J Shoulder Elbow Surg (2019) 28, 296–303

ELSEVIER

Journal of Shoulder and Elbow Surgery

www.elsevier.com/locate/ymse

Development and validation of a new elbow-specific scoring system for patients with elbow stiffness: the Shanghai Elbow Dysfunction Score

Ziyang Sun, MD[a], Wenjun Liu, MD[b], Wei Wang, MD[a,b,*], Cunyi Fan, MD, PhD[a,b,*]

[a]*Department of Orthopedics, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, China*
[b]*Department of Orthopedics, Shanghai Sixth People's Hospital East Affiliated to Shanghai University of Medicine & Health Sciences, Shanghai, China*

Background: Clinical scoring systems are increasingly important and popular for the evaluation of orthopedic patients. Elbow stiffness commonly causes functional impairment and upper-limb disability. The purpose of this study was to develop and validate a new elbow-specific assessment score to evaluate joint function in patients with elbow stiffness.

Methods: The new system, the Shanghai Elbow Dysfunction Score (SHEDS), was developed in 3 portions: elbow motion capacities, elbow-related symptoms, and patient satisfaction level. A total of 73 patients with elbow stiffness were prospectively included. Intraclass correlation coefficients and Cronbach α values were calculated for test-retest reliability and internal consistency, respectively. Construct validity was assessed by correlating the SHEDS with previously validated scoring systems. Effect sizes (ES) and standardized response means (SRMs) were calculated for responsiveness.

Results: Positive reliability with an intraclass correlation coefficient of 0.83 and adequate homogeneity with a Cronbach α value of 0.74 were found for the SHEDS. Good to excellent validity using Spearman correlation coefficients (SCCs) were determined for the total (0.51-0.82), motion (0.65-0.89), and symptom (0.35-0.53) scores. Responsiveness was large for the total ES, 3.48; SRM, 2.96), motion (ES, 2.54; SRM, 2.08), and symptom (ES, 1.26; SRM, 1.14) scores. There were no ceiling or floor effects. Significant positive correlations were found between patient satisfaction levels and the final scores (SCC, 0.62), as well as the score changes of the SHEDS (SCC, 0.42).

Conclusion: Our results suggest that the newly developed SHEDS is an excellent, comprehensive, valid scoring system to evaluate joint function in patients with elbow stiffness.

Level of evidence: Development or Validation of Outcome Instruments

Keywords: Clinical scoring system; elbow stiffness; validation; reliability; validity; responsiveness

The Ethics Committee of Shanghai Jiao Tong University Affiliated Sixth People's Hospital East Campus approved this prospective study at the start of this research (study No. 2016-030).

*Reprint requests: Cunyi Fan, MD, PhD, and Wei Wang, MD, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, 600 Yishan Road, Shanghai 200233, China.

E-mail addresses: cyfan@sjtu.edu.cn (C. Fan); wahaa1987@163.com (W. Wang).

The elbow is a constrained, complex, hinged joint, which positions the hand inside a sphere, created by shoulder movement, with a radius the length of the upper arm and forearm. Normal elbow motion ranges between approximately 0° (full extension) and 140°-150° of flexion. Forearm rotation varies but averages about 75° in pronation and 85° in supination. Morrey et al[25] estimated that the minimum motion required

1058-2746/$ - see front matter
https://doi.org/10.1016/j.jse.2018.08.007

for relatively normal function was an extension-flexion arc of 100° (between 30° and 130°) and forearm rotation of 100° of combined pronation and supination. Elbow stiffness is a well-recognized disabling condition that causes functional impairment in the upper limb and interferes with daily activities, and it is a very common complication after elbow injury or owing to arthropathy. Both bony and soft-tissue factors are important etiologies.[10,18,30]

Clinical scoring instruments are increasingly used as functional measurement tools for orthopedic patients. They help estimate the severity of dysfunction, evaluate treatment effectiveness, and compare different treatment methods. In most elbow-related scoring systems, pain accounts for most of the overall score almost uniformly whereas neurologic function is inadequately assessed. There is also wide variability in strength, instability, and deformity measurements among systems.[21,38] In clinical practice, however, the main concern of patients with elbow stiffness is often movement restoration rather than pain relief. Neurologic (especially ulnar nerve) symptoms, weakness, and instability are commonly seen in patients after elbow trauma and related surgical procedures.[4] Furthermore, the Oxford Elbow Score[9] (OES) has been reported to be the only elbow-specific scoring system validated for reliability, validity, and responsiveness using high-quality methodology.[41,42] Even a well-established, validated instrument may not be applicable for all different populations, the need for published scoring systems for elbow stiffness is clear. However, for this special population, there have been no published validation studies, and existing systems have some inadequacies.

Therefore, the purpose of this study was to develop and validate a new elbow-specific functional assessment score for patients with elbow stiffness. We called this new scoring system the Shanghai Elbow Dysfunction Score (SHEDS) and hypothesized that the SHEDS would have positive reliability, adequate homogeneity, good validity, and significant responsiveness for our target population. In addition, we hypothesized that both the final scores and the score changes of the SHEDS would show positive correlations with patient satisfaction levels.

Materials and methods

Development

Item extraction and screening

Our purpose was to evaluate joint function in a target population of patients with elbow stiffness. On the basis of clinical practice experience, we have observed that elbow motion capacities and symptoms are the most important considerations for patients with elbow disorders. Almost all elbow-related scoring systems assess elbow extension-flexion and forearm rotation, whereas evaluations of symptoms, including pain, muscular strength, stability, and nerve function, vary widely.

Besides the crucial elbow motion capacities, the following 5 targets, in our experience, were important for evaluating elbow stiffness: pain, ulnar nerve function, stability, strength, and satisfaction. (1) Pain has a strong impact on elbow function and health status measures,[12] and pain is weighted at 30%-50% in most elbow-related scoring systems.[21] (2) Delayed-onset ulnar neuropathy is common after elbow trauma, and elbow surgery carries risks of neurovascular complications, most commonly ulnar neuritis.[4] This underscores the importance of evaluating ulnar nerve function before and after elbow surgery.[5,7,37] Unfortunately, ulnar nerve function is neglected in most elbow-related scoring systems. (3) Instability may indicate collateral ligament dysfunction and/or structural abnormalities and thus prove important in planning operative options and postoperative rehabilitation.[4] (4) Along with joint instability, reduced muscle strength, often observed clinically after elbow trauma and surgery, is evaluated in several elbow scoring systems.[21] Both long-term disuse atrophy and elevation or incision of the muscle during surgery may lead to muscular strength reduction.[8,40] (5) No existing elbow-related scoring system records patient satisfaction, which is often influenced by patient expectations.[35] Objective parameters have been found by some researchers to have no correlation with patient satisfaction.[6] The addition of patient satisfaction, as in the Knee Society Scoring System developed in 2012,[26] may reconcile discrepancies between clinician- and patient-derived assessment tools.

Item pattern determination appropriately for follow-up

The SHEDS was gradually refined into 3 portions: elbow motion capacities (EMC), elbow-related symptoms (ERS), and patient satisfaction level. Clinical experience suggests that patients with elbow stiffness are more concerned with elbow function in daily activities than with absolute increase in range of motion (ROM)[27]; thus, patient-rated questionnaires were chosen for EMC assessment. An innovative 8-item daily activity form (Table I) was developed. Normal elbow motion was split into 8 independent parts (p1-p4 and s1-s4), in which forearm rotation was divided into pronation (p) and supination (s) and extension was divided as follows: (1) 0°-30°, (2) 30°-60°, (3) 60°-90°, and (4) greater than 90°. By use of 3-dimensional motion-capture technology in 3 different planes simultaneously,[16,44] critical positional and functional tasks along with various contemporaneous tasks were identified for each elbow ROM by reference to the work of Sardelli et al[33] and other studies.[17,29] After discussion by our clinical team, p1, p2, p3, and p4 were assigned to the representative tasks of tying shoelaces, riding a bike, using a computer mouse and keyboard, and combing hair, respectively. For s1, s2, s3, and s4, mopping the floor, opening a door, reading a magazine, and washing one's face, respectively, were selected. Finally, each item was graded as follows: not difficult, somewhat difficult, or unable. We then randomly selected 20 patients with elbow stiffness that was exactly at each of the 4 ranges of extension ROM for pretesting (eg, the extension-flexion arc of 35°-50° was chosen for group 2) in the outpatient service. The results confirmed that ratings of not difficult or somewhat difficult were assigned by

Table I Eight-item activity form for EMC portion

Item	Pronation	Supination
Extension 1 (0°-30°)	Tying shoelaces (p1)	Mopping the floor (s1)
Extension 2 (30°-60°)	Riding a bike (p2)	Opening a door (s2)
Extension 3 (60°-90°)	Using a computer mouse and keyboard (p3)	Reading a magazine (s3)
Extension 4 (>90°)	Combing hair (p4)	Washing one's face (s4)

EMC, elbow motion capacities.

patients for doing expected activities corresponding to their preserved range, whereas ratings of somewhat difficult or unable were assigned as expected to the other 3 ranges, confirming the instrument's appropriate use to evaluate elbow mobility.

The ERS portion was assessed as follows: Pain expression was evaluated by a patient-rated visual analog scale (VAS) questionnaire from 0 to 10, owing to established strong influences of psychological and sociologic factors on pain.[11] To statistically strengthen the subjective VAS score, pain was graded and defined as none (0), mild (1-3), moderate (4-6), or severe (7-10). Ulnar nerve symptoms were graded as none, sensory, motor with no disability, or motor with disability, as in the Liverpool Elbow Score.[34] Strength was initially thought to be best measured using quantitative testing instruments,[36] but manual muscle strength testing (MMT) proved more practical to implement and eliminated concerns of quantitative testing such as arm dominance and defining abnormal strength values. Strength was graded in a bimodal manner: MMT of 5 (level V) or MMT lower than 5. Elbow stability was graded simply as stable or unstable. The validated, self-administered scale (very satisfied, somewhat satisfied, neutral, somewhat dissatisfied, or very dissatisfied) of Mahomed et al[22] was used for patient satisfaction.

Point allocation

A score range between 0 and 100 points (with lower scores representing greater severity) was selected for the SHEDS for simplicity of calculation (Table II). There were 3 grades (an x point was assigned for grade difference) in each item of the EMC portion; therefore, the total possible score for EMC was as follows: $2x \times 8 = 16x$ points. Pain is weighted as 30%-50% in most elbow-related scoring systems,[23,38] and ulnar nerve assessment is very important as mentioned earlier.[4,5,7,37] Therefore, we agreed that the point weightings for pain and ulnar nerve function should exceed those for stability and strength. As the main concern of patients with elbow stiffness is movement restoration rather than pain relief, we chose to give the same score for pain as for ulnar nerve items, in contrast to most other systems. Therefore, a y point was assigned uniformly for the grade difference in the ERS portion, with a total possible score of $3y + 3y + y + y = 8y$ points. A total score of $4z$ points could be added, in which z represents the grade difference in satisfaction level. Thus, we ended with a potential total of $16x + 8y + 4z = 100$ points.

Considering that achieving a functional ROM was the most important consideration for patients with elbow stiffness, yet the importance of improving elbow symptoms should not be ignored, we agreed to further refine the instrument by changing it so that the weighted values of the EMC portion were balanced to comprise nearly half of the total score while maintaining a score ratio of about 1:1 between the EMC and ERS portions. As an additional refinement, it was determined to be more appropriate to assign a score of 3 points for x, 5 points for y, and 3 points for z, that is, 6 total points for each movement (48 total points for the EMC portion), 15 points for pain and ulnar nerve symptoms, 5 points for stability and muscle strength (40 total points for the ERS portion), and 12 points for satisfaction. Finally, we subtracted 5 points from the SHEDS in cases in which radial or median nerve symptoms were present (Table III).

Table II Point allocation for each item of SHEDS

	Score
EMC portion (total score = $2x \times 8 = 16x$ points)	
Each item	
Not difficult	$2x$ points
Somewhat difficult	x points
Unable	0 point
ERS portion (total score = $3y + 3y + y + y = 8y$ points)	
Pain	
None	$3y$ points
Mild	$2y$ points
Moderate	y points
Severe	0 point
Ulnar nerve	
None	$3y$ points
Sensory	$2y$ points
Motor with no disability	y points
Motor with disability	0 point
Strength	
MMT = 5	y points
MMT < 5	0 point
Stability	
Stable	y points
Unstable	0 point
Patient satisfaction level (total score = $4z$ points)	
Very satisfied	$4z$ points
Somewhat satisfied	$3z$ points
Neutral	$2z$ points
Somewhat dissatisfied	z points
Very dissatisfied	0 point

SHEDS, Shanghai Elbow Dysfunction Score; *EMC*, elbow motion capacities; *ERS*, elbow-related symptoms; *MMT*, manual muscle strength testing. The total score for the SHEDS is $16x + 8y + 4z = 100$ points.

Patients

This was a prospective study that assessed patients who presented to our institution for elbow arthrolysis because of elbow stiffness from December 2016 to April 2017. The exclusion criteria were (1) unwillingness to participate in or cooperate with follow-up, (2) illiteracy or inability to understand the contents of questionnaires for cognitive reasons, and (3) mental illness. During the study period, 82 patients underwent surgery for elbow stiffness at our institution. Of these, 73 met the inclusion and exclusion criteria (Table IV). All 73 completed the SHEDS, Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire, OES, Mayo Elbow Performance Score (MEPS), and Short Form 36 (SF-36) preoperatively; and 68 completed these at final follow-up (5 refused or were lost to follow-up). Data were analyzed anonymously; all patients approved the results of this study by oral consent. The oral consent approval was documented in the patients' files. All clinical investigations were conducted in accordance with the guidelines of the Declaration of Helsinki.

Testing and evaluation of measurement qualities

Reliability

Intraclass correlation coefficients (ICCs) were calculated for test-retest reliability. Initial testing was performed 14 days before hospitalization by patients and doctors from clinical teams in the outpatient service. The second examination was performed just before surgery by patients and doctors from clinical teams in inpatient departments. The Cronbach α (CA) was calculated for internal consistency.

Validity

"Construct validity" is defined as the degree to which the scores of a particular instrument correspond with a gold standard. Unfortunately,

Table III Shanghai Elbow Dysfunction Score

	Score, points
EMC	
Tying shoelaces	
Not difficult	6
Somewhat difficult	3
Unable	0
Mopping the floor	
Not difficult	6
Somewhat difficult	3
Unable	0
Riding a bike	
Not difficult	6
Somewhat difficult	3
Unable	0
Opening a door	
Not difficult	6
Somewhat difficult	3
Unable	0
Using a computer mouse and keyboard	
Not difficult	6
Somewhat difficult	3
Unable	0
Reading a magazine	
Not difficult	6
Somewhat difficult	3
Unable	0
Combing hair	
Not difficult	6
Somewhat difficult	3
Unable	0
Washing one's face	
Not difficult	6
Somewhat difficult	3
Unable	0
ERS	
Pain* (visual analog scale)	
None (0)	15
Mild (1-3)	10
Moderate (4-6)	5
Severe (7-10)	0
Ulnar nerve†	
None	15
Sensory	10
Motor with no disability	5
Motor with disability	0
Strength	
MMT = 5	5
MMT < 5	0
Stability	
Stable	5
Unstable	0
Patient satisfaction level	
Very satisfied	12
Somewhat satisfied	9
Neutral	6
Somewhat dissatisfied	3
Very dissatisfied	0

EMC, elbow motion capacities; *ERS*, elbow-related symptoms; *MMT*, manual muscle strength testing.

* Pain is scored as the degree of pain on a visual analog scale.

† We subtract 5 points from the Shanghai Elbow Dysfunction Score in cases in which radial or median nerve symptoms are present.

Table IV Demographics and clinical characteristics of patients

	Data
Characteristic	
No. of patients	73
Male	42 (58)
Age, yr	37 ± 12 (8-71)
BMI, kg/m^2	23.3 ± 3.0 (16.4-31.1)
Dominant arm	41 (56)
Disease duration,* mo	26 ± 55 (2-360)
Previous elbow ORIF	59 (81)
Immobilization time, weeks	2 ± 3 (0-19)
Follow-up time,† mo	14 ± 1 (12-16)
Etiology (traumatic and atraumatic)	
Synovialis	2 (3)
Simple elbow dislocation	5 (7)
Distal humeral fracture	25 (34)
Radial head fracture	6 (8)
Coronoid fracture	1 (3)
Olecranon fracture	15 (21)
Galeazzi fracture	1 (3)
Monteggia fracture	4 (5)
Floating elbow fracture	5 (7)
Terrible triad injury	9 (12)

BMI, body mass index; *ORIF*, open reduction and internal fixation. Categorical variables are presented as number (percentage). Continuous variables are presented as mean ± standard deviation (range).

* Length of time with stiff elbow.

† Postoperative period after elbow release.

no gold-standard test has been previously established for pre- and post-arthrolysis elbow status. The 2-part DASH questionnaire can be used to measure disability for any upper-limb region and was shown to be valid and responsive.[3] The 3-part OES is a valid, reliable, and responsive self-administered instrument, useful for several types of elbow function measurements.[9] The DASH score and OES have been shown to correlate with general health measures such as the SF-36.[9,39] Consisting of physician-rated pain, ROM, and stability, as well as patient-rated daily function, the MEPS[24] was the most widely used elbow functional assessment in a large systematic review.[13] In this project, the disability part of the DASH questionnaire, the elbow function part of the OES, and the ROM and daily activity function parts of the MEPS were consolidated as the EMC portion of the SHEDS. The symptom section of the DASH questionnaire and the pain and stability portions of the MEPS were combined in the ERS portion. Construct validity was then assessed by correlating the SHEDS with the DASH score, OES, MEPS, and SF-36 score (physical component summary [PCS] and mental component summary [MCS]) in total scores and the EMC portion, ERS portion, and pain portion using Spearman correlation coefficients (SCCs).

Responsiveness

We calculated the effect size (ES)[19] and standardized response mean (SRM)[28] to quantify responsiveness for all systems. Possible floor and ceiling effects were scrutinized. Last, the final scores and score changes of the SHEDS were correlated with patient satisfaction levels.

Translation

The SHEDS was first discussed, developed, and improved in a simplified Chinese (mainland)–language version by the expert committee

led by C.F., including most of our clinical team members who were experienced in research and treatment of elbow stiffness. The translation from simplified Chinese (mainland) into the English-language version was performed using the cross-cultural adaptation process of Guillemin et al,[14,15] maintaining content validity by proceeding in 6 steps: translation, synthesis, back translation, expert committee review, pretesting, and submission for appraisal. At last, a final English-language version (Table III) was produced after consensus from the committee on semantic, idiomatic, experiential, and conceptual equivalence between the English-language and original simplified Chinese (mainland)–language versions.

Statistical analyses

All statistical analyses were performed using SPSS software (version 22.0; IBM, Armonk, NY, USA). Categorical data are presented as number (percentage). Continuous data are presented as mean ± standard deviation (range). $P < .05$ was considered statistically significant. Positive reliability and sufficient homogeneity were assumed when the ICC and CA were at least 0.70.[41] Construct validity or correlation was considered strong if the SCC was greater than 0.5; moderate, 0.35-0.5; or weak, less than 0.35.[20] Responsiveness was considered large if the ES or SRM was greater than 0.8; moderate, 0.5-0.8; or small, less than 0.5.[31] Floor or ceiling effects were determined to exist when more than 15% of a patient group achieved the highest or lowest possible score of the SHEDS.[23]

Results

All patients completed the SHEDS with no difficulties and no missing or multiple responses. The ICCs were 0.83, 0.81, and 0.89 for the total score, EMC portion, and ERS portion, respectively; the CAs were 0.74, 0.78, and 0.62, respectively (Table V). All the SCCs were positive, except the relationship with the DASH score, which is scored inversely (Table VI). The SHEDS overall scores correlated well with the compared total scores ($P < .001$ for all SCCs) for the DASH score (0.82), OES (0.76), MEPS (0.51), and SF-36 score (PCS, 0.55; MCS, 0.70). The EMC and ERS portions of the SHEDS correlated either strongly or moderately with similar parts of the DASH questionnaire, OES, and MEPS. In general, the different SHEDS portions also correlated well with the DASH score, OES, MEPS, and SF-36 score, with either high or moderate correlations in total scores (0.51-0.82) and the EMC (0.65-0.89), ERS (0.35-0.53), and pain (0.43-0.57) portions of the SHEDS.

Table V Test-retest reliability and internal consistency of SHEDS

	Test, mean (SD)	Retest, mean (SD)	ICC	CA
Total score	49 (13)	45 (10)	0.83	0.74
EMC portion	18 (12)	13 (8)	0.81	0.78
ERS portion	31 (6)	31 (6)	0.89	0.62

SHEDS, Shanghai Elbow Dysfunction Score; *SD,* standard deviation; *ICC,* intraclass correlation coefficient; *CA,* Cronbach α; *EMC,* elbow motion capacities; *ERS,* elbow-related symptoms.

Table VI Construct validity: SCCs between SHEDS and DASH score, OES, MEPS, and SF-36 score

	Total	EMC portion	ERS portion	Pain portion
DASH score	0.82*	0.89*	0.53*	0.57*
OES	0.76*	0.82*	—	0.43*
MEPS	0.51*	0.65*	0.35†	0.57*
SF-36 PCS score	0.55*	—	—	(SF-36 BP)
SF-36 MCS score	0.70*	—	—	(0.47*)

SCC, Spearman correlation coefficient; *SHEDS,* Shanghai Elbow Dysfunction Score; *DASH,* Disabilities of Arm, Shoulder and Hand questionnaire; *OES,* Oxford Elbow Score; *MEPS,* Mayo Elbow Performance Score; *SF-36,* Short Form 36; *EMC,* elbow motion capacities; *ERS,* elbow-related symptoms; *PCS,* physical component summary; *MCS,* mental component summary; *BP,* body pain.

* $P < .001$.

† $P < .01$.

The SHEDS showed large (ES > 0.8 or SRM > 0.8) responsiveness for total score (ES, 3.48; SRM, 2.96; $P < .001$) and both the EMC and ERS portions (Table VII). This analysis also showed that the SHEDS was more responsive than the DASH score, with an ES of 2.83 and SRM of 2.33 ($P < .001$); OES, with 2.55 and 2.00, respectively ($P < .001$); MEPS, with 2.21 and 1.80, respectively ($P < .001$); SF-36 PCS score, with 1.64 and 1.33, respectively ($P < .001$); and SF-36 MCS score, with 1.14 and 0.97, respectively ($P < .001$). No floor or ceiling effects were found (Table VIII). A significant positive correlation was found between final scores (0.62, $P < .001$), as well as score changes of the SHEDS (0.42, $P < .001$), and satisfaction levels (Table IX), which was higher than that found using other systems.

Discussion

With an activated myofibroblast–mast cell–neuropeptide axis[10,18] and a local osteoinductive microenvironment,[30] the elbow is more vulnerable to motion loss than any other joint after trauma or arthropathy. This lost motion is often accompanied by pain, numbness, weakness, or instability and may interfere with leisure activities and basic activities of living. However, existing elbow-related assessment scores have defects in evaluation content or validation methodology for elbow stiffness; therefore, a new instrument with high-quality validation is needed.

The SHEDS, a newly developed and positively validated functional measurement tool for elbow stiffness, is simple enough to be rapidly administered in clinics, with a scale range from 0 to 100 points (with higher scores indicating better function) and a content of 13 items, comprising physician-assessed ulnar nerve function, muscular strength, and stability, as well as patient-assessed motion capacities, pain, and satisfaction.

The increased importance and use of self-assessment scores as additional measurement tools to the physician-based objective evaluation can be attributed to their advantages regarding financial and logistical concerns and elimination

Table VII Responsiveness of SHEDS compared with DASH score, OES, MEPS, and SF-36 score

Questionnaires	Mean (SD)			*P* value	ES*	SRM*
	Preoperative	Follow-up	Change			
SHEDS						
Total score	49 (13)	88 (9)	38 (13)	<.001	3.48 (L)	2.96 (L)
EMC portion	18 (11)	41 (6)	23 (11)	<.001	2.54 (L)	2.08 (L)
ERS portion	31 (6)	36 (4)	5 (6)	<.001	1.26 (L)	1.14 (L)
Pain portion	12 (4)	14 (2)	2 (4)	<.001	0.81 (L)	0.64 (M)
DASH score						
Total score	52 (17)	11 (11)	41 (18)	<.001	2.83 (L)	2.33 (L)
EMC portion	44 (16)	9 (9)	35 (16)	<.001	2.75 (L)	2.22 (L)
ERS portion	8 (3)	3 (3)	6 (3)	<.001	1.98 (L)	1.64 (L)
Pain	3 (2)	1 (1)	2 (2)	<.001	1.29 (L)	0.99 (L)
OES						
Total score	50 (17)	87 (12)	37 (19)	<.001	2.55 (L)	2.00 (L)
EMC portion	36 (22)	86 (14)	51 (23)	<.001	2.73 (L)	2.19 (L)
Pain	75 (23)	92 (12)	17 (25)	<.001	0.92 (L)	0.67 (M)
MEPS						
Total score	64 (15)	92 (9)	28 (15)	<.001	2.21 (L)	1.80 (L)
EMC portion	18 (11)	43 (4)	25 (11)	<.001	3.03 (L)	2.18 (L)
ERS portion	46 (9)	49 (7)	3 (9)	.013	0.34 (S)	0.31 (S)
Pain	37 (9)	40 (7)	3 (9)	.008	0.35 (S)	0.33 (S)
SF-36 score						
PCS	53 (17)	80 (16)	27 (20)	<.001	1.64 (L)	1.33 (L)
MCS	51 (24)	75 (18)	24 (25)	<.001	1.14 (L)	0.97 (L)
Pain	70 (15)	83 (10)	13 (15)	<.001	1.02 (L)	0.85 (L)

SHEDS, Shanghai Elbow Dysfunction Score; *DASH,* Disabilities of Arm, Shoulder and Hand questionnaire; *OES,* Oxford Elbow Score; *MEPS,* Mayo Elbow Performance Score; *SF-36*, Short Form 36; *SD,* standard deviation; *ES,* effect size; *SRM,* standardized response mean; *L*, large; *EMC,* elbow motion capacities; *ERS,* elbow-related symptoms; *PCS,* physical component summary; *MCS,* mental component summary.

* For ES and SRM, large responsiveness (L) is defined as greater than 0.8; moderate responsiveness (M), 0.5 to 0.8; and small responsiveness (S), less than 0.5.

Table VIII Floor and ceiling effects of SHEDS

Component	SHEDS, n* (%)		Floor effect, %†	Ceiling effect, %†
	Lower limit	Upper limit		
Total score	25 (1)	82 (1)	0	0
EMC portion	0 (2)	45 (1)	2.7	0
ERS portion	15 (1)	40 (8)	0	11.0

SHEDS, Shanghai Elbow Dysfunction Score; *EMC,* elbow motion capacities; *ERS,* elbow-related symptoms.

* Number of patients showing lowest or highest values in various parts.

† Percent of patients achieving lowest or highest values in various parts.

Table IX Correlations of final scores and score changes with patient satisfaction

Questionnaire	SHEDS	DASH score	OES	MEPS	SF-36 PCS score	SF-36 MCS score
SCC_1	0.62*	0.49*	0.52*	0.49*	0.27†	0.26†
SCC_2	0.42*	0.27†	0.30†	0.29†	0.20	0.20

SHEDS, Shanghai Elbow Dysfunction Score; *DASH,* Disabilities of Arm, Shoulder and Hand questionnaire; *OES,* Oxford Elbow Score; *MEPS*, Mayo Elbow Performance Score; *SF-36*, Short Form 36; *PCS,* physical component summary; *MCS,* mental component summary; SCC_1, Spearman correlation coefficient between final score and patient satisfaction; SCC_2, Spearman correlation coefficient between score change and patient satisfaction.

* $P < .001$.

† $P < .05$.

of observer bias.[21] Among the existing elbow-related scoring systems, self-assessment questionnaires such as the OES and Patient-Rated Elbow Evaluation assess only pain in their symptom domains and rely on only highly subjective VAS scores. Physician-assessed systems such as the Broberg and Morrey rating systems and the Hospital for Special Surgery Assessment Scale have the deficiency of using only objective parameters to assess elbow function, and the MEPS includes only extension-flexion ROM, neglecting forearm rotation. The American Shoulder and Elbow Surgeons–Elbow (ASES-E) questionnaire is an elbow-specific and both patient- and physician-completed questionnaire; however, with more than 50 responses, it is somewhat unwieldy.

Our study group consisted of 73 patients (mean age, 37 years) with a male-female ratio of almost 1:1, comparable with other validation studies' numbers of patients, ages, and sexes.[9,34] The minimum sample size of respondents while developing a new scoring system is 3 times the number of items in the system.[2] Therefore, with a total of 13 items, a total sample size of 73 was considered sufficient. The presented patients represent a wide spectrum of elbow stiffness etiologies, including varied fracture patterns and atraumatic factors (Table II); this heterogeneous collection allows for a universal validation.

The statistical evaluation for the SHEDS included the assessment of test-retest reliability, internal consistency, validity, responsiveness, and floor and ceiling effects. The assessment of test-retest reliability resulted in ICCs of at least 0.81, which indicates positive reliability. In the current literature, the optimal time for the retest assessment is missing, but in most cases, a time period of 1 or 2 weeks is considered appropriate.[41] In addition, a CA of 0.74 represents high internal consistency for the SHEDS, with the highest value of 0.78 (for the EMC portion) not exceeding the value of 0.95 that might indicate item redundancy.[42] In the literature, no gold standard exists for comparison of construct validity between elbow scores. Therefore, the decision was made to correlate the SHEDS with previously validated scoring systems (DASH score, OES, MEPS, and SF-36 score). Our innovative 8-item activity form (EMC portion) correlated well with the DASH score, OES, and MEPS, with SCCs of at least 0.65, as well as the total score (≥0.51) and the ERS portion (≥0.35), which indicated good to excellent validity.[21,38] These 8 activities, with 3 levels of severity, are not only independent but also complementary to one another, which makes them appropriate for evaluation of full elbow mobility. Although pain comprised only 15% of the SHEDS, in clinical practice, the main concern of patients is movement restoration rather than pain relief, and pain in the SHEDS correlated well with the DASH score (0.57), OES (0.43), and SF-36 score (0.47), all validated as strongly correlated with pain levels.[9,39] In the current literature, various statistics to determine responsiveness are available; however, the best method remains unestablished.[1] Thorborg et al[43] proved that the determination of the ES and SRM in addition to the Global Perceived Effect Score is a considerable improvement in assessing responsiveness. It has also been suggested that distribution-based methods are solely useful when suitable external anchors are not available for use, and anchor-based methods should be given greater weight than distribution-based methods for converging on a single value.[32] However, no reported gold standard of external anchors has been established for our data. Under this condition, distribution-based methods such as the ES and SRM are still the first choice for validation.[9,20] In this study, the ES, SRM, and correlations between final scores, as well as score changes of the SHEDS, and patient satisfaction levels were calculated to assess responsiveness. A large value of the ES and SRM was found in total scores (ES, 3.48; SRM, 2.96), the EMC portion (ES, 2.54; SRM, 2.08), and the ERS portion (ES, 1.26; SRM, 1.14), with significant positive correlations (ES, 0.62; SRM, 0.42) between final scores, as well as score changes of the SHEDS, and patient satisfaction levels.

This study had some weaknesses. First, it was a single-center study; a cross-cultural adaptation of the SHEDS into other languages and determination of its clinometric properties should be conducted before it can be used worldwide. Second, anchor-based analysis methods could be applied for responsiveness if suitable external anchors were identified. Third, the etiologies of elbow stiffness include trauma, synovitis, primary osteoarthritis, and congenital anomalies, whereas the vast majority of our patients (97%) had post-traumatic elbow stiffness.

Conclusions

On the basis of the present data, our results suggest that the newly developed SHEDS may prove to be an excellent, comprehensive, valid scoring system to evaluate joint function in patients with elbow stiffness. Further research with a larger population, from multiple clinical centers, with multiple etiologies, and using different external anchors in a prospective way is required. In addition, the applicability of the SHEDS to other elbow surgical procedures such as open reduction and internal fixation, arthroplasty, and joint replacement could be validated in the future.

Acknowledgment

This study was supported by Project of Key Discipline Group of Shanghai Pudong New Area Health and Family Planning Commission (PWZxq2017-03), Three-year Action Plan of Clinical Skills and Innovation of Shanghai Hospital Development Center (16CR3029A), and Project of Medical Guidance of Foundation of Shanghai Committee of Science and Technology (17411966900).

Disclaimer

The authors, their immediate families, and any research foundations with which they are affiliated have not received any financial payments or other benefits from any commercial entity related to the subject of this article.

References

1. Angst F. The new COSMIN guidelines confront traditional concepts of responsiveness. BMC Med Res Methodol 2011;11:152, author reply 52. http://dx.doi.org/10.1186/1471-2288-11-152
2. Barrett P, Kline P. The observation to variable ratio in factor analysis; Personality Study & Group Behaviour 1981;1:23-33.
3. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. J Hand Ther 2001;14:128-46.
4. Cai J, Wang W, Yan H, Sun Y, Chen W, Chen S, et al. Complications of open elbow arthrolysis in post-traumatic elbow stiffness: a systematic review. PLoS One 2015;10:e0138547. http://dx.doi.org/10.1371/journal.pone.0138547
5. Cai J, Zhou Y, Chen S, Sun Y, Yuanming O, Ruan H, et al. Ulnar neuritis after open elbow arthrolysis combined with ulnar nerve subcutaneous transposition for post-traumatic elbow stiffness: outcome and risk factors. J Shoulder Elbow Surg 2016;25:1027-33. http://dx.doi.org/10.1016/j.jse.2016.01.013
6. Capuano L, Poulain S, Hardy P, Longo UG, Denaro V, Maffulli N. No correlation between physicians administered elbow rating

systems and patient's satisfaction. J Sports Med Phys Fitness 2011;51:255-9.
7. Chen NC, Liu A. The role of prophylactic ulnar nerve release during elbow contracture release. J Hand Surg Am 2014;39:1837-9. http://dx.doi.org/10.1016/j.jhsa.2014.04.009
8. Chen W, Wang W, Li Z, Qian Y, Song J, Liu J, et al. Effect on muscle strength of the upper extremities after open elbow arthrolysis. J Shoulder Elbow Surg 2017;1:63-71. http://dx.doi.org/10.1016/j.jses.2017.06.006
9. Dawson J, Doll H, Boller I, Fitzpatrick R, Little C, Rees J, et al. The development and validation of a patient-reported questionnaire to assess outcomes of elbow surgery. J Bone Joint Surg Br 2008;90:466-73. http://dx.doi.org/10.1302/0301-620x.90b4.20290
10. Doornberg JN, Bosse T, Cohen MS, Jupiter JB, Ring D, Kloen P. Temporary presence of myofibroblasts in human elbow capsule after trauma. J Bone Joint Surg Am 2014;96:e36. http://dx.doi.org/10.2106/jbjs.m.00388
11. Doornberg JN, Ring D, Fabian LM, Malhotra L, Zurakowski D, Jupiter JB. Pain dominates measurements of elbow function and health status. J Bone Joint Surg Am 2005;87:1725-31. http://dx.doi.org/10.2106/jbjs.d.02745
12. Elliott TE, Renier CM, Palcher JA. Chronic pain, depression, and quality of life: correlations and predictive value of the SF-36. Pain Med 2003;4:331-9.
13. Evans JP, Smith CD, Fine NF, Porter I, Gangannagaripalli J, Goodwin VA, et al. Clinical rating systems in elbow research—a systematic review exploring trends and distributions of use. J Shoulder Elbow Surg 2018;27:e98-106. http://dx.doi.org/10.1016/j.jse.2017.12.027
14. Guillemin F. Cross-cultural adaptation and validation of health status measures. Scand J Rheumatol 1995;24:61-3.
15. 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;46:1417-32.
16. Hassan EA, Jenkyn TR, Dunning CE. Direct comparison of kinematic data collected using an electromagnetic tracking system versus a digital optical system. J Biomech 2007;40:930-5. http://dx.doi.org/10.1016/j.jbiomech.2006.03.019
17. Henmi S, Yonenobu K, Masatomi T, Oda K. A biomechanical study of activities of daily living using neck and upper limbs with an optical three-dimensional motion analysis system. Mod Rheumatol 2006;16:289-93. http://dx.doi.org/10.1007/s10165-006-0499-x
18. Hildebrand KA, Zhang M, Befus AD, Salo PT, Hart DA. A myofibroblast-mast cell-neuropeptide axis of fibrosis in post-traumatic joint contractures: an in vitro analysis of mechanistic components. J Orthop Res 2014;32:1290-6. http://dx.doi.org/10.1002/jor.22676
19. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989;27:S178-89.
20. Kim SJ, Basur MS, Park CK, Chong S, Kang YG, Kim MJ, et al. Crosscultural adaptation and validation of the Korean Version of the New Knee Society Knee Scoring System. Clin Orthop Relat Res 2017;475:1629-39. http://dx.doi.org/10.1007/s11999-017-5307-8
21. Longo UG, Franceschi F, Loppini M, Maffulli N, Denaro V. Rating systems for evaluation of the elbow. Br Med Bull 2008;87:131-61. http://dx.doi.org/10.1093/bmb/ldn023
22. Mahomed N, Gandhi R, Daltroy L, Katz JN. The self-administered patient satisfaction scale for primary hip and knee arthroplasty. Arthritis 2011;2011:591253. http://dx.doi.org/10.1155/2011/591253
23. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? Qual Life Res 1995;4:293-307.
24. Morrey BF, Adams RA. Semiconstrained arthroplasty for the treatment of rheumatoid arthritis of the elbow. J Bone Joint Surg Am 1992;74:479-90.
25. Morrey BF, Askew LJ, Chao EY. A biomechanical study of normal functional elbow motion. J Bone Joint Surg Am 1981;63:872-7.
26. Noble PC, Scuderi GR, Brekke AC, Sikorskii A, Benjamin JB, Lonner JH, et al. Development of a new Knee Society scoring system. Clin Orthop Relat Res 2012;470:20-32. http://dx.doi.org/10.1007/s11999-011-2152-z
27. Park MJ, Kim HG, Lee JY. Surgical treatment of post-traumatic stiffness of the elbow. J Bone Joint Surg Br 2004;86:1158-62.
28. Peolsson A, Vavruch L, Hedlund R. Long-term randomised comparison between a carbon fibre cage and the Cloward procedure in the cervical spine. Eur Spine J 2007;16:173-8. http://dx.doi.org/10.1007/s00586-006-0067-2
29. Pieniazek M, Chwala W, Szczechowicz J, Pelczar-Pieniazek M. Upper limb joint mobility ranges during activities of daily living determined by three-dimensional motion analysis—preliminary report. Ortop Traumatol Rehabil 2007;9:413-22.
30. Ranganathan K, Loder S, Agarwal S, Wong VW, Forsberg J, Davis TA, et al. Heterotopic ossification: basic-science principles and clinical correlates. J Bone Joint Surg Am 2015;97:1101-11. http://dx.doi.org/10.2106/jbjs.n.01056
31. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61:102-9. http://dx.doi.org/10.1016/j.jclinepi.2007.03.012
32. Revicki DA, Erickson PA, Sloan JA, Dueck A, Guess H, Santanello NC. Interpreting and reporting results based on patient-reported outcomes. Value Health 2007;10(Suppl 2):S116-24. http://dx.doi.org/10.1111/j.1524-4733.2007.00274.x
33. Sardelli M, Tashjian RZ, MacWilliams BA. Functional elbow range of motion for contemporary tasks. J Bone Joint Surg Am 2011;93:471-7. http://dx.doi.org/10.2106/jbjs.i.01633
34. Sathyamoorthy P, Kemp GJ, Rawal A, Rayner V, Frostick SP. Development and validation of an elbow score. Rheumatology (Oxford) 2004;43:1434-40. http://dx.doi.org/10.1093/rheumatology/keh367
35. Scuderi GR, Bourne RB, Noble PC, Benjamin JB, Lonner JH, Scott WN. The new Knee Society Knee Scoring System. Clin Orthop Relat Res 2012;470:3-19. http://dx.doi.org/10.1007/s11999-011-2135-0
36. Shahgholi L, Bengtson KA, Bishop AT, Shin AY, Spinner RJ, Basford JR, et al. A comparison of manual and quantitative elbow strength testing. Am J Phys Med Rehabil 2012;91:856-62. http://dx.doi.org/10.1097/PHM.0b013e31825f14f9
37. Shuai C, Hede Y, Shen L, Yuanming O, Hongjiang R, Cunyi F. Is routine ulnar nerve transposition necessary in open release of stiff elbows? Our experience and a literature review. Int Orthop 2014;38:2289-94. http://dx.doi.org/10.1007/s00264-014-2465-0
38. Smith MV, Calfee RP, Baumgarten KM, Brophy RH, Wright RW. Upper extremity-specific measures of disability and outcomes in orthopaedic surgery. J Bone Joint Surg Am 2012;94:277-85. http://dx.doi.org/10.2106/jbjs.j.01744
39. SooHoo NF, McDonald AP, Seiler JG III, McGillivary GR. Evaluation of the construct validity of the DASH questionnaire by correlation to the SF-36. J Hand Surg Am 2002;27:537-41.
40. Sun Z, Xiong H, Fan C. Impact of different glucose metabolism status on clinical outcomes of open arthrolysis for post-traumatic elbow stiffness. J Shoulder Elbow Surg 2018;27:1072-7. http://dx.doi.org/10.1016/j.jse.2018.01.015
41. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34-42. http://dx.doi.org/10.1016/j.jclinepi.2006.03.012
42. The B, Reininga IH, El Moumni M, Eygendaal D. Elbow-specific clinical rating systems: extent of established validity, reliability, and responsiveness. J Shoulder Elbow Surg 2013;22:1380-94. http://dx.doi.org/10.1016/j.jse.2013.04.013
43. Thorborg K, Holmich P, Christensen R, Petersen J, Roos EM. The Copenhagen Hip and Groin Outcome Score (HAGOS): development and validation according to the COSMIN checklist. Br J Sports Med 2011;45:478-91. http://dx.doi.org/10.1136/bjsm.2010.080937
44. Wu G, van der Helm FC, Veeger HE, Makhsous M, Van Roy P, Anglin C, et al. ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—part II: shoulder, elbow, wrist and hand. J Biomech 2005;38:981-92.