ELSEVIER

Scientific/Clinical Article

Using Carpal Tunnel Questionnaire in clinical practice: A systematic review of its measurement properties

Check for updates

Saurabh P. Mehta PT, PhD [a,b,*], Gwen Weinstock-Zlotnick PhD, OTR/L, CHT [c], Kayla L. Akland PT, DPT [d], Margaret M. Hanna PT, DPT [e], Kayla J. Workman PT, DPT [f]

[a] School of Physical Therapy, Marshall University, Huntington, WV, USA
[b] Department of Orthopedic Surgery, Joan C. Edwards School of Medicine, Marshall University, Huntington, WV, USA
[c] Department of Hand Therapy, Hospital for Special Surgery, New York, NY, USA
[d] Fleming County Hospital Rehabilitation, Flemingsburg, KY, USA
[e] OrthoCarolina, Charlotte, NC, USA
[f] EmergeOrtho Physical Therapy, Jacksonville, NC, USA

ARTICLE INFO

ABSTRACT

Introduction: Carpal Tunnel Questionnaire (CTQ) is widely used for assessing condition-specific impairments in individuals with carpal tunnel syndrome (CTS) or for assessing outcomes after carpal tunnel surgery (carpal tunnel release [CTR]). A systematic review of its measurement properties can greatly facilitate its evidence-based use in clinical practice. The purpose of this study was to systematically locate, appraise, and synthesize the evidence concerning the reliability, responsiveness, validity, minimal detectable change (MDC), and minimal clinically important difference (MCID) for the CTQ and its scales.
Study Design: This is a systematic review of measurement properties.
Methods: Using predefined keywords, PubMed, CINAHL, PsychInfo, and ProQuest were searched to locate primary studies that assessed measurement properties of the CTQ. The methodological quality of the included studies was assessed using a standardized tool. Data concerning the measurement properties were extracted and synthesized. The pooled estimates for the indices of test-retest reliability, standard error of measurement, responsiveness, MDC, and MCID were calculated from the included studies.
Results: A total of 34 articles were deemed eligible and included in this review. The methodological quality of these 34 studies was generally good. Most studies suggested that the CTQ and its scales had good test-retest reliability and internal consistency. However, few studies found that the Symptom Severity Scale had more than one factor. The responsiveness of the CTQ and its scales was excellent across the studies. The pooled estimates for the MDC_{90} and MCID for Symptom Severity Scale/Functional Status Scale were 0.72/0.79 and 1.05/1.13, respectively.
Discussion: The results of this review support the use of CTQ and its scales in assessing conditions-specific impairments in individuals with CTS or after CTR. However, an effort should be made to review and modify the content of the symptom severity scale due to multiple reports challenging its unidimensional structure.
Conclusions: The totality of evidence emerging from this systematic review suggests that the CTQ and its scales provide reliable and valid estimate of impairments resulting from CTS or after CTR.

© 2020 Hanley & Belfus, an imprint of Elsevier Inc. All rights reserved.

Introduction

Carpal tunnel syndrome (CTS) is one of the most common peripheral nerve injuries resulting from compression of the median nerve in the carpal tunnel affecting up to 7.8% of the working population in the United States.[1] Individuals suffering from the CTS report a unique set of symptoms that include burning, tingling, and numbness over the thumb, index, and middle fingers of the affected hand along with muscle weakness. The symptoms of CTS lead to significant personal, societal, and health care burden with the average time lost from work being 27 days per claimant.[1]

Measures that provide a valid assessment of impairments resulting from CTS can provide the foundation for clinical decision making while managing such patients in hand therapy practice. Like any

other musculoskeletal condition, a myriad of self-reported and objective measures are purported for assessing impairments related to CTS. Measurement properties of performance-based measures such as grip strength[2] and Jebsen-Taylor Hand Function Test[3] have been assessed, and these measures have been found to be extremely useful in patients with CTS. On the other hand, self-reported outcome measures (SROMs) such as the Patient Evaluation Measure (PEM),[4] the Disabilities of Arm, Shoulder, and Hand (DASH),[5] Michigan Hand Questionnaire,[6] Historical-Objective Scale (Hi-Ob scale),[7] and Upper Extremity Function Scale (UEFS),[8] are believed to provide valid subjective reporting of impairments resulting from CTS.

The Carpal Tunnel Questionnaire (CTQ) is a condition-specific SROM primarily developed to assess the specific impairments experienced by patients with CTS.[9] Since its inception, the CTQ has often been used as a reference standard when validating other measures in the CTS population.[4,5,7] The CTQ assesses CTS-specific impairments across the domains of the Internal Classification of Functioning, Disability and Health through its two scales: an 11-item Symptom Severity Scale (SSS) (body function) and 8-item Functional Status Scale (FSS) (activity and participation). Each item, across both the scales, has 5 possible response options ranging from 1 to 5 with 1 being no concerns and 5 being the worst status. Individual studies have assessed the measurement properties such as the reliability, validity, and responsiveness as well as comparative advantages of using the CTQ over other measures in patients with CTS.[9,10] However, a rigorously conducted systematic review of the existing literature that outlines the usefulness of the CTQ in different clinical contexts can serve as a useful evidence-based summary for hand therapists while using the CTQ in their practice.

In the past, reviews[11,12] as well as a review of reviews[13] have been conducted for examining the evidence of the measurement properties of the CTQ. While being useful sources of information, the primary literature from which the clinical recommendations were provided in these reviews did not undergo critical appraisal to examine the potential risk of bias in their methodology. Leite et al[11] as well as Changulani et al[12] provided a balanced summary of the results on using the CTQ in clinical practice but did not weigh the quality of the literature included in the review. This raises a concern about the clinical recommendations emanating from such reviews and their applicability to guide clinical practice. Critical appraisal of primary studies can facilitate a balanced overview of the evidence from these studies. A review that adopts a systematic approach in locating the evidence, uses standardized tools in appraising the evidence, and provides clinically meaningful interpretation of the evidence is highly desirable in addressing the knowledge-to-action gap related to using the CTQ in hand therapy practice.

Purpose of the study

This study systematically located the literature examining the measurement properties of the CTQ, evaluated the methodological quality of this literature, and provided recommendations for using the CTQ in hand therapy practice. Specifically, the review summarized the literature to provide an overview of reliability, internal consistency, validity, and responsiveness of the CTQ in different clinical contexts. The study also provided the estimates of standard error of measurement (SEM), minimal detectable change (MDC), and minimal clinically important difference (MCID) in help interpreting the scores of the CTQ in patients with CTS.

Material and methods

Eligibility criteria

Studies published in English where at least one measurement property was assessed for the CTQ or its scales were included in this review. No search restrictions were placed related to publication date of the studies. Studies where the CTQ was used as an outcome measure to examine the impairment were excluded.

Literature search

Four databases (PubMed, CINAHL, PsychInfo, and ProQuest for Conference papers) were searched in January 2019 to locate the studies published on the measurement properties of the CTQ using a predefined list of keywords. The search strategy used the following combination of keywords: (Boston Carpal Tunnel Questionnaire, CTQ, Carpal Tunnel Questionnaire), (psychometric properties, reliability, validity, responsiveness, clinically important change, minimal detectable change, Rasch analysis, cross-cultural adaptation). Bibliography of the primary studies included in the review and previous reviews was also screened for locating potentially eligible studies.

Study selection

First, the titles of all the articles retrieved from the search process were reviewed and studies that did not appear to meet the inclusion criteria were removed. Second, the abstracts of the published studies were reviewed to further determine the eligibility of the studies for this review. Finally, the remaining studies underwent full-text review, and the final list of studies that were eligible for this review was prepared.

Data collection

The data for each measurement property were synthesized and compiled from the included studies. The pooled estimates for test-retest reliability (intraclass correlation coefficient [ICC]), measurement error [SEM], responsiveness [effect size {ES} and standardized response means {SRM}], true change [MDC at 90% confidence level {MDC_{90}}]), and patient-important change (MCID) were derived from these studies. If the values for MDC were not provided, they were calculated using the SEM values (MDC90 = SEM * the square root of 2 * 1.65; MDC 95 = SEM * the square root of 2 * 1.96). If the SEM values were not provided, they were calculated using the values for the ICC (SEM = s*the square root of $1-r$, where r is the ICC value of the scale and s is the standard deviation [SD] of the measurement made on the first occasion).

To derive pooled estimates, the studies were weighted based on their sample sizes in that studies with larger sample were weighted higher. For example, the ICC for the test-retest reliability of the CTQ was reported in seven studies. First, the total sample size for which the ICC was reported for CTQ and its scales was calculated by summing the sample sizes for these seven studies. Subsequently, each study was assigned a percentage weight based on its sample size considering the total sample across all seven studies as 100%. This approach has been adopted in the past in similar reviews.[14,15]

Assessment of risk of bias

Two independent reviewers (SPM and GWZ) completed the critical appraisals of the primary studies. The methodological quality of the studies included in this review was assessed using a standardized evaluation appraisal checklist developed.[16] The appraisal checklist has 12 criteria, each with possible score between 0 and 2, to quantitatively assess the methodological quality of the studies. The total appraisal score can range from 0 to 24, with 0 designating the lowest quality study. The scores of 0 to 24 were normalized to a range of 0% to 100% with 100% indicating the best quality study. The disagreements that arose between the two reviewers during critical appraisal were resolved by achieving consensus through discussions.

The agreement between the reviewers in conducting appraisal was examined using the weighted kappa, where kappa value of ≥0.8 was considered to be strong agreement.[17]

Results

Figure 1 illustrates the flow diagram highlighting the results of the search process, screening the search results to locate the eligible studies, and the summary of appraisal for the included studies. A preliminary search retrieved one hundred thirty-three articles of which four were duplicate citations. Of the remaining 129 citations that underwent the review of their title and abstract, 39 articles were considered for the full-text review. Following the full-text review, 3 were excluded because they assessed measurement properties for other outcome measures and not CTQ[18-20] and 2 studies were excluded because they were published in other languages.[21,22] This resulted in 34 articles that met all the eligibility criteria and were selected for this review.[5,9,10,23-53] Table 1 illustrates the characteristics of the studies included in this review.

Table 2 illustrates the results of the critical appraisal in rank order. The quality of the studies ranged from 38% to 88%, with 14 articles receiving a score greater than 70%. Failure to perform sample size estimation or rationalization, incomplete documentation of study protocol raising concerns about the bias in the administration of test and measures, and narrow of scope of psychometric properties assessed were the major methodological concerns among the studies. The unweighted kappa between the two raters in assessing the methodological quality of the studies was 0.84 (95% confidence interval 0.79-0.89).

Cross-cultural adaptation

Several studies had assessed the psychometric properties of the CTQ or its scales in other languages such as Swedish,[26] Spanish,[45,46]

Hong Kong Chinese,[32] Chinese,[41] Japanese,[34,48] Korean,[35,40,44] Greek,[51] Persian,[52] Polish,[53] and Turkish.[19,47] In general, the versions of the CTQ in these languages were created by adhering to the suggested guidelines for performing cross-cultural adaptation and translation of the SROM.[54] De Smet et al[31] assessed the concurrent validity of the Dutch version of the DASH against the CTQ in patients with CTS. However, it was not clear whether they used the Dutch version of the CTQ for this purpose. Atroshi et al[26] assessed the psychometric properties of the Swedish version of the CTQ with 14 additional items added to the original version of the CTQ to determine the improvement after the carpal tunnel release (CTR). Of the 14 new items, 2 were related to pain and functional limitations due to palmar pain, 8 inquired about the level of satisfaction with recovery in different CTS-related symptoms, and the other 4 asked patients to comment on satisfaction with the CTR, improvement in quality of life after CTR, and whether they would undergo CTR in future or would recommend it to close friend.[26]

Kim et al[40] claimed to perform the translation of the CTQ into Korean for the first time. The results of their adaptation process showed that the item "opening of jar" from the FSS of the CTQ was culturally not relevant to the target population; therefore, it was changed to "opening of screw-topped bottles" in the Korean version of the CTQ. Also, the item "household chores" was changed to "household cleaning" in the Korean version because no equivalent word for "chores" could be found in the Korean language.[40] Interestingly, Park et al[44] also claimed to have translated the CTQ into Korean for the first time but did not report modifications to any of the items of the CTQ.

Administrative burden

Two studies reported the time to complete the CTQ. Imaeda et al[34] reported that participants took an average of 5 min and 40 s

Located Citations (n = 133)
PubMed, n = 114
CINAHL, n = 13
ProQuest, n = 4
PsychInfo, n = 2

Removal of Duplicates
After removal of exact duplicates, n = 129

Included for Full Text Review, n = 39

Excluded at Title/Abstract Review
Not adequate data or not relevant n = 90

Included
Full-text critical appraisal, n = 34

Excluded at Full Text Review, n = 5
Articles published in other language, n = 2
Measurement properties assessed for a different measure, n = 3

Quality summary of appraised primary papers, n = 34
Poor (<50%): n = 7 (20.6%)
Good (51%-70%): n = 13 (38.2%)
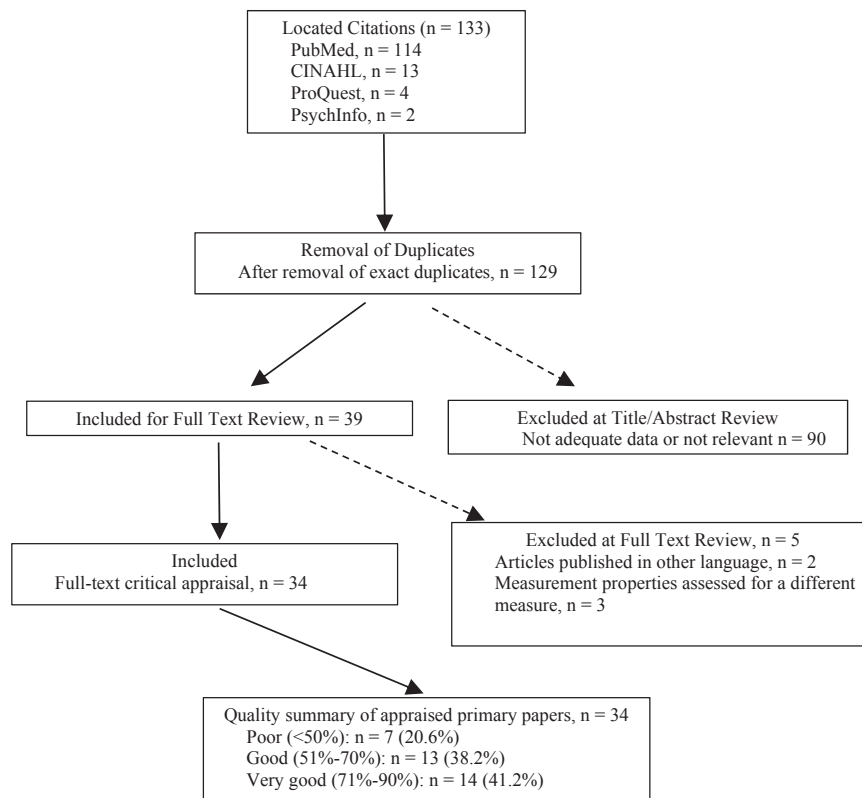Very good (71%-90%): n = 14 (41.2%)

Figure 1. Flow diagram illustrating the search process.

Table 1
Characteristics of the studies included in this review

Study	Population and measurements	Measurement properties examined	Study results
Amadio et al 1996[23]	22 patients (12 women and 9 men; mean age 60 [range 33-80] years) completed the CTQ, SF-36, and AIMS2 a day before CTR and 3 months after the surgery	Concurrent validity Responsiveness	FSS scores showed high correlations and therefore convergent validity with AIMS2, SF-36—physical role, grip, and pinch strength; SSS and FSS scores were the most responsive to change compared to the other self-report or objective measures
Amirjani et al 2011[24]	190 patients with CTS (141 women and 49 men; between 20 and 86 years of age) and 122 healthy participants (91 women and 31 men; between 20 and 89 years of age) completed CTQ and Purdue Pegboard Test	Concurrent validity	The CTQ scores showed divergent relationship with the Purdue Pegboard Test with low to moderate correlation coefficients
Atroshi et al 2007[25]	100 patients with CTS (141 women and 49 men; mean age 52 ± 15 years) completed used CTQ and SF-36	Concurrent validity Responsiveness	Both SSS and FSS showed superior responsiveness compared to SF-36 scales; CTQ scales had moderate relationships with health utility index of SF-6D
Atroshi et al 1998[26]	102 patients with CTS (67 women and 35 men; mean age 52 [range 21-88] years) completed Swedish version of CTQ and SF-36	Test-retest reliability Concurrent validity Responsiveness	Both the SSS and FSS showed acceptable reliability and concurrent relationships between SF-36 scales that assess similar constructs (eg, physical role); both SSS and FSS had high responsiveness indices compared to SF-36 scales with an exception of SF-36 bodily pain scale
Atroshi et al 2009[27]	213 patients with CTS (145 women and 68 men; mean age of 52 ± 17 and 55 ± 16 years, respectively) of whom 187 were scheduled for CTR-completed 11-item SSS, brief 6-item SSS, and the QuickDASH	Test-retest reliability Internal consistency Concurrent validity	The brief 6-item SSS showed excellent test-retest reliability and internal consistency. The brief SSS also showed excellent concurrent validity in assessing disability when compared with QuickDASH.
Bakhsh et al 2012[28]	48 patients with CTS (37 women and 11 men; mean age of 60 ± 14.46 years for the sample) completed the BQ, DASH, and M²DASH questionnaires before CTR and twice over the 6-8 weeks after CTR.	Validity Reliability Responsiveness	Test-retest reliability of the SSS and FSS was deemed to be good and superior compared to the DASH and M²DASH. The responsiveness of the SSS and FSS was good as assessed by comparing pre- and post-CTR (P < .0001).
Bessette et al 1998[29]	196 patients (69% women and 31% men; mean age 49 ± 15 years) completed CTQ, SF-36, SF-12, and Quality of Life Rating Scale before and 6 months after CTR.	Responsiveness	Condition-specific CTQ had much higher responsiveness compared to generic measures of health status such as SF-36, SF-12, and Quality of Life Rating Scale.
Bougea et al 2017[51]	90 patients (75 women and 15 men; mean age 57.3 ± 13.8 years) were classified on severity level of CTS using electrophysiological grading. All the patients completed the Greek version of CTQ, of which half of patients repeated CTQ within 1 week.	Translation in Greek Test-retest reliability Construct validity	The Greek CTQ has high test-retest reliability and acceptable concurrent validity based on the severity level of the CTS.
Chatterjee et al 2009[30]	42 patients (34 women and 8 men; mean age 59 [range 21-88] years) completed CTQ and MHQ before and 6 months after CTR.	Responsiveness	CTQ as well as SSS and FSS had much higher responsiveness assessed using SRM compared to the MHQ.
De Smet et al 2007[31]	119 patients (98 women and 21 men; mean age 51 [range 26-60] years) completed Dutch version of the CTQ and the DASH preoperatively than the DASH was completed at 1 year postoperatively.	Concurrent validity	The CTQ scales demonstrated expected concurrent relationships with high correlations between the FSS and the DASH and moderate correlations between the SSS and the DASH.
Fok et al 2007[32]	50 patients with CTS (42 women and 8 men; mean age 66 [range 32-81] years) completed the Hong Kong Chinese version of the CTQ before their scheduled appointment with orthopedic specialist and again an hour later.	Translation in Hong Kong Chinese Reliability Internal consistency	The SSS and FSS scales of the CTQ showed acceptable test-retest reliability (assessed using Pearson correlation coefficients) and internal consistency.
Gay et al 2003[33]	42 patients with CTS (24 women and 18 men; mean age 55 ± 13 years) completed the CTQ, DASH, and SF-36 before surgery, and of them, 34 completed these measures again at 6 and 12 weeks after CTR.	Responsiveness	Of the three measures, the CTQ and its scales had superior responsiveness indices compared to the DASH and the SF-36 scales at 6 and 12 weeks after CTR.
Greenslade et al 2004[5]	57 patients with CTS (41 women and 16 men; mean age 59 [range 53-64] years) completed the CTQ and DASH preoperatively and 3 months after CTR. Of these, 31 also completed the CTQ and DASH 14 days after the first data collection session for assessing reliability.	Test-retest reliability Responsiveness	The DASH had slightly better responsiveness compared to the FSS but SSS had higher responsiveness compared to the DASH. The CTQ scales showed good test-retest reliability.
Hassankhani et al 2018[52]	142 patients with CTS (123 women and 19 men; age range [18-70] years) completed Persian CTQ and the Persian QuickDASH. The CTQ was again completed 2 to 6 days later.	Translation into Persian Test-retest reliability Construct validity	The Persian CTQ showed acceptable validity and reliability. Reliability in SSS was affected by confusion in language translation of pain, numbness, and tingling.
Imaeda et al 2007[34]	87 patients with CTS (mean age 58.3 ± 13 years) completed the Japanese versions of the CTQ, DASH, and SF-36. Of these, 72 completed the CTQ 1-2 weeks later and 45 completed the CTQ and DASH 3 months after CTR.	Test-retest reliability Factor structure Concurrent validity Responsiveness	The CTQ scales showed acceptable reliability and expected concurrent relationships with the DASH and SF-36, and superior responsiveness indices compared to the DASH. However, the results suggested that the SSS had two factors versus hypothesized unidimensionality.
Jeon et al 2011[35]	56 patients with CTS (50 women and 6 men; mean age 55 [range 32-77] years) completed the Korean versions of the CTQ and DASH before surgery and 6 months after CTR.	Responsiveness	While the CTQ scales as well as the DASH showed good responsiveness in detecting change after CTR, the CTQ scales had much higher responsiveness indices compared to the DASH.

Table 1 (*continued*)

Study	Population and measurements	Measurement properties examined	Study results
Jerosch-Herold et al 2011[36]	63 patients with CTS (32 women and 31 men; mean age 60.4 ± 13.5 years) completed the CTQ and several clinician-administered tests before surgery and at 4 and 8 months after CTR.	Responsiveness	The results suggested that the SSS had excellent responsiveness and FSS had moderate responsiveness at 4 and 8 months after CTR; both the scales being more responsive compared to the clinician-administered tests.
Katz et al 1994[37]	104 patients (70% women and 30% men; mean age 55 [range 25-87] years) completed subset of questions from SSS and FSS, activities of daily living questionnaire, and surgery satisfaction questionnaire as well as underwent testing for grip and pinch strength, static 2-point discrimination, and pressure sensibility tests before surgery and 6 weeks and 3 months after CTR.	Responsiveness	The results suggested that the SSS and FSS showed excellent responsiveness compared to grip/pinch strength or pressure sensibility tests.
Katz et al 1996[38]	216 patients (women:men ratio provided for those receiving worker's compensation versus those who are not; mean age ranged from 37.9 to 42.7 years for different subgroups) with CTS-completed CTQ, satisfaction questionnaire, and grip at baseline and again at 6-month follow-up.	Internal consistency Concurrent validity	The results suggested that FSS and SS were internally consistent and showed expected concurrent relationships with grip strength (low) and satisfaction questionnaire (moderate).
Kim et al 2013[39]	66 patients (57 women and 9 men; mean age 54 ± 10.4 years) completed the CTQ before surgery and 3 months after CTR.	Minimal clinically important difference	Minimal clinically important differences were defined for the CTQ (0.92), SSS (1.14), FSS (0.74).
Kim et al 2015[40]	53 patients (45 women and 8 men; mean age 59 ± 12 years) completed the Korean versions of the CTQ and the DASH before surgery and 3 months after CTR.	Translated in Korean Reliability Concurrent validity Responsiveness	The Korean version of the CTQ demonstrated excellent reliability (ICC>0.90) and responsiveness indices that were superior to those of the DASH. However, the CTQ scales demonstrated moderate relationships with the DASH.
Levine et al 1993[9]	67 patients with CTS (50 women and 17 men; mean age 57 years [range 19-88 years]) completed the CTQ, grip and pinch strength, and two-point discrimination testing before surgery and at different follow-up interval after CTR.	Reliability Internal consistency Responsiveness	The CTQ scales showed excellent reliability, concurrent relationships with the grip strength, and high responsiveness to change in those who underwent CTR.
Lue et al 2015[50]	123 patients with CTS (104 women and 19 men; mean age 50.1 ± 8.6 years) completed CTQ.	Confirmatory factor analysis	The hypothesized two-factor (symptoms and functions) model had poor overall fit compared to three-factor models. The three-factor model of daytime pain, nocturnal numbness/tingling, and hand function consisting of 11 items had the best overall fit and superior reliability and validity.
Lue et al 2014[41]	99 patients with CTS (85 women and 14 men; mean age 49.5 ± 9.5 years) completed Chinese versions of the CTQ, DASH, and SF-36 and were tested for grip and pinch strength. Of these, 51 patients completed the CTQ within 1 week and 23 completed the CTQ within 3 months after initial assessment.	Acceptability Test-retest reliability Responsiveness Construct validity Minimal detectable change	The Chinese version of the CTQ scales were found to be acceptable, had good reliability, and showed expected concurrent relationships with the DASH and SF-36 scales. The SSS showed excellent responsiveness but the FSS had moderate responsiveness.
Ozer et al 2013[42]	114 patients with CTS, of whom 87 were nondiabetic (58 women and 29 men; mean age 46.7 [range 22-72] years); 27 diabetic patients (21 women and 6 men; mean age 50.9 years [range 33-73]) completed the CTQ before surgery, as well as 3 and 6 months after CTR	Minimal clinically important difference Effect size	The SSS and FSS showed excellent responsiveness in diabetic as well as nondiabetic patients. However, diabetic patients required much larger change in both SSS and FSS to deem themselves "better" compared to nondiabetic patients with CTS.
Ozyurekoglu et al 2006[43]	28 patients with CTS (19 women and 9 men; mean age 38 years [range 28-77]) completed SSS scales of the CTQ and Katz-Stirrat Hand Diagram	Minimal clinically important difference Responsiveness	SSS excellent responsiveness and a change of 1.04 was considered indicative of clinically important difference.
Park et al 2013[44]	54 patients with CTS (50 women and 4 men; mean age 50 years [range 18-65]) completed the Korean versions of CTQ, DASH, and ED-5D at baseline, 2 weeks later, and again at 3 months after local corticosteroid injection.	Construct validity Reliability Responsiveness	SSS and FSS showed excellent test-retest reliability significant, expected correlations between K-CTQ (high) and EQ-5D (low to moderate) but moderate effect sizes for responsiveness.
Rosales et al 2002[45]	42 patients with CTS (36 women and 6 men; mean age 54 [range 34-63] years); completed the Spanish translations of the CTQ and the DASH	Internal consistency Reliability	SSS and FSS both showed good test-retest reliability and internal consistency albeit DASH has superior indices for these measurement properties.
Rosales et al 2009[46]	42 patients with CTS (36 women and 6 men; mean age 54 [range 25-87] years) completed the Spanish versions of the CTQ, DASH, SF-36, grip and pinch strength	Responsiveness	SSS and FSS demonstrated high responsiveness compared to the DASH and SF-36.
Sezgin et al 2006[47]	67 patients with CTS (62 women and 5 men; mean age 49.8 + 8.1 years) completed Turkish version of the CTQ, SF-36, Visual Analog Scale for pain, and pinch and grip strength measures and went on to complete CTQ again within 7 days.	Reliability Internal consistency Construct validity	The SSS and FSS showed acceptable reliability and internal consistency. The SSS and FSS showed expected concurrent validity as shown by low to moderate correlations with pain scale and different subscales of the SF-36.
Trybus et al 2019[53]	218 patients with CTS (173 women and 45 men; mean age 56.8 ± 13.7 years) completed the Polish versions of the CTQ, DASH, and MHQ. After 14 days, 189 patients completed CTQ again.	Translated to Polish Test-retest reliability Construct validity and reliability	The Polish CTQ exhibited excellent test-retest reliability, internal consistency. Concurrent relationships between the CTQ with the DASH and MHQ were high.

Table 1 (continued)

Study	Population and measurements	Measurement properties examined	Study results
Uchiyama et al 2007[48]	60 patients with CTS (49 women and 11 men; mean age 60 [range 21-86] years) completed the Japanese versions of the CTQ, DASH, and SF-36 and also had grip and pinch strength assessments before surgery and within 3 months after CTR.	Responsiveness	The SSS and FSS showed much higher responsiveness compared to the DASH, SF-36, or grip/pinch strength.
Upatham et al 2008[49]	31 patients with CTS (4 male and 27 female; ranging between 30 and 70 years of age) completed Thai version of the CTQ.	Internal consistency	FSS and SSS had high internal consistency.
Zyluk et al 2011[10]	50 patients with (43 women and 7 men; mean age 56 [range 34-84] years) completed the CTQ, DASH, and PEM and also underwent testing for grip, key-pinch, and filament test before surgery and 1 and 6 months after CTR.	Internal consistency Responsiveness Validity	The CTQ, DASH, and PEM exhibited high internal consistency. The CTQ and PEM demonstrated highest responsiveness to change compared to the DASH at one-month and six-month assessments. The CTQ showed expected concurrent relationships with the DASH and PEM as well as grip and pinch strength assessments.

SF-36 = Short-Form 36; AIMS2 = Arthritis Impact Measurement Scale; CTR = carpal tunnel release; CTQ = Carpal Tunnel Questionnaire; SSS = Symptom Severity Scale; FSS = Functional Status Scale; DASH = Disabilities of Arm, Shoulder, and Hand; QuickDASH = shorter version of the DASH; M²DASH = Manchester Modified DASH; SRM = standardized response means.

(range 2-30 min) to complete the Japanese-CTQ, whereas Greenslade et al[5] reported this time to be 5.6 (±3.5) minutes for the English CTQ. No study reported time to complete specific subscales of SSS and FSS.

Reliability and internal consistency

Table 3 illustrates the summary of the test-retest and absolute reliability as well as internal consistency of the CTQ and its subscales in individuals with CTS or those who had CTR. For the test-retest reliability, the data have only been summarized where it

Table 2
Methodological quality of the studies included in this review

Study	1	2	3	4	5	6	7	8	9	10	11	12	Total (%)
Atroshi et al 2009[27]	2	2	2	2	1	1	2	2	2	2	1	2	88
Lue et al 2014[41]	2	2	2	2	0	2	2	1	2	2	2	2	88
Katz et al 1994[37]	2	2	1	1	1	2	2	2	2	1	1	2	79
Lue et al 2015[50]	2	2	2	1	1	0	2	1	2	2	2	2	79
Rosales et al 2009[46]	2	2	2	1	0	2	2	2	2	1	1	2	79
Amirjani et al 2011[24]	2	2	2	1	1	0	2	2	2	1	1	2	75
Bakhsh et al 2012[28]	2	2	1	2	0	2	1	1	2	2	1	2	75
Trybus et al 2018[53]	1	2	1	2	1	2	2	1	2	1	1	2	75
Bougea et al 2017[51]	2	2	2	2	0	0	1	1	2	2	1	2	71
Greenslade et al 2004[5]	2	2	2	1	1	1	1	2	1	1	1	2	71
Jerosch-Herold et al 2011[36]	2	1	1	1	1	1	2	2	2	1	1	2	71
Katz et al 1996[38]	2	1	0	1	1	1	2	2	2	2	1	2	71
Levine et al 1993[9]	2	1	1	1	0	2	2	2	2	1	1	2	71
Sezgin et al 2006[47]	2	2	0	1	0	2	2	2	1	1	1	2	71
Kim et al 2015[40]	1	1	1	2	0	1	1	2	2	2	1	2	67
Uchiyama et al 2007[48]	1	1	1	1	0	2	2	2	2	1	1	2	67
Atroshi et al 1998[26]	1	1	0	1	1	2	2	1	2	1	2	1	63
Bessette et al 1998[29]	1	2	1	1	1	1	2	1	2	1	1	1	63
Imaeda et al 2007[34]	2	1	0	2	0	2	2	1	2	1	1	2	63
Kim et al 2013[39]	2	1	1	1	0	1	2	1	2	1	1	2	63
Park et al 2013[44]	2	2	0	1	0	2	2	1	1	1	1	2	63
Rosales et al 2002[45]	2	1	1	1	0	2	2	1	1	1	1	2	63
Chatterjee et al 2009[30]	2	2	0	1	2	0	2	1	1	1	1	1	58
Gay et al 2003[33]	2	2	1	1	0	1	1	1	2	1	1	1	58
Ozer et al 2013[42]	2	2	0	0	1	2	2	1	1	1	1	1	58
Upatham et al 2008[49]	2	1	0	1	0	2	1	2	1	1	1	2	58
Ozyurekoglu et al 2006[43]	2	2	0	1	0	2	1	1	1	1	1	1	54
Atroshi et al 2007[25]	1	1	1	0	0	2	2	1	1	1	1	1	50
Hassankhani et al 2018[52]	1	1	1	1	1	0	1	1	1	2	1	1	50
Jeon et al 2011[35]	2	1	1	1	0	1	1	1	1	1	1	2	50
Zyluk et al 2011[10]	2	0	0	1	0	1	1	1	1	1	1	2	46
Fok et al 2007[32]	1	1	0	1	0	2	1	1	0	1	1	1	42
Amadio et al 1996[23]	1	1	1	0	0	0	1	1	1	1	1	1	38
De Smet et al 2007[31]	0	1	0	0	1	1	1	1	1	1	1	1	38

was examined using ICC. The subscales of the CTQ showed good test-retest reliability with ICC values ranging from 0.75 to 0.95 for the SSS and 0.79 to 0.94 for the FSS when reassessments were conducted either in a short term (≤7 days) or somewhat longer term (>7 days).[28,34,40,41,44,51-53]

Atroshi et al[26] assessed the reliability of the Swedish version of the CTQ by assessing differences in the mean scores of SSS and FSS obtained at an interval of 14 days. These differences were not significant suggesting adequate reliability of these subscales.[26] The early research by Levine et al[9] that proposed CTQ as a disease-specific measure assessed the reliability using Pearson correlation coefficient (PCC) and found that the PCC values were 0.91 and 0.93 for SSS and FSS, respectively, indicating excellent reproducibility.

SEM values for the SSS and FSS were 0.31 and 0.27, respectively, for the Chinese version of the CTQ.[41] SEM values were calculated from the ICC values for the SSS and FSS provided in six studies.[28,34,40,41,51,53]

Internal consistency for the CTQ or its scales was examined using Cronbach's alpha (CA) in most studies. The internal consistency for the CTQ and its scales was deemed to be high with CA values ≥0.80 in individuals treated conservatively for CTS[32,34,44,47,49,51-53] or those who underwent CTR.[9,10,26,28,38,40,45]

Validity

Table 4 shows the data for validity of the CTQ and its scales across the published literature. A summary of these data is provided below.

Content validity

Levine et al[9] developed the content of the English version of the CTQ by consulting a group of hand surgeons, rheumatologists, and patients. Several studies indicated that the SSS has two or more factors. Lue et al[50] found that the factor structure of SSS was unstable in that 3 items ("Do you have numbness in hand", "How severe is numbness or tingling at night", and "How often did hand numbness or tingling wake you up during a typical night in the last two weeks?") had poor factor loadings. They proposed a two-factor model (nocturnal numbness/tingling and daytime pain) consisting of 11 items in total for improved precision of measurement. Similarly, the SSS was found to have three-factor structure (nighttime symptoms, daytime pain symptoms, and functions) in another study, prompting the researchers to propose a brief 6-item symptom scale.[27] Finally, the Japanese version of the SSS also showed two factors.[34] Ceiling and floor effects were also examined to ensure the content validity of the Japanese,[34] Korean,[40] Chinese,[41]

Table 3
Summary of literature for the reliability and internal consistency of the CTQ scores

Type of reliability	Data extracted from the included studies
Test-retest reliability (ICC)	
Short term (1-7 days)	
	0.92 in patients requesting CTR for Korean SSS[40]
	0.94 in patients requesting CTR for Korean FSS[40]
	0.93 in patients requesting CTR for total score of Korean CTQ[40]
	0.81 in patients with CTS for Chinese SSS[41]
	0.83 in patients with CTS for Chinese FSS[41]
	0.75 in patients with CTS for Greek SSS[51]
	0.79 in patients with CTS for Greek FSS[51]
	0.54 in patients with CTS for Persian SSS[52]
	0.77 in patients with CTS for Persian FSS[52]
Moderate/long term (after 7 days)	
	0.97 in patients with CTS for total score of CTQ[24]
	0.93 in patients with CTS for Korean SSS[44]
	0.84 in patients with CTS for Korean FSS[44]
	0.85 in patients with CTS for Polish SSS[53]
	0.87 in patients with CTS for Polish FSS[53]
	0.82 in patients with CTS for Japanese SSS[34]
	0.83 in patients with CTS for Japanese FSS[34]
Undefined reassessment period	
	0.95 in patients with CTR for SSS[28]
	0.92 in patients with CTR for FSS[28]
Standard error of measurement	
	0.31 in patients with CTS for Chinese SSS[41]
	0.27 in patients with CTS for Chinese FSS[41]
	0.15 in patients with CTR for SSS[28]
	0.26 in patients with CTR for FSS[28]
	0.28 in patients with CTR for Korean SSS[40]
	0.24 in patients with CTR for Korean FSS[40]
	0.31 in patients with CTS for Japanese SSS[34]
	0.33 in patients with CTS for Japanese FSS[34]
	0.25 in patients with CTS for Korean SSS[44]
	0.49 in patients with CTS for Korean FSS[44]
	0.39 in patients with CTS for Greek SSS[51]
	0.39 in patients with CTS for Greek FSS[51]
	0.32 in patients with CTS for Polish SSS[53]
	0.34 in patients with CTS for Polish FSS[53]
Internal consistency (Cronbach's alpha)	
	0.93 in patients with CTS for Korean SSS[44]
	0.95 in patients with CTS for Korean FSS[44] 0.84 in patients with CTS for Japanese SSS[34]
	0.90 in patients with CTS for Japanese FSS[34]
	0.84 in patients with CTS for Hong Kong Chinese SSS[32]
	0.90 in patients with CTS for Hong Kong Chinese FSS[32]
	0.82 in patients with CTS for Turkish SSS[47]
	0.88 in patients with CTS for Turkish FSS[47]
	0.82-0.86 in patients with CTS for Thai SSS[49]
	0.81-0.84 in patients with CTS for Thai CTQ-FSS[49]
	0.96 in patients with CTR for SSS[28]
	0.91 in patients with CTR for FSS[28]
	0.94 in patients with CTR for total scores of the CTQ[10]
	0.92 in patients 3 months after CTR for Swedish SSS[26]
	0.93 in patients 3 months after CTR for Swedish FSS[26]
	0.89 in patients requesting CTR for Korean SSS[40]
	0.90 in patients requesting CTR for Korean FSS[40]
	0.90 in patients requesting CTR for Korean CTQ[40]
	0.90 in patients with CTS for Spanish SSS[45]
	0.91 in patients with CTS for Spanish FSS[45]
	0.89 in patients with CTS SSS[38]
	0.89 in patients with CTS FSS[38]
	0.89 in mixed sample of those with CTS/CTR for SSS[9]
	0.91 in those with CTS/CTR for FSS[9]

CTQ = Carpal Tunnel Questionnaire; CTR = carpal tunnel release; CTS = carpal tunnel syndrome; FSS = Functional Status Scale; ICC = intraclass correlation coefficient; SSS = Symptom Severity Scale.

and Polish[53] versions of the CTQ. Of 87 individuals with CTS recruited in the study to validate Japanese version of CTQ, 1 and 5 individuals had ceiling effect on SSS and FSS, respectively, with the score of 0.[34] The Chinese version showed floor effect (24.2%) in FSS in individuals with CTS.[41] The Korean and Polish versions did not show ceiling/floor effects.

Construct validity

The convergent validity of the CTQ and its scales was assessed by examining their concurrent relationships with self-reported as well as objective performance measures across published literature. In general, the CTQ scales showed expected convergent or divergent relationships with measures assessing similar constructs. The FSS consistently showed high correlations ($r > 0.70$) with the DASH in patients with CTS who were managed conservatively or had undergone CTR,[28,40,41,44,52,53] whereas the relationships between the SSS and the DASH were moderate to high ($r > 0.60$) in these patients.[28,40,41,44,52,53] Both SSS and FSS showed low correlations ($r \leq 0.40$) with other self-reports that examine diverse constructs such as mental health or emotional domains of Short-Form 36 (SF-36),[25,26,34,41,47] EuroQol,[44] or Visual Analogue Scale for pain.[34,47] The CTQ scales also showed low correlations ($r \leq 0.40$) with most measures assessing wrist/hand performance. For example, both SSS and FSS had low correlations with grip strength (r values between 0.29 and 0.38) in patients with CTS.[9,38,41,47] Both scales had low correlations with pinch strength (r values between 0.15 and 0.26) with an exception of the inception study of the CTQ by Levine et al[9] where SSS showed moderate correlations with the FSS ($r = 0.60$). The correlations of the SSS and FSS were low with other performance measures such as two-point discrimination (r values of 0.15 and 0.42),[9] Semmes Weinstein Monofilament testing (r values of 0.17 and 0.24),[9] or Purdue Pegboard test (r values of −0.10 and −0.45).[24]

Known-group validity

Amadio et al[23] assessed known-group validity of the SSS and FSS by comparing their presurgery and 3 months after CTR. These within-group comparisons revealed significant differences in the scores for both these scales ($P < .01$). The score of 1.95 on the Greek version of SSS discriminated those patients with CTS with low grading versus those with high grading on electrophysiological studies with sensitivity/specificity of 75.5%/68.3%.[51]

Responsiveness

The responsiveness of the CTQ and its scales assessed using ES and/or SRM is synthesized in Table 5. The responsiveness of the CTQ was examined in the published literature in patients who underwent CTR. Most studies supported the responsiveness for the total scores of CTQ in patients who underwent CTR with ES/SRM exceeding 0.8 with assessments conducted before surgery to as little as 1-month follow-up[10] to up to 1-year follow-up after CTR.[31] In particular, the ES for the total score of the CTQ ranged between 0.92 and 1.98[10,29,33,40] and the SRM ranged between 1.10 and 1.56.[10,29,30,33,40] The responsiveness of the SSS was also found to be good as shown by the ES/SRM consistently >0.08 (the highest being 2.4) across 17 studies with assessments conducted before surgery and up to 1-year follow-up interval.[5,23,25,26,29-31,33-36,40-42,46,48] The reported ES and SRM for the FSS varied widely. Although many studies deemed the FSS to be highly responsive with ES/SRM values between 0.80 to 1.26,[25,26,33,35,37,42] several found FSS to be less responsive with low ES/SRM low (<0.80).[31,33,34,40,41]

MDC and MCID

Table 6 shows the extracted information concerning the MDC and MCID for the CTQ and its subscales across the published literature. Kim et al[39] found that the minimal meaningful change 3 months after CTR was 0.92 for the total score, 1.14 for the SSS subscale, and 0.74 for the FSS subscale. In addition, Ozer et al[42] examined MCID for the CTQ subscales in individuals with or without diabetes who had underwent CTR. The results showed

Table 4
Summary of literature for the validity of the CTQ scores

Type of validity	Data extracted from the included studies
Content validity	
	The Japanese[34] and Chinese[50] versions of the version of the SSS were found to have 2 factors, whereas FSS was deemed to be unidimensional.
	Marked item redundancy observed on the item response analysis and SSS was found to have three factors resulting in authors suggesting 6-item SSS[27]
Convergent or divergent construct validity (PCC or Spearman rank correlation)	
With DASH/QuickDASH	High correlation ($r = 0.77$) with SSS in patients with CTS[28]
	High correlation ($r = 0.87$) with FSS in patients with CTS[28]
	Moderate correlation ($r = 0.63$) for Chinese SSS in patients with CTS[41]
	High correlation ($r = 0.75$) for Chinese FSS in patients with CTS[41]
	Moderate correlation ($r = 0.54$) for Japanese SSS in patients with CTS[34]
	High correlation ($r = 0.80$) for Japanese FSS in patients with CTS[34]
	High correlations for Korean SSS in patients with CTS before ($r = 0.87$) and after ($r = 0.80$) receiving corticosteroid injections[44]
	High correlations for Korean FSS in patients with CTS before ($r = 0.87$) and after ($r = 84$) receiving corticosteroid injections[44]
	Moderate correlations ($r = 0.61$) for Korean SSS in patients requesting CTR[40]
	Moderate correlations ($r = 0.67$) for Korean FSS in patients requesting CTR[40]
	Moderate correlations ($r = 0.67$) for the total scores of the Korean CTQ in patients requesting CTR[40]
	Moderate correlation ($r = 0.64$) for Persian SSS with QuickDASH in patients with CTS[52]
	High ($r = 0.70$) for Persian FSS with QuickDASH in patients with CTS[52]
	High correlation ($r = 0.71$) for Polish SSS with DASH in patients with CTS[53]
	High ($r = 0.81$) for Polish FSS with DASH in patients with CTS[53]
With MHQ	High correlations for Polish SSS (-0.72) and FSS (-0.76) in patients with CTS[53]
With SF-36	Correlations ranging from as low as $r = 0.20$ with mental health scale to as high as $r = 0.64$ for bodily pain scale of the SF-36 for Swedish SSS in patients with CTS scheduled for CTR[26]
	Correlations ranging from as low as $r = 0.29$ with emotional role scale to as high as $r = 0.70$ for physical role scale of the SF-36 for Swedish FSS in patients with CTS scheduled for CTR[26]
	Correlations ranging from as low as $r = 0.33$ with physical functioning scale to as high as $r = 0.64$ for bodily pain scale of the SF-36 for Swedish SSS in patients after CTR[26]
	Correlations ranging from as low as $r = 0.34$ with mental health scale to as high as $r = 0.67$ for bodily pain scale of the SF-36 for Swedish FSS in patients after CTR[26]
	Correlations ranging from as low as $r = -0.23$ with physical function scale to as high as $r = -0.66$ for bodily pain scale of the SF-36 for Japanese SSS in patients with CTS[34]
	Correlations ranging from as low as $r = -0.19$ with physical function scale to as high as $r = -0.63$ for bodily pain scale of the SF-36 for Japanese FSS in patients with CTS[34]
	Moderate correlations with physical functioning scale for Turkish SSS ($r = -0.55$) and FSS ($r = -0.54$) in patients with CTS[47]
	Moderate correlations with physical functioning scale for Chinese FSS ($r = -0.48$) in patients with CTS[41]
	Moderate correlations with physical role scale for Turkish SSS ($r = -0.54$) and FSS ($r = -0.40$) in patients with CTS[47]
	Moderate correlation with emotional role scale for Turkish SSS ($r = -0.40$) and low correlation for FSS ($r = -0.29$) in patients with CTS[47]
	Low correlations with mental health scale for Chinese SSS ($r = -0.28$) and FSS ($r = -0.22$) in patients with CTS[41]
	Moderate correlations with bodily pain scale for Turkish SSS ($r = -0.63$) and FSS ($r = -0.44$) in patients with CTS[47]
	High correlation with bodily pain scale for Chinese SSS ($r = 0.72$) in patients with CTS[41]
With VAS	Moderate correlations with Japanese SSS ($r = 0.40$) and low correlation for FSS ($r = 0.23$) in patients with CTS[34]
	Moderate correlation for Turkish SSS ($r = 0.51$) and low correlation for FSS ($r = 0.38$) in patients with CTS[47]
With EQ-5D	Moderate correlation for Korean SSS in patients with CTS before ($r = -0.64$) but high correlation after ($r = -0.70$) receiving corticosteroid injections[44]
	Moderate correlations for Korean FSS in patients with CTS before ($r = -0.49$) and after ($r = -0.43$) receiving corticosteroid injections[44]
With grip strength	Poor correlation for SSS ($r = -0.38$) but moderate correlation for FSS ($r = -0.50$) in patients with CTS[9]
	Poor correlations in patients with CTS receiving ($r = -0.32$) or not receiving ($r = -0.30$) workers' compensation[38]
	Poor correlations for Chinese FSS (-0.35) in patients with CTS[41]
	Poor correlations for Turkish SSS ($r = -0.29$) and FSS ($r = -0.36$) in patients with CTS[47]
With pinch strength	Moderate correlations for SSS ($r = -0.47$) and for FSS ($r = -0.60$) in patients with CTS[9]
	Poor correlations for Turkish SSS ($r = -0.226$) and FSS ($r = -0.15$) in patients with CTS[47]
	Poor correlations for Chinese FSS (-0.22) in patients with CTS[41]
Known-group validity	
	SSS (mean difference of 1.33 ± 0.76; $P < .01$) and FSS (mean difference of 1.33 ± 0.76; $P < .01$) were significantly discriminative of the subgroups of patients before CTR and 3 months after CTR[23]
	Greek SSS score of 1.95 discriminated those CTS patients with low grading versus those with high grading on electrophysiological studies with sensitivity/specificity of 75.5%/68.3%[51]

CTQ = Carpal Tunnel Questionnaire; CTR = carpal tunnel release; CTS = carpal tunnel syndrome; FSS = Functional Status Scale; DASH = Disabilities of Arm, Shoulder, and Hand; EQ-5D = MHQ, Michigan Hand Questionnaire; PCC = Pearson Correlation Coefficient; SF-36 = Short-Form 36; SSS = Symptom Severity Scale; VAS = Visual Analogue Scale for pain.

that the MCID for SSS/FSS was 1.45/1.95 and 1.55/2.05 for 3-month and 6-month follow-up, respectively, in individuals with diabetes who had undergone CTR.[42] Similarly, MCID for SSS/FSS was 0.8/1.25 and 1.6/1.45 for 3-months and 6-month follow-up, respectively, in individuals without diabetes who had undergone CTR.[42]

Pooled estimates of measurement properties

Table 7 shows the pooled estimates of indices for test-retest reliability (ICC), standard error of measurement, responsiveness (ES and SRM), true change (MDC_{90} and MDC_{95}), and clinical meaningful change (MCID). The pooled estimates showed that the

Table 5
Summary of literature for the responsiveness of the CTQ scores

Measure of responsiveness	Data extracted from the included studies
Effect size	
Large (≥0.8)	Total score of CTQ
	1.98 for 0- to 3-month assessment interval after CTR[29]
	1.3 and 1.71 for assessments conducted at 0-6 weeks and 0-12-weeks after CTR[33]
	1.02 for 0- to 3-month assessment interval after CTR[40]
	0.92 for 0- to 1-month reassessment interval and 1.29 for 0- to 1-month reassessment interval after CTR[10]
	SSS
	2.4 for 0- to 3-month assessment interval after CTR[25]
	2.1 for Swedish version for 0- to 3-month assessment interval after CTR[26]
	2.03 for 0- to 3-months assessment interval after CTR[37]
	1.07 for Korean version for 0- to 3-month assessment interval after CTR[40]
	1.74 and 1.96 for assessments conducted at 0-6 weeks and 0-12 weeks after CTR[33]
	1.41 for Spanish version for 0- to 12-weeks assessment interval after CTR[46]
	1.4 for Korean version for 0- to 6-month assessment interval after CTR[35]
	1.27 and 1.28 for assessments conducted at 0-4 months and 0-8 months after CTR[36]
	1.12 for assessment intervals that ranged from 0-3 months to 0-9 months after CTR[41]
	1.08 for Japanese version for 0- to 3 month assessment interval after CTR[48]
	0.99 for Japanese version for 0- to 3-month assessment interval after CTR[34]
	>0.8 for 0-3 months and 0- to 6-month assessment intervals after CTR[42]
	FSS
	1.05 for 0- to 6-month assessment interval after CTR[33]
	1.0 for 0- to 3-months assessment interval after CTR[25]
	0.94 for Swedish version for 0- to 3-month assessment interval after CTR[26]
	0.9 for Korean version for 0- to 6-month assessment interval after CTR[35]
	0.86 for 0- to 3-month assessment interval after CTR[37]
	>0.8 for 0-3 months and 0- to 6-month assessment intervals after CTR[42]
Medium (0.4-0.79)	Total score of CTQ
	0.41 for 1- to 6-month assessment interval after CTR[10]
	FSS
	0.77 for Korean version for 0- to 6-month assessment interval after CTR[40]
	0.71 and 0.68 for assessments conducted at 0-4 months and 0-8-months after CTR[36]
	0.7 for Spanish version for 0- to 12-week assessment interval[46]
	0.63 for Japanese version for 0- to 3-month assessment interval after CTR[48]
	0.56 for assessment interval that ranged from 0-3 months to 0-9 months after CTR[41]
	0.48 for 0- to 6-month assessment interval after CTR[33]
Standardized response means	
Large (≥0.8)	Total score of CTQ
	1.56 for 0- to 6-month assessment interval after CTR[29]
	1.26 for 0- to 6-month assessment interval after CTR[10]
	1.22 for 0- to 6-month assessment interval after CTR[30]
	1.21 and 1.66 for assessments conducted at 0-6 weeks and 0-12-weeks after CTR[33]
	1.10 for Korean version for 0- to 3-month assessment interval after CTR[40]
	SSS
	2.02 for Korean version for 0- to 6-month assessment interval after CTR[40]
	1.9 for 0- to 3-month assessment interval after CTR[25]
	1.75 for 0- to 3-month assessment interval after CTR[23]
	1.75 for Spanish version for 0- to 12-week assessment interval[46]
	1.7 for Swedish version for 0- to 12-week assessment interval[26]
	1.67 and 2.01 for assessments conducted at 0-6 weeks and 0-12-weeks after CTR[33]
	1.5 for Korean version for 0- to 6-month assessment interval after CTR[35]
	1.33 for 0- to 6-month assessment interval after CTR[30]
	1.07 for 0- to 3-month assessment interval after CTR[5]
	1.05 for Dutch version for 0- to 1-year assessment interval after CTR[31]
	1.03 for Chinese version for assessment interval that varied between 3 and 9 months after CTR[41]
	1.00 for Japanese version for 0- to 3-month assessment interval after CTR[48]
	0.85 for Japanese version for 0- to 3-month assessment interval after CTR[34]
	FSS
	1.26 for 0- to 3-month assessment interval after CTR[23]
	1.1 for Korean version for 0- to 6-month assessment interval after CTR[40]
	0.98 for 0- to 6-month assessment interval after CTR[30]
	0.96 for Korean version for 0- to 3-month assessment interval after CTR[40]
	0.94 for Swedish version for 0- to 3-month assessment interval after CTR[26]
	0.9 for 0- to 3-month assessment interval after CTR[25]
Medium (0.4-0.79)	Total score of CTQ
	0.74 for 0-1 month and 1- to 6-month assessment intervals after CTR[10]
	FSS
	0.76 for Japanese version for 0- to 3-month assessment intervals after CTR[48]
	0.74 for Korean version for 0- to 3-month assessment interval after CTR[40]
	0.70 for Japanese version for 0- to 3-month assessment intervals after CTR[34]
	0.62 for 0- to 3-month assessment interval after CTR[5]
	0.62 for Chinese version for assessment interval that varied between 3 and 9 months after CTR[41]
	0.51 for Spanish version for 0- to 12-week assessment intervals after CTR[46]
	0.46 for 0- to 6-week assessment interval after CTR[33]

CTQ = Carpal Tunnel Questionnaire; CTR = carpal tunnel release; CTS = carpal tunnel syndrome; FSS = Functional Status Scale; SSS = Symptom Severity Scale.

Table 6
Evidence regarding the minimal detectable and clinically important changes for the CTQ and its scales on the score ranging between 0 and 5 points

Measure of change	Data extracted from the included studies
MDC	
MDC95	0.86 for Korean version of SSS and 0.75 for FSS at 1-week retest interval in patients with CTS[39]
MCID	Total score
	0.83 in patients assessed 6 months after CTR[29]
	0.92 in Korean version in patients assessed 3 months after CTR[39]
	1.04 in patients assessed 3 weeks after steroid injection for CTS[43]
	SSS
	0.8 in patients with CTS without diabetes and 1.45 in patients with CTS with diabetes 3 months after CTR[42]
	1.55 in patients with CTS without diabetes and 1.6 in patients with CTS with diabetes 6 months after CTR[42]
	1.14 in Korean version in patients assessed 3 months after CTR[39]
	FSS
	1.25 in patients with CTS without diabetes and 1.95 in patients with CTS with diabetes 3 months after CTR[42]
	1.45 in patients with CTS without diabetes and 2.05 in patients with CTS with diabetes 6 months after CTR[42]
	0.74 in Korean version in patients assessed 3 months after CTR[39]

CTQ = Carpal Tunnel Questionnaire; CTR = carpal tunnel release; CTS = carpal tunnel syndrome; FSS = Functional Status Scale; SSS = Symptom Severity Scale.

CTQ as well as both its scales had good test-retest reliability (ICC of 0.95, 0.85, and 0.86, respectively, for the total score, SSS, and FSS) and responsiveness (both ES and SRM exceeding 0.8 for the total score, SSS, and FSS). The pooled MDC90 and MDC95 were 0.71 and 0.85, respectively, for the SSS scores and they were 0.79 and 0.94, respectively, for the FSS scores.

Discussion

This review provides a comprehensive summary of the measurement properties of the CTQ across the published literature. The results of this review suggest that the CTQ is a robust measure for assessing condition-specific impairments in individuals with CTS or those who are recovering from CTR and has excellent body of evidence supporting its measurement properties. A few studies raised concerns about the multidimensional structure of the SSS, especially the items in the SSS that inquire about nocturnal as well as daytime symptoms had two-factor structure.[27,34,50] Clinicians should consider MCID of 1.05 and 1.13, respectively, for SSS and FSS while evaluating change at 3 months after CTR. Being a condition-specific measure, there was relatively less diversity in patient populations in which the measurement properties of the CTQ were examined. This enhances the generalizability of the results of this review.

The CTQ has been recommended as a preferred measure in individuals with CTS or CTR in previous reviews.[11,12] Leite et al[11] adopted a systematic search process to locate the literature on the measurement properties of the CTQ and, however, did not pursue critical appraisal of the literature to balance their recommendations in view of the quality of primary studies. Changulani

Table 7
Pooled estimates for selected measurement properties for the scores CTQ and its scales

Measure	ICC[a]	SEM[a]	ES[b]	SRM[c]	MDC90[d]	MDC95[d]	MCID[e]
CTQ	0.95	–	1.41	1.33	–	–	
SSS	0.85	0.30	1.51	1.39	0.72	0.85	1.05
FSS	0.86	0.34	0.81	0.82	0.79	0.94	1.13

CTQ = Carpal Tunnel Questionnaire; ES = effect size; FSS = Functional Status Scale; ICC = intraclass correlation coefficient; MCID = minimal clinically important difference; MDC90 = minimal detectable change at 90% confidence interval; MDC95 = minimal detectable change at 95% confidence interval; SSS = Symptom Severity Scale; SEM = standard error of measurement; SRM = standardized response means.
[a] ICC and SEM.[28,34,40,41,44,51,53]
[b] ES.[25,26,33-37,41,42,46,48]
[c] SRM.[5,23,25,26,30,31,33-35,40,41,46,48]
[d] MDC90 and MDC95.[28,34,40,41,44,51,53]
[e] MCID.[29,39,42,43]

et al[12] provided a narrative summary for using the CTQ in clinical practice without conducting a systematic search process to locate literature or performing appraisal to determine the methodological quality of the literature on the CTQ. Poor methodological quality of primary studies threatens internal validity of their results, thereby impacts generalizability of the evidence to clinical practice. Second, both these reviews were conducted over a decade ago and there has been significant proliferation of literature concerning the measurement properties of the CTQ.[10,24,25,27,28,30-32,34-36,39-42,44,46,48-50] Finally, Hoang-Kim[13] addressed the methodological quality of the literature while conducting their review but did not provide estimates of measurement properties from across the literature such that they facilitate decision making while using the CTQ in clinical practice. On all these counts, our review provides a comprehensive and updated summary of evidence concerning the measurement properties of the CTQ in view of the methodological quality of primary studies and provides pooled estimates for measurement properties which should be helpful while using the CTQ.

The English as well as other language versions of the CTQ did not reveal any concerns with comprehension. Barring few changes to ensure cultural and semantic equivalence,[40] most translated versions retained the content of the English CTQ which shows the relevance of its items across different cultural contexts. Atroshi et al[26] added several items to the Swedish CTQ to examine the satisfaction in symptoms after CTR. However, these added items have not been validated in the English or any other language versions of the CTQ. While using the conventional 15% threshold for considering ceiling or floor effect of the CTQ,[55] only the FSS of the CTQ-Chinese demonstrated floor effect (24.2%).[41] It is logical to assume that individuals with long-standing CTS would have greater concern with wrist/hand functions. The average duration of CTS-related impairments was 2.5 years in individuals recruited in the study[41] which might be a possible reason for floor effect observed in the FSS of the Chinese version. Of note, no study has examined the ceiling/floor effects for the English version of the CTQ and its scales. This is an important knowledge gap that needs to be addressed in future.

While we acknowledged that the initial research by Levine et al[9] that conceived the CTQ used PCC as an index of reliability, we have only extracted the data where ICC was used for assessing test-retest reliability. The PCC only examines linearity of the relationship between two sets of scores and in turn quantifies this linear relationship. Therefore, the PCC does not address the variability in the scores of the CTQ obtained at two separate time points which induces a systematic error. The ICC addresses this variability by computing the ratio of between-subject and within-subject variability between two scores obtained by different raters or different

time points.[56] The pooled estimates of the ICC for the total score of the CTQ as well as SSS and FSS were >0.85 suggesting that the CTQ and its scales have good test-retest reliability in assessing CTS-related symptoms and functional deficits across a variety of clinical contexts. Most studies assessed the reliability of the CTQ over a relatively short retest interval of 1 to 7 days[28,40,41] with few other with retest interval of up to 2 weeks.[34,44] Shorter retest intervals ensure a level of consistency in patients' reported experiences of their symptoms and functional impairments, thereby minimizing systematic error in data acquisition.

Standard error of measurement is a commonly used statistic for assessing absolute reliability for a given score of an outcome measure. The SEM is a function of test-retest reliability (ICC value) and variability (SD) of scores for that measure. The lower the ICC and/or greater the variability, the larger the measurement error is for a given score. Most studies, with an exception of one,[41] did not examine the SEM for CTQ or its scales. However, we were able to calculate the SEM for SSS and FSS using the indices of ICC and SD provided in few of the studies.[28,34,40,44] Our pooled estimates of SEM for SSS/FSS were 0.30/0.34 and were comparable to the only study that provided estimates of SEM for SSS (0.31) and FSS (0.27).[41] Similar to most measurement properties, SEM is dependent on clinical context and sample for which it was established. Given the paucity of research outlining the SEM separately for those with CTS or the ones being assessed after CTR, we propose using the SEM values of 0.30 and 0.34 for SSS and FSS, respectively, as further research refines these indices.

Most studies suggested that the SSS and FSS are internally consistent with coherent items as indicated by the CA > 0.80. The CA values of >0.90 are viewed as a concern because they indicate item redundancy thereby providing narrow mapping of domain of interest.[57] Interestingly, the CA values for both scales were <0.90 when assessed in subgroups of patients with CTS managed conservatively[26,32,34,47,49]; however, the values were consistently >0.90 when assessing patients who were assessed after CTR.[10,26,28,40] It is difficult to explain this trend because individuals in both these subgroups likely experience similar set of impairments albeit the symptoms and functional deficits might be somewhat less intense in those with CTR. Nonetheless, the CTQ and its scales have adequate internal consistency for use in individuals with CTS or those after CTR.

All three studies that conducted factor analysis clearly identified two-factor structure for SSS.[27,34,41] Two of those studies assessed the factor structure of Japanese[34] and Chinese versions[41] of the CTQ. The recommended guideline is to recruit at least 10 participants per item for a multi-item scale to precisely assess factor structure of a scale.[58] Both these studies recruited much lower sample (87 and 99 participants, respectively) for assessing the factor structure of 19 composite items across the two scales of CTQ. Atroshi et al[27] recruited requisite number of participants (N = 213) while assessing factor structure of the CTQ. Despite the limitation of sample size across two of these three studies, there was a consensus across three studies that the SSS has a two-factor structure with one factor mapping the nocturnal symptoms and the other assessing daytime symptoms. Atroshi et al[27] even offered a shorter 6-item SSS that adheres to the tenet of unidimensionality; however, there has not been much uptake on the shortened version at least in the realm of outcomes research concerning CTS. An optimal solution would be to validate the findings of multidimensionality of the English version of the SSS, identify redundant items within it, and examine the consistency of these findings with those postulated by Atroshi et al.[27]

Most articles that assessed the construct validity of the CTQ and its scales with other self-reported measures examined their relationships with the DASH[28,34,40,41,44] and the SF-36.[25,26,34,41,47]

Expectedly, the FSS showed high correlations ($r > 0.7$) with the DASH given that both assess disability resulting from CTS symptoms. Given the high concordance between the DASH and the FSS, the DASH likely provides adequate mapping of disability resulting from CTS. Considering that the SF-36 assesses multiple domains to determine overall health status, moderate to low correlations ($r \leq 0.7$) observed between the SF-36 and CTQ scales were not surprising. Mainly, the emotional role and mental health scales had low correlations with the CTQ scales (r values between 0.22 and 0.40) which clearly highlighted the divergence in the domains of interest between scales of the SF-36 and CTQ. The scores for CTQ and its scales had low correlations ($r < 0.40$) with measures assessing functional performance such as grip strength,[9,38,41,47] pinch strength,[9,41,47] or Purdue Pegboard Test.[24] Perceived impairment and actual performance often do not match. Therefore, it is important for rehabilitation practitioners to examine both the self-reported impairment and functional performance to obtain a broader overview of disability and address specific treatment needs.

Measures that are responsive to change in patient's status facilitate decision making regarding treatment effectiveness, prognosis, and resumption of activities/occupations that were restricted due to disease symptoms. The scores for CTQ and its scales demonstrated high responsiveness with ES and SRM values consistently exceeding 0.80 in patients with CTS or those after CTR.[5,23,25,26,29-31,33-36,40-42,46,48] This is important considering that the CTQ is a condition-specific measure with presumably high mapping of patient-centered concerns in those with CTS or CTR. High responsiveness yields a further advantage to CTQ when decision regarding choice of measure needs to be made to determine CTS-related impairments.

Some studies have provided the MCID values for total score of the CTQ.[29,39,43] Nonetheless, clinicians should use the MCID values for the two scales while determining change in status due to clearly pre-established factors based on two scales. Our pooled estimates of MCID for the SSS (1.05 points) and FSS (1.13 points) should assist clinicians in designing short-term goals relating to patient's symptoms and functional status. MCID is derived from anchor-based assessment of change where patients indicate whether minimal change has occurred in their status. MDC provides distribution-based assessment of change derived from variability of the recruited sample making it context dependent. The inference of MDC may not be accurate when used in a different clinical context. Therefore, clinicians using the CTQ should prefer MCID over MDC when designing treatment goals for individuals with CTS or after CTR.

Although we used a standardized systematic review design for conducting this review, our study has a few limitations that need to be outlined. First, we cannot be certain whether our search captured all the relevant articles despite adopting a comprehensive search process (eg, predefined search terms, hand search of relevant professional journals, hand search of bibliography of articles included in the review). Second, the relevant data from the primary studies were extracted by entry-level physical therapy students. To overcome this, we trained these students to conduct the data extraction at the outset and also had them complete pilot data extractions for practice and review. Both the senior authors (SM and GWZ) checked the accuracy of data extraction by matching the content of data table with the relevant articles. We hope this process minimized any systematic bias in the data extraction process. Finally, we only accepted articles that were published in English. This may have resulted in omission of some relevant articles published in other languages.

Though the results of this review suggest that there is excellent evidence to support the measurement properties of the CTQ, there

are important knowledge gaps that can clearly raise some caution. Several studies have identified two factors (nighttime symptoms and daytime symptoms) for the SSS, but still there has been limited attempt to address this measurement issue. Atroshi et al[27] provided a shorter 6-item version of the SSS but this brief version has not been tested in different context nor it has made into routine clinical practice. Future research should either build a stronger evidence for this shortened SSS or consider fragmenting the SSS into two separate scales and assess their measurement properties. Along the same lines, no study has conducted Rasch analysis of the CTQ. Rasch procedures can identify the items that are misfit especially in the SSS, provide an assessment of factor structure of the CTQ, and facilitate calculation of scaled score for the CTQ scales which might be more accurate compared to the raw score. Furthermore, MCID may provide a useful matrix to set short-term goals; scores for the CTQ scale that suggest Patient Acceptable Symptom State (PASS) are yet to be determined. The PASS score indicates the level beyond which patients consider themselves well and dissatisfied even as they may have mild residual symptoms.[59] This forms the end point of care in most patients. The PASS values for the CTQ scales can be valuable especially because CTS is a chronic condition where management of symptoms is often the primary goal versus looking for absolute resolution of impairments.

Conclusion

The CTQ is a condition-specific self-reported measure that is commonly used for assessing impairments in individuals with CTS or those who undergo CTR. In general, a large proportion of studies (27 of 34) were of good to very good quality with quality rating of >50%. The CTQ and its scales provide reliable and valid assessments of condition-specific impairments in those with CTS and are recommended for use in clinical practice. This review has also provided pooled estimates for SEM, MDC, and MCID for the CTQ and its scales which will be useful for clinicians in integrating CTQ in clinical practice.

References

1. Dale AM, Harris-Adamson C, Rempel D, et al. Prevalence and incidence of carpal tunnel syndrome in US working populations: pooled analysis of six prospective studies. *Scand J Work Environ Health*. 2013;39(5):495–505.
2. Geere J, Chester R, Kale S, Jerosch-Herold C. Power grip, pinch grip, manual muscle testing or thenar atrophy-which should be assessed as a motor outcome after carpal tunnel decompression? A systematic review. *BMC Musculoskelet Disord*. 2007;8:114.
3. Sears ED, Chung KC. Validity and responsiveness of the Jebsen-Taylor hand function test. *J Hand Surg*. 2010;35(1):30–37.
4. Hobby JL, Watts C, Elliot D. Validity and responsiveness of the patient evaluation measure as an outcome measure for carpal tunnel syndrome. *J Hand Surg Br*. 2005;30(4):350–354.
5. Greenslade JR, Mehta RL, Belward P, Warwick DJ. Dash and Boston questionnaire assessment of carpal tunnel syndrome outcome: what is the responsiveness of an outcome questionnaire? *J Hand Surg Br*. 2004;29(2):159–164.
6. Chung KC, Pillsbury MS, Walters MR, Hayward RA. Reliability and validity testing of the Michigan hand outcomes questionnaire. *J Hand Surg Am*. 1998;23(4):575–587.
7. Giannini F, Cioni R, Mondelli M, et al. A new clinical scale of carpal tunnel syndrome: validation of the measurement and clinical-neurophysiological assessment. *Clin Neurophysiol*. 2002;113(1):71–77.
8. Pransky G, Feuerstein M, Himmelstein J, Katz JN, Vickers-Lahti M. Measuring functional outcomes in work-related upper extremity disorders. Development and validation of the upper extremity function scale. *J Occup Environ Med*. 1997;39(12):1195–1202.
9. 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. *J Bone Joint Surg Am*. 1993;75(11):1585–1592.
10. Zyluk A, Piotuch B. A Comparison of DASH, PEM and Levine questionnaires in outcome measurement of carpal tunnel release. *Handchir Mikrochir Plast Chir*. 2011;43(3):162–166.
11. Leite JC, Jerosch-Herold C, Song F. A systematic review of the psychometric properties of the Boston Carpal Tunnel Questionnaire. *BMC Musculoskelet Disord*. 2006;7:78.
12. Changulani M, Okonkwo U, Keswani T, Kalairajah Y. Outcome evaluation measures for wrist and hand: which one to choose? *Int Orthop*. 2008;32(1):1–6.
13. Hoang-Kim A, Pegreffi F, Moroni A, Ladd A. Measuring wrist and hand function: common scales and checklists. *Injury*. 2011;42(3):253–258.
14. Roy JS, MacDermid JC, Woodhouse LJ. Measuring shoulder function: a systematic review of four questionnaires. *Arthritis Rheum*. 2009;61(5):623–632.
15. Mehta SP, Fulton A, Quach C, Thistle M, Toledo C, Evans NA. Measurement properties of the lower extremity functional scale: a systematic review. *J Orthop Sports Phys Ther*. 2016;46(3):200–216.
16. Macdermid JC. Quality appraisal for clinical measurement studies: evaluation form and guidelines. In: Law MM JC, ed. *Evidence-Based Rehabilitation: A Guide to Practice*. 3rd ed. Thorofare, NJ: SLACK Incorporated; 2014:325–338.
17. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276–282.
18. Amirjani N, Ashworth NL, Olson JL, Morhart M, Chan KM. Discriminative validity and test-retest reliability of the Dellon-modified Moberg pick-up test in carpal tunnel syndrome patients. *J Peripher Nerv Syst*. 2011;16(1):51–58.
19. Koldas Dogan S, Ay S, Evcik D, Baser O. Adaptation of Turkish version of the questionnaire Quick disability of the arm, shoulder, and hand (Quick DASH) in patients with carpal tunnel syndrome. *Clin Rheumatol*. 2011;30(2):185–191.
20. Mondelli M, Ginanneschi F, Rossi S, Reale F, Padua L, Giannini F. Inter-observer reproducibility and responsiveness of a clinical severity scale in surgically treated carpal tunnel syndrome. *Acta Neurol Scand*. 2002;106(5):263–268.
21. de Campos CC, Manzano GM, de Andrade LB, Castelo Filho A, Nobrega JA. [Translation and validation of an instrument for evaluation of severity of symptoms and the functional status in carpal tunnel syndrome]. *Arq Neuropsiquiatr*. 2003;61(1):51–55.
22. Oteo-Alvaro A, Marin MT, Matas JA, Vaquero J. [Spanish validation of the Boston Carpal Tunnel Questionnaire]. *Med Clin (Barc)*. 2016;146(6):247–253.
23. 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. *J Hand Surg*. 1996;21(3):338–346.
24. Amirjani N, Ashworth NL, Olson JL, Morhart M, Chan KM. Validity and reliability of the Purdue Pegboard Test in carpal tunnel syndrome. *Muscle Nerve*. 2011;43(2):171–177.
25. Atroshi I, Gummesson C, McCabe SJ, Ornstein E. The SF-6D health utility index in carpal tunnel syndrome. *J Hand Surg Eur Vol*. 2007;32(2):198–202.
26. Atroshi I, Johnsson R, Sprinchorn A. Self-administered outcome instrument in carpal tunnel syndrome. Reliability, validity and responsiveness evaluated in 102 patients. *Acta Orthop Scand*. 1998;69(1):82–88.
27. Atroshi I, Lyren PE, Gummesson C. The 6-item CTS symptoms scale: a brief outcomes measure for carpal tunnel syndrome. *Qual Life Res*. 2009;18(3):347–358.
28. Bakhsh H, Ibrahim I, Khan W, Smitham P, Goddard N. Assessment of validity, reliability, responsiveness and bias of three commonly used patient-reported outcome measures in carpal tunnel syndrome. *Ortop Traumatol Rehabil*. 2012;14(4):335–340.
29. 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. *Med Care*. 1998;36(4):491–502.
30. Chatterjee JS, Price PE. Comparative responsiveness of the Michigan hand outcomes questionnaire and the Carpal Tunnel Questionnaire after carpal tunnel release. *J Hand Surg*. 2009;34(2):273–280.
31. De Smet L, De Kesel R, Degreef I, Debeer P. Responsiveness of the Dutch version of the DASH as an outcome measure for carpal tunnel syndrome. *J Hand Surg Eur Vol*. 2007;32(1):74–76.
32. Fok M, Leung HB, Lee WM. Evaluation of a Hong Kong Chinese version of a self-administered questionnaire for assessing symptom severity and functional status of carpal tunnel syndrome: cross-cultural adaptation and reliability. *Hong Kong Med J*. 2007;13(5):342–347.
33. 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. *J Hand Surg*. 2003;28(2):250–254.
34. Imaeda T, Uchiyama S, Toh S, et al. Validation of the Japanese society for surgery of the hand version of the carpal tunnel syndrome instrument. *J Orthop Sci*. 2007;12(1):14–21.
35. Jeon SH, Lee JH, Chung MS, et al. Responsiveness of the Korean version of the disabilities of the arm, shoulder and hand questionnaire (K-DASH) after carpal tunnel release. *Clin Orthop Surg*. 2011;3(2):147–151.
36. Jerosch-Herold C, Shepstone L, Miller L, Chapman P. The responsiveness of sensibility and strength tests in patients undergoing carpal tunnel decompression. *BMC Musculoskelet Disord*. 2011;12:244.
37. Katz JN, Gelberman RH, Wright EA, Lew RA, Liang MH. Responsiveness of self-reported and objective measures of disease severity in carpal tunnel syndrome. *Med Care*. 1994;32(11):1127–1133.
38. 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. *Am J Public Health*. 1996;86(1):52–56.
39. Kim JK, Jeon SH. Minimal clinically important differences in the Carpal Tunnel Questionnaire after carpal tunnel release. *J Hand Surg Eur Vol*. 2013;38(1):75–79.
40. Kim JK, Lim HM. The Korean version of the Carpal Tunnel Questionnaire. Cross cultural adaptation, reliability, validity and responsiveness. *J Hand Surg Eur Vol*. 2015;40(2):200–205.

41. Lue Y-J, Lu Y-M, Lin G-T, Liu Y-F. Validation of the Chinese version of the Boston Carpal Tunnel Questionnaire. *J Occup Rehabil*. 2014;24(1):139–145.
42. Ozer K, Malay S, Toker S, Chung KC. Minimal clinically important difference of carpal tunnel release in diabetic and nondiabetic patients. *Plast Reconstr Surg*. 2013;131(6):1279–1285.
43. Ozyurekoglu T, McCabe SJ, Goldsmith LJ, LaJoie AS. The minimal clinically important difference of the carpal tunnel syndrome symptom severity scale. *J Hand Surg*. 2006;31(5):733–738. discussion 739-740.
44. Park DJ, Kang JH, Lee JW, et al. Cross-cultural adaptation of the Korean version of the Boston Carpal Tunnel Questionnaire: its clinical evaluation in patients with carpal tunnel syndrome following local corticosteroid injection. *J Korean Med Sci*. 2013;28(7):1095–1099.
45. Rosales RS, Delgado EB, Diez de la Lastra-Bosch I. Evaluation of the Spanish version of the DASH and carpal tunnel syndrome health-related quality-of-life instruments: cross-cultural adaptation process and reliability. *J Hand Surg*. 2002;27(2):334–343.
46. Rosales RS, Diez de la Lastra I, McCabe S, Ortega Martinez JI, Hidalgo YM. The relative responsiveness and construct validity of the Spanish version of the DASH instrument for outcomes assessment in open carpal tunnel release. *J Hand Surg Eur Vol*. 2009;34(1):72–75.
47. Sezgin M, Incel NA, Serhan S, Camdeviren H, As I, Erdogan C. Assessment of symptom severity and functional status in patients with carpal tunnel syndrome: reliability and functionality of the Turkish version of the Boston Questionnaire. *Disabil Rehabil*. 2006;28(20):1281–1285.
48. Uchiyama S, Imaeda T, Toh S, et al. Comparison of responsiveness of the Japanese Society for Surgery of the Hand version of the carpal tunnel syndrome instrument to surgical treatment with DASH, SF-36, and physical findings. *J Orthop Sci*. 2007;12(3):249–253.
49. Upatham S, Kumnerddee W. Reliability of Thai version Boston questionnaire. *J Med Assoc Thai*. 2008;91(8):1250–1256.
50. Lue YJ, Wu YY, Liu YF, Lin GT, Lu YM. Confirmatory factor Analysis of the Boston Carpal Tunnel Questionnaire. *J Occup Rehabil*. 2015;25(4):717–724.
51. Bougea A, Zambelis T, Voskou P, et al. Reliability and validation of the greek version of the Boston Carpal Tunnel Questionnaire. *Hand (N Y)*. 2018;13:593–599.
52. Hassankhani GG, Moradi A, Birjandinejad A, Vahedi E, Kachooei AR, Ebrahimzadeh MH. Translation and validation of the Persian version the Boston carpal tunnel syndrome questionnaire. *Arch Bone Jt Surg*. 2018;6(1):71–77.
53. Trybus M, Koziej M, Belka M, Bednarek M, Banach M. The Polish version of the Boston Carpal Tunnel Questionnaire: Associations between patient-rated outcome measures and nerve conduction studies. *J Plast Reconstr Aesthet Surg*. 2019;72:924–932.
54. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993;46(12):1417–1432.
55. 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(1):34–42.
56. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163.
57. Streiner DL. Starting at the beginning: an introduction to coefficient alpha and internal consistency. *J Pers Assess*. 2003;80(1):99–103.
58. Norris M, Lecavalier L. Evaluating the use of exploratory factor analysis in developmental disability psychological research. *J Autism Dev Disord*. 2010;40(1):8–20.
59. Tubach F, Ravaud P, Baron G, et al. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. *Ann Rheum Dis*. 2005;64(1):34.

JHT Read for Credit
Quiz: # 713

\# 1. The study design is
    a. RCTs
    b. a systematic review
    c. qualitative
    d. a case series

\# 2. The authors sought to appraise the _____ of the CTQ
    a. MCID and MDC
    b. validity and reliability
    c. responsiveness
    d. all of the above

\# 3. The CTQ is typically used to assess
    a. outcomes after carpal tunnel release
    b. impairments from carpal tunnel syndrome
    c. a and b above
    d. none of the above

\# 4. Almost _____ articles were found
    a. 25
    b. 35
    c. 75
    d. 95

\# 5. The authors endorse the clinical use of the CTQ
    a. true
    b. false

When submitting to the HTCC for re-certification, please batch your JHT RFC certificates in groups of 3 or more to get full credit.