

Development and Validation of a Simplified Chinese Version and a Face-Scale Version of the Oxford Shoulder Score

A 2-Center Prospective Study

Wu Xu,* MD, Kailun Wu,[†] MD, Stephen Roche,[‡] MD, PhD, Weili Fu,[§] MD, PhD, Lixin Huang,* MD, PhD, Michael F.G. Held,[‡] MD, PhD, and Jiong Jiong Guo,*^{||} MD, PhD

Investigation performed at the Department of Orthopedics, The First Affiliated Hospital of Soochow University, Suzhou, PR China

Background: There has not yet been a pictorial version of a patient-reported outcome measure for shoulder pain.

Purpose: To translate the English version of the Oxford Shoulder Score (OSS) to a simplified Chinese version (SC-OSS) and to validate a new face-scale version of the OSS (FS-OSS), while investigating cross-cultural adaptation, validation, and reproducibility of both versions in patients with shoulder pain.

Study Design: Cohort study (diagnosis); Level of evidence, 2.

Methods: The translation and cross-cultural adaptation of the SC-OSS was performed using a forward-backward translation method. The FS-OSS was developed on the basis of the SC-OSS, using the Wong-Baker FACES Pain Rating Scale for reference. Participants were asked to complete the SC-OSS, FS-OSS, Simple Shoulder Test (SST), Constant-Murley score (CMS), and 36-Item Short Form Health Survey (SF-36). Validation and reproducibility were tested by calculating Cronbach α values for internal consistency as well as by intraclass correlation coefficients. Time needed to complete the scores was used to test cross-cultural adaption.

Results: A total of 312 respondents participated in the research and completed all outcome measures. The internal consistency was strong, with a Cronbach α of .94 and .91 for the FS-OSS and SC-OSS, respectively. High intraclass correlation coefficient values for the FS-OSS score (0.95) and SC-OSS (0.92) were obtained, which indicated excellent test-retest reliability. The Pearson correlation coefficients of the SC-OSS and FS-OSS with the SST ($r = 0.67$ and 0.65 , respectively), CMS ($r = 0.62$ and 0.66 , respectively), and SF-36 ($r = 0.52$ and 0.57 , respectively) indicated good construct validity. The time needed to complete the FS-OSS was less than that needed for the SC-OSS and SST.

Conclusion: The FS-OSS and SC-OSS were validated as reliable instruments for patients with shoulder pain. For Chinese patients, the face-scale version was easier to understand than the cross-cultural text version.

Keywords: Oxford Shoulder Score (OSS); cross-cultural adaptation; face scale; validation

Shoulder pain is the third most common musculoskeletal problem encountered in orthopaedic practice after low back pain and neck pain.²⁵ The Oxford Shoulder Score (OSS) is an internationally recognized assessment instrument to assess the pain perception and quality of life in patients with shoulder pain and to evaluate the effectiveness of different treatments.^{23,24} After its original English version was published in 1996,⁶ it has been translated into many languages in different cultural settings to communicate

internationally.^{2,3,8-12,17,21,22,29-32} In 2015, a simplified Chinese version of OSS (SC-OSS) was developed by Xu et al.³⁹

Even so, high illiteracy and low income in the developing countries may limit the applicability of the OSS. Facial expression drawings (face scales) are a popular method of assessing pain severity in pediatric populations; they are easier to understand and more suitable for children and illiterate people. The Wong-Baker FACES Pain Rating Scale (Wong-Baker scale)³⁸ is one of several scales that have been used for pain assessment in multiple pediatric settings.^{4,5}

To evaluate shoulder function in Chinese patients, we translated the English version of the OSS to a simplified

The Orthopaedic Journal of Sports Medicine, 9(9), 23259671211023751
DOI: 10.1177/23259671211023751
© The Author(s) 2021

This open-access article is published and distributed under the Creative Commons Attribution - NonCommercial - No Derivatives License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits the noncommercial use, distribution, and reproduction of the article in any medium, provided the original author and source are credited. You may not alter, transform, or build upon this article without the permission of the Author(s). For article reuse guidelines, please visit SAGE's website at <http://www.sagepub.com/journals-permissions>.

TABLE 1
Questions From the Original OSS^a

1)	During the past 4 weeks, how would you describe the worst pain you had from your shoulder? <i>None/Mild/Moderate/Severe/Unbearable</i>
2)	During the past 4 weeks, have you had any trouble dressing yourself because of your shoulder? <i>No trouble at all/A little bit of trouble/Moderate trouble/Extreme difficulty/Impossible to do</i>
3)	During the past 4 weeks, have you had any trouble getting in or out of a car or using public transport because of your shoulder? <i>No trouble at all/A little bit of trouble/Moderate trouble/Extreme difficulty/Impossible to do</i>
4)	During the past 4 weeks, have you had been able to use a knife and fork at the same time? <i>Yes, easily/With little difficulty/With moderate difficulty/With extreme difficulty/No, impossible</i>
5)	During the past 4 weeks, could you do household shopping on your own? <i>Yes, easily/With little difficulty/With moderate difficulty/With extreme difficulty/No, impossible</i>
6)	During the past 4 weeks, could you carry a tray containing a plate of food across a room? <i>Yes, easily/With little difficulty/With moderate difficulty/With extreme difficulty/No, impossible</i>
7)	During the past 4 weeks, could you brush/comb your hair with the affected arm? <i>Yes, easily/With little difficulty/With moderate difficulty/With extreme difficulty/No, impossible</i>
8)	During the past 4 weeks, how would you describe the pain you usually had from your shoulder? <i>None/Very mild/Mild/Moderate/Severe</i>
9)	During the past 4 weeks, could you hang your clothes up in a wardrobe, using the affected arm? <i>Yes, easily/With little difficulty/With moderate difficulty/With extreme difficulty/No, impossible</i>
10)	During the past 4 weeks, have you been able to wash and dry yourself under both arms? <i>Yes, easily/With little difficulty/With moderate difficulty/With extreme difficulty/No, impossible</i>
11)	During the past 4 weeks, how much has pain from your shoulder interfered with your usual work (including housework)? <i>Not at all/A little bit/Moderately/Greatly/Totally</i>
12)	During the past 4 weeks, have you been troubled by pain from your shoulder in bed at night? <i>No nights/Only 1 or 2 nights/Some nights/Most nights/Every night</i>

^aOSS, Oxford Shoulder Score. (From Dawson et al.⁶)

Chinese version (SC-OSS) by partially referencing the Xu et al version,³⁹ and we also developed a face-scale version of the OSS (FS-OSS). The purpose of this study was to validate the SC-OSS and FS-OSS using cross-cultural adaptation, validation, and reproducibility of these 2 versions in Chinese patients with shoulder pain. We hypothesized that both versions of the OSS could overcome language and literacy barriers.

METHODS

Translation and Cultural Adaptation Procedure

All study participants signed informed consent forms before inclusion, and the clinical research ethics committee of our hospital approved the study protocol. The translation and cross-cultural adaptation of the SC-OSS (Appendix 1) was performed according to the guidelines reported by

Beaton et al,¹ using a “forward-backward translation” method. First, the original OSS (Table 1) was translated by an expert committee that consisted of 1 orthopaedic surgeon (J.J.G.), 1 rehabilitation physician, 1 physical therapist, and 1 language expert, and an initial Chinese-version OSS was created. In this process, some of the translation referenced a previous study.³⁹ After the first process, a mediation session between the expert committee and our research team was performed to obtain a modified Chinese version. Backward translation was made by a native English speaker and the research team, then the original version and modified Chinese version were compared.

In the cognitive debriefing step, a cohort of 6 patients with shoulder pain (proficient in Mandarin Chinese) was asked to test the initial version, after which the SC-OSS was finalized. The FS-OSS was then created on the basis of the SC-OSS. A flowchart of the translation and testing procedures is shown in Figure 1.

^{||}Address correspondence to Jiong Jiong Guo, MD, PhD, Department of Orthopedics, The First Affiliated Hospital of Soochow University, Suzhou 215006, PR China (email: drjguo@163.com or guojiongjiong@suda.edu.cn).

^{*}Department of Orthopedics, The First Affiliated Hospital of Soochow University, Suzhou, PR China.

[†]Department of Orthopedics, Dushu Lake Hospital Affiliated to Soochow University, Suzhou, PR China.

[‡]Department of Orthopaedic Surgery, Groote Schuur Hospital and Red Cross War Memorial Children’s Hospital, University of Cape Town, Cape Town, South Africa.

[§]Department of Orthopedics, West China Hospital, Sichuan University, Chengdu, PR China.

W.X. and K.W. contributed equally to this article.

Final revision submitted January 27, 2021; accepted February 24, 2021.

The authors declared that there are no conflicts of interest in the authorship and publication of this contribution. AOSSM checks author disclosures against the Open Payments Database (OPD). AOSSM has not conducted an independent investigation on the OPD and disclaims any liability or responsibility relating thereto.

Ethical approval for this study was obtained from The First Affiliated Hospital of Soochow University.

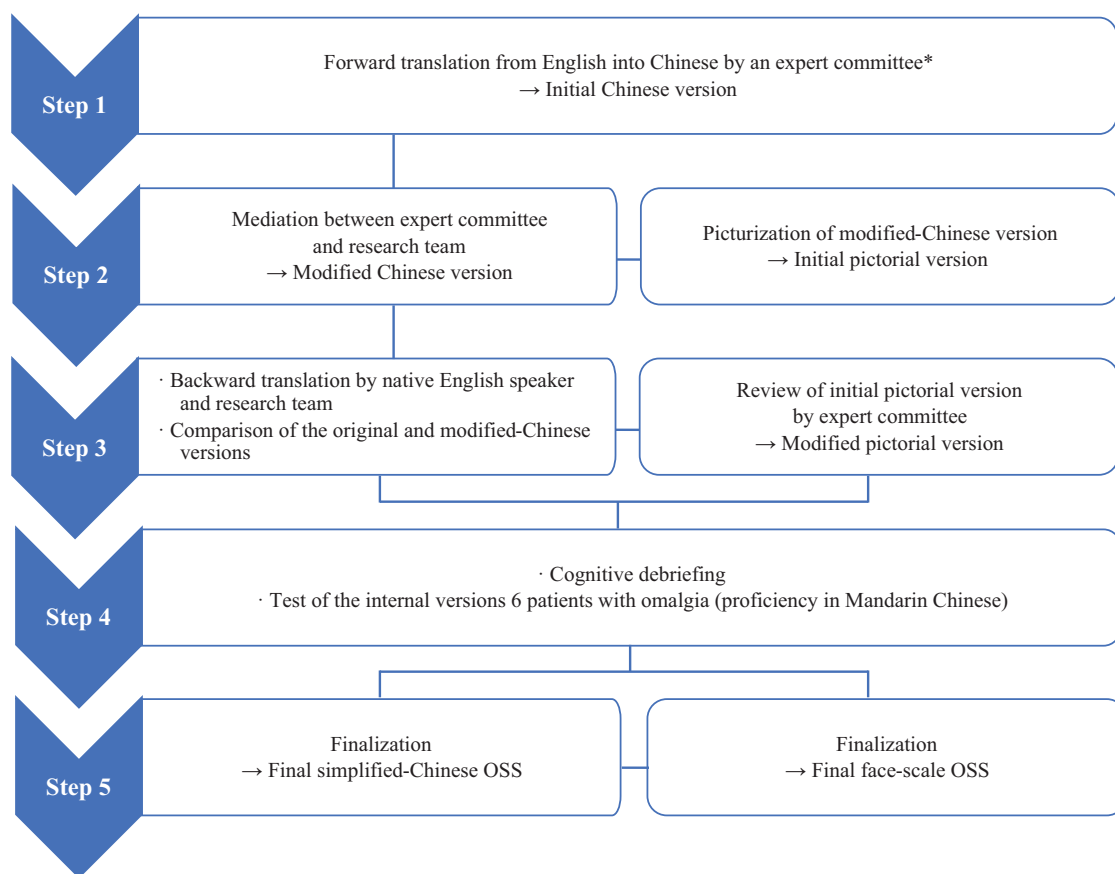


Figure 1. Flowchart of the translation and cultural adaptation process of the OSS from English into SC-OSS and face-scale version. *Expert committee consisting of 1 orthopaedic surgeon, 1 rehabilitation physician, 1 physical therapist, and 1 language expert. OSS, Oxford Shoulder Score; SC-OSS, simplified Chinese version of OSS.

Participants

The finalized SC-OSS and FS-OSS were administered to a sample of 312 consecutive patients affected by shoulder pain who visited the outpatient clinic of The First Affiliated Hospital of Soochow University and West China Hospital, Sichuan University from April 2016 to December 2017. Patients with glenohumeral instability and fracture of the shoulder were excluded. Patient characteristics, including age, age groups, sex, affected side, duration of symptoms, education, and clinical diagnosis, were collected during the first visit to our outpatient clinic.

Psychometric Assessments

All participants were asked to complete the SC-OSS, FS-OSS, Simple Shoulder Test (SST), Constant-Murley score (CMS), and 36-Item Short Form Health Survey (SF-36) during their first visit. Within an interval of 5 to 7 days after the first visit, they completed the SC-OSS and FS-OSS for a second time to evaluate test-retest reliability. The time required to complete the questionnaires and any difficulty encountered in answering a question were recorded. For illiterate or low-literate participants, researchers were available to

help with the questionnaires if needed; these situations were also recorded to assess the ability of communication.

OSS and SC-OSS. The OSS is a valid and reliable questionnaire for shoulder injuries (excluding instability) and comprises 12 items on pain and disability (see Table 1). Each question has 5 possible responses with scores from 0 (worst pain and maximal limitation) to 4 (no pain and no functional limitation), for a total possible score of 48.³⁹ The SC-OSS as developed for this study is shown in Appendix 1.

FC-OSS Scoring. After developing the SC-OSS, a pictorial version was developed using the Wong-Baker scale³⁸ as a reference (Moola et al, unpublished data, [2020]). The questions, possible responses, item scores, and total score were modeled on the OSS with corresponding pictures placed after each question. Each question had a scale of 5 facial options from no pain (4 points) to worst pain (0 points). Lower scores indicate higher levels of pain and disability. The FC-OSS is shown in Chinese and English in Appendices 2 and 3.

Simple Shoulder Test. The SST is an internationally used, simple self-report questionnaire, which was developed for measuring functional limitations of the affected shoulder in patients with shoulder dysfunction.¹⁸ It consists of 12 questions (yes = 1/no = 0). The items of the SST are about function-related pain (2 items), function/strength

(7 items), and range of motion (3 items). The total scores range from 0 (worst function) to 12 (best function).

TABLE 2
Participant Characteristics^a

Characteristics	Total Sample (N = 312)	Male (n = 147)	Female (n = 165)
Age, y, mean ± SD	52.2 ± 10.8	51.0 ± 10.8	53.3 ± 10.8
Age group			
<30 y	26 (8)	14 (10)	12 (7)
30-49 y	121 (39)	57 (39)	64 (39)
50-70 y	119 (38)	53 (36)	66 (40)
>70 y	46 (15)	23 (16)	23 (14)
Affected side			
Right	171 (55)	81 (55)	90 (55)
Left	115 (37)	51 (35)	64 (39)
Bilateral	24 (8)	13 (9)	11 (7)
Duration of symptoms			
<3 mo	92 (29)	42 (29)	50 (30)
3 to 6 mo	95 (30)	47 (32)	48 (29)
6 to 12 mo	75 (24)	36 (24)	39 (24)
>1 y	50 (16)	22 (15)	28 (17)
Education level			
Illiterate	22 (7)	5 (3)	17 (10)
Primary school	91 (29)	21 (14)	70 (42)
Secondary school	156 (50)	92 (63)	64 (39)
Tertiary education	43 (14)	29 (20)	14 (8)
Clinical diagnosis			
Subacromial bursitis/ impingement without RCT	81 (26)	31 (21)	50 (30)
RCT	69 (22)	41 (28)	28 (17)
Biceps tenosynovitis	31 (10)	14 (10)	17 (10)
Frozen shoulder	78 (25)	32 (22)	46 (28)
Osteoarthritis	22 (7)	9 (6)	13 (8)
Calcifying tendinopathy	9 (3)	4 (3)	5 (3)
Other	22 (7)	16 (11)	6 (4)

^aData are reported as n (%) unless otherwise indicated. RCT, rotator cuff tear.

Constant-Murley Score. The CMS is widely used for evaluating the outcomes for the treatment of shoulder disorders and measures pain perception, functional assessment, range of motion, and strength. The Chinese version of the CMS has been found to have excellent validity and reliability.²⁰

36-Item Short Form Health Survey. The SF-36 is a generic questionnaire for assessing the health status of patients, and it consists of 8 domains: physical function (PF), bodily pain (BP), general health (GH), vitality (VT), social function (SF), role-physical (RP), role-emotional (RE), and mental health (MH).³⁶ Scores for each dimension range from 0 (poor health) to 100 (good health). The SF-36 has also been translated and culturally adapted into Chinese.¹⁶

Statistical Analysis

All statistical analyses were performed using SPSS Version 26 (IBM). Descriptive statistics of means, standard deviations, proportions, and percentages were used to describe the baseline characteristics of the participants. The internal consistency was determined by calculating Cronbach α . The intraclass correlation coefficient (ICC) was used to calculate test-retest reliability. The standard error of measurement (SEM) and the minimal detectable change (MDC) were calculated to determine the measurement errors. The Pearson correlation coefficient (r) was used to assess the construct validity. For all analyses, $P < .05$ was considered statistically significant.

RESULTS

Translations and Cross-Cultural Adaptation

After the forward and backward translation process, certain cross-cultural differences were noted and resolved. Chinese usually use chopsticks and spoons; therefore, the “knife and fork” were replaced in question 4. “Bowl” was substituted for “tray” for a similar reason. In addition, considering the education level of the patients, the translators used simple words as much as possible for clear expression.

TABLE 3
Descriptive Analysis of the FS-OSS and SC-OSS^a

Score	Mean ± SD	Cronbach α	ICC (95% CI) ^b	Median (IQR)	SEM (SEM%) ^c	MDC (MDC%) ^d
FS-OSS		.94	0.95 (0.91-0.98)		2(8)	6 (23)
Test	24.23 ± 9.11			24 (18-32)		
Retest	24.04 ± 8.75			24 (17-32)		
SC-OSS		.91	0.92 (0.86-0.95)		3(11)	8 (31)
Test	25.56 ± 9.98			25 (16-33)		
Retest	24.78 ± 9.51			24 (16-32)		

^aFS-OSS, face-scale version of Oxford Shoulder Score; ICC, intraclass correlation coefficient; IQR, interquartile range; MDC, minimum detectable change; SC-OSS, simplified Chinese version of Oxford Shoulder Score; SEM, standard error of measurement; SEM%, error percentage of standard error of measurement.

^bThe first test was conducted at the beginning of this research, and the second test was conducted 5 to 7 days later to calculate test-retest reliability.

^cThe SEM was calculated as $SD \times \sqrt{1 - ICC}$; the SEM% was calculated as $(SEM/mean) \times 100\%$.

^dThe MDC was calculated as $(\sqrt{2} \times 1.96) \times SEM$; the MDC% was calculated as $(MDC/mean) \times 100\%$.

TABLE 4
Internal Consistency of the SC-OSS and FS-OSS^a

Functional Domain Demonstrated by Pictures	SC-OSS		FS-OSS	
	<i>r</i> ^b	α	<i>r</i> ^b	α
Q1: Ability to play sport (worst pain)	0.60	.90	0.73	.93
Q2: Ability to dress independently	0.69	.91	0.78	.94
Q3: Ability to use public transport	0.67	.92	0.72	.96
Q4: Ability to eat independently	0.66	.90	0.69	.94
Q5: Ability to shop independently	0.69	.91	0.64	.93
Q6: Ability to carry a plate of food	0.71	.92	0.74	.94
Q7: Ability to comb own hair	0.68	.91	0.82	.95
Q8: Amount of pain from shoulder	0.58	.90	0.57	.93
Q9: Ability to hang clothes up	0.74	.91	0.81	.94
Q10: Ability to wash under arms	0.75	.92	0.80	.95
Q11: Ability to perform usual work	0.70	.90	0.68	.94
Q12: Ability to sleep comfortably	0.62	.89	0.70	.94

^aFS-OSS, face-scale version of the Oxford Shoulder Score; SC-OSS, simplified Chinese version of Oxford Shoulder Score.

^bPearson correlation coefficient.

TABLE 5
Construct Validity of the SC-OSS and FS-OSS^a

	Pearson Correlation Coefficient (<i>r</i>) ^b	
	SC-OSS	FS-OSS
SC-OSS	—	0.88
FS-OSS	0.88	—
SST	0.67	0.65
CMS	0.62	0.66
SF-36 subscale		
PF	0.71	0.76
RP	0.52	0.49
BP	0.71	0.68
GH	0.35	0.56
VT	0.43	0.54
SF	0.49	0.68
RE	0.37	0.30
MH	0.39	0.50
PCS	0.72	0.70
MCS	0.46	0.52

^aBP, bodily pain; CMS, Constant-Murley score; FS-OSS, face-scale version of the Oxford Shoulder Score; GH: general health; MCS, mental component score; MH mental health; PCS, physical component score; PF, physical function; RE, role-emotional; RP, role-physical; SC-OSS, simplified Chinese version of Oxford Shoulder Score; SF, social function; SF-36, 36-Item Short Form Health Survey; SST: Simple Shoulder Test; VT, vitality. Dashes indicate no data.

^b*r* > 0.80, excellent; 0.61 to 0.80, very good; 0.41 to 0.60, good; 0.21 to 0.40, fair; and 0 to 0.20, poor correlation.

There were no major problems in the graphical translation. Some discrepancies were caused by cultural differences and daily habits. For example, in question 1, “worst pain” in the OSS was interpreted by graphics showing pain during typical Chinese competitive sports, such as badminton and swimming. In question 3 of the OSS, “car” and

TABLE 6
Time Needed to Complete the Scores^a

Education Level	Time Needed to Complete, s		
	FS-OSS	SC-OSS	SST
Illiterate	416 ± 146	724 ± 182	1128 ± 444
Primary school	234 ± 101	298 ± 122	448 ± 190
Secondary school	141 ± 34	143 ± 42	165 ± 44
Tertiary education	101 ± 18	92 ± 16	98 ± 16
Overall	182 ± 107	222 ± 180	306 ± 307

^aData are reported as mean ± SD. FS-OSS, face-scale version of the Oxford Shoulder Score; SC-OSS, simplified Chinese version of Oxford Shoulder Score.

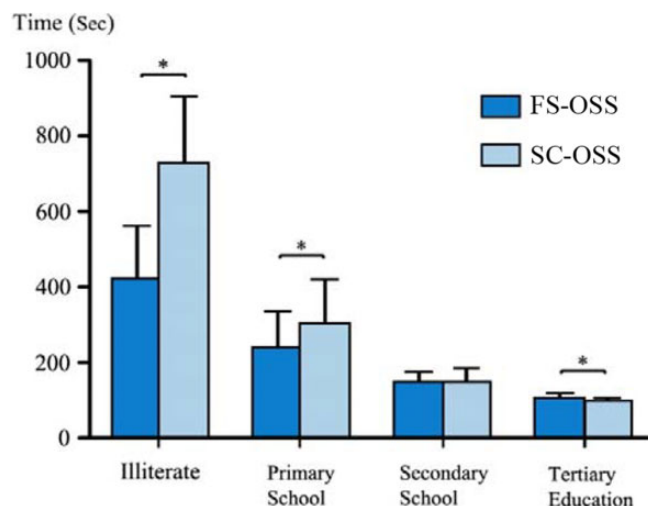


Figure 2. Comparison of the time needed to complete the score of the simplified Chinese version and the face-scale version of the OSS. **P* < .05. FS-OSS, face-scale version of the Oxford Shoulder Score; SC-OSS, simplified Chinese version of Oxford Shoulder Score.

“public transport” have abundant meanings in Chinese. The illustrator used the most representative elements—a Chinese bus and a handhold—which are familiar images for more elderly patients. The term “usual work” in OSS question 11 was represented by illustrations of office work, hoeing, and housekeeping being performed, which were indicative of activities associated with omalgia in the population as a whole.

Participant Characteristics

All 312 participants completed the questionnaires. The average age was 52.2 ± 10.8 years (range, 18-75 years). Educational qualification was divided into 4 levels: tertiary education (43; 14%), secondary school (156; 50%), primary school (91; 29%), and illiterate (22; 7%). The demographic characteristics of the patients are summarized in Table 2.

Ceiling and Floor Effects

There were no missing data for individual items in both the SC-OSS and the FC-OSS. No patient scored the minimum or maximum possible scores; therefore, no significant ceiling or floor effects were observed.

Reliability

The average FS-OSS score was 24.23 ± 9.11 initially and 24.04 ± 8.75 in the retest taken 5 to 7 days later, with the median (interquartile range) being 24 (18-32) and 24 (17-32). The average initial SC-OSS was 25.56 ± 9.98 , and it was 24.78 ± 9.51 during retesting, with the median (interquartile range) values of 25 (16-33) and 24 (16-32), respectively. The Cronbach α was .91 and the ICC was 0.92 (95% CI, 0.86-0.95) for the SC-OSS, proving high internal consistency and excellent test-retest reliability. The Cronbach α was 0.94 and the ICC was 0.95 (95% CI, 0.91-0.98) for the FS-OSS, which indicated better internal consistency and test-retest reliability than the SC-OSS. The measurement errors expressed in SEM were 2 (8%) and 3 (11%) for the FC-OSS and SC-OSS, respectively, while the MDC values were 6 (23%) and 8 (31%), respectively (Table 3). Table 4 shows the Cronbach α distributions.

Construct Validity

Table 5 shows the correlation among the SC-OSS, FS-OSS, SST, CMS, and each domain of the SF-36 to assess the construct validity. The Pearson r showed an excellent relationship between the SC-OSS and FS-OSS ($r = 0.88$). There was a very good correlation between the SC-OSS and the SST ($r = 0.67$) and CMS ($r = 0.62$) as well as the PF ($r = 0.71$), BP ($r = 0.71$), and physical component score (PCS) ($r = 0.72$) sections of the SF-36. The SC-OSS showed good correlation with the RP ($r = 0.52$), VT ($r = 0.43$), and mental component score (MCS) ($r = 0.46$) of the SF-36. Fair correlation was observed between the SC-OSS and the GH ($r = 0.35$), RE ($r = 0.37$), and MH ($r = 0.39$) subscales of the SF-36.

The FS-OSS was marginally better than the SC-OSS, with very good correlation between it and the SST ($r = 0.65$), CMS ($r = 0.66$), and the PF ($r = 0.76$), BP ($r = 0.68$), SF ($r = 0.68$), and PCS ($r = 0.70$) subscales of the SF-36. Good correlation was observed with the RP ($r = 0.49$), GH ($r = 0.56$), VT ($r = 0.54$), MH ($r = 0.50$), and MCS ($r = 0.52$) subscales of the SF-36. The RE subscale ($r = 0.30$) of the SF-36 had the weakest relationship to the FS-OSS (Table 5).

Time to Completion

Table 6 shows the time needed to complete the FS-OSS, SC-OSS, and SST, used as a surrogate marker of comprehension. Patients with a higher education level needed a shorter time to complete the questionnaires. The mean time needed to complete the FS-OSS (182 ± 107 seconds) was notably less than that needed for the SC-OSS (222 ± 180 seconds) and the SST (306 ± 307 seconds). Figure 2 presents

comparisons of the time spent completing the FS-OSS and SC-OSS by education level. Illiterate and primary school groups needed significantly less time to complete the FS-OSS compared with the SC-OSS ($P < .05$), whereas the tertiary education group spent significantly more time completing the FS-OSS ($P < .05$).

DISCUSSION

The present findings are generally consistent with our hypotheses and suggest that FS-OSS and SC-OSS are valid and reliable instruments for patients with shoulder pain in China. The present findings also suggest that the FS-OSS is easier to understand.

Various questionnaires evaluate shoulder symptoms, such as Western Ontario Osteoarthritis of the Shoulder Index,¹⁹ the Shoulder Pain and Disability Index,²⁸ the OSS, and the Western Ontario Rotator Cuff Index.^{14,15} The OSS was originally created to evaluate patients with chronic shoulder complaints by Dawson et al in 1996,⁶ and it has been proven to be a valid and reliable instrument, which is well-accepted and easily completed by patients. It can be used to evaluate most types of shoulder pain, including rotator cuff tears⁷ and frozen shoulder.²⁶ Table 7 summarizes the cross-cultural adaptation versions of the OSS.

TABLE 7
Description of Cross-Cultural Adaptation Versions
for the OSS^a

Language-Population	Year	Sample Size	Time Interval of Test-Retest Reliability, d		Cronbach α
			ICC		
English (original)	1996	111	2	NA	.89
German	2004	94	1-3	0.98	.94
Norwegian	2008	74	7	0.83	.87
Italian	2010	110	2	0.97	.95
Dutch	2010	103	1-3	0.98	.92
Turkish	2011	84	2	0.99	.92
Danish	2011	102	3	0.98	.93
Korean	2012	105	4	0.95	.91
Spanish	2015	120	2	0.97	.95
Persian	2015	100	3	0.93	.93
Simplified Chinese	2015	121	3-5	0.97	.92
Brazilian	2016	30	7-15	0.92	.93
French	2016	144	3	0.91	.93
Portuguese	2018	111	2-4	0.92	.90
Romanian	2018	125	2	0.95	.95
Polish	2019	69	7-14	0.99	.96
Current study					
SC-OSS	2020	312	5-7	0.92	.91
FS-OSS	2020	312	5-7	0.95	.94

^aFS-OSS, face-scale version of the Oxford Shoulder Score; ICC, intraclass correlation coefficient; NA, not available; OSS, Oxford Shoulder Score; SC-OSS, simplified Chinese version of Oxford Shoulder Score.

Translation, cross-cultural adaptation, and validation are necessary for direct international comparisons and cultural equivalence. Currently, the Western Ontario Rotator Cuff Index,³⁴ the Western Ontario Osteoarthritis of the Shoulder Index,¹³ the Rotator Cuff Quality of Life Index,³⁵ and the Shoulder Pain and Disability Index³³ are available for application in Chinese patients, as well as the simplified Chinese version of the OSS developed in 2015.³⁹ Unevenly distributed wealth and educational resources as well as high illiteracy rates are prevailing challenges in most developing countries, and simplified patient-reported outcome measures are needed to overcome some of the resulting cultural and language barriers. In the current study, scores on the SC-OSS and FS-OSS achieved conceptual, semantic, idiomatic, and experiential equivalence during the translation and cross-cultural adaptation process. The short time to complete the newly developed questionnaires suggests they were acceptable to the participants and easily understood.

In terms of reliability, the ICC values of SC-OSS (0.92) and FS-OSS (0.95) assessed at 5 to 7 days were high and indicated excellent test-retest reliability. Our version of the SC-OSS had comparable ICC values at 5 to 7 days when compared with the previously published version³⁹ (0.97 at 3-5 days). The ICC value for the SC-OSS was higher than that of Norwegian (0.83)⁸ and French (0.91)³² versions and equal to the Portuguese (0.92)¹⁰ and Brazilian (0.92)¹⁷ versions. Regarding the FS-OSS, the ICC value was higher than that of the SC-OSS and Persian (0.93)²² versions and comparable with that obtained for the Korean (0.95)²⁹ and Romanian (0.95)¹¹ versions. The test-retest interval (5-7 days) of our study was longer than that of most of the other studies and marginally shorter than versions validated for Brazil (7-15 days)¹⁷ and Poland (7-14 days).² It also showed good internal consistency, with a Cronbach α of .70 to .95. This was higher than the Norwegian (0.87)⁸ and Portuguese (0.90)¹⁰ versions and comparable with the Korean (0.91)⁹ version. For the FS-OSS, the Cronbach α was equal to the German (0.94) version¹² and higher than most versions of the OSS.

The SST, CMS, and SF-36 subscales demonstrated satisfactory construct validity of the SC-OSS and FS-OSS. The SC-OSS showed fair correlation with the GH, RE, and MH subscales of the SF-36 and good correlation with the remaining SF-36 subscales, the SST, and the CMS. Only fair correlation between the FS-OSS and the SF-36 RE subscale was found. According to the results, the SC-OSS and FS-OSS had good convergent and divergent validity.

Pictorial scores have been used in obstetrics and gynecology,²⁷ as well as in health-related quality of life instruments.³⁷ We used the time needed for participants to complete our instruments as a surrogate marker of comprehension. The results suggest that the pictorial version (FS-OSS) was more easily understood than SC-OSS and SST, especially for low-literacy participants, and might therefore be more applicable for countries and regions with high illiteracy rates.

Limitations

Several limitations exist in the present study. First, all participants included in our study were patients who

visited our hospital, which could not fully represent the entirety of our country. Second, the responsiveness of SC-OSS and FS-OSS were not evaluated accurately, as we used the time needed to complete the instruments as a surrogate marker of responsiveness and understanding. Further research and more accurate assessment are still needed. Third, although we divided the participants into 4 groups according to the level of education, the sample size was too small to compare literate versus illiterate patients; this parameter was included mainly to describe the population adequately. Further studies should focus on the comparison of reliability between literate and illiterate populations with a large-powered study sample. In addition, we did not have a standardized way to introduce the image-only scoring system, and all illiterate patients needed explanation. Therefore, investigators may have had an influence on the understanding of the questionnaire by illiterate patients, which may have introduced bias.

CONCLUSION

The FS-OSS and SC-OSS were validated as reliable instruments for patients with shoulder pain. For Chinese patients, the face-scale version was easier to understand than the cross-cultural text version.

ACKNOWLEDGMENT

The authors thank Mr An Liu, medical student at Soochow University School of Medicine, for collecting and analyzing the data for this study. They also acknowledge all the translators and expert committee members (Professor Ling Qin from The Chinese University of Hong Kong, Professor Tiansi Tang, Professor Huilin Yang, and Mr Ju-Sheng Shieh) who assisted with the translation of the OSS into Chinese, the patients who participated in the study, and the illustrators who assisted in the process of making the pictorial version.

REFERENCES

1. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 2000;25:3186-3191. doi:10.1097/00007632-200012150-00014
2. Bejer A, Szczepanik M, Płocki J, Szymczyk D, Kulczyk M, Pop T. Translation, cross-cultural adaptation and validation of the Polish version of the Oxford Shoulder Score in patients undergoing arthroscopic rotator cuff repair. *Health Qual Life Outcomes*. 2019;17:191. doi:10.1186/s12955-019-1263
3. Berendes T, Pilot P, Willems J, Verburg H, te Slaa R. Validation of the Dutch version of the Oxford Shoulder Score. *J Shoulder Elbow Surg*. 2010;19(6):829-836. doi:10.1016/j.jse.2010.01.017
4. Chambers CT, Giesbrecht K, Craig KD, Bennett SM, Huntsman E. A comparison of faces scales for the measurement of pediatric pain: children's and parents' ratings. *Pain*. 1999;83(1):25-35. doi:10.1016/s0304-3959(99)00086-x
5. Chambers CT, Hardial J, Craig KD, Court C, Montgomery C. Faces scales for the measurement of postoperative pain intensity in children following minor surgery. *Clin J Pain*. 2005;21(3):277-285. doi:10.1097/00002508-200505000-00011

6. Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. *J Bone Joint Surg Br.* 1996;78(4):593-600.
7. Dawson J, Hill G, Fitzpatrick R, Carr A. The benefits of using patient-based methods of assessment. Medium-term results of an observational study of shoulder surgery. *J Bone Joint Surg Br.* 2001;83(6):877-882. doi:10.1302/0301-620x.83b6.11316
8. Ekeberg OM, Bautz-Holter E, Tveitå EK, Keller A, Juel NG, Brox JL. Agreement, reliability and validity in 3 shoulder questionnaires in patients with rotator cuff disease. *BMC Musculoskelet Disord.* 2008;9:68. doi:10.1186/1471-2474-9-68
9. Frich LH, Noergaard PM, Brorson S. Validation of the Danish version of Oxford Shoulder Score. *Dan Med Bull.* 2011;58(11): A4335.
10. Gonçalves RS, Caldeira CQ, Rodrigues MV, Felícia SC, Cavalheiro LM, Ferreira PL. Cross-cultural adaptation and validation of the Portuguese version of the Oxford Shoulder Score (OSS). *Acta Reumatol Port.* 2018;43:102-108.
11. Haragus H, Prejbeanu R, Patrascu J, et al. Cross-cultural adaptation and validation of the Romanian Oxford Shoulder Score. *Medicine (Baltimore).* 2018;97:e10926. doi:10.1097/MD.00000000000010926
12. Huber W, Hofstaetter JG, Hanslik-Schnabel B, Posch M, Wurnig C. The German version of the Oxford Shoulder Score—cross-cultural adaptation and validation. *Arch Orthop Trauma Surg.* 2004;124(8):531-536. doi:10.1007/s00402-004-0716-z
13. Jia ZY, Zhang C, Cui J, Xue CC, Xu WD. Translation and validation of the Simplified Chinese version of Western Ontario Osteoarthritis of the Shoulder Index (WOOS). *Medicine (Baltimore).* 2018;97(13):e8691. doi:10.1097/MD.00000000000008691
14. Kirkley A, Alvarez C, Griffin S. The development and evaluation of a disease-specific quality-of-life questionnaire for disorders of the rotator cuff: the Western Ontario Rotator Cuff Index. *Clin J Sport Med.* 2003;13(2):84-92. doi:10.1097/00042752-200303000-00004
15. Lee M, Chen JY, Liow MHL, Chong HC, Chang P, Lie D. Critical shoulder angle and acromial index do not influence 24-month functional outcome after arthroscopic rotator cuff repair. *Am J Sports Med.* 2017;45(13):2989-2994. doi:10.1177/0363546517717947
16. Li L, Wang HM, Shen Y. Chinese SF-36 Health Survey: translation, cultural adaptation, validation, and normalisation. *J Epidemiol Community Health* 2003;57:259-263. doi:10.1136/jech.57.4.259
17. Lima Eda S, Natour J, Moreira E, Jones A. Translation, cultural adaptation and reproducibility of the Oxford Shoulder Score questionnaire for Brazil, among patients with rheumatoid arthritis. *Sao Paulo Med J.* 2016;134(1):40-46. doi:10.1590/1516-3180.2015.00800108
18. Lippitt SB, Harryman DT, Matsen FA. A practical tool for evaluation of function: the Simple Shoulder Test. In: Matsen FA, Fu FH, Hawkins RJ, eds. *The Shoulder: A Balance of Mobility and Stability.* Rosemont: The American Academy of Orthopedic Surgeons;1993:501-559.
19. Lo IK, Griffin S, Kirkley A. The development of a disease-specific quality of life measurement tool for osteoarthritis of the shoulder: the Western Ontario Osteoarthritis of the Shoulder (WOOS) index. *Osteoarthritis Cartilage.* 2001;9(8):771-778. doi:10.1053/joca.2001.0474
20. Ma T, Kao MJ, Lin IH, et al. A study on the clinical effects of physical therapy and acupuncture to treat spontaneous frozen shoulder. *Am J Chin Med.* 2006;34:759-775. doi:10.1142/S0192415X06004272
21. Murena L, Vulcano E, D'Angelo F, Monti M, Cherubino P. Italian cross-cultural adaptation and validation of the Oxford Shoulder Score. *J Shoulder Elbow Surg.* 2010;19(3):335-341. doi:10.1016/j.jse.2009.07.068
22. Naghdi S, Nakhostin Ansari N, et al. Simple Shoulder Test and Oxford Shoulder Score: Persian translation and cross-cultural validation. *Arch Orthop Trauma Surg.* 2015;135(12):1707-1718. doi:10.1007/s00402-015-2330-7
23. Nicholson JA, Searle HKC, MacDonald D, McBirnie J. Cost-effectiveness and satisfaction following arthroscopic rotator cuff repair: does age matter? *Bone Joint J.* 2019;101(7):860-866. doi:10.1302/0301-620X.101B7.BJJ-2019-0215.R1
24. Norman JG, Brealey S, Keding A, Torgerson D, Rangan A. Does time to surgery affect patient-reported outcome in proximal humeral fractures? A subanalysis of the PROFHER randomized clinical trial. *Bone Joint J.* 2020;102(1):33-41. doi:10.1302/0301-620X.102B1.BJJ-2020-0546.R1
25. Ostor AJK, Richards CA, Prevost AT, Hazleman BL, Speed CA. Inter-rater reproducibility of clinical tests for rotator cuff lesions. *Ann Rheum Dis.* 2004;63:1288-1292. doi:10.1136/ard.2003.014712
26. Othman A, Taylor G. Is the constant score reliable in assessing patients with frozen shoulder? 60 shoulders scored 3 years after manipulation under anaesthesia. *Acta Orthop Scand.* 2004;75(1):114-116. doi:10.1080/00016470410001708230
27. Reid PC, Coker A, Coltart R. Assessment of menstrual blood loss using a pictorial chart: a validation study. *BJOG.* 2000;107(3):320-322. doi:10.1111/j.1471-0528.2000.tb13225.x
28. Roach KE, Budiman-Mak E, Songsiridej N, Lertratanakul Y. Development of a shoulder pain and disability index. *Arthritis Care Res.* 1991;4(4):143-149.
29. Roh YH, Noh JH, Kim W, Oh JH, Gong HS, Baek GH. Cross-cultural adaptation and validation of the Korean version of the Oxford shoulder score. *Arch Orthop Trauma Surg.* 2012;132(1):93-99. doi:10.1007/s00402-011-1393-3
30. Torres-Lacomba M, Sánchez-Sánchez B, Prieto-Gómez V, et al. Spanish cultural adaptation and validation of the shoulder pain and disability index, and the oxford shoulder score after breast cancer surgery. *Health Qual Life Outcomes.* 2015;13:63. doi:10.1186/s12955-015-0256-y
31. Tuğay U, Tuğay N, Gelecek N, Özkan M. Oxford Shoulder Score: cross-cultural adaptation and validation of the Turkish version. *Arch Orthop Trauma Surg.* 2011;131(5):687-694. doi:10.1007/s00402-010-1242-9
32. Tuton D, Barbe C, Salmon JH, Dramé M, Nérot C, Ohl X. Transcultural validation of the Oxford Shoulder Score for the French-speaking population. *Orthop Traumatol Surg Res.* 2016;102:555-558. doi:10.1016/j.otsr.2016.05.010
33. Wang W, Jia ZY, Liu J, et al. Cross-cultural adaptation and validation of the Chinese version of the shoulder pain and disability index in patients with symptomatic shoulder pain: a prospective case series. *Medicine (Baltimore).* 2018;97(26):e11227. doi:10.1097/MD.00000000000011227
34. Wang W, Xie QY, Jia ZY, Cui L, Liu D, Wang CR. Cross-cultural translation of the Western Ontario Cuff Index in Chinese and its validation in patients with rotator cuff disorders. *BMC Musculoskelet Disord.* 2017;18(1):178. doi:10.1186/s12891-017-1536-y
35. Wang W, Zhang C, Cui L, Xie QY, Jia ZY, Zheng W. Reliability, validity and responsiveness of the Chinese version of the Rotator Cuff Quality of Life Index (RC-QOL) in patients with rotator cuff disorders. *PLoS One.* 2018;13(11):e0206347. doi:10.1371/journal.pone.0206347
36. Ware JE, Sherbourne . The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care.* 1992;30(6):473-483.
37. Whitehurst DGT, Latimer NR, Kagan A, et al. Developing accessible, pictorial versions of health-related quality-of-life instruments suitable for economic evaluation: a report of preliminary studies conducted in Canada and the United Kingdom. *Pharmacoecoon Open.* 2018;2(3):225-231. doi:10.1007/s41669-018-0083-2
38. Wong-Baker FACES Pain Rating Scale. Wong-Baker FACES Foundation. Accessed January 2020. <http://wongbakerfaces.org>
39. Xu X, Wang F, Wang X, Wei X, Wang Z. Chinese cross-cultural adaptation and validation of the Oxford shoulder score. *Health Qual Life Outcomes.* 2015;13:193. doi:10.1186/s12955-015-0383-5

APPENDIX

术前/术后 左/右

牛津肩关节功能评分 (OSS)

在每一个问题的选框内打√

在过去 4 周内您的肩关节情况				
1、请形容您肩关节最疼痛时的程度，请选择：				
无 <input type="checkbox"/>	轻度 <input type="checkbox"/>	中度 <input type="checkbox"/>	重度 <input type="checkbox"/>	无法忍受 <input type="checkbox"/>
2、肩关节疼痛是否影响您着装（穿衣服、穿鞋等），请选择：				
不影响 <input type="checkbox"/>	稍有困难 <input type="checkbox"/>	困难 <input type="checkbox"/>	非常困难 <input type="checkbox"/>	无法做到 <input type="checkbox"/>
3、肩关节疼痛是否影响您自行乘坐交通工具（如上下车、拉扶手、拉吊环等动作），请选择：				
不影响 <input type="checkbox"/>	稍有困难 <input type="checkbox"/>	困难 <input type="checkbox"/>	非常困难 <input type="checkbox"/>	无法做到 <input type="checkbox"/>
4、肩关节疼痛是否影响您双手使用餐具就餐（如使用筷子、刀叉、汤勺等动作），请选择：				
不影响 <input type="checkbox"/>	稍有困难 <input type="checkbox"/>	困难 <input type="checkbox"/>	非常困难 <input type="checkbox"/>	无法做到 <input type="checkbox"/>
5、肩关节疼痛是否影响您家用购物（如提购物袋、搬物品等动作），请选择：				
不影响 <input type="checkbox"/>	稍有困难 <input type="checkbox"/>	困难 <input type="checkbox"/>	非常困难 <input type="checkbox"/>	无法做到 <input type="checkbox"/>
6、肩关节疼痛是否影响您在室内端餐盘（如从厨房端到餐厅），请选择：				
不影响 <input type="checkbox"/>	稍有困难 <input type="checkbox"/>	困难 <input type="checkbox"/>	非常困难 <input type="checkbox"/>	无法做到 <input type="checkbox"/>
7、肩关节疼痛是否影响您梳头，请选择：				
不影响 <input type="checkbox"/>	稍有困难 <input type="checkbox"/>	困难 <input type="checkbox"/>	非常困难 <input type="checkbox"/>	无法做到 <input type="checkbox"/>
8、肩关节疼痛是否影响您日常生活，请选择：				
不影响 <input type="checkbox"/>	稍有困难 <input type="checkbox"/>	困难 <input type="checkbox"/>	非常困难 <input type="checkbox"/>	无法做到 <input type="checkbox"/>
9、肩关节疼痛是否影响您晾晒或挂衣物，请选择：				
不影响 <input type="checkbox"/>	稍有困难 <input type="checkbox"/>	困难 <input type="checkbox"/>	非常困难 <input type="checkbox"/>	无法做到 <input type="checkbox"/>
10、肩关节疼痛是否影响您洗澡时自行擦洗全身各部位，请选择：				
不影响 <input type="checkbox"/>	稍有困难 <input type="checkbox"/>	困难 <input type="checkbox"/>	非常困难 <input type="checkbox"/>	无法做到 <input type="checkbox"/>
11、肩关节疼痛是否影响您平时工作，请选择：				
不影响 <input type="checkbox"/>	稍有困难 <input type="checkbox"/>	困难 <input type="checkbox"/>	非常困难 <input type="checkbox"/>	无法做到 <input type="checkbox"/>
12、肩关节疼痛是否影响您夜晚睡眠，请选择：				
不影响 <input type="checkbox"/>	偶尔影响 <input type="checkbox"/>	有时影响 <input type="checkbox"/>	经常影响 <input type="checkbox"/>	每晚 <input type="checkbox"/>

最后，请检查您是否完成所有问题，谢谢。

Appendix 1. Simplified Chinese Oxford Shoulder Score as developed for this study.

牛津大学肩关节评分(OSS)

在过去的4周内
你的肩关节情况……

左
右

✓ 在每个问题的选框内打勾

<p>Q1 请形容您肩关节最疼痛时的程度： 请选择</p>  <p><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>	<p>Q4 肩关节疼痛是否影响您双手使用餐具就餐 (如使用刀叉、筷子、汤勺等动作)：请选择</p>  <p><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>
<p>Q2 肩关节疼痛是否影响您着装(如穿衣裤、穿鞋等)：请选择</p>  <p><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>	<p>Q5 肩关节疼痛是否影响您家用购物 (提购物袋、搬物品等动作)：请选择</p>  <p><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>
<p>Q3 肩关节疼痛是否影响您自行乘坐交通工具 (如上下车、拉扶手、拉吊环等动作)： 请选择</p>  <p><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>	<p>Q6 肩关节疼痛是否影响您在室内端餐盘 (如从厨房端到餐厅)：请选择</p>  <p><input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/></p>

Appendix 2. Face-scale Oxford Shoulder Score as developed for this study: Chinese version.

牛津大学肩关节评分 (OSS)

在过去的4周内
你的肩关节情况……

左
右

✓ 在每个问题的选框内打勾

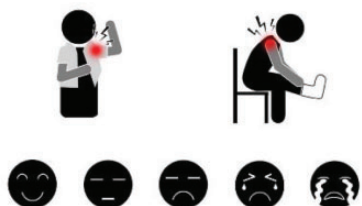
<p>Q7 肩关节疼痛是否影响您梳头；请选择</p>	<p>Q10 肩关节疼痛是否影响您洗澡时自行全身各部位擦洗；请选择</p>
 	 
<p>Q8 肩关节疼痛是否影响您日常生活；请选择</p>	<p>Q11 肩关节疼痛是否影响您平时工作；请选择</p>
 	 
<p>Q9 肩关节疼痛是否影响您晾晒或挂衣物；请选择</p>	<p>Q12 肩关节疼痛是否影响您夜晚睡眠；请选择</p>
 	 

During the past 4 weeks:

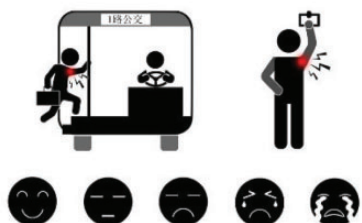
1. How would you describe the worst pain you had from your shoulder?



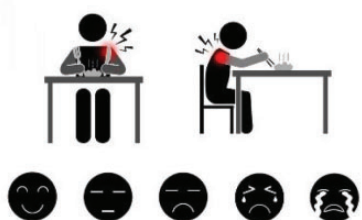
2. Have you had any trouble dressing yourself because of your shoulder?



3. Have you had any trouble getting in and out of a car or using public transport because of your shoulder?



4. Have you been able to use a knife and fork at the same time?

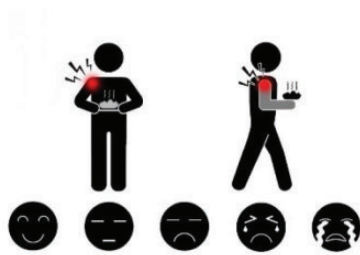


5. Could you do the household shopping on your own?

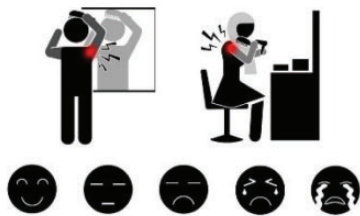


During the past 4 weeks:

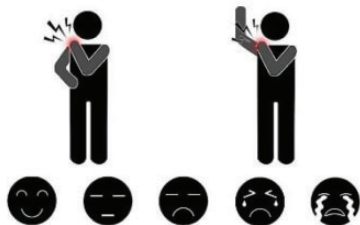
6. Could you carry a tray containing a plate of food across a room?



7. Could you brush/comb your hair with the affected arm?



8. How would you describe the pain you usually had from your shoulder?



9. Could you hang your clothes up in a wardrobe using the affected arm?

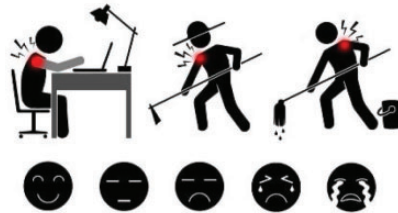


10. Have you been able to wash and dry yourself under both arms?

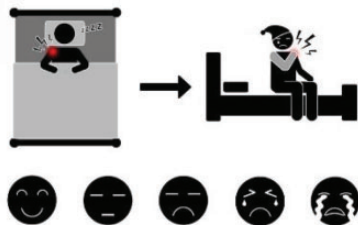


During the past 4 weeks:

11. How much has pain from your shoulder interfered with your usual work (including housework)?



12. Have you been troubled by pain from your shoulder in bed at night?



Appendix 3. (continued).