

The Shanghai Elbow Dysfunction Score: Psychometric Properties, Reliability and Validity Study of the Turkish Version

Evaluation & the Health Professions
2023, Vol. 0(0) 1–8
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/01632787231183089
journals.sagepub.com/home/ehp



Tansu Birinci, PT, PhD¹ , Suleyman Altun, MD², Nezh Ziroğlu, MD³, and Ebru Kaya Mutlu, PT, PhD⁴

Abstract

The Shanghai Elbow Dysfunction Score (SHEDS) is a self-reported assessment of post-traumatic elbow stiffness that measures elbow-related symptoms and elbow motion capacities. This study aimed to (1) translate and cross-culturally adapt the SHEDS into Turkish and (2) investigate the psychometric properties of the Turkish version in patients with post-traumatic elbow stiffness. The psychometric properties of the Turkish version of the SHEDS (SHEDS-T) were tested in 108 patients (72 male; mean age, 43.2 ± 11.2 years) with post-traumatic elbow stiffness. Cronbach's alpha was used to assess internal consistency. The intraclass correlation coefficients were used to estimate test-retest. Construct validity was analyzed with the Turkish version of the Disabilities Arm, Shoulder and Hand (DASH), the Mayo Elbow Performance Score (MEPS), and the Short Form-12 (PCS-12 and MCS-12). The SHEDS-T showed sufficient internal consistency (Cronbach's α coefficient = 0.83) and test-retest reliability (ICC = .96). The correlation coefficients between the SHEDS-T, the DASH, and the MEPS were .75 and .54, respectively ($p < .001$). There was a moderate correlation between the SHEDS-T and PCS-12 ($r = .65, p = .01$) and a weak positive correlation between the SHEDS and the MCS-12 ($r = .40, p = .03$). The SHEDS-T has sufficient reliability and validity to measure elbow-related symptoms and elbow motion capacities for Turkish-speaking individuals with post-traumatic elbow stiffness.

Keywords

elbow, immobilization, questionnaire, pain, psychometrics, range of motion

Introduction

Post-traumatic elbow stiffness is one of the common complications of elbow fracture (Akhtar et al., 2021). A systematic and comprehensive approach toward elbow assessment is key to evidence-based management of post-traumatic elbow stiffness. Patient-reported outcome measures (PROMs) are frequently used to assess hand and arm function for post-traumatic elbow stiffness (Badalamente et al., 2013; Evans et al., 2018; The et al., 2013). PROMs include the Disabilities Arm, Shoulder and Hand, the Mayo Elbow Performance Score, the Liverpool Elbow Score, the Oxford Elbow Score, the Patient Rated Elbow Evaluation, and the American Shoulder and Elbow Surgeons-Elbow (Evans et al., 2018; The et al., 2013).

A systematic review assessing the clinical rating systems in elbow research concluded that the Mayo Elbow Performance Score (MEPS) and the Disabilities Arm, Shoulder and Hand (DASH) are the most frequently used clinical rating systems for elbow trauma (Evans et al., 2018). It is pointed out that the Oxford Elbow Score (OES) is the only elbow-specific rating system validated using a high-quality methodology (The

et al., 2013). Although both the DASH and MEPS are highly preferred, and the OES was developed using a high-quality methodology, the best instrument for evaluation of the elbow appears to be more dependent on the specific research question because the various scales perform differently for various conditions (Badalamente et al., 2013; Evans et al., 2018; The et al., 2013). Therefore, an elbow stiffness-specific scale might be more sensitive to changes related to treatment

¹Faculty of Health Sciences, Division of Physiotherapy and Rehabilitation, Istanbul Medeniyet University, Istanbul, Turkey

²Clinic of Orthopaedics and Traumatology, University of Health Sciences, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Istanbul, Turkey

³Clinic of Orthopaedics, Traumatology and Sports Surgery, Acibadem Atakent Hospital, Istanbul, Turkey

⁴Faculty of Health Sciences, Division of Physiotherapy and Rehabilitation, Bandırma Onyedi Eylül University, Balıkesir, Turkey

Corresponding Author:

Tansu Birinci, PT, PhD, Istanbul Medeniyet University, Faculty of Health Sciences, Division of Physiotherapy and Rehabilitation, Kartal/Istanbul 34862, Turkey.

Email: tansubirinci@hotmail.com

for post-traumatic elbow stiffness. The Shanghai Elbow Dysfunction Score (SHEDS) seems to be able to meet the demand.

The SHEDS, developed as an elbow-specific scoring system for patients with elbow stiffness, has good internal consistency, interrater reliability, and construct validity compared with the DASH, the OES, and MEPS (Sun et al., 2019). The SHEDS provides the opportunity to evaluate patients with post-traumatic elbow stiffness from the perspective of the International Classification of Functioning, Disability, and Health (ICF), a common language to describe health and disability (World Health Organization, 2001). It is aligned with the critical categories of the ICF, including d450-dressing, d4750-driving human-powered transport, d3601-using writing machines, d5202-carrying for hair, d6402-cleaning living area, d445-hand and arm use, d166-reading, d5100-washing body parts, b280-sensation of pain, s1201-spinal nerves, b730-muscle power functions, and b715-stability of joint functions (WHO, 2001; Sun et al., 2019). Due to the lack of validated, reliable elbow-specific functional assessment scales in Turkish that assess functioning related to post-traumatic elbow stiffness, this study aimed to cross-culturally adapt the SHEDS instrument into Turkish and investigate its psychometric properties.

Methods

One hundred eight consecutive patients with post-traumatic elbow stiffness were admitted to the Department of Orthopedics and Traumatology at the University of Health Sciences, Istanbul Bakirkoy Dr. Sadi Konuk Training and Research Hospital, and the Department of Orthopedics and Traumatology at the Beylikduzu State Hospital. The study was conducted between June 2021 and March 2022. Ethical approval was obtained from the Non-invasive Clinical Research Ethics Committee of Istanbul University-Cerrahpasa (Approval number: 2021/35) and conducted according to the Declaration of Helsinki. Verbal and written explanations were provided to patients about the study, and each provided written informed consent. This study was registered on [ClinicalTrials.gov](https://www.clinicaltrials.gov) (Registration number: NCT04872205).

Participants

The inclusion criteria were as follows: (1) aged between 18 and 65 years; (2) having undergone surgery for an elbow fracture; (3) loss of movement that is greater than 30° in elbow extension and movement that is less than 120° in elbow flexion; (4) ability to read and write in Turkish, and (5) having no pathology in visual ability and hearing. The exclusion criteria were as follows: (1) history of malunion or nonunion; (2) presence of non-healing wound or infection; (3) the occurrence of complex regional pain syndrome, heterotopic ossification, myositis ossification, or post-traumatic ankylosing; (4) having any cardiovascular diseases, neurological

disorders, rheumatic diseases or psychiatric diseases; (5) shoulder, elbow or wrist movement limitation in contralateral upper extremity or absence of limbs in the contralateral upper extremity; and (6) previously received physiotherapy for elbow limitation.

Sample Size

It is recommended that at least 50 patients are included in validity and responsiveness studies, where correlation coefficients are calculated, but larger samples (e.g. over 100 patients) are preferred (de Vet et al., 2011). Therefore, assuming a 5–10% of dropout rate, 110 patients were invited to the present study.

Measures

Shanghai Elbow Dysfunction Score (SHEDS) is a 13-item questionnaire specific to the elbow region. The questionnaire is divided into three subscales: elbow motion capacities (8-item daily activity), elbow-related symptoms (4-item clinical symptoms), and patient satisfaction level (1 item). Total scores range between 0 and 100 points, where the higher scores indicate a better outcome. Five points are subtracted from the SHEDS in patients with radial or median nerve symptoms. A score of 100 points is possible in a patient with the ability to perform daily living activities without difficulty, no pain, no ulnar nerve symptoms, normal muscle strength, stable elbow joint, and very satisfied (Sun et al., 2019).

The Disabilities of the Arm, Shoulder and Hand (DASH) Questionnaire is a 30-item (optional additional eight questions for sport and work assessment) patient-reported outcome measure specific to the upper limb function. The questionnaire scored on a 5-point Likert scale (no difficulty – unable). The cumulative DASH score is ranged from 0 to 100, where the higher scores indicate an increasing degree of disability (Düger et al., 2006; Hudak et al., 1996).

Mayo Elbow Performance Score (MEPS) is an 8-item patient-reported outcome measure specific to the elbow region. The questionnaire is divided into four subscales: pain (1 item, 45 points), range of motion (1 item, 20 points), stability (1 item, 10 points), and daily functional tasks (5 items, 25 points). Total scores range between 0 and 100 points, where the higher scores indicate a better outcome. A total score between 90 and 100 points can be considered excellent; between 75 and 89 points, good; between 60 and 74 points, fair; less than 60 points, poor (Celik, 2015; Longo et al., 2008).

Short Form-12 (SF-12) is the shortened form of Short Form-36 and evaluates the health-related quality of life perception. It consists of 12 items: seven dealing with the physical components (PCS-12) and five related to the mental components (MCS-12) of Short Form-12. The range of both scores is 0–100, with higher scores indicating a better health-related quality of life (Kocuyigit et al., 1999; Ware et al., 1996).

Study Protocol

The present study consisted of two phases: (1) Translation and cross-cultural adaptation and (2) psychometric testing (Prinsen et al., 2018; Mokkink et al., 2020). For the translation and cross-cultural adaptation, the SHEDS was translated into Turkish and culturally adapted following the five stages recommended by Beaton et al. (Beaton et al., 2000). In stage 1, the English version of the SHEDS was translated into Turkish by two translators (a physiotherapist with 8-years of experience and a blinded independent researcher) whose first language was Turkish. In stage 2, both translations were compared and reviewed by a bilingual person. In stage 3, once the first Turkish translation had been agreed upon, the translated Turkish version of the SHEDS was translated back into English by two translators. In stage 4, the back-translated version of the SHEDS was compared to the initial English version of the SHEDS by a committee consisting of four translators. Any diversities or discrepancies were discussed with the translators. The pre-final version of the Turkish version of the SHEDS (SHEDS-T) was developed for field testing. In the final stage, the pre-final version was administered to 30 eligible patients with post-traumatic elbow stiffness (9 females; mean age 40.2 ± 10.8 years). Patients were questioned about the questions and words that were difficult to understand after completing the form (Willis, 2015). Any difficulty in understanding an item was recorded and discussed among the authors, adjustments were made, as necessary, and the scale was finalized (Supplementary File 1).

For the psychometric testing, test-retest reliability, inter-rater reliability, internal consistency, and measurement error were used to assess reliability. Structural validity was assessed by Confirmatory factor analysis (CFA). Construct validity was assessed by hypothesis testing, where the strength of the relationship between SHEDS-T and the DASH, the MEPS, PCS-12, and MCS-12. It was hypothesized that (1) the total score from SHEDS-T would show a positive correlation (correlation coefficient of 0.70 or greater) with the DASH score due to measuring similar constructs, (2) the total score from SHEDS-T would show a positive correlation (correlation coefficient of 0.70 or greater) with the MEPS score due to measuring similar constructs, (3) the total score from SHEDS-T would show a positive correlation (correlation coefficient between .50 and .70) with the PCS-12 score due to the measuring related, but dissimilar constructs, and (4) the total score from SHEDS-T would show a positive correlation (correlation coefficient of .30 and .50) with the MCS-12 score due to the measuring unrelated constructs.

A physiotherapist with an 8-years of experience administered all questionnaires. In addition, the elbow-related symptoms subscale of SHEDS-T was also assessed by another physiotherapist to determine inter-rater reliability. The second assessment, in which the patients were asked to complete the SHEDS-T again, was

performed within a 5-to-7-day period after the first assessment to determine the test-retest reliability of the SHEDS-T. No treatment was provided during this period to minimize the risk of short-term clinical change. Only those patients who reported having “no clinical change” were included in the reliability analysis. The magnitude of change was assessed in a subgroup of patients who had conservative treatment for 6 weeks (Birinci et al., 2019). The patients were assessed at baseline and after a 6-week treatment.

Data Analysis

All statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) 20.0 (SPSS Inc., Chicago, IL, USA). The level of significance considered was $p < .05$. The intraclass correlation coefficient (ICC) was calculated to estimate the test-retest reliability. An ICC greater than .75 was considered excellent reliability (Cicchetti, 2001). The inter-rater reliability was analyzed through Cohen's kappa for elbow-related symptoms of the SHEDS-T (McHugh, 2012). The internal consistency of the SHEDS-T was analyzed through Cronbach's alpha (α) coefficient at the first completion of the form. An α value ranging from .70 to .95 was considered acceptable reliability (Tavakol & Dennick, 2011). Measurement error was assessed with the standard error of measurement (SEM). The SEM was calculated as the standard deviation of the scores times the square root of (1-ICC). MDC₉₅ was calculated as the SEM multiplied by 1.96 multiplied by the square root of 2 (Mokkink et al., 2020). Construct validity was investigated through predefined hypotheses and correlation statistics (Pearson correlation coefficient). The correlation strength was interpreted as lower than .50 weak correlation; .50–.70 moderate correlation; and higher than .70 strong correlation (Mukaka, 2012). CFA was conducted to test the factor structure. Maximum likelihood with robust standard errors (i.e., MLR) was used as an estimator in the following analysis to avoid problems related to data nonnormality. Various fit indices were used to evaluate the following; CFA: Comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), the probability of close fit associated with the RMSEA (Cfit of RMSEA), and the standardized root mean square residual (SRMR). A model adequately explains the data when the CFI and TLI are higher than .90 (better if > 0.95), the RMSEA is lower than .08 (better if lower than .05), the Cfit of RMSEA is nonsignificant (better if higher than .50), if the SRMR is lower than .08 (Brown, 2015). The floor (score 0–10) and ceiling effects (score 90–100) of the SHEDS-T at the first completion of the form were assessed by calculating the proportion of patients scoring the minimum or maximum values on the scale relative to the total number of patients. A threshold of $>15\%$ was defined as a floor or ceiling effect (Mokkink et al., 2010). The magnitude of change was assessed using the effect size (ES). The ES was determined by calculating the differences in the means of baseline and post-treatment data divided by the

Table 1. Demographics and Clinical Characteristics of Participants.

Characteristics (n = 108)	n (%)
Age (years)	43.2 ± 11.2
Body Mass index (kg/m ²)	26.5 ± 4.2
Sex (Female/Male)	36 (33.3)/72 (66.7)
Education (years)	10.2 ± 3.6
Involved side (Right/Left)	37 (40.7)/54 (59.3)
Postoperative time (weeks)	5.2 ± 3.6
Immobilization time (days)	18.5 ± 4.6
Etiology	
Coronoid fracture	6 (5.5)
Distal humeral fracture	32 (29.6)
Galeazzi fracture	3 (2.7)
Monteggia fracture	5 (4.62)
Olecranon fracture	27 (25)
Radial head fracture	32 (29.6)
Terrible triad injury	3 (2.7)
Disability arm shoulder and hand questionnaire	45.6 ± 16.2
Mayo elbow performance score	55.1 ± 13.0
Short Form-12	
Physical component summary	32.2 ± 5.8
Mental component summary	43.9 ± 9.9

Note. Continuous variables are presented as mean ± standard deviation.

standard deviation at baseline. A value between .20 and .50 was considered small effects, between .51 and .80 moderate effects, and between higher than .80 large effects (de Vet et al., 2003).

Results

Translation and Cross-Cultural Adaptation

The translators had no difficulties translating the items into Turkish during the translation process, and the back translation corresponded very well with the original version. The preliminary testing showed that some points needed to be clarified for the cultural adaptation of SHEDS-T. To clarify these particular activities (opening a door, using a computer mouse and keyboard, and combing hair) have to be done with the affected arm, we added this point as a footnote. Furthermore, “reading a book” or “reading a newspaper” is more often than “reading a magazine” in Turkey. To retain the reference to the same joint motion as reading a magazine, this item was modified to reading a magazine/book or newspaper. In the present study, one hundred and ten patients with post-traumatic elbow stiffness were screened for possible inclusion. Two patients were excluded for various reasons, and 108 patients with a mean age of 43.2 ± 11.2 years participated in the first assessment (Table 1).

Reliability

Ninety-one patients (55 males; mean age: 41.2 ± 9.3 years) participated in the second assessment for the test-retest reliability

analysis (Table 2). The internal consistency of the SHEDS-T was adequate, with Cronbach’s α coefficient of .83. The ICC_{2,1} was .96 (.93–.97) for the total score of the SHEDS-T, indicating excellent test-retest reliability. The inter-rater reliability of the SHEDS-T was almost perfect for pain, ulnar nerve, and strength (Cohen’s Kappa > .90, and strong for stability (Cohen’s Kappa = .87) (Table 3). The SEM and MDC₉₅ were determined as 8.18 and 22.67 for the Elbow Motion Capacities subscale, 6.85 and 18.98 for the Elbow-Related Symptoms subscale, and 12.33 and 34.17 for the total score of the SHEDS-T.

Validity

A priori criterion of a strong positive correlation was met for the DASH score ($r = .75, p = .001$; Hypothesis 1). However, a priori criterion of a strong positive correlation was not met for the MEPS score ($r = .54, p = .009$, Hypothesis 2). There was a moderate correlation between the total score of the SHEDS-T and PCS-12 ($r = .65, p = .01$, Hypothesis 3). In addition, a weak positive correlation was found between the total score of the SHEDS-T and MCS-12 subscale ($r = .40, p = .03$, Hypothesis 4) (Table 4). At least 75% of the findings were in accordance with the hypotheses, indicating good construct validity. The CFA showed good fit indices (RMSEA = .02, Cfit of RMSEA = .19, SRMR = .06). The CFI and the TLI reached the recommended level (CFI = .92 and TLI = .91). The standardized item loadings ranged from .55 to .88 (Supplementary File 2). All the loadings are found significant at the $p < .001$ level.

Table 2. Test-Retest Reliability and Internal Consistency of the Turkish Version of the Shanghai Elbow Dysfunction Score.

	Mean \pm SD (95% CI) (n = 91)	Test-retest Reliability / ICC (95% CI) (n = 91)	Internal Consistency / Cronbach Alpha (n = 108)
Elbow motion capacities			
First Assessment	16.5 \pm 10.5 (14.3–18.6)	0.94 (0.91–0.96)	0.85
Second Assessment	15.3 \pm 11.8 (13.0–17.6)		
Elbow-related symptoms			
First Assessment	22.1 \pm 9.6 (20.2–24.3)	0.95 (0.93–0.97)	0.87
Second Assessment	22.4 \pm 9.3 (20.4–24.2)		
Patient satisfaction level			
First Assessment	8.2 \pm 6.0 (7.3–9.8)	0.86 (0.81–0.90)	—
Second Assessment	8.5 \pm 2.6 (7.9–9.0)		
Total score			
First Assessment	46.9 \pm 19.5 (43.1–50.9)	0.96 (0.93–0.97)	0.83
Second Assessment	45.2 \pm 20.1 (41.1–49.7)		

Notes. CI, Confidence interval; ICC, Intraclass correlation coefficient; SD, Standard deviation.

Table 3. Inter-Rater Reliability of the Turkish Version of the Shanghai Elbow Dysfunction Score for Elbow-Related Symptoms.

Elbow-Related Symptoms	Rater I (n = 91)	Rater II (n = 91)	Inter-rater Reliability (Cohen's kappa)
	n (%)	n (%)	
Pain			
None	15 (16.5)	15 (16.5)	I
Mild	22 (24.2)	22 (24.2)	
Moderate	20 (22.0)	20 (22.0)	
Severe	34 (37.4)	34 (37.4)	
Ulnar nerve			
None	62 (68.1)	62 (68.1)	I
Sensory	9 (9.9)	9 (9.9)	
Motor with no disability	12 (13.2)	12 (13.2)	
Motor with disability	8 (8.8)	8 (8.8)	
Strength			
MMT = 5	27 (29.7)	24 (26.3)	0.92
MMT < 5	64 (70.3)	67 (73.7)	
Stability			
Stable	79 (86.8)	76 (83.5)	0.87
Unstable	12 (13.2)	15 (16.5)	

Note. MMT, Manual muscle testing.

The floor and ceiling effects and the number of items answered were identical during the test and retest assessments. Floor and ceiling effects were calculated as 1% and 4% in the first administration of the SHEDS-T. The magnitude of change was analyzed in the 54 patients (32 males; age: 42.3 \pm 5.1 years) diagnosed with post-traumatic elbow stiffness. The baseline and after a 6-week treatment mean scores of the SHEDS-T were 45.4 \pm 18.2 and 90.5 \pm 11.0, respectively, which resulted in a large effect (ES of 2.47, 95% CI 1.21–4.22).

Discussion

This study aimed to translate and cross-culturally adapt the SHEDS into Turkish and provide reliability, validity, and

psychometric properties for the SHEDS-T in Turkish-speaking individuals with post-traumatic elbow stiffness. The findings pointed out that the SHEDS-T has sufficient test-retest reliability (ICC = .96), internal consistency (Cronbach's α coefficient = .83), and validity to be used as an outcome measure for Turkish-speaking individuals with post-traumatic elbow stiffness. Based on the present findings, the SHEDS-T does not exhibit a ceiling or floor effect, and the MDC₉₅ values were determined as 34.17 for the total score of the SHEDS-T. Any change of less than these values of MDC₉₅ on repeated administrations of the SHEDS-T might reflect measurement errors rather than an actual change in the elbow function.

Elbow functioning cannot be summarized as just one domain. Pain, muscle strength, stability, activity and

Table 4. Correlation Between the Turkish Version of the Shanghai Elbow Dysfunction Score and Patient-Reported Outcome Measures.

Variables (<i>n</i> = 108)	Shanghai Elbow Dysfunction Score		
	Elbow Motion Capacities	Elbow-Related Symptoms	Total Score
Disability arm shoulder and hand questionnaire	0.78 (0.001)**	0.48 (0.002)**	0.75 (0.001)**
Mayo elbow performance score	0.52 (0.003)**	0.57 (0.004)*	0.54 (0.009)**
Short Form-12			
Physical component summary	0.31 (0.02)*	0.28 (0.03)*	0.65 (0.01)*
Mental component summary	0.28 (0.01)*	0.27 (0.01)*	0.40 (0.03)*

Note. Pearson correlation test $p < 0.05^*$, $p < 0.01^{**}$

participation, and emotional function should be assessed in the ICF framework (Nuttall et al., 2010; WHO, 2001). A recent study reported that patients with elbow injuries are primarily concerned about the activity and participation section of the ICF (Vincent et al., 2021). The SHEDS might be recommended as a key tool for assessing outcomes in post-traumatic elbow stiffness because it can capture important aspects of functioning such as dressing, carrying for hair, cleaning living area, hand and arm use, washing body parts, using writing machines, participating recreational and leisure activities (WHO, 2001; Sun et al., 2019). However, the SHEDS does not capture an important category, of which d550-eating. Safaee-Rad et al. examined three feeding activities and found that drinking from a cup, eating with a fork, and eating with a spoon all required an arc of flexion from 70° to 130° and 40° of pronation to 60° of supination (Safaee-Rad et al., 1990). Henmi et al. reported that flexion values ranged from a mean of 140° ± 5° for washing one's face, and the mean pronation and supination value was 75° and 100°, respectively, for eating a meal, shampooing one's hair, and washing one's face (Henmi et al., 2006). These motion arcs and positional values cover the activities questioned in the SHEDS; therefore, it may provide an indirect interpretation regarding eating activity functioning.

Pain at the end ranges of motion, which is a factor limiting elbow movements, strongly impacts elbow function and health status in patients with post-traumatic stiffness (Dávila & Johnston-Jones, 2006). VAS is used to assess pain intensity in the SHEDS, and the score is defined as none (0), mild (1–3), moderate (4–6), or severe (7–10). Grading the VAS score is compatible with the substantial clinical benefit values of 2.2–2.6 for the VAS-pain score after upper extremity surgery (Randall et al., 2022). The SHEDS, in which pain is weighted at 15%, has a more appropriate content for patients with post-traumatic stiffness compared to other elbow-related scoring systems in which pain is weighted at 30–50% (Longo et al., 2008; Sun et al., 2019). However, it is not specified as pain intensity at rest or during activity in the SHEDS, so we recommend that patients should be asked to indicate their pain during activity.

The SHEDS consists of patient-rated and physician-rated questions: ulnar nerve dysfunction, strength, and stability. Caution should be used in interpreting the physician-rated

questions as these may have inherent observer bias (Turchin et al., 1998). The present study showed that the inter-rater reliability of the SHEDS-T was sufficient for ulnar nerve dysfunction, strength, and stability assessment. Manual muscle testing (MMT) is used to assess the muscle strength in the SHEDS, and the score is classified as MMT = 5 or MMT < 5. MMT is one of the most commonly used methods for assessing muscle strength. However, measuring strength with a hand-held dynamometer would be a more objective and responsive measure to differentiate among patients with various degrees of muscle weakness compared with the MMT (Quinn & Gordon, 2015; Schrama et al., 2014). Since hand-held dynamometer measurement should be done with an elbow position of about 90 degrees which is mid-range, and is often achievable for patients whose end ranges are more a problem; therefore, MMT can provide a global idea in the early stages. (Schrama et al., 2014). After the elbow joint ROM increases, the hand-held dynamometer could be measured over time and compared to normative data.

Patient satisfaction is an essential and commonly used indicator for measuring healthcare quality (Ahmed et al., 2014). Dawson et al. stated that patient-reported and condition-specific measures for particular anatomic locations (predominantly joints) are closely aligned with patient satisfaction (Dawson et al., 2012). The SHEDS is the only elbow-related scoring system that assesses patient satisfaction, often influenced by patient expectations. The Global Rating of Change (GRC) scale is used to assess patients' satisfaction in the SHEDS, and the score is classified as very satisfied to very dissatisfied. The GRC scale provides a method to question whether the patient has improved or deteriorated, and the information gained is fundamental to clinical practice and used in making decisions regarding prognosis, treatment, and ongoing management (Kamper et al., 2009).

This study has some limitations that should be highlighted. A cross-cultural adaptation of the SHEDS into other languages is needed to compare the result of the present study. However, evidence for the reliability and validity of the SHEDS- T is similar to the psychometric properties of the English version of the SHEDS. Second, anchor-based analysis methods could be applied for responsiveness if suitable external anchors were identified. Third, content validity was not assessed.

Conclusion

The present study provides evidence that the Turkish version of the SHEDS has semantically and linguistically sufficient reliability and validity to be used as an outcome measure to evaluate the joint function of Turkish-speaking patients with post-traumatic elbow stiffness. The Turkish version of the SHEDS is also comprehensive, brief, and easy to administer and interpret, with a minimal investment of time required in research or clinical practice and evaluative application. Future studies should determine the MCID for the Turkish version of the SHEDS for patients with post-traumatic elbow stiffness.

Acknowledgments

The authors would like to thank all the patients who participated in the study.

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: The authors, their immediate family, and any research foundation with which they are affiliated did not receive any financial payments or other benefits from any commercial entity related to the subject of this article.

Funding

The author(s) did not receive financial support for the research, authorship, and/or publication of this article.

Informed Consent

Written informed consent was obtained from the patient(s) for their anonymized information to be published in this article.

Institutional Review Board Approval

Ethical approval for this study was obtained from the Non-invasive Clinical Research Ethics Committee of Istanbul University-Cerrahpasa (Approval Number: 2021/35).

Clinical Trials

This study was registered on [ClinicalTrials.gov](https://clinicaltrials.gov) (Registration number: NCT04872205).

ORCID iD

Tansu Birinci  <https://orcid.org/0000-0002-7993-3254>

Supplemental Material

Supplemental material for this article is available online.

References

- Ahmed, F., Burt, J., & Roland, M. (2014). Measuring patient experience: Concepts and methods. *Patient*, 7(3), 235–241. <https://doi.org/10.1007/s40271-014-0060-5>
- Akhtar, A., Hughes, B., & Watts, A. C. (2021). The post-traumatic stiff elbow: A review. *Journal of Clinical Orthopaedics and Trauma*, 19(6), 125–131. <https://doi.org/10.1016/j.jcot.2021.05.006>
- Badalamente, M., Coffelt, L., Elfar, J., Gaston, G., Hammert, W., Huang, J., Lattanza, L., Macdermid, J., Merrell, G., Netscher, D., Panthaki, Z., Rafijah, G., Trczinski, D., & Graham, B., American society for surgery of the hand clinical trials and outcomes committee. (2013). Measurement scales in clinical research of the upper extremity, part 2: Outcome measures in studies of the hand/wrist and shoulder/elbow. *The Journal of Hand Surgery*, 38(2), 407–412. <https://doi.org/10.1016/j.jhsa.2012.11.029>
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186–3191. <https://doi.org/10.1097/00007632-200012150-00014>
- Birinci, T., Razak Ozdincler, A., Altun, S., & Kural, C. (2019). A structured exercise programme combined with proprioceptive neuromuscular facilitation stretching or static stretching in posttraumatic stiffness of the elbow: A randomized controlled trial. *Clinical Rehabilitation*, 33(2), 241–252. <https://doi.org/10.1177/0269215518802886>
- Brown, T. A. (2015). Confirmatory factor analysis for applied research (2nd ed.). New York: Guilford Press
- Celik, D. (2015). Psychometric properties of the Mayo elbow performance score. *Rheumatology International*, 35(6), 1015–1020. <https://doi.org/10.1007/s00296-014-3201-1>
- Cicchetti, D. V. (2001). The precision of reliability and validity estimates re-visited: Distinguishing between clinical and statistical significance of sample size requirements. *Journal of Clinical and Experimental Neuropsychology*, 23(5), 695–700. <https://doi.org/10.1076/jcen.23.5.695.1249>
- Dávila, S. A., & Johnston-Jones, K. (2006). Managing the stiff elbow: Operative, nonoperative, and postoperative techniques. *Journal of Hand Therapy*, 19(2), 268–281. <https://doi.org/10.1197/j.jht.2006.02.017>
- Dawson, J., Doll, H., Boller, I., Fitzpatrick, R., Little, C., Rees, J., & Carr, A. (2012). Specificity and responsiveness of patient-reported and clinician-rated outcome measures in the context of elbow surgery, comparing patients with and without rheumatoid arthritis. *Orthopaedics & Traumatology Surgery & Research*, 98(6), 652–658. <https://doi.org/10.1016/j.otsr.2012.05.011>
- de Vet, H. C., Terwee, C. B., & Bouter, L. M. (2003). Current challenges in clinimetrics. *Journal of Clinical Epidemiology*, 56(12), 1137–1141. <https://doi.org/10.1016/j.jclinepi.2003.08.012>
- de Vet, H. C. W., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in Medicine - A practical guide*: U.S.A., Cambridge University Press, New York.
- Düger, T., Yakut, E., Öksüz, Ç., Yörük, S., Semin Bilgütay, B., Ayhan, Ç., Leblebicioğlu, G., Kayihan, H., Kırdı, N., Yakut, Y., & Güler, Ç. (2006). Kol, Omuz ve El sorunları (Disabilities of the Arm, Shoulder and Hand - DASH) Anketi Türkçe uyarlamasının güvenilirliği ve geçerliliği. *Fizyoterapi Rehabilitasyon*, 17(3), 99–107. <https://dergipark.org.tr/en/pub/tfrd/issue/12972/156608>
- Evans, J. P., Smith, C. D., Fine, N. F., Porter, I., Gangannagaripalli, J., Goodwin, V. A., & Valderas, J. M. (2018). Clinical rating

- systems in elbow research—a systematic review exploring trends and distributions of use. *Journal of Shoulder and Elbow Surgery / American Shoulder and Elbow Surgeons*, 27(4), Article e98–e106. <https://doi.org/10.1016/j.jse.2017.12.027>
- Henmi, S., Yonenobu, K., Masatomi, T., & Oda, K. (2006). A biomechanical study of activities of daily living using neck and upper limbs with an optical three-dimensional motion analysis system. *Modern Rheumatology*, 16(5), 289–293. <https://doi.org/10.1007/s10165-006-0499-x>
- Hudak, P. L., Amadio, P. C., & Bombardier, C. (1996). Development of an upper extremity outcome measure: The DASH (disabilities of the arm, shoulder and hand) [corrected]. The upper extremity collaborative group (UECG). *American Journal of Industrial Medicine*, 29(6), 602–608. [https://doi.org/10.1002/\(SICI\)1097-0274\(199606\)29:6<602::AID-AJIM4>3.0.CO;2-L](https://doi.org/10.1002/(SICI)1097-0274(199606)29:6<602::AID-AJIM4>3.0.CO;2-L)
- Kamper, S. J., Maher, C. G., & Mackay, G. (2009). Global rating of change scales: A review of strengths and weaknesses and considerations for design. *The Journal of Manual & Manipulative Therapy*, 17(3), 163–170. <https://doi.org/10.1179/jmt.2009.17.3.163>
- Kocyigit, H., Aydemir, O., Fisek, G., Olmez, N., & Memis, A. (1999). Reliability and validity of the Turkish version of Short-Form-36 (SF-36). *Ilac ve Tedavi Derg*, 12, 102–106. [in Turkish].
- Longo, U. G., Franceschi, F., Loppini, M., Maffulli, N., & Denaro, V. (2008). Rating systems for evaluation of the elbow. *British Medical Bulletin*, 87(1), 131–161. <https://doi.org/10.1093/bmb/ldn023>
- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276–282. <https://doi.org/10.11613/BM.2012.031>
- Mokkink, L. B., Boers, M., van der Vleuten, C. P. M., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2020). COSMIN risk of bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: A delphi study. *BMC Medical Research Methodology*, 20(1), 293. <https://doi.org/10.1186/s12874-020-01179-5>
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., Bouter, L. M., & de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63(7), 737–745. <https://doi.org/10.1016/j.jclinepi.2010.02.006>
- Mukaka, M. M. (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal*, 24(3), 69–71. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3576830/>
- Nuttall, D., Birch, A., Trail, I. I., Espag, M. P., Clark, D. I., & Stanley, J. K. (2010). Assessing elbow assessment, past, present and future. *Shoulder Elbow*, 2(1), 43–54. <https://doi.org/10.1111/j.1758-5740.2009.00046.x>
- Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27(5), 1147–1157. <https://doi.org/10.1007/s11136-018-1798-3>
- Quinn, L., & Gordon, J. (2015). *Documentation for rehabilitation-E-book: A guide to clinical decision making in physical therapy*. Elsevier Health Sciences.
- Randall, D. J., Zhang, Y., Li, H., Hubbard, J. C., & Kazmers, N. H. (2022). Establishing the minimal clinically important difference and substantial clinical benefit for the pain visual analog scale in a post-operative hand surgery population. *The Journal of Hand Surgery*, 47(7), 645–653. <https://doi.org/10.1016/j.jhsa.2022.03.009>
- Safaei-Rad, R., Shwedyk, E., Quanbury, A. O., & Cooper, J. E. (1990). Normal functional range of motion of upper limb joints during performance of three feeding activities. *Archives of Physical Medicine and Rehabilitation*, 71(7), 505–509. <https://pubmed.ncbi.nlm.nih.gov/2350221/>
- Schrama, P. P., Stenneberg, M. S., Lucas, C., & van Trijffel, E. (2014). Intraexaminer reliability of hand-held dynamometry in the upper extremity: A systematic review. *Archives of Physical Medicine and Rehabilitation*, 95(12), 2444–2469. <https://doi.org/10.1016/j.apmr.2014.05.019>
- Sun, Z., Liu, W., Wang, W., & Fan, C. (2019). Development and validation of a new elbow-specific scoring system for patients with elbow stiffness: The Shanghai elbow dysfunction score. *Journal of Shoulder and Elbow Surgery*, 28(2), 296–303. <https://doi.org/10.1016/j.jse.2018.08.007>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53–55. <https://pubmed.ncbi.nlm.nih.gov/28029643/>
- The, B., Reininga, I. H., El Moumni, M., & Eygendaal, D. (2013). Elbow-specific clinical rating systems: Extent of established validity, reliability, and responsiveness. *Journal of Shoulder and Elbow Surgery*, 22(10), 1380–1394. <https://doi.org/10.1016/j.jse.2013.04.013>
- Turchin, D. C., Beaton, D. E., & Richards, R. R. (1998). Validity of observer-based aggregate scoring systems as descriptors of elbow pain, function, and disability. *The Journal of Bone and Joint Surgery. American Volume*, 80(2), 154–162. <https://doi.org/10.2106/00004623-199802000-00002>
- Vincent, J. I., MacDermid, J. C., King, G. J. W., & Grewal, R. (2021). The Patient-rated elbow evaluation and the American shoulder and elbow surgeons-elbow form capture aspects of functioning that are important to patients with elbow injuries. *Journal of Hand Therapy*, 34(3), 415–422. <https://doi.org/10.1016/j.jht.2020.02.001>
- Ware, J. Jr, Kosinski, M., & Keller, S. D. (1996). A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Med Care*, 34(3), 220–233. <https://doi.org/10.1097/00005650-199603000-00003>
- Willis, G. B. (2015). The practice of cross-cultural cognitive interviewing. *Public Opinion Quarterly*, 79(S1), 359–395. <https://doi.org/10.1093/poq/nfu092>
- World Health Organization. (2001). *International classification of functioning, disability and health: ICF*. World Health Organization.