

University of Groningen

Translation, cross-cultural adaptation, validity, and reliability of the Indonesian version of the Oxford Shoulder Score for patients with shoulder pain

Deviandri, Romy; Daud, Afrianto; Aminata, Iman W.; Octarina, Putri; Mecca, Nasywa D.; van der Veen, Hugo C.; van den Akker-Scheek, Inge

Published in:
JSES international

DOI:
[10.1016/j.jseint.2024.08.175](https://doi.org/10.1016/j.jseint.2024.08.175)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2024

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Deviandri, R., Daud, A., Aminata, I. W., Octarina, P., Mecca, N. D., van der Veen, H. C., & van den Akker-Scheek, I. (2024). Translation, cross-cultural adaptation, validity, and reliability of the Indonesian version of the Oxford Shoulder Score for patients with shoulder pain. *JSES international*, 8(6), 1268-1274. <https://doi.org/10.1016/j.jseint.2024.08.175>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



Translation, cross-cultural adaptation, validity, and reliability of the Indonesian version of the Oxford Shoulder Score for patients with shoulder pain



Romy Deviadri, MD, PhD^{a,b,c,*}, Afrianto Daud, PhD^d, Iman W. Aminata, MD^e, Putri Octarina, BMed^a, Nasywa D. Mecca^c, Hugo C. van der Veen, MD, PhD^b, Inge van den Akker-Scheek, MD, PhD^b

^aDepartment of Surgery, Faculty of Medicine, Universitas Riau, Arifin Achmad Hospital, Pekanbaru, Indonesia

^bDepartment of Orthopedics, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

^cDivision of Sports Injury, Fit Centre Institute, Pekanbaru, Indonesia

^dFaculty of Teachers Training and Education, Universitas Riau, Pekanbaru, Indonesia

^eDepartment of Orthopaedic and Traumatology-Fatmawati Hospital, Jakarta, Indonesia

ARTICLE INFO

Keywords:

Outcome
Psychometrics
Patient-reported
Questionnaire
Shoulder
Quality of life

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

Background: No questionnaire is currently available for use in patients with shoulder pain in an Indonesian-speaking population. This study aimed to translate the Oxford Shoulder Score (OSS) into Indonesian and assess its validity and reliability for use in Indonesian-speaking patients with shoulder pain.

Methods: After a forward and backward translation procedure, the validity and reliability of the questionnaire were investigated. All patients who were treated in a hospital in Indonesia for shoulder pain during the inclusion period were asked to complete 3 questionnaires: the Indonesia-OSS (I-OSS), the Medical Outcomes Study 12-Item Short-Form Health Survey, and the American Shoulder and Elbow Surgeons questionnaire. Participants were asked to complete the I-OSS a second time after a 1-week interval. Following Consensus-Based Standards for the Selection of Health Measurement Instruments guidelines, construct validity, test-retest reliability, internal consistency, floor and ceiling effects, and measurement error were determined. The Bland-Altman method was used to explore systematic bias.

Results: Data of 100 patients could be used to determine validity, and data of 87 patients to determine test-retest reliability. Construct validity can be considered good, as more than 75% of the predefined hypotheses on correlations between the I-OSS and the other questionnaires could be confirmed. An intraclass correlation coefficient value of 0.99 was found, indicating good test-retest reliability. A Cronbach's α of 0.95 implied good internal consistency, and no floor or ceiling effects were found. The standard error of measurement was 1.8, with minimal detectable change at the individual level was 5.1, and at the group level was 0.5. Bland-Altman analysis showed no systematic bias.

Conclusion: The I-OSS can be considered a valid and reliable questionnaire for Indonesian-speaking patients with shoulder pain.

© 2024 The Author(s). Published by Elsevier Inc. on behalf of American Shoulder and Elbow Surgeons. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

This cohort prospective study protocol was approved by the Institutional Review Board of Riau University (B-089-1/UN19.5.1.1.8/UEPKK/2021). Participants provided written informed consent prior to participation.

The datasets used and/or analyzed during the present study are available from the corresponding author on reasonable request.

*Corresponding author: Romy Deviadri, MD, PhD, Faculty of Medicine, Universitas Riau, Diponegoro street no 1, Pekanbaru 28133, Indonesia.

E-mail addresses: r.deviandri@umcg.nl, romydeviadri@lecturer.unri.ac.id (R. Deviadri).

<https://doi.org/10.1016/j.jseint.2024.08.175>

2666-6383/© 2024 The Author(s). Published by Elsevier Inc. on behalf of American Shoulder and Elbow Surgeons. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Shoulder pain with overhead activity and limited range of motion are common symptoms of shoulder disease reducing the quality of life of patients. The community prevalence of shoulder pain varies widely across countries, with a median of 16% (range 0.67%-55.2%). The number of persons with shoulder pain is quite vast, with the incidence of shoulder pain ranging from 7.7 to 62 per 1000 persons per year.¹⁸ Rotator cuff disease is one of the common causes of persistent shoulder complaints.^{8,20} It is associated with pain and decreased shoulder mobility, functional

abilities, work activities, and quality of life.² Measurement tools are essential and critical in patient management to appraise subjective health and evaluate and improve therapeutic measures.^{15,33}

Many measurement methods have been developed over the years to assess the outcome of treatment of shoulder diseases; among these are patient-reported outcome measures (PROMs).^{6,21} PROMs can be general or investigating a specific disease, whereby the patient's perspective of their pain, function, daily life problems, and quality of life, for example, is assessed.^{15,31} As most questionnaires reflect characteristics of the population in which they were established, such as language and social culture, translation and cultural adaptation of PROMs are vital for use in other communities.^{7,33}

Published in 1996 and modified in 2009 by Dawson et al, the Oxford Shoulder Score (OSS) is a widely used shoulder-specific PROM for assessing patients with shoulder pain.^{9,10} OSS is a self-assessment of pain and function of the shoulder, which is considered quick, simple, and reliable.²⁹ Because the OSS was originally developed in an English-speaking country, it needs a translation and cross-cultural adaptation process before being used in populations with other languages and cultures. Moreover, the clinimetric properties of these versions should again be assessed. The OSS questionnaire has already been translated and adapted into different languages but not into Indonesian,^{3,4,8,11,15,22,23,27–29,33} whereas the Oxford hip and knee scores have already been validated in Indonesian.²⁴

This study aims to translate and cross-culturally adapt the original OSS into an Indonesian-language version following internationally accepted guidelines¹⁹ and assess the validity and reliability of this Indonesia-OSS (I-OSS).

Methods

An institutional review board approved the study protocol, and patients provided informed consent before participating in this study. The developers of the OSS approved the translation procedure.

Translation procedure

The I-OSS translation follows the guidelines for both forward and backward translation, which were first introduced by Guillemain et al and modified by Beaton et al.^{2,12} The primary process comprises 5 stages: initial translation, synthesis, back-translation, expert committee review, and pretesting. The initial translation stage was started by 2 independent Indonesian-English licensed translators who completed the conceptual and literal translation of OSS into Indonesian (T1 and T2). Next, the synthesized version (T12) was created using feedback from the 2 initial translations. T12 was then independently translated back into English by 2 licensed Indonesian-English translators (BT1 and BT2) to identify any inconsistencies with the original English version. Subsequently, a committee of experts, comprising 3 orthopedic surgeons, 1 methodologist, and 1 translator, reviewed the Indonesian translation. Following the committee's review, one of the authors (R. D.) edited the questionnaire to create a prefinal translated version. Then, the pretesting stage was conducted by distributing the prefinal translated version to 10 patients with shoulder problems. One researcher (R. D.) documented difficulties experienced by the patients when completing the questionnaire. Documentation was reviewed to modify the questionnaire into the final version of the I-OSS (Fig. 1).

Patients and procedure

We included all Indonesian-speaking patients treated for shoulder pain—diagnosis validated by an orthopedic surgeon who has experienced more than 5 years in this field—between July 2022 and June 2023 at a hospital in Indonesia. Subjects were selected after obtaining formal informed consent and checking the inclusion and exclusion criteria. To be included in this validation study, subjects had to speak Indonesian as a primary language, complain of shoulder pain, and were treated for shoulder pain in the hospital. The patients with neuropathic pain, neoplasms, systemic inflammatory diseases, and shoulder instability were excluded. After informed consent, patients were given a first set of questionnaires (part A) to be filled out at the outpatient clinic. Then, patients were asked to come again after 1 week and fill out a second set of questionnaires (part B).

Part A consisted of the I-OSS and 2 complementary questionnaires, the Medical Outcomes Study 12-Item Short-Form Health Survey (SF-12) and the American Shoulder and Elbow Surgeons (ASES) questionnaire, to be used to assess validity. Part B consisted of only a single questionnaire, the I-OSS, to assess test-retest reliability. Additionally, to determine whether the patient's health status and shoulder function remained stable between the completion of parts A and B, patients were provided at the start of the part B questionnaire with a Global Rating of Change (GRC) question: "Has your status changed since filling out the initial questionnaire?" There were 3 potential responses: (1) no change, (2) improvement in the problem, and (3) worsening of the problem. Only patients who reported no change in their shoulder symptoms were included in the test-retest analysis. Moreover, those patients who completed the second set of questionnaires more than 1 month later were excluded.

Patient-reported outcome measures

The OSS is a shoulder-specific questionnaire for assessing patients with shoulder pain. It has 12 items, consisting of 4 items about pain and 8 questions about daily shoulder function. Each item is assessed based on 5 categories of answers, which are 0 (worst) to 4 (best condition), and a total score ranging from 0 to 48.^{9,10}

The ASES score is a shoulder-specific questionnaire. This questionnaire comprised of 11 questions, which are separated into 2 sections. The first section, focusing on pain, assesses the patient's level of pain using a 10-cm visual analog scale. This scale ranges from 0, indicating no pain, to 10, representing the most severe pain. The second section, function, comprises 10 questions that assess the capacity to carry out everyday tasks, from simple, such as putting a coat, to more demanding, such as tossing a ball above one's head or raising a 10-pound item. Answers are on a 4-point Likert scale from 0, unable to do, to 3, not difficult. The scores derived from the pain and function subsections are converted into percentages, with each accounting for 50% of the total score. Scores on the ASES range from a minimum of 0 points to a maximum of 100 points. We translated the ASES into Indonesian following the international guidelines.³²

The SF-12 is a generic PROM assessing health-related quality of life. It consists of 8 subscores: physical functioning, physical role functioning, bodily pain, general health perceptions, vitality, social function, emotional role functioning, and mental health. In addition, the sum of the physical functioning, physical role functioning, bodily pain, and general health perceptions subscales generates a physical component summary score (PCS), and the sum of the vitality, social function, emotional role functioning, and mental health subscales generates a mental component summary score

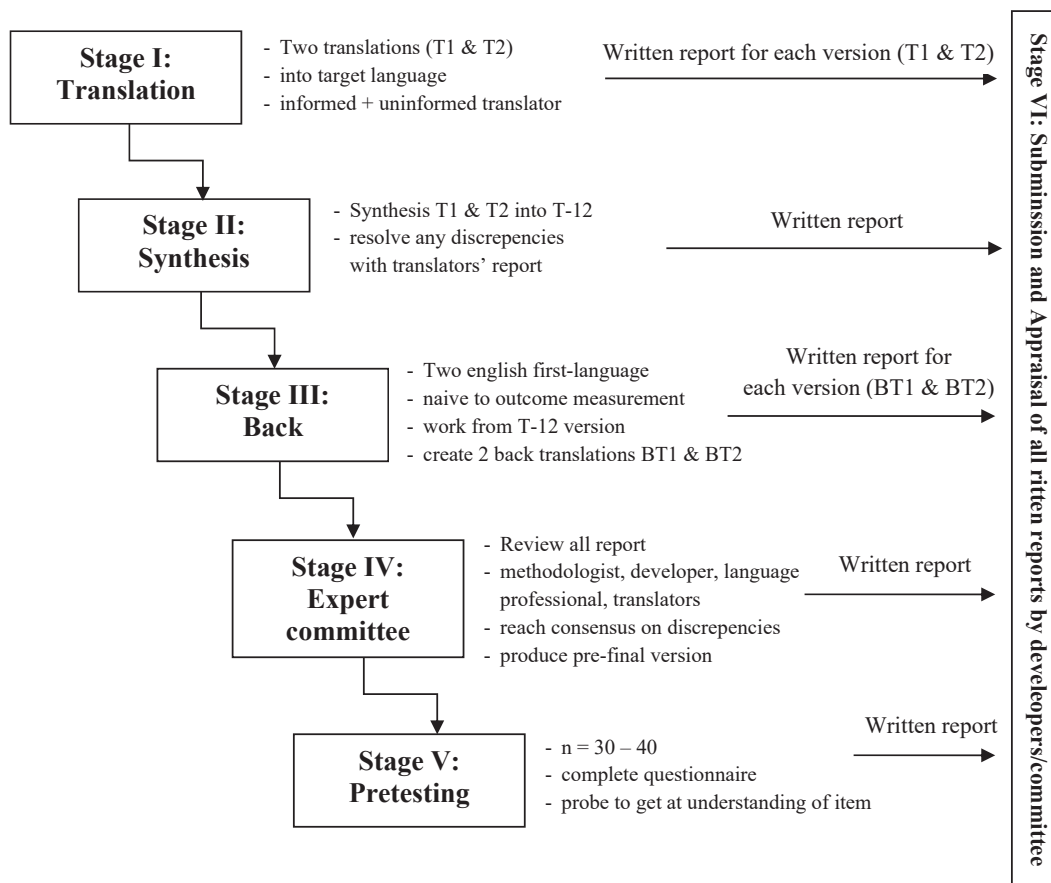


Figure 1 Flowchart of the translation procedure.^{2,12}

(MCS). Standardized scores range from 0 to 100, with higher scores indicating better health status.³⁰ SF-12 was translated to Indonesian version and shown a reliable and valid measure of health-related quality of life in Indonesian middle-aged and older adults.¹

Validity

Validity refers to the degree to which a measurement instrument accurately measures what it intends to measure, ensuring it aligns with the proposed interpretation. The construct validity of the I-OSS was assessed by determining the correlation of the I-OSS total score with the total score on the ASES, the PCS, and the MCS. In line with the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) guidelines, predetermined hypotheses were defined regarding the strength of the correlation between the I-OSS and the ASES, PCS, and MCS.¹⁹ Since both I-OSS and ASES questionnaires are shoulder-specific PROMs, we hypothesized that the I-OSS would correlate better with the ASES than with the PCS and MCS. The correlation between the original OSS and the ASES was 0.91 in a previous study.¹³ Based on the correlations found in that study, correlations of 0.6 or more between the I-OSS and the ASES were hypothesized. The Turkish version of OSS showed a correlation of 0.63 with the PCS and 0.38 with the MCS. Based on this, it was hypothesized to find correlations of 0.5 or more between the I-OSS and the PCS and the correlation less than 0.5 between the I-OSS and the MCS. The I-OSS was designed primarily to assess physical functioning, as opposed to social or emotional aspects. Consequently, stronger correlations were anticipated between the I-OSS and the PCS compared to those

between the I-OSS and the MCS. Good construct validity is indicated when a minimum of 75% of these hypotheses are verified.²⁶

Floor and ceiling effects

The presence of floor and ceiling effects was evaluated. These effects are deemed to exist if more than 15% of the participants attain either the lowest or highest possible score on the I-OSS.²⁶

Reliability

Reliability reflects the degree to which individuals can be differentiated from one another in spite of errors in measurement. In accordance with the COSMIN guidelines, reliability was evaluated through 3 aspects: internal consistency, test-retest reliability, and measurement error.^{19,26} Internal consistency is about how closely related the items of a questionnaire are to each other. Test-retest reliability focuses on how consistent patients' scores remain across repeated measurements in a period where no actual change in the construct that is being measured is expected. Measurement error, on the other hand, quantifies the systematic inaccuracies in a patient's score that are not due to actual changes in the construct being measured. To assess systematic bias, the Bland-Altman method was employed.⁵

Statistical analysis

The characteristics of the study population and scores on the questionnaires are described using means and standard

Table I
Demographic characteristics of the participants (N = 100).

Characteristic	Mean ± SD or n (%)
Age (y)	50.3 ± 12.9
Sex	
Male	28 (28)
Female	72 (72)
Diagnosis	
Adhesive capsulitis	23 (23)
Rotator cuff tear	68 (68)
Shoulder impingement	9 (9)
Affected site	
Right	59 (59)
Left	41 (41)
Education	
Primary school	1 (1)
Middle school	5 (5)
High school	63 (63)
University	31 (31)

y, year; SD, standard deviation.

deviations, or frequencies and percentages. To determine construct validity, Spearman rho correlation coefficients were calculated between the scores on the I-OSS and the other questionnaires. The Spearman rho values were interpreted as high ($r > 0.6$), moderate ($0.6 < r < 0.3$), or low ($r < 0.3$).¹³ To determine internal consistency, Cronbach's α was calculated. Values between 0.70 and 0.95 indicate good internal consistency.²⁶

To evaluate the test-retest reliability, the intraclass correlation coefficient (ICC) between the initial and subsequent I-OSS scores was calculated. Values < 0.5 , $0.5-0.75$, $0.75-0.9$, and > 0.90 indicate poor, moderate, good, and excellent reliability, respectively.^{16,17} To assess measurement error, the standard error of measurement (SEM) and the minimal detectable change (MDC) were calculated. SEM was determined by multiplying the pooled standard deviation by $\sqrt{(1-r)}$, where r is the ICC. The MDC at the individual level (MDC_{ind}) was determined using the formula $1.96 \times SEM \times \sqrt{2}$ and MDC at the group level (MDC_{grp}) by dividing MDC_{ind} by \sqrt{n} .^{14,26}

Absolute reliability was evaluated using Bland-Altman plots. The absence of systematic bias was indicated when 0 fell within the 95% confidence interval (CI) of the mean difference between the first and second administration of the I-OSS. The 95% limits of agreement were measured with the formula mean difference $\pm 1.96 \times SD_{diff}$, where SD_{diff} is the standard deviation of the mean difference between the first and second administration of the I-OSS.⁴ Statistical analyses were conducted using SPSS Statistics version 26.0 (IBM Corp., Armonk, NY, USA), with a level of significance set at 5%.

Results

Patient characteristics

In total, 100 patients completed part A, and their responses were used to study validity and internal consistency. Ninety-two patients (92%) additionally also filled in part B. Five patients were excluded as they reported better shoulder function when they filled out part B compared to when they completed part A. No patients were excluded due to missing data. Thus, to assess test-retest reliability, data from 87 patients were used. Most of the patients graduated from high school, and the mean age was 50.3 ± 12.9 years. The details of the demographic characteristics of the patients are provided in Table I.

Table II
Scores on patient-reported outcome measures.

Measure	Mean ± SD
I-OSS	
Part A	33.9 ± 11.0
Part B	33.8 ± 11.4
ASES	49.4 ± 16.3
SF-12	
PCS	32.7 ± 3.5
MCS	33.8 ± 4.5

SD, standard deviation; I-OSS, Indonesia-Oxford Shoulder Score; ASES, American Shoulder and Elbow Surgeons score; SF-12, Medical Outcomes Study 12-Item Short-Form Health Survey; PCS, physical component summary score; MCS, mental component summary score.

Table III
Spearman rho correlation coefficients among I-OSS, ASES, PCS, and MCS.

	Correlation	Hypothesized	Found	Confirmed
1	I-OSS and ASES	> 0.6	0.7	Yes
2	I-OSS and PCS	≥ 0.5	0.3	No
3	I-OSS and MCS	< 0.5	0.1	Yes
4	I-OSS and ASES correlation stronger than that between I-OSS and PCS		0.7:0.3	Yes
5	I-OSS and ASES correlation stronger than that between I-OSS and MCS		0.7:0.1	Yes
6	I-OSS and PCS correlation stronger than that between I-OSS and MCS		0.3:0.1	Yes

I-OSS, Indonesia-Oxford Shoulder Score; ASES, American Shoulder and Elbow Surgeons score; MCS, mental component summary score; PCS, physical component summary score.

Mean scores of the first and second assessments of the I-OSS were 33.9 ± 11.0 and 33.8 ± 11.4 , respectively. All scores on the PROMs can be found in Table II.

Validity

Of the 6 predefined hypotheses on the magnitude of associations between the I-OSS and the ASES, PCS, or MCS, 5 hypotheses were confirmed (a total of 83%). The I-OSS showed strong correlations with the ASES ($r = 0.7$), low correlation with the PCS ($r = 0.3$), and low correlation with MCS ($r = 0.1$). As hypothesized, the I-OSS was more strongly related to the ASES than to the PCS and MCS and more strongly related to PCS than to MCS (Table III). There were no floor or ceiling effects.

Internal consistency, test-retest reliability, and measurement error

The Cronbach's α was 0.95, indicating good internal consistency. The ICC had a value of 0.99 ($P < .001$), and the 95% CI ranged from 0.997 to 0.999. The SEM, MDC_{ind} , and MDC_{grp} were determined to be 1.8, 5.1, and 0.5, respectively. The Bland-Altman analysis showed that the average difference of the 2 I-OSS assessments was 0.2 (95% CI, -2.2 to 2.6 , 95% limits of agreement, -21.4 to 21.8) (Fig. 2). No systematic bias was present because the value of 0 was in the 95% CI of the mean difference between the test and retest scores.

Discussion

This study aimed to translate the OSS into Indonesian and to gain insight into the validity and reliability of this questionnaire in an Indonesian-speaking population with shoulder pain. Based on

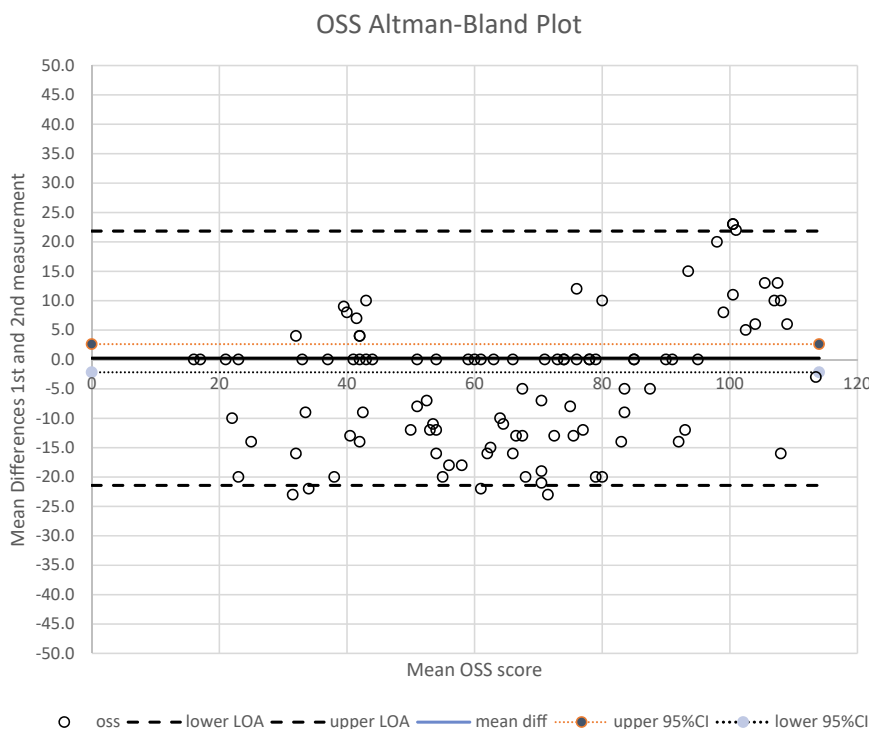


Figure 2 The Bland-Altman plot illustrates the mean difference between the first and second measurements of the Indonesian version of the OSS. In this plot, *open circles* depict individual data points. The mean difference is shown by a *solid black line*. The *dotted lines* indicate the 95% confidence intervals (CIs), and the *dashed lines* represent the limits of agreement. OSS, Oxford Shoulder Score.

the results, OSS can be considered valid and reliable for use in Indonesian-speaking patients.

The construct validity of I-OSS can be considered good, with more than 75% of predefined hypotheses confirmed. As hypothesized, I-OSS showed a strong correlation with ASES ($r = 0.7$). This is comparable with a previous study that showed a good correlation between the 2 ($r = 0.91$).¹³ As hypothesized, I-OSS showed a low correlation with the MCS ($r = 0.1$). This result is in line with the study on the Turkish version of the OSS ($r = 0.38$).²⁸ Moreover, the I-OSS was more strongly correlated to the PCS domains than to the MCS domains, again confirming our hypothesis. However, contrary to our hypothesis, the correlation of the I-OSS with PCS was less than 0.5. This was probably caused by the difference in characteristics of the participants and also the variation in diseases of the subjects among those studies. A cultural difference also might be correlated with the relatively poor validity noted between the PCS and I-OSS. According to prior study, there is a cultural difference between European population and Indonesian population regarding the PCS component of SF-12 where PCS is negatively affected more than MCS as we age in Indonesian population than in European population.¹ However, I-OSS was more strongly correlated to the PCS domains than to the MCS domains, confirming our hypothesis.

No floor or ceiling effects were observed. Theoretically, a ceiling effect could have occurred, especially in patients whose adhesive capsulitis was healing and had taken physiotherapy adequately, but this was not the case. No floor or ceiling effects were found in other-language versions, like Persian and Turkish.^{22,28} This result indicates a good content validity of I-OSS. Also, the internal consistency of I-OSS can be considered excellent, with a Cronbach's α of 0.95. This is comparable with the Chinese (0.92), Dutch (0.92), German (0.94), Turkish (0.92), and Persian (0.91) versions.^{4,15,22,28,33}

The I-OSS showed excellent test-retest reliability (ICC, 0.99). This is comparable with the Polish (0.99), Chinese (0.97), Dutch (0.98), Turkish (0.99), and Persian (0.90) versions [3,5,6,16,17].^{4,8,15,28,33} All showed excellent test-retest reliability with the test-retest time interval of 1 to 2 weeks. In accordance with COSMIN guidelines, the interval between the test and retest should be sufficiently long to avoid recall bias, yet short enough to ensure that the patients' characteristics relevant to the construct being measured remain unchanged.^{18,25} In this study, a week of time interval was used. Also, this study used the GRC question as an exclusion for patients that actually changed, which may have resulted in a high ICC value compared to some other studies. The application of GRC in the reliability study is a good practice and results in better insight into reliability.

SEM of I-OSS was 1.8, MDC_{grp} was 0.5, and MDC_{ind} was 5.1. These values are comparable to those presented in the Polish study of the OSS (SEM 1.14, MDC_{ind} 3.15).⁸ The results indicated sufficient capability of the I-OSS for comparisons at group level, as only small values are needed to detect change. As only values greater than the SEM can be reliably differentiated from the measurement error, a difference should exceed 1.8 to detect a statistically significant change in scores on I-OSS at group level. The difference between the 2 measurements should be greater than the MDC_{ind} value in individual patients and SEM, so it can be distinguished from the measurement error and to confirm a real change occurred. Given that the MDC_{ind} is quite low (5.1), this indicates sufficient capability of I-OSS as an appropriate tool for monitoring individual patients over time.

Limitations

There are some limitations to this study. There were some patients who did not come in the second time, yet the total number of

participants is deemed adequate: following COSMIN guidelines, at least 100 patients are required to assess validity and 50 patients to examine test-retest reliability.²⁵ Furthermore, this population is only comprised of patients with rotator cuff disease, adhesive capsulitis, and shoulder impingement. Other possible populations, such as glenohumeral arthritis diagnoses, were not included in the present study. This may make the findings in this study less generalizable to other populations.

Another potential limitation is that the comparison of the I-OSS was made with only one other disease-specific questionnaire, the ASES, and a general health-related quality of life questionnaire (SF-12). However, no other Indonesian versions of a PROM are available to assess shoulder function or patients' quality of life until now. Consequently, the number of hypotheses was relatively less.

Next, future research should explore the responsiveness of I-OSS as well as minimal clinically important difference (MCID). MCID should be determined to examine further whether a measured difference is also clinically important as perceived by the patient. For the Danish OSS version, the reported MCID was 6.0 with an effect size of 0.96.^{16,25} These findings are not transferable to the Indonesian version though, so the responsiveness and MCID of I-OSS should be assessed in future research. The information on responsiveness and capability of I-OSS to detect change over time is required for the interpretation of these scores when used in longitudinal research and in clinical practice to monitor Indonesian-speaking patients over time.

Conclusion

I-OSS can be considered a valid and reliable questionnaire for use in Indonesian patients with shoulder disease. This questionnaire enables us to measure patient-perceived symptoms, function, and outcome of treatment in the Indonesian population with shoulder disease. More research is needed into MCID and responsiveness of the I-OSS.

Acknowledgments

The authors thank all Indonesian Orthopedic Society for Sports Medicine and Arthroscopy (IOSSMA) members who participated in the study.

Disclaimers:

Funding: No funding was disclosed by the authors.

Conflicts of interest: The authors report there are no competing interests to declare from all author, their immediate family, and any research foundation with which they are affiliated did not receive any financial payments or other benefits from any commercial entity related to the subject of this article.

References

- Arovah NI, Heesch KC. Assessment of the validity and reliability of the Indonesian version of short form 12 (SF-12). *J Prev Med Hyg* 2021;62:E421-9. <https://doi.org/10.15167/2421-4248/jpmh2021.62.2.1878>.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25:3186-91.
- 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. <https://doi.org/10.1186/s12955-019-1263-1>.
- 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;19:829-36. <https://doi.org/10.1016/j.jse.2010.01.017>.
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-10.
- Bot SD, Terwee CB, van der Windt DA, Bouter LM, Dekker J, de Vet HC. Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. *Ann Rheum Dis* 2004;63:335-41. <https://doi.org/10.1136/ard.2003.007724>.
- Celik D, Atalar AC, Demirhan M, Dirican A. Translation, cultural adaptation, validity and reliability of the Turkish ASES questionnaire. *Knee Surg Sports Traumatol Arthrosc* 2013;21:2184-9. <https://doi.org/10.1007/s00167-012-2183-3>.
- Cook KF, Roddey TS, Olson SL, Gartsman GM, Valenzuela FF, Hanten WP. Reliability by surgical status of self-reported outcomes in patients who have shoulder pathologies. *J Orthop Sports Phys Ther* 2002;32:336-46. <https://doi.org/10.2519/jospt.2002.32.7.336>.
- Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. *J Bone Joint Surg Br* 1996;78:593-600.
- Dawson J, Rogers K, Fitzpatrick R, Carr A. The Oxford shoulder score revisited. *Arch Orthop Trauma Surg* 2009;129:119-23. <https://doi.org/10.1007/s00402-007-0549-7>.
- Ekeberg OM, Bautz-Holter E, Tveit EK, Keller A, Juel NG, Brox JJ. Agreement, reliability and validity in 3 shoulder questionnaires in patients with rotator cuff disease. *BMC Musculoskelet Disord* 2008;9:68. <https://doi.org/10.1186/1471-2474-9-68>.
- 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.
- Hapuarachchi KS, Poon PC. A correlation study of the American shoulder and elbow society score and the Oxford shoulder score with the use of regression analysis to predict one score from the other in patients undergoing reverse shoulder joint arthroplasty for cuff tear arthropathy. *Shoulder Elbow* 2014;6:81-9. <https://doi.org/10.1177/1758573213518499>.
- Hinkle DE, Wiersma W, Jurs SG. Applied statistics for the behavioral sciences. 4th ed. Boston, MA: Houghton Mifflin; 1998.
- 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;124:531-6. <https://doi.org/10.1007/s00402-004-0716-z>.
- Jones IA, Togashi R, Heckmann N, Vangness CT Jr. Minimal clinically important difference (MCID) for patient-reported shoulder outcomes. *J Shoulder Elbow Surg* 2020;29:1484-92. <https://doi.org/10.1016/j.jse.2019.12.033>.
- Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 2016;15:155-63. <https://doi.org/10.1016/j.jcm.2016.02.012>.
- Lucas J, van Doorn P, Hegedus E, Lewis J, van der Windt D. A systematic review of the global prevalence and incidence of shoulder pain. *BMC Musculoskelet Disord* 2022;23:1073. <https://doi.org/10.1186/s12891-022-05973-8>.
- Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63:737-45. <https://doi.org/10.1016/j.jclinepi.2010.02.006>.
- Motta Gda R, Amaral MV, Rezende E, Pitta R, Vieira TC, Duarte ME, et al. Evidence of genetic variations associated with rotator cuff disease. *J Shoulder Elbow Surg* 2014;23:227-35. <https://doi.org/10.1016/j.jse.2013.07.053>.
- 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;19:335-41. <https://doi.org/10.1016/j.jse.2009.07.068>.
- Naghdi S, Nakhostin Ansari N, Rustaie N, Akbari M, Ebadi S, Senobari M, et al. Simple shoulder test and Oxford shoulder score: Persian translation and cross-cultural validation. *Arch Orthop Trauma Surg* 2015;135:1707-18. <https://doi.org/10.1007/s00402-015-2330-7>.
- 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;132:93-9. <https://doi.org/10.1007/s00402-011-1393-3>.
- Silitonga J, Djaja YP, Dilogo IH, Pontoh LAP. Cross-cultural adaptation and psychometric validation of the Indonesian version of the Oxford Hip Score. *Bone Jt Open* 2021;2:765-72. <https://doi.org/10.1302/2633-1462.29.BJO-2021-0111.R1>.
- Sørensen L, van Tulder M, Johannsen HV, Ovesen J, Oestergaard LG. Responsiveness and minimal important change of the Oxford shoulder score, EQ-5D, and the fear-avoidance belief questionnaire physical activity subscale in patients undergoing arthroscopic subacromial decompression. *JSES Int* 2021;5:869-74. <https://doi.org/10.1016/j.jseint.2021.05.008>.
- 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. <https://doi.org/10.1016/j.jclinepi.2006.03.012>.
- Torres-Lacomba M, Sánchez-Sánchez B, Prieto-Gómez V, Pacheco-da-Costa S, Yuste-Sánchez MJ, Navarro-Brazález B, 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. <https://doi.org/10.1186/s12955-015-0256-y>.

28. 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;131:687-94. <https://doi.org/10.1007/s00402-010-1242-9>.
29. 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;102:555-8. <https://doi.org/10.1016/j.otsr.2016.05.010>.
30. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36): I. conceptual framework and item selection. *Med Care* 1992;30:473-83.
31. Wright RW, Baumgarten KM. Shoulder outcomes measures. *J Am Acad Orthop Surg* 2010;18:436-44. <https://doi.org/10.5435/00124635-201007000-00006>.
32. Wylie JD, Beckmann JT, Granger E, Tashjian RZ. Functional outcomes assessment in shoulder surgery. *World J Orthop* 2014;5:623-33. <https://doi.org/10.5312/wjo.v5.i5.623>.
33. 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. <https://doi.org/10.1186/s12955-015-0383-5>.