JSES International 5 (2021) 328–333

Contents lists available at ScienceDirect

JSES International

journal homepage: www.jsesinternational.org

Validation of the cross-cultural adapted Italian version of the Oxford Elbow Score

Sara Padovani, MD [a,b,*], Maurizia Capuzzo, MD [c], Leo Massari, MD [b], Gaetano Caruso, MD [b], Paolo Arrigoni, MD [d,e,f], Carlo Zaolino, MD [d], Davide Cucchi, MD [g], Giorgia Valpiani, MSc, PhD [h], Alessandra Colozza, MD [a]

[a] *Department of Orthopedics and Traumatology, Ospedale degli Infermi di Faenza, Faenza (RA), Italy*
[b] *Orthopedics and Traumatology Unit, Sant Anna University Hospital of Ferrara, Ferrara, Italy*
[c] *Department of Morphology, Experimental Medicine and Surgery, Section of Anesthesia and Intensive Care, Azienda Ospedaliero-Universitaria S. Anna, Ferrara, Italy*
[d] *U.O.C. 1a Clinica Ortopedica, ASST Centro Specialistico Ortopedico Traumatologico Gaetano Pini-CTO, Milan, Italy*
[e] *Laboratory of Applied Biomechanics, Department of Biomedical Sciences for Health, Università degli Studi di Milano, Milan, Italy*
[f] *Research Center for Adult and Pediatric Rheumatic Diseases (RECAP-RD), Department of Biomedical Sciences for Health, Università degli Studi di Milano, Milan, Italy*
[g] *Department of Orthopaedics and Trauma Surgery, Universitätsklinikum Bonn, Bonn, Germany*
[h] *Research Innovation Quality and Accreditation Unit, S. Anna University Hospital of Ferrara, Ferrara, Italy*

ARTICLE INFO

Keywords:
Oxford Elbow Score
scoring system
patient-reported outcome measures
elbow joint
validation study
Italian translation

Level of evidence: Basic Science Study; Validation of Outcome Instruments

Background: The Oxford Elbow Score (OES) is an English-language questionnaire specifically designed to evaluate surgical elbow outcomes. This scoring system has been translated into other languages. Given the lack of an Italian version of the OES, the present study was designed to establish, culturally adapt, and validate the Italian version.

Methods: The OES questionnaire was culturally adapted to Italian patients in accordance with the literature guidelines with a pilot phase including seven patients with elbow problems and seven healthy subjects. The study includes 110 participants from three hospitals, who underwent elbow surgery for acute (70%) or chronic diseases. At least one month after elbow surgery, at the "index visit", the physician completed the Mayo Elbow Performance Index and patients completed the following questionnaires: the Italian OES, the shortened version of the Disability of Arm, Shoulder and Hand Questionnaire (*Quick*DASH) and the Short-Form 36 Health Survey. Internal consistency was evaluated using Cronbach's alpha. Reproducibility was assessed using the intraclass correlation coefficient in ten patients who completed the OES again two-three days after the index visit. Construct validity was assessed using Spearman correlation coefficients. Responsiveness was evaluated in 68 patients who answered the questionnaires four months after the index visit, using the Wilcoxon signed-rank test, the effect size and the standardized response mean calculation.

Results: Cronbach's alpha was excellent: 0.86 (0.82-0.90) for OES pain, 0.92 (0.90-0.94) for OES function, and 0.90 (0.87-0.93) for OES social/psychological. The intraclass correlation coefficient was 0.94 (0.78-0.98) for OES pain, 0.91 (0.71-0.97) for OES function, 0.95 (0.83-0.98) for OES social-psychological and 0.93 (0.76-0.98) for OES total. The Spearman ρ was >0.7 for OES pain and *Quick*DASH, for OES function and both *Quick*DASH and Mayo Elbow Performance Index, and for OES social-psychological and *Quick*-DASH. Regarding responsiveness, the mean of the changes between the two visits ranged from 33.9 for OES pain, to 44 points for OES function and OES social/psychological. The effect size and the standardized response mean were >0.8 for all OES domains.

Conclusion: This study demonstrates that the Italian version of the OES, translated in accordance with the international standardized guidelines, is reliable, valid, and responsive in patients who have undergone elbow surgery.

Institutional review board approval was received from the Coordinator Center: Comitato Etico Area Vasta Emilia Centro (AVEC-CE) (no. 95/2018). Participating Centers: Comitato Etico Romagna (CEROM) (no. 2131) and Comitato Etico Milano Area 2 (no. 1068).

* Corresponding author: Sara Padovani, MD, Department of Orthopedic and Traumatology, Ospedale degli Infermi di Faenza, Viale Stradone 9, 48018 Faenza, Ravenna, Italy.
E-mail address: sara.padovani1987@gmail.com (S. Padovani).

https://doi.org/10.1016/j.jseint.2020.10.025

Clinical rating systems represent the current standards to measure general health, regional, joint- and disease-specific outcomes, and they can be divided into physician-rated and patient-rated questionnaires.[18] Physician-rated questionnaires are objective tools which use clinical and functional measurements, whereas patient-reported outcome measures focus on a patient's subjective experience in relation to a health condition and its therapy. These different perspectives may lead to similar results and should be complementary to assess any clinical outcomes. However, any measuring instrument used should prove to be reliable, valid and responsive (sensitive to change), as well as acceptable to patients. Furthermore, these properties should be tested on reference groups of patients similar to those being studied, thereby ensuring the validity of a tool from both a linguistic and a cultural perspective.[10]

In the field of elbow-specific rating systems, many questionnaires have been developed in recent decades. A systematic review of the literature identified 12 of them, and the authors concluded that the only elbow-specific rating system validated using high-quality methodology is the Oxford Elbow Score (OES).[27] The OES is a patient-administered patient-reported outcome measure, specific for elbow surgery, validated on heterogeneous populations.[6] Some authors demonstrated that a single-assessment numeric evaluation[24] or a subjective elbow value[25] (both ranging from 0 to 100) correlates highly with the OES[24] and the Mayo Elbow Performance Index.[25] Nevertheless, a single indicator may be difficult to interpret and clinicians need to know which dimensions are affected by the pathology to choose a suitable treatment for a specific patient.

Cross-cultural adaptation and validation of the OES have already been done in five languages: Dutch,[7] Danish,[22] Persian,[9] Turkish,[28] and German.[20] An adaptation to the Italian language and culture has not yet been done. The purpose of the present study is to translate and adapt the OES questionnaire to Italian culture and to test it for psychometric proprieties such as reliability, validity, and responsiveness.

Materials and methods

This multicenter, prospective cohort study was performed in the Orthopedic Department of three hospitals located in Northern Italy, from July 2018 to February 2020. The OES copyright owners have granted us the license to use the original version of the OES. The ethics committees approved the study and all participants gave their informed written consent. The study was registered in ClinicalTrial.gov (registry no. NCT03727516).

Translation and adaptation process

The translation and cross-cultural adaptation of the original version of the OES to the Italian language and culture was done in accordance with the current literature guidelines[3] and the standards provided by the Isis Outcomes Translation and Linguistic Validation Process (https://innovation.ox.ac.uk/clinical-outcomes/services/translation-linguistic-validation/).

For the forward translation, two native Italian speakers, one of them being a physician experienced in medical translation, translated the OES into Italian. A consensus committee made up of two orthopedic surgeons, a psychologist, an epidemiologist, a physiatrist, an anesthesiologist and a representative of a patients' association analyzed the two different Italian translations. The committee discussed the translations in accordance with Italian culture, resolved any discrepancies by consensus and created a reconciled version (version 1) of the OES. This consensus version was translated back into English by two others nonmedical professional native English translators. Neither of the translators was aware of the OES original version and of the concepts investigated. No major language problems were found in the forward- and back-translations of the OES. The only minor problem encountered was the different synonyms used by the two different translators, for example, “Ha avuto difficoltà nel sollevare oggetti in casa, come ad esempio gettare la spazzatura, a causa del Suo problema al gomito?”; “putting out the rubbish” could be translated into Italian as “gettare la spazzatura” or “portare fuori la spazzatura” or “buttare la spazzatura”. In terms of cross-cultural adaptation, the consensus committee chose the first option because the term “gettare” recalls an action performed by the elbow joint. At the end of this phase, the consensus committee provided a new consensus version (version 2) which was semantically and grammatically equivalent to the original version of the OES. At this stage, the experts finalized the Italian version of the OES.

The Italian OES version was tested in a pilot phase, where it was administered to seven patients (four women and three men; average age of 68.8 years, standard deviation (SD) 11.4; education 9.8 years, SD 4.5 years) with elbow pathologies, and to seven sex- and age-matched healthy volunteers (four women and three men, average age of 62.7 years, SD 11.8; education 10.1 years, SD 5.0 years). The time taken to answer the Italian version of the OES, the number of missing items and the perceived difficulty, measured using the numerical rating scale (NRS) ranging from 0 to 100, were collected. All the participants completed the questionnaire and the project manager (SP) reviewed the comments, discussed them with the consensus committee and delivered the final version of the OES ready for the present study.

The patients

The orthopedic department of two hospitals began collecting data in July 2018, whereas the orthopedic department of the third hospital started the study in November 2018. All centers finalized patient recruitment in February 2020.

The study involved all adult patients who underwent at least one follow-up visit at the outpatient clinics more than one month after elbow surgery. During that visit, the patients were asked to participate in the study and, if they consented, the visit was considered an “index visit”. The inclusion criteria were as follows: $\geq$18 years of age, ability to speak and understand the Italian language, and to have undergone elbow surgery due to trauma (distal humerus, radial head, proximal ulna fractures, terrible triad, elbow dislocation, distal biceps rupture) or osteoarthritis (post-traumatic and degenerative osteoarthritis, heterotopic ossifications, and stiffness after previous elbow surgery), epicondylitis or rheumatoid arthritis. Five occupational diseases were included, while bilateral elbow disorders and multiple-trauma patients were excluded. The variables collected included the following: age (years), gender, body mass index, education (number of years), employment (employed, retired, unemployed, or student), dominant upper limb and operated limb, diagnosis, dates of surgery, and outpatient follow-up visits.

Those consenting to participate in the study received the patient questionnaires at the time of the index visit. They either personally completed them or were called by the investigator to provide answers by phone, if they preferred this latter option. To assess the

test-retest reliability, ten patients were asked to answer the OES by phone two to three days after the scheduled follow-up visit. Moreover, the patients who had received any prescriptions from the orthopedic specialist and had scheduled a subsequent follow-up visit were requested to answer the study questionnaires again at the new visit, where they received a new copy of the Italian version of the OES.

Research instruments

The OES was originally developed and validated by Dawson et al[6] in 2008 to assess the outcomes of elbow surgery. It is currently the only rating scale that can be recommended with confidence for use in a clinical setting, as stated by The et al.[27] The questionnaire refers to the period of "the past four weeks" and consists of 12 items grouped in the following three domains: pain, function and social/psychological. Each domain includes four items, each of which is scored from 0 to 4, with higher scores denoting greater severity. The total OES ranges from 0 to 48 and the score of 48 (normal elbow) is independent of age, sex, and hand dominance.[14] Individual domain scores can be converted into a metric score of 0-100 (lower scores representing greater severity).[6] Usually, the patient completes a hard-copy questionnaire of the OES on their own. Cronbach's alpha coefficient, used to assess reliability of the OES domains in the original study, was >0.8 for all of them.[6] The minimal clinically important difference was around 10 on average for the OES function scale, and around 18 for both the OES pain and social-psychological scales.[5] The validity of the OES was assessed using good or excellent quality methods and all of its measurement properties were favorable.[27]

The Mayo Elbow Performance Index (MEPI)[19] is an instrument introduced in 1993, used to assess the limitations caused by pathologies of the elbow during daily activities. It is completed by the physician and consists of the following four parts: pain (scored 0, 15, 30, or 45 points), ulnohumeral motion (scored 5, 15, or 20 points), stability (scored 0, 5, or 10 points), and ability to perform five functional tasks (each scored 0 or 5 points). The total score ranges from 5 to 100 points, with lower scores indicating greater severity. The function of the elbow is considered excellent (total score $\geq$ 90), good (75-89), modest (60-74), and poor (<60).

The shortened version of the Disability of Arm, Shoulder and Hand Questionnaire (*Quick*DASH)[13] consists of 11 items used to measure physical function and symptoms in patients with any musculoskeletal disorders of the upper limb. The questionnaire refers to the previous week and each item has five response options (1 no difficulty; 2 mild difficulty; 3 moderate difficulty; 4 severe difficulty; 5 unable). A formula makes it possible to calculate the final score which ranges from 0 (no disability) to 100 (the greatest possible disability). The minimal important change for the *Quick*-DASH is 11.7.[15] We administered the version of *Quick*DASH validated for the Italian context.[21]

The Short-Form 36 Health Survey (SF-36) is a widely used questionnaire to assess health-related quality of life.[29] The questionnaire refers to the four previous weeks, and measures eight dimensions of health reporting a score from 0 (worst) to 100 (best) for each dimension. Higher scores mean better functioning. Specifically, the dimensions are as follows: physical functioning with ten items, role physical (RP) with four items, bodily pain with two items, general health with five items, vitality with four items, social functioning (SF) with two items, role emotional with three items, and mental health with five items. Differences in SF-36 scores of more than five points were reported as clinically meaningful.[30] We administered the validated Italian version of SF-36.[2]

Statistical analyses

Categorical data are presented as a percentage (%). Descriptive statistics including means, SDs and ranges were calculated for the demographic and clinical characteristics of the patients.

A type-I error in two-tailed tests was considered significant ($P < .05$). Statistical analysis was performed using the following software packages: SPSS (released 2011. IBM SPSS Statistics for Windows, Version 20.0; Armonk, NY, USA), STATA (StataCorp. 2015, Stata Statistical Software: Release 14, College Station, TX, USA), and MedCalc (MedCalc for Windows, version 12.7.8.0, MedCalc Software, Ostend, Belgium). The Shapiro-Wilk test showed that variables of OES, MEPI, and *Quick*DASH were not normally distributed. Therefore, nonparametric tests were used.

As far as the reliability of the Italian version of the OES is concerned, we assessed the internal consistency of the questionnaire using Cronbach's alpha coefficient with 95% confidence interval (CI) to analyze the answers given by the patients at the index visit. Cronbach's α was considered acceptable when it exceeded 0.70[26] and excellent when it exceeded 0.80.[1]

To assess the test-retest (intraobserver) reliability, we used the intraclass correlation coefficient (ICC) with 95% CI in the group of patients who answered the OES twice, that is, once at the index visit and once two to three days later. The ICC value was considered excellent when $\geq$ 0.75.[11]

The construct validity was tested by determining the relationship between the Italian version of the OES and the scores of MEPI, *Quick*DASH, and SF-36 questionnaires. The Spearman correlation coefficient (ρ) was used and ρ values were interpreted as high (0.70-0.89) and very high (0.90-1.00).[8]

Responsiveness was evaluated in the patients who had answered the study questionnaires at the index visit and again at a follow-up visit scheduled after any prescriptions. To assess the ability of the tool to find a statistically significant difference, we used the paired Wilcoxon signed-rank test. To assess the ability of the tool to reflect clinically significant changes over time, we calculated the effect size (ES) and the standardized response mean (SRM). The ES was computed as the average difference between the follow-up score and the index visit score (average change score) divided by the SD of the index visit score.[16] The SRM was calculated as the average change score divided by the SD of the change scores. The ES and SRM scores were interpreted as follows: a value of 0.2-0.4 was considered a small effect, 0.5-0.7 a moderate effect, and $\geq$ 0.8 a large effect.[4]

Moreover, we evaluated floor and ceiling effects by calculating the number of participants who had the best or worst OES possible values at the index visit.

Sample size calculation

In accordance with the general rule of thumb, we calculated that Cronbach alpha required ten patients per item, thus suggesting a sample size of 120 patients.[23] At the end of February 2020, the outbreak of COVID-19 strongly affected the country, and especially the regions where the study was performed. Therefore, we had to interrupt the study after the collection of data from 110 patients.

As regards the test-retest reliability of the Italian version of the OES, to detect a correlation coefficient $\geq$ 0.80, with alpha 0.05 and power 0.80, a sample size of ten cases is required. Concerning responsiveness, we used the values reported by Dawson et al (OES function 56.2 $\pm$ 24.8, OES pain 45.4 $\pm$ 25.3, and OES social/psychological 43.3 $\pm$ 23.6) for the calculation.[6] The sample size needed to detect a difference of 20% between values of OES at the two different times, with $P < .05$ and power >0.80, is of 38, 61, and 59

patients, respectively. Considering a dropout of 10% of cases, for this analysis we decided to collect data from 68 patients.

Results

In the pilot phase, which was performed to test the OES Italian consensus version, the average time taken to answer the OES was four minutes and there were no missing items for either of the groups, that is, patients and healthy subjects. Four patients and four healthy subjects out of all the participants rated perceived difficulty as zero (NRS 0-100). The NRS reported by the remaining patients ranged from 30 to 40 and those reported by the remaining healthy subjects ranged from 10 to 20. In addition, the patients did not ask for supplementary instructions and explanations to answer the questions.

One hundred ten patients participated in the study to validate the Italian version of the OES (Fig. 1). They answered the study questionnaires at the index visit, an average of 3.1 months (SD 2.2, median 2 months) after surgery on the elbow. Eighty-two of them (75%) personally completed the hard-copy questionnaires and 28 (25%) answered by phone when called by the investigator. However, the physician completed the MEPI at the time they examined the patient. There were no missing items in the patient questionnaires. The demographic and clinical characteristics of the study participants are reported in Table I.

Regarding the reliability of the Italian OES version, Cronbach's alpha was 0.86 (95% CI 082-0.90) for OES pain, 0.92 (95% CI 0.90-0.94) for OES function, and 0.90 (95% CI 0.87-0.93) for OES social-psychological. The test-retest reliability was assessed in ten patients (six men and four women), with an average age of 54.5 years (SD 21.4, median 61 years), and education 11.6 years (SD 3.5, median 13 years). These patients answered the OES twice, at an interval of two to three days. The ICC was 0.94 (95% CI 0.78-0.98) for OES pain, 0.91 (95% CI 0.71-0.97) for OES function, 0.95 (95% CI 0.83-0.98) for OES social-psychological, and 0.93 (95% CI 0.76-0.98) for OES total.

Construct validity is shown in Table II, where the Spearman correlation coefficient was >0.7 for OES pain and *Quick*DASH, for OES function and both *Quick*DASH and MEPI, and for OES social-psychological and *Quick*DASH. OES pain was highly related to none of the SF-36 domains, OES function with physical functioning, RP, SF, and role emotional domains, and OES social-psychological with RP, bodily pain, and SF domains.

Responsiveness was assessed in 68 patients, for whom the orthopedic specialist, at the index visit, had scheduled a subsequent follow-up visit. The average interval elapsed between the index visit and the follow-up visit was four months (SD 0.8, median four months). Of the 68 patients, 36 were men and 32 women, their average age was 47.8 years (SD 16.4, median 46.5 years), and education 13.5 years (SD 3.0, median 13 years). Forty-nine patients (72%) were employed, 14 (20%) retired, and only five were unemployed. However, 49 patients (72%) used to perform manual activities regularly, regardless of whether they were employed or retired. Globally, 44 patients (64%) had their dominant limb affected. The change in the OES values at the follow-up visit, in comparison with those at the index visit, was much greater than 0.8 for all the OES domains (Table III).

Of the 110 patients considered at the index visit, nobody rated zero value (floor effect) for OES pain and OES function, and six patients (0.5%) rated zero value for OES social-psychological. Fourteen patients (12.7%), 30 (27.2%), and seven patients (0.6%) rated 100 value (ceiling effect) for OES pain, for OES function, and for OES social-psychological, respectively.

Discussion

The present study shows the steps of translation and cross-cultural Italian adaptation of the original version of the OES questionnaire to ensure that respondents understand the questions as intended. We performed this in accordance with the standardized literature guidelines. Moreover, we assessed the measurement properties of the Italian version of the OES to validate it formally.

The internal consistency was assessed by Cronbach's alpha, which is >0.8 for all OES domains. This finding demonstrates excellent reliability, even though the size of the sample (110 patients) is slightly smaller than the one originally planned (120 patients). Moreover, our Cronbach's alpha coefficients are in line with the highest values reported in the literature: OES pain ranges from 0.76[28] to 0.89,[6] OES function from 0.79[28] to 0.95,[9] and OES social-psychological from 0.80[15] to 0.90.[7,20]

The test-retest reliability is high (ICC > 0.9) for all the domains of the Italian version of the OES. The values of ICC reported in literature range from 0.76[9] to 0.99[22] for pain, from 0.79[28] to 0.99[22] for function, and from 0.75[9] to 0.99[22] for social/psychological. Plaschke et al,[22] in the validation study of the Danish version of OES, found the highest values of the ICC for all the subscales. This finding could

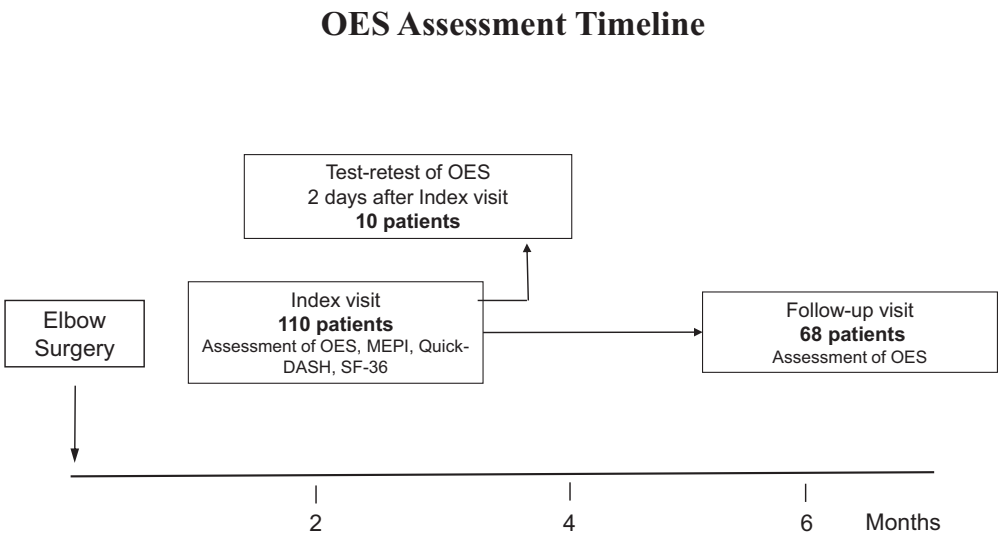

Figure 1 OES timeline assessment. *OES*, Oxford Elbow Score.

Table I
Demographic and clinical characteristics of the patients included in the study

	N	%	Mean ± SD	Range
Patients	110			
Gender				
Female	56	50.9		
Male	54	49.1		
Age (yr)			53.6 ± 18.3	18-87
Body mass index (kg/m^2)			25.1 ± 3.9	17.3-35.1
Education (yr)			12.5 ± 3.4	5-17
Employment				
Employed	62	53.4		
Retired	37	33.6		
Non employed*	11	10.0		
Side affected				
Right	62	56.4		
Left	48	43.6		
Dominant limb affected	65	59.1		
Diagnosis				
Fractures/fracture-dislocations	69	62.7		
Epicondylitis	8	7.3		
Secondary osteoarthritis	12	11		
Distal biceps tendon rupture	8	7.3		
Others	13	11.7		
OES				
Pain			67.9 ± 23.2	19.0-100
Function			62.8 ± 31.9	12.5-100
Social-psychological			50.0 ± 30.4	0-100
MEPI			81.7 ± 15.9	40-100
*Quick*DASH			39.6 ± 26.4	0-86.4

SD, standard deviation; *OES*, Oxford Elbow Score (range 0-100, with higher score better clinical picture); *MEPI*, Mayo Elbow Performance Index (range 5-100, with higher scores indicating better clinical picture); Quick*DASH*, Disability of the Arm, Shoulder and Hand, with range between 0 (no disability) and 100 (the greatest possible disability).
* Nonemployed including students.

Table II
Spearman's correlation coefficients between the three domains of the Oxford Elbow Score (OES), the Mayo Elbow Performance Index (MEPI), the *Quick* Disability of Arm, Shoulder and Hand Questionnaire (*Quick*DASH) total and the domains of Short-Form 36 Health Survey (SF-36)

	OES pain	OES function	OES social-psychological
OES pain			
OES function	0.770		
OES social-psychological	0.735	0.817	
MEPI	0.655	0.710	0.675
*quick*DASH	−0.782	**−0.902**	−0.87
SF-36*			
PF	0.591	0.749	0.627
RP	0.611	0.729	0.736
BP	0.595	0.695	0.767
GH	0.479	0.589	0.512
VT	0.556	0.628	0.588
SF	0.626	0.750	0.748
RE	0.622	0.749	0.692
MH	0.433	0.421	0.497

* SF-36, Short-Form 36 Health Survey: PF, physical functioning; RP, role physical; BP, bodily pain; GH, general health, VT, vitality; SF, social functioning; RE, role emotional; MH, mental health. Correlation coefficients between 0.70 and 0.89 are underlined and those ≥0.90 in bold.

Table III
OESs at the index visit and at the follow-up visit, mean of the changes in scores, effect size, and standardized response mean

	Index visit	Follow-up visit	Mean change	*P* value	Effect size	SRM
OES pain	61.9 (22.7)	95.8 (8.6)	33.9 (19.6)	<.001	1.49	1.72
OES function	51.4 (32.1)	95.4 (10.8)	44.1 (30.8)	<.001	1.37	1.43
OES social	41.3 (28.9)	85.9 (18.8)	44.5 (25.5)	<.001	1.54	1.74

OES, Oxford Elbow Score; *SRM*, standardized response mean.
Values are expressed as mean (standard deviation). *P* value in accordance with the paired Wilcoxon signed-rank test.

be due to the uniformity of the study population (all had undergone total elbow arthroplasty) and the high number of patients assessed for test-retest reliability (45[22] vs ten in the present study).

The Italian version of the OES shows an excellent correlation with *Quick*DASH. OES function is highly correlated with MEPI, whereas OES pain and OES social-psychological are moderately correlated with MEPI. These results are in agreement with previous studies by other researchers.[6,7,9,15,28] Dawson et al found a Spearman's correlation coefficient value of 0.381 for OES social-psychological and MEPI,[6] whereas Plaschke et al[22] reported a high correlation between all three domains of OES and MEPI, but they used the Pearson correlation coefficient. As far as the relationship between OES and SF-36 domains is concerned, the limited correlation coefficients found are similar to those reported by others.[9,28] These coefficients suggest a limited effect played by the elbow pathology on the individual's general quality of life.

The Italian version of the OES is able to detect the change four months after the index visit. The mean of the changes between the two visits ranges from 33.9 for OES pain to 44 points for OES function and social/psychological (Table III). These changes are much greater than the average values of Minimal Clinically Important Difference (10 for OES function scale and 18 for both OES pain and OES social-psychological).[5] Accordingly, both the ES and the SRM values are > 0.8 for all OES domains. They are larger than those reported in literature, where ES ranges from 0.49[7] to 1.14[6] for OES pain, from 0.56[7] to 0.93[15] for OES function and from 0.54[7] to 1.42[15] for OES social/psychological. However, the low values found by de Haan et al for both ES and SRM may be due to the small sample size (43 patients) and the short time elapsed between the two assessments (52 days).[7] The excellent responsiveness found in our study may be due, at least in part, to the high percentage of patients diagnosed with acute elbow lesions.

We found a floor effect only for OES social-psychological in 0.5% of patients at the index visit. The ceiling effect is present in 12.7%, 27.2%, and 0.6% for OES pain, function, and social-psychological, respectively. The finding concerning OES function is in line with Iordens et al who found a ceiling effect for OES function in 20% of the patients six weeks after a minor injury such as elbow dislocation.[15] However, in our study the ceiling effect did not prevent us from detecting a relevant change in the OES between the index visit and the follow-up visit.

In our opinion, our study has the following strengths. First, the study is multicenter, and the centers participated voluntarily, thus ensuring good data collection, as evidenced by the lack of missing items. Second, the standardized translation and cross-cultural Italian adaptation of the original version of the OES was performed in accordance with the literature. Third, the number of patients assessed for internal consistency (110) and for responsiveness (68), is higher than those of other authors.[6,7,9,15,20,28] Fourth, the responsiveness for OES domains was evaluated at a fixed time (four months) after the index visit.

The study, however, also has some limitations. First, our study population is not homogeneous for diagnosis and, consequently, for surgical treatment. However, the OES has been validated for the assessment after elbow surgery for both acute injury and chronic diseases (osteoarthritis, post-traumatic stiffness, rheumatoid arthritis, and epicondylitis).[5,6] The OES validation in Danish was performed on patients who had undergone total elbow arthroplasty, but the authors did not specify whether the elbow injury was acute or chronic and did not assess responsiveness.[22] Moreover, Iordens et al validated the Dutch version of the OES in patients treated nonoperatively for traumatic elbow dislocations.[15] Second, some patients (25%) answered the OES, *Quick*DASH, and SF-36 by

phone (generally, they were willing to participate in the study but did not have the time to answer the questionnaires at the time of the index visit). However, the verbal (phone) administration of *Quick*DASH replicates clinically relevant scores of the written *Quick*DASH,[17] and the telephone-administration mode of SF-36 is equivalent to and as valid as the self-administered mode.[12] We do not have information regarding the phone administration of the OES. However, Chronbach's alpha of the 82 self-answered OES questionnaires was >0.8 for all OES domains (0.85 for pain, 0.92 for function and 0.91 for social-psychological).

Conclusion

Our study demonstrates that the Italian version of the OES, translated in accordance with the international standardized guidelines, is reliable, valid, and responsive in patients who underwent elbow surgery. The OES questionnaire can now be used with confidence also with Italian-speaking patients after elbow surgery.

Acknowledgments

The authors acknowledge Oxford and Isis Outcomes and part of Isis Innovation Limited for providing the Oxford Elbow Score questionnaire. The authors acknowledge Francesca Bravi (epidemiologist), Lina Pavanelli (representative of a patients' association), Elena Rossi (psychologist), and Fernando De Filippo (physiatrist) for their contribution in the translation and cross-cultural adaptation of the Italian version of the Oxford Elbow Score. The authors also acknowledge Chiara Morotti for her support in the statistical analysis. The authors acknowledge Centro Linguistico Cesena for English editing.

Disclaimer

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

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jseint.2020.10.025.

References

1. Andresen EM. Criteria for assessing the tools of disability outcomes research. Arch Phys Med Rehabil 2000;81(12 suppl 2):s15-20.
2. Apolone G, Mosconi P. The Italian SF-36 Health Survey: translation, validation and norming. J Clin Epidemiol 1998;51:1025-36.
3. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000;25:3186-91.
4. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers; 1988.
5. Dawson J, Doll H, Boller I, Fitzpatrick R, Little C, Rees J, et al. Comparative responsiveness and minimal change for the Oxford Elbow Score following surgery. Qual Life Res 2008;17(10):1257-67. https://doi.org/10.1007/s11136-008-9409-3.
6. Dawson J, Doll H, Boller I, Fitzpatrick R, Little C, Rees J, et al. The development and validation of patient-reported questionnaire to assess outcomes of elbow surgery. J Bone Joint Surg 2007;90-B:466-73. https://doi.org/10.1302/0301-620X.90B4.20290.
7. de Haan J, Goei H, Schep NW, Tuinebreijer WE, Patka P, den Hartog D. The reliability, validity and responsiveness of the Dutch version of the Oxford Elbow Score. J Orthop Surg Res 2011;6:39. https://doi.org/10.1186/1749-799X-6-39.
8. Domholdt E. Physical therapy research: principles and application. 2nd ed. Philadelphia: W.B. Saunders Company; 2000. p. 78-267.
9. Ebrahimzadeh MH, Birjandinejad A, Razi S, Mardani-Kivi M, Reza Kachooei A. Oxford shoulder score: a cross-cultural adaptation and validation study of the Persian version in Iran. Iran J Med Sci 2015;40:404-10.
10. 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. https://doi.org/10.1016/j.jse.2017.12.027.
11. Fleiss JL. The design and analysis of clinical experiments. New York: John Wiley Sons; 1986. p. 1-32.
12. García M, Rohlfs I, Vila J, Sala J, Pena A, Masiá R, et al., REGICOR Investigators. Comparison between telephone and self-administration of short form health survey questionnaire (SF-36). Gac Sanit 2005;19:433-9. https://doi.org/10.1016/s0213-9111(05)71393-5.
13. Gummesson C, Atroshi I, Ekdahl C. The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. BMC Musculoskelet Disord 2003;4:11. https://doi.org/10.1186/1471-2474-4-11.
14. Guyver PM, Cattell AE, Hall MJ, Brinsden MD. Oxford Elbow Scores in an asymptomatic population. Ann R Coll Surg Engl 2013;95:415-7. https://doi.org/10.1308/003588413X13629960048352.
15. Iordens GIT, Den Hartog D, Tuinebreijer WE, Eygendaal D, Schep NWL, Verhofstad MHJ, et al. Minimal important change and other measurement properties of the Oxford Elbow Score and the quick disabilities of the arm, shoulder, and hand in patients with a simple elbow dislocation; validation study alongside the multicenter FuncSiE trial. PLoS One 2017;12(9):e0182557. https://doi.org/10.1371/journal.pone.0182557.
16. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. Med Care 1989;27(3 suppl):S178-89.
17. London DA, Stepan JG, Boyer MI, Calfee RP. Performance characteristics of the verbal QuickDASH. J Hand Surg Am 2014;39:100-7. https://doi.org/10.1016/j.jhsa.2013.09.041.
18. Longo UG, Franceschi F, Loppini M, Maffulli N, Denaro V. Rating systems for evaluation of the elbow. Br Med Bull 2008;87:131-61. https://doi.org/10.1093/bmb/ldn023.
19. MacDermid JC, Michlovitz SL. Examination of the elbow: linking diagnosis, prognosis, and outcomes as a framework for maximizing therapy interventions. J Hand Ther 2006;19:82-97. https://doi.org/10.1197/j.jht.2006.02.018.
20. Marquardt J, Schöttker-Königer T, Schäfer A. Validation of the German version of the Oxford Elbow Score: a cross-sectional study. Orthopade 2016;45:695-700. https://doi.org/10.1007/s00132-016-3292-3.
21. Padua R, Ceccarelli e, Romanini G, Zanoli G, Amadio PC, Campi A. Italian validation of the disability of the arm, shoulder and hand (DASH) questionnaire. Cross-cultural adaptation and validation. J Hand Surg Br 2003;28 B:179-86. https://doi.org/10.1016/s0266-7681(02)00303-0.
22. Plaschke HC, Jørgensen A, Thillemann TM, Brorson S, Olsen BS. Validation of the Danish version of the Oxford Elbow Score. Dan Med J 2013;60:A4714.
23. Preacher KJ, MacCallum RC. Exploratory factor analysis in behavior genetics research: factor recovery with small sample sizes. Behav Genet 2002;32:153-61. https://doi.org/10.1023/a:1015210025234.
24. Razaeian S, Wiese B, Zhang D, Krettek C, Meller R, Hawi N. Correlation between Oxford Elbow Score and single assessment numeric evaluation: is one simple question enough? J Shoulder Elbow Surg 2020;29:1223-9. https://doi.org/10.1016/j.jse.2020.01.067.
25. Schneeberger AG, Kösters MC, Steens W. Comparison of the subjective elbow value and the Mayo elbow performance score. J Shoulder Elbow Surg 2014;23:308-12. https://doi.org/10.1016/j.jse.2013.11.018.
26. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. Oxford, United Kingdom: Oxford University Press; 2003.
27. The B, Reinninga IHF, El Moumni M, Eygendaal D. Eboe-specific clinical rating systems: extent of established validity,reliability, and responsiveness. J Shoulder Elbow Surg 2013;22:1380-94. https://doi.org/10.1016/j.jse.2013.04.013.
28. Yosmaoglu HB, Doğan D, Sonmezer E. The reliability and validity of the Turkish version of the Oxford Elbow Score. J Orthop Surg Res 2016;11:95. https://doi.org/10.1186/s13018-016-0429-3.
29. Ware JE, Sherbourne CD. The MOS 36-item Short-Form health status survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30:473.
30. Ware JE, Snow KK, Kosinski M, Gandek B. SF-36 health survey: manual and interpretation guide. Boston: The Health Institute; 1993.