. Weakness	0.67	0.73	0.74
6. Numbness/tingling in fingers	0.52	0.47	0.54
7. Cold sensitivity	0.44	0.46	0.52
8. Ability to perform daily activities	0.72	0.79	0.76

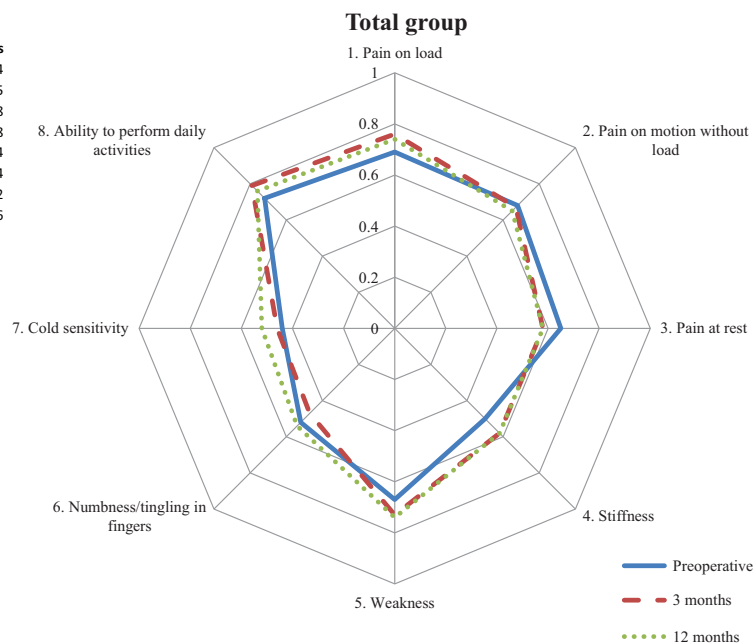


Figure 2. Spearman's correlation coefficient between HQ-8 questions and total QuickDASH score for the total group, preoperatively and at 3 and 12 months follow-up. The main axis represents the correlation coefficient (r_s 0–1).

and total Quick-DASH score in the pre-operative analysis. Detailed information for the total group and the 12 diagnostic subgroups is available in Figures 2 to 5, and in the Online Supplementary Tables S1–S8.

Hypothesis 1: The a priori criterion of a high positive correlation ($r_s \geq 0.70$) was met in the total group for the question concerning the ability to perform daily activities (item 8) and the total Quick-DASH score at all follow-ups. Although high positive r_s values were seen, the expected a priori level was only met at the postoperative follow-ups for the subgroup thumb osteoarthritis, and for Dupuytren's contracture and for trigger finger it was only reached at the 3 months follow-up (Figures 2 to 5 and Online Supplementary Table S8).

Hypothesis 2: The a priori criterion of high positive correlations was met in the total group and at all follow-ups for pain on load, pain on motion without load, pain at rest and the total Quick-DASH score. Pain on load showed the strongest correlation ($r_s > 0.70$) at 3 and 12 months follow-up. However, a weaker r_s was noted preoperatively for all pain-related questions and total Quick-DASH scores for the subgroup thumb osteoarthritis (Figures 2 and 5 and Online Supplementary Tables S1–S3).

Hypothesis 3: A stronger positive correlation in the total group was seen for weakness ($r_s > 0.70$) and total Quick-DASH score at all follow-ups, compared to r_s values for stiffness, numbness and cold sensitivity as hypothesised a priori (Figure 2 and Online

Supplementary Tables S4–S7). For the thumb osteoarthritis group, r_s values for weakness and total Quick-DASH scores preoperatively and at 3 and 12 months follow-up were 0.36, 0.63 and 0.70, respectively (Figure 5 and Online Supplementary Table S5).

Floor and ceiling effects

No ceiling effect (score >95) was noted in any of the HQ-8 questions or at pre, 3 and 12 months follow-up in the total group or in the 12 diagnostic subgroups. A floor effect (score <5) was seen in all HQ-8 questions and follow-ups in the total group (Table 1). Differences related to HQ-8 questions in the 12 subgroups are presented in the Online Supplementary Tables S1–S8.

Data completeness

The percentage of missing item responses in the responded questionnaires and for all HQ-8 questions was $<2.6\%$, except for the question concerning cold sensitivity (item 7) where 6.2%, 7.6% and 2.9% missing responses were noted at pre, 3 and 12 months follow-up, respectively (Table 2). Detailed information for each of the diagnostic subgroups is available in the Online Supplementary Tables S1–S8.

Magnitude of change

A varied pattern in effect sizes (ES), small to large, for the HQ-8 items 1–8 was seen for subgroups of elective hand-related diagnosis (Table 3). A large ES was seen

HQ-8 questions	Preoperative	3 months	12 months
1. Pain on load	0.65	0.71	0.72
2. Pain on motion without load	0.63	0.68	0.67
3. Pain at rest	0.52	0.61	0.63
4. Stiffness	0.52	0.57	0.63
5. Weakness	0.63	0.7	0.71
6. Numbness/tingling in fingers	0.43	0.55	0.6
7. Cold sensitivity	0.43	0.53	0.54
8. Ability to perform daily activities	0.68	0.77	0.76

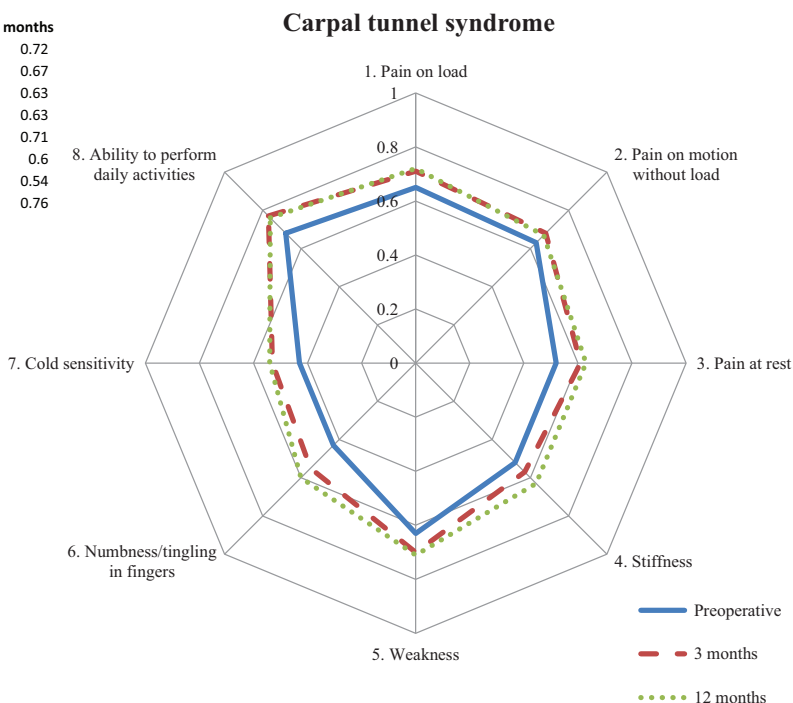


Figure 3. Spearman's rank correlation coefficient between questions in HQ-8 and total QuickDASH score for the subgroup carpal tunnel syndrome, preoperatively and at 3 and 12 months follow-up. The main axis represents the correlation coefficient (r_s 0–1).

HQ-8 questions	Preoperative	3 months	12 months
1. Pain on load	0.54	0.66	0.65
2. Pain on motion without load	0.5	0.5	0.53
3. Pain at rest	0.45	0.4	0.45
4. Stiffness	0.33	0.6	0.6
5. Weakness	0.55	0.62	0.66
6. Numbness/tingling in fingers	0.48	0.47	0.46
7. Cold sensitivity	0.39	0.46	0.47
8. Ability to perform daily activities	0.56	0.65	0.64

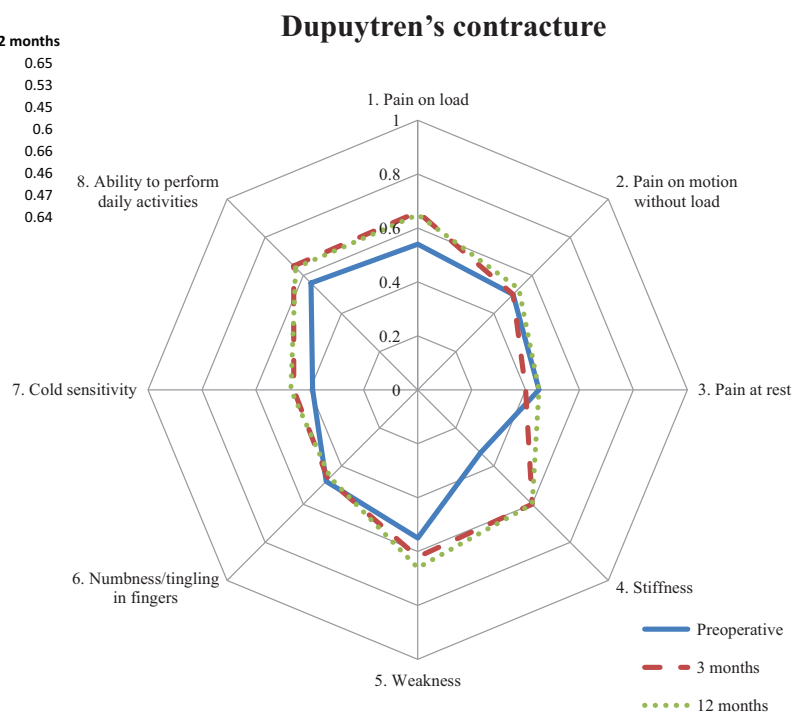


Figure 4. Spearman's rank correlation coefficient between questions in HQ-8 and total QuickDASH score for the subgroup Dupuytren's contracture, preoperatively and at 3 and 12 months follow-up. The main axis represents the correlation coefficient (r_s 0–1).

HQ-8 questions	Preoperative	3 months	12 months
1. Pain on load	0.43	0.72	0.76
2. Pain on motion without load	0.46	0.69	0.63
3. Pain at rest	0.47	0.65	0.54
4. Stiffness	0.35	0.56	0.51
5. Weakness	0.36	0.63	0.7
6. Numbness/tingling in fingers	0.4	0.45	0.43
7. Cold sensitivity	0.31	0.45	0.44
8. Ability to perform daily activities	0.53	0.77	0.75

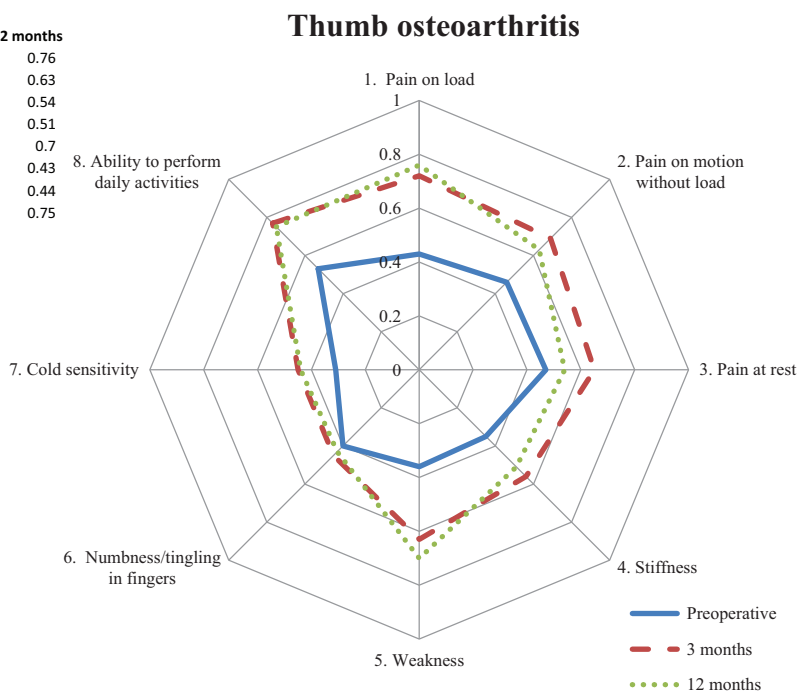


Figure 5. Spearman's rank correlation coefficient between questions in HQ-8 and total QuickDASH score for the subgroup thumb osteoarthritis, preoperatively and at 3 and 12 months follow-up. The main axis represents the correlation coefficient (r_s 0–1).

Table 1. Floor and ceiling effects.

	#1	#2	#3	#4	#5	#6	#7	#8
<i>Pre-surgery follow-up (12,384–12,991)</i>								
<5	21%	31%	38%	27%	24%	43%	45%	19%
>95	5%	1%	1%	4%	4%	4%	3%	5%
<i>3-months follow-up (n = 9891–10,528)</i>								
<5	23%	42%	57%	24%	23%	51%	48%	33%
>95	1%	0%	0%	2%	2%	2%	3%	1%
<i>12-months follow-up (n = 9701–9835)</i>								
<5	34%	53%	62%	34%	32%	53%	46%	43%
>95	1%	0%	0%	2%	2%	2%	4%	1%

= Item number.

Note: Percent of responded questionnaires, score <5 (floor) and score >95 (ceiling) for the HQ-8 questions at pre, 3 and 12 months follow-up (0 representing no problem and 100 worst problem imaginable).

Table 2. Descriptive statistics for item-response in individual HQ-8 questions (0 representing no problem and 100 worst problem imaginable).

	#1	#2	#3	#4	#5	#6	#7	#8
<i>Pre-surgery follow-up</i>								
Number*	13,197	13,197	13,197	13,197	13,197	13,197	13,197	13,197
Missing	244	235	247	328	342	312	813	206
Missing %	1.8%	1.8%	1.9%	2.5%	2.6%	2.4%	6.2%	1.6%
<i>3 months follow-up</i>								
Number*	10,702	10,702	10,702	10,702	10,702	10,702	10,702	10,702
Missing	234	181	175	174	202	210	811	207
Missing %	2.2%	1.7%	1.6%	1.6%	1.9%	2.0%	7.6%	1.9%
<i>12 months follow-up</i>								
Number*	9986	9986	9986	9986	9986	9986	9986	9986
Missing	188	151	158	158	179	180	285	170
Missing %	1.9%	1.5%	1.6%	1.6%	1.8%	1.8%	2.9%	1.7%

* = Total sample. # = item number.

Note: Number and percent of missing item responses for each HQ-8 question (%) in total group at pre, 3 and 12 months follow-up.

Table 3. Effect sizes at 3 and 12 months follow-up in subgroups of elective hand-related diagnosis. Detailed information is given in the Online Supplementary Tables S9–15.

		Effect sizes (ES) at 3 months follow-up/ES at 12 months follow-up.						
		Thumb osteo-arthritis 3/12 months follow-up	Dupuytren's contracture 3/12 months follow-up	De Quervain 3/12 months follow-up	Ganglion 3/12 months follow-up	Trigger finger 3/12 months follow-up	Carpal tunnel syndrome 3/12 months follow-up	Ulnar nerve entrapment 3/12 months follow-up
HQ-8								
1	Pain on load	M/L	S/S	M/L	S/M	S/L	S/M	S/S
2	Pain on motion without load	M/L	S/S	S/M	S/S	S/M	S/M	S/S
3	Pain at rest	M/L	S/S	S/S	S/S	S/M	M/M	S/S
4	Stiffness	S/M	M/M	S/S	S/S	S/M	S/M	S/S
5	Weakness	S/M	S/S	S/M	S/S	S/M	S/M	S/S
6	Numbness/tingling in fingers	S/S	S/S	S/S	S/S	S/S	M/L	S/S
7	Cold sensitivity	S/S	S/S	S/S	S/S	S/S	S/S	S/S
8	Ability to perform daily activities	M/L	M/M	M/M	S/S	M/L	S/M	S/S

Note: Mean paired change between pre- and 3 months follow-up/SD on pre- months scores and mean change between pre- and 12 months follow-up/SD on the pre- months scores expressed as small (S), medium (M) or large (L) effect size (ES). According to Cohen's criteria, an effect size of 0.20 is considered small, 0.50 is medium and 0.80 is large.⁴⁷

for all pain questions (ES: -1.17 , -0.95 , -0.75) and the ability to perform daily activities (ES: -0.96) in the subgroup thumb osteoarthritis at the 12 months follow-up. This was also seen for pain on load for the subgroups de Quervain and trigger finger at the same follow-up. Furthermore, the question about the ability to perform daily activities also showed a large ES for the diagnosis trigger finger and numbness/tingling in fingers showed a large ES for the diagnosis carpal tunnel syndrome at the 12 months follow-up. A small to medium ES was noted in all other analyses. Detailed information is given in Online Supplementary Tables S9–15.

For acute hand injuries (digital nerve injuries, extensor and flexor tendon injuries, thumb ligament injuries and finger fractures), effect sizes were evaluated between 3 and 12 months follow-up. The effect size for all diagnoses and HQ-8 questions was small (see the Online Supplementary Tables S16–20).

Discussion

This study provides evidence of construct validity of the HQ-8 questionnaire for the evaluation of patient-reported outcomes in a registry-based setting. Furthermore, we have shown a lack of ceiling effect, an expected floor effect, good data completeness and an ability to detect change over time for HQ-8.

To include both single-item questions (HQ-8) and the Quick-DASH provides the opportunity to evaluate both separate symptoms and overall self-reported disability. The HQ-8 questions mainly focus on

impairments on the level of body function in the affected/operated hand, while most items in the Quick-DASH evaluate the overall ability to perform certain activities, regardless of which hand/arm that is used.^{30,48} Since the occurrence and importance of different symptoms vary between different hand-related diagnoses, a conscious choice was to present the included HQ-8 items individually, and not as a summed score. The collected information about individual symptoms can also give better guidance to patients on the expected results after a surgical procedure. The main goal for a quality registry is to facilitate quality improvement work as well as clinical research by collecting treatment results for a large number of patients. By investigating patient-reported outcomes, more targeted interventions can be offered to our patients. Patients with diagnoses such as Dupuytren's contracture, in which pain is an uncommon symptom, usually score within normal population values for the Quick-DASH,⁴⁹ which makes this instrument less valuable for evaluating treatment results in these patients. The questions in the HQ-8 regarding stiffness and ability to perform daily activities may serve as more relevant and sensitive outcome measures, which was supported by the effect size calculations seen in the present study.

The development of HQ-8 and the primary investigation of content validity were performed prior to the present study and included evaluation of *relevance*, *comprehensiveness* and *comprehensibility* of the included items. Rating of problems on cold exposure was added to increase *comprehensiveness* of included

items. The decision to limit the included items to physical symptoms and a general question concerning ability to perform daily activities was made, despite the fact that personal factors or psychological consequences of a hand trauma are well known mediators of outcome.^{14,50} To combine HQ-8 with other PROMs, covering these areas is therefore a possibility for future research projects. Although several steps were taken to evaluate and confirm content validity of HQ-8 prior to the present study, the process was not part of a planned research study and one has to bear in mind that the process of establishing content validity is an ongoing process. A systematic literature search on qualitative research focusing on experienced symptoms in various subgroups of hand-related diagnoses may contribute to the investigation of content validity of the HQ-8. Further studies are therefore needed to establish content validity regarding the comprehensiveness of the included items in HQ-8.

The use of an 11 increment NRS question has been recommended as it is a responsive scale with good completion rate, easy to use and similar psychometric properties compared to a visual analogue scale.^{41,43,51–54} However, the definition of anchors has been commented on for pain ratings⁵⁵ but may be true also for ratings of other impairments. The upper anchor used for HQ-8 (worst problem imaginable) may depend on patients' imagination of the worst case scenario, whereas the alternative 'worst pain ever experienced' is only interpretable when having knowledge about patients' pain history.⁵⁵ The idea of accuracy when measuring pain has also been challenged since an unequivocal reference standard does not exist or cannot be obtained.⁵⁶ Self-reports of pain can be influenced by previous experiences, behavioural, affective or cognitive factors and vary depending on the context in which the pain is experienced. When measuring change over time, it is therefore of importance to use mean paired comparisons to limit the variability between person to person. Multidimensional pain scales have been recommended offering a broader understanding of pain experience than NRS,⁵⁷ but are not applicable in a total population of hand surgery patients. To capture different levels of pain intensity, ranking of both pain on load, on motion without load and at rest was included in HQ-8. Furthermore, it is also possible to include other PROMs in research projects.

Construct validity for HQ-8 was good, showing a broader range in r_s between the single items in HQ-8 and Quick-DASH when analysed in subgroups compared to the total group analyses. A weaker r_s for stiffness, numbness and cold sensitivity with the Quick-DASH may indicate that some clinical outcome aspects are not captured by the Quick-DASH

instrument. A symptom questionnaire such as the HQ-8 may therefore provide additional value.

A high level of positive correlation between item 8 (ability to perform daily activities) and the total Quick-DASH score was found (hypothesis 1). This was expected since the main focus on items included in Quick-DASH reflects the ability to perform certain activities. Question 8 reflects the ability to perform daily activities with the affected/operated hand, whereas Quick-DASH does not distinguish between hand use, hand dominance or affected hand. This may be part of an explanation why a 'perfect' correlation close to 1 is not present and confirms the relevance of both outcome measures.

Strong positive correlations were seen in the total group for all HQ-8 questions concerning pain and the total Quick-DASH score (hypothesis 2). In our experience, pain is one of the most limiting factors for satisfactory performance of activities. This is also confirmed in previous studies.^{58–60} A weaker r_s was, however, noted preoperatively for all pain-related questions and total Quick-DASH scores for the subgroup thumb osteoarthritis. A possible explanation may be that patients with thumb osteoarthritis rank their problems with pain as more problematic compared to performance in activities preoperatively. Access to compensatory strategies may also be reflected in a proportionally lower Quick-DASH score. Postoperatively, the perceived pain improves and the scores on pain and Quick-DASH are therefore reflected in stronger r_s values.

In the total group, a stronger positive correlation was seen for weakness with total Quick-DASH score than for stiffness, numbness and cold sensitivity (hypothesis 3). Hand strength has previously been shown to correlate strongly with DASH scores in a variety of hand-related diagnoses.^{58,60–62} For the thumb osteoarthritis group, higher r_s values were seen postoperatively for weakness and total Quick-DASH score. This may indicate that an increased strength after surgery is important for performance of activities.

A lack of content validity may be indicated if a floor and ceiling effect occurs since the ability to discriminate between subjects at the extremes of the scale is lost.²⁵ A floor effect²⁷ was seen in the total group and in all questions (HQ-8) with the largest floor effect seen at 12-months follow-up. This, however, was not observed in all of the selected subgroups since certain symptoms are neither relevant before, nor after surgery for some diagnoses. It is also expected that treatments should decrease or completely diminish symptoms that are relevant for different subgroups.

Data completeness was very good and surpassed the broad guidelines on the defined threshold.⁴⁴ One exception was cold sensitivity (item 7) with a slightly higher percentage of missing item responses at pre and

3 months follow-up. This may be explained by the time-point when the question was answered since the relevance of responding may vary during different seasons. Although cold sensitivity has been described as a frequent problem in mixed groups of hand injuries,^{63,64} it is not present for all hand surgery patients. This may influence the motivation to reply.

Interesting to note is that the effect sizes in various elective subgroups were consistent with clinical experience. For instance, pain is the symptom causing disability for patients with thumb osteoarthritis and is therefore the main indication for surgery.¹⁷ A large ES at 12 months follow-up for all included pain questions (item 1-3) in HQ-8, as well as the question concerning ability to perform daily activities, was therefore as expected. Another example of expected ES is the improvement of numbness/tingling in fingers (item 6) following a carpal tunnel release. The small ES following surgery for ulnar nerve entrapment indicates poorer treatment results in this patient group.

The small ES between 3 and 12 months follow-up for the selected acute hand-related injuries is also as expected. The main change in scores probably occurs within three months after the injury and when the rehabilitation usually is completed. The preoperative scores are completed retrospectively and recall bias especially in an acute injury situation may exist. We therefore chose to prioritise ES calculations on the mean paired change between 3 and 12 months. Even though the results of the ES calculations were consistent with clinical experience following surgery in the selected subgroups, one has to remember that ES only measures the magnitude of change and not the clinical importance.²³

According to the COSMIN guidelines, responsiveness refers to the validity of a change score²³ and it has been defined as the ability to detect minimal important change (MIC) over time, even if these changes are small. Furthermore, instruments should be able to distinguish MIC from measurement error.²⁵ The present study included longitudinal data at pre, 3 and 12 months follow-up. In a future study, we plan to collect test-retest data at a time-interval (e.g. 1-3 weeks) relevant for analyses of agreement and measurement error.

Methodological considerations

The response rate was 45% for 3 months and 47% for 12 months postoperative questionnaires. It has been reported that on-line surveys in general are less likely than paper surveys to receive high response rates.⁶⁵ This was initially also the case for the HAKIR surveys, but after improving web-functionality and sending out a reminder phone text messages, the response rate is now similar between web and paper. Response rates around 30–40% have commonly been reported for

large on-line surveys.⁶⁵ The response rate for the HAKIR questionnaire could therefore be considered as acceptable, even though efforts should be made to increase it further.

In conclusion, the present study provides evidence of the construct validity of the HQ-8, including good data completeness, expected floor effect, lack of ceiling effect and an ability to detect changes in scores over time. The different associations between HQ-8 and Quick-DASH in different hand diagnoses indicate that the HQ-8 measures unique aspects of disability. Additional strengths of the HQ-8 are the specific focus on the affected hand and a more detailed reporting of symptoms that patients usually regard as important, such as pain, stiffness and cold sensitivity. The HQ-8 may serve as an important complement to the ‘broader’ content in Quick-DASH with its main focus on overall performance on activity level. The initial content validity process performed during the registry set-up and prior to the present study included several steps confirming the relevance, comprehensiveness and comprehensibility of the included items in HQ-8. However, more studies are needed to further confirm content validity, evaluate test-retest reliability and responsiveness including clinically meaningful change of the HQ-8 in subgroups of hand diagnoses.

Acknowledgements

The authors want to thank Lars Hagberg, MD, PhD, Stockholm for invaluable professional contribution during the early development of HQ-8.

Pre-Print version

An earlier preprint version of the manuscript is available on BMC's preprint repository Research Square (<https://doi.org/10.21203/rs.2.18115/v1>).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received financial support for the research, authorship, and/or publication of this article: The HAKIR registry is funded by the Swedish Association of Local Authorities and Regions and the Swedish Ministry of Health and Social Affairs.

Ethical approval

The study was approved by the Regional Ethics Committee in Stockholm (Dnr: 2017/2023-31). The HAKIR registry is web-based with a secure logon function. The legal requirements for healthcare registries in Sweden do not demand active

consent from the participants. There are however strict demands on information to all participants before registration, for instance about the option of not being registered (opt-out) and the possibility of having all personal data erased at any time. When patients are scheduled for hand surgery, they are informed orally about HAKIR by staff, as well as being informed through an information brochure which is available in six languages (https://hakir.se/wp-content/uploads/2018/06/Eng_Patientinformation_2018.pdf).

Availability of data and materials

Public access to data on individual patients is restricted by the Swedish Authorities (Public Access to Information and Secrecy Act; <http://www.government.se/information-material/2009/09/public-access-to-information-and-secrecy-act/>) but data can be made available for researchers after a special review that includes approval of the research project by both an Ethics Committee and the authorities' data safety committees. For HAKIR, the Central Personal Data Controller is the South General Hospital in Stockholm and the Data Protection Officer can be reached at GDPR.sodersjukhuset@ssl.se. We have supplied all aggregated data in supplementary tables.

Guarantor

IKC

Contributorship

MA and IKC were responsible for the design and planning of the study. EE, MÅ and IKC performed the statistical analyses and interpreted the data together with MA. IKC did the drafting of the manuscript with contributions from the other authors. All authors read and approved the final manuscript.

ORCID iD

Ingela K Carlsson  <https://orcid.org/0000-0002-2746-8346>

Supplemental material

Supplemental material for this article is available online.

References

- Hailer NP, Garellick G and Karrholm J. Uncemented and cemented primary total hip arthroplasty in the Swedish hip arthroplasty register. *Acta Orthop* 2010; 81: 34–41.
- Robertsson O, Ranstam J, Sundberg M, et al. The Swedish knee arthroplasty register: a review. *Bone Joint Res* 2014; 3: 217–222.
- Makela KT, Matilainen M, Pulkkinen P, et al. Countrywise results of total hip replacement. An analysis of 438,733 hips based on the Nordic arthroplasty register association database. *Acta Orthop* 2014; 85: 107–116.
- Robertsson O, Bizjajeva S, Fenstad AM, et al. Knee arthroplasty in Denmark, Norway and Sweden. A pilot study from the Nordic arthroplasty register association. *Acta Orthop* 2010; 81: 82–89.
- Arner M. Developing a national quality registry for hand surgery: challenges and opportunities. *EFORT Open Rev* 2016; 1: 100–106.
- Kane RL and Chung KC. Establishing a national registry for hand surgery. *J Hand Surg Am* 2020; 45: 57–61.
- Ashwood M, Jerosch-Herold C and Shepstone L. Learning to live with a hand nerve disorder: a constructed grounded theory. *J Hand Ther* 2019; 32: 334–344.
- Barbier O, Penta M and Thonnard JL. Outcome evaluation of the hand and wrist according to the international classification of functioning, disability, and health. *Hand Clin* 2003; 19: 371–378.
- Bialocerkowski AE, Grimmer KA and Bain GI. Validity of the patient-focused wrist outcome instrument: do impairments represent functional ability? *Hand Clin* 2003; 19: 449–455.
- Carlsson IK, Edberg AK and Wann-Hansson C. Hand-injured patients' experiences of cold sensitivity and the consequences and adaptation for daily life: a qualitative study. *J Hand Ther* 2010; 23: 53–61.
- Chemnitz A, Dahlin LB and Carlsson IK. Consequences and adaptation in daily life – patients' experiences three decades after a nerve injury sustained in adolescence. *BMC Musculoskelet Disord* 2013; 14: 252.
- Elliot D and Harris SB. The assessment of flexor tendon function after primary tendon repair. *Hand Clin* 2003; 19: 495–503.
- Engstrand C, Kvist J and Krevers B. Patients' perspective on surgical intervention for Dupuytren's disease – experiences, expectations and appraisal of results. *Disabil Rehabil* 2016; 38: 2538–2549.
- Gustafsson M, Persson LO and Amilon A. A qualitative study of stress factors in the early stage of acute traumatic hand injury. *J Adv Nurs* 2000; 32: 1333–1340.
- Jerosch-Herold C, Mason R and Chojnowski AJ. A qualitative study of the experiences and expectations of surgery in patients with carpal tunnel syndrome. *J Hand Ther* 2008; 21: 54–61.
- Schuind FA, Mouraux D, Robert C, et al. Functional and outcome evaluation of the hand and wrist. *Hand Clin* 2003; 19: 361–369.
- Wilcke M, Roginski M, Astrom M, et al. A registry based analysis of the patient reported outcome after surgery for trapeziometacarpal joint osteoarthritis. *BMC Musculoskelet Disord* 2020; 21: 63.
- Zimmerman M, Anker I, Karlsson A, et al. Ulnar nerve entrapment in diabetes: patient-reported outcome after surgery in national quality registries. *Plast Reconstr Surg Glob Open* 2020; 8: e2740.
- Zimmerman M, Dahlin E, Thomsen NOB, et al. Outcome after carpal tunnel release: impact of factors related to metabolic syndrome. *J Plast Surg Hand Surg* 2017; 51: 165–171.
- Gummesson C, Ward MM and Atroshi I. The shortened disabilities of the arm, shoulder and hand questionnaire (QuickDASH): validity and reliability based on responses within the full-length DASH. *BMC Musculoskelet Disord* 2006; 7: 44.

21. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; 63: 737–745.
22. Kennedy CA, Beaton DE, Smith P, et al. Measurement properties of the QuickDASH (disabilities of the arm, shoulder and hand) outcome measure and cross-cultural adaptations of the QuickDASH: a systematic review. *Qual Life Res* 2013; 22: 2509–2547.
23. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol* 2010; 10: 22.
24. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res* 2018; 27: 1159–1170.
25. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60: 34–42.
26. Streiner DL and Norman GR. *Health measurements scales: a practical guide to their development and use*. 3rd ed. Oxford: Oxford University Press, 2003.
27. McHorney CA and Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1995; 4: 293–307.
28. No author listed. *Classification of surgical procedures*. NOMESKO 4619996.
29. World Health Organization. *The international statistical classification of diseases and health related problems*. 10th ed. Geneva: WHO, 1996.
30. Beaton DE, Wright JG and Katz JN; Upper Extremity Collaborative Group. Development of the QuickDASH: comparison of three item-reduction approaches. *J Bone Joint Surg Am* 2005; 87: 1038–1046.
31. Chung KC, Pillsbury MS, Walters MR, et al. Reliability and validity testing of the Michigan hand outcomes questionnaire. *J Hand Surg Am* 1998; 23: 575–587.
32. Dias JJ, Bhowal B, Wildin CJ, et al. Assessing the outcome of disorders of the hand. Is the patient evaluation measure reliable, valid, responsive and without bias? *J Bone Joint Surg Br* 2001; 83: 235–240.
33. Dias JJ, Rajan RA and Thompson JR. Which questionnaire is best? The reliability, validity and ease of use of the patient evaluation measure, the disabilities of the arm, shoulder and hand and the Michigan hand outcome measure. *J Hand Surg Eur* 2008; 33: 9–17.
34. Gummesson C, Atroshi I and Ekdahl C. The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. *BMC Musculoskelet Disord* 2003; 4: 11.
35. Hudak PL, Amadio PC, Bombardier C, et al. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The upper extremity collaborative group (UECG). *Am J Ind Med* 1996; 29: 602–608.
36. Rabin R and de Charro F. EQ-5D: a measure of health status from the EuroQol group. *Ann Med* 2001; 33: 337–343.
37. Drennan J. Cognitive interviewing: verbal data in the design and pretesting of questionnaires. *J Adv Nurs* 2003; 42: 57–63.
38. Willis GB and Artino AR, Jr. What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. *J Grad Med Educ* 2013; 5: 353–356.
39. Stihl K. Content validity and internal consistency of the hand surgery self assessment questionnaire. 2010; Inst neurobiologi, vårdvetenskap och samhälle. One-year master thesis, 15 academic points, <https://hakil.se/wp-content/uploads/2020/06/Magisteruppsats-Kerstin-Stihl-2010.pdf> (accessed 25 June 2020).
40. Graneheim UH and Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Educ Today* 2004; 24: 105–112.
41. Williamson A and Hoggart B. Pain: a review of three commonly used pain rating scales. *J Clin Nurs* 2005; 14: 798–804.
42. Mukaka MM. Statistics corner: a guide to appropriate use of correlation coefficient in medical research. *Malawi Med J* 2012; 24: 69–71.
43. Hjermstad MJ, Fayers PM, Haugen DF, et al. Studies comparing numerical rating scales, verbal rating scales, and visual analogue scales for assessment of pain intensity in adults: a systematic literature review. *J Pain Symptom Manage* 2011; 41: 1073–1093.
44. D, Vet HC, Terwee CB, Mokkink LB, et al. *Measurement in medicine: a practical guide*. New York: Cambridge University Press, 2011.
45. Kazis LE, Anderson JJ and Meenan RF. Effect sizes for interpreting changes in health status. *Med Care* 1989; 27: S178–189.
46. Stratford PW, Binkley FM and Riddle DL. Health status measures: strategies and analytic methods for assessing change scores. *Phys Ther* 1996; 76: 1109–1123.
47. Cohen J. *Statistical power analysis for the behaviour sciences*. New York: Academic Press, 1977.
48. Coenen M, Kus S, Rudolf KD, et al. Do patient-reported outcome measures capture functioning aspects and environmental factors important to individuals with injuries or disorders of the hand? *J Hand Ther* 2013; 26: 332–342.
49. Engstrand C, Krevers B, Nylander G, et al. Hand function and quality of life before and after fasciectomy for Dupuytren contracture. *J Hand Surg Am* 2014; 39: 1333–1343.
50. Koestler AJ. Psychological perspective on hand injury and pain. *J Hand Ther* 2010; 23: 199–210.
51. Ferreira-Valente MA, Pais-Ribeiro JL and Jensen MP. Validity of four pain intensity rating scales. *Pain* 2011; 152: 2399–2404.
52. Flaherty SA. Pain measurement tools for clinical practice and research. *Aana J* 1996; 64: 133–140.
53. Gagliese L, Weizblit N, Ellis W, et al. The measurement of postoperative pain: a comparison of intensity scales in

- younger and older surgical patients. *Pain* 2005; 117: 412–420.
54. Litcher-Kelly L, Martino SA, Broderick JE, et al. A systematic review of measures used to assess chronic musculoskeletal pain in clinical and randomized controlled clinical trials. *J Pain* 2007; 8: 906–913.
 55. Walton DM, Elliott JM, Salim S, et al. A reconceptualization of the pain numeric rating scale: anchors and clinically important differences. *J Hand Ther* 2018; 31: 179–183.
 56. Feinstein A. *Clinimetrics*. New Haven: Yale University Press, 1987.
 57. Badalamente M, Coffelt L, Elfar J, et al. Measurement scales in clinical research of the upper extremity, part 2: outcome measures in studies of the hand/wrist and shoulder/elbow. *J Hand Surg Am* 2013; 38: 407–412.
 58. Ali M, Brogren E, Wagner P, et al. Association between distal radial fracture malunion and patient-reported activity limitations: a long-term follow-up. *J Bone Joint Surg Am* 2018; 100: 633–639.
 59. Montero A, Mulero JF, Tornero C, et al. Pain, disability and health-related quality of life in osteoarthritis-joint matters: an observational, multi-specialty trans-national follow-up study. *Clin Rheumatol* 2016; 35: 2293–2305.
 60. Swart E, Nellans K and Rosenwasser M. The effects of pain, supination, and grip strength on patient-rated disability after operatively treated distal radius fractures. *J Hand Surg Am* 2012; 37: 957–962.
 61. MacDermid JC, Wessel J, Humphrey R, et al. Validity of self-report measures of pain and disability for persons who have undergone arthroplasty for osteoarthritis of the carpometacarpal joint of the hand. *Osteoarthritis Cartil* 2007; 15: 524–530.
 62. Wilcke MK, Abbaszadegan H and Adolphson PY. Patient-perceived outcome after displaced distal radius fractures. A comparison between radiological parameters, objective physical variables, and the DASH score. *J Hand Ther* 2007; 20: 290–298.
 63. Carlsson IK and Dahlin LB. Self-reported cold sensitivity in patients with traumatic hand injuries or hand-arm vibration syndrome – an eight year follow up. *BMC Musculoskelet Disord* 2014; 15: 83.
 64. Craigen M, Kleinert JM, Crain GM, et al. Patient and injury characteristics in the development of cold sensitivity of the hand: a prospective cohort study. *J Hand Surg* 1999; 24: 8–15.
 65. Nulty DD. The adequacy of response rates to online and paper surveys: what can be done? *Assessment Eval High Educ* 2008; 33: 301–314.