J Orthop Sci (2011) 16:732–736
DOI 10.1007/s00776-011-0141-4

ORIGINAL ARTICLE

Reliability and validity of the Western Ontario Shoulder Instability Index (WOSI) in the Japanese population

Taku Hatta · Nobuhisa Shinozaki · Rei Omi · Hirotaka Sano · Nobuyuki Yamamoto · Akira Ando · Hiroyuki Sugaya · Toshitake Aizawa · Shinichi Kuriyama · Eiji Itoi

Received: 6 May 2011 / Accepted: 28 July 2011 / Published online: 25 August 2011

Abstract
Background The Western Ontario Shoulder Instability Index (WOSI) questionnaire is a disease-specific self-evaluated measurement tool used for patients with shoulder instability. The purpose of this study was to evaluate the reliability and validity of the WOSI Japanese version for the Japanese population.
Methods The WOSI Japanese version was prepared by use of several steps including translation, back translation, and verification by the author and owner of the WOSI. Reliability (internal consistency and reproducibility) and validity were then tested for 85 patients with shoulder instability. All patients completed the WOSI Japanese version, Quick DASH, and SF-36, and underwent assessment of the Rowe rating scale for the purpose of internal consistency and validity. Fifty-nine of the patients repeated the WOSI 2 weeks after the first examination for assessment of test–retest reproducibility.
Results Both internal consistency (0.84 for Cronbach's alpha) and reproducibility (Pearson's $r = 0.91$, intraclass correlation coefficient = 0.91) indicated the reliability of the WOSI Japanese questionnaire was good. Comparison with the Rowe rating scale ($p < 0.001$) also showed validity was good. The correlation coefficient between the WOSI Japanese version and the Quick DASH (Pearson's $r = 0.62$) was higher than that between the WOSI Japanese version and the SF-36 (Pearson's $r = 0.10$–0.36 for subscales).
Conclusion The WOSI Japanese questionnaire is a reliable self-assessment tool that is comparable with the original English version. We therefore recommend use of the WOSI Japanese version when evaluating patients with shoulder instability in the Japanese population.

Electronic supplementary material The online version of this article (doi:10.1007/s00776-011-0141-4) contains supplementary material, which is available to authorized users.

T. Hatta · N. Shinozaki · R. Omi · H. Sano · N. Yamamoto · E. Itoi (✉)
Department of Orthopaedic Surgery, Tohoku University School of Medicine, 1-1 Seiryo-machi, Aoba-ku, Sendai 980-8574, Japan
e-mail: itoi-eiji@med.tohoku.ac.jp

A. Ando · H. Sugaya
Shoulder and Elbow Service, Funabashi Orthopaedic Sports Medicine Center, Funabashi, Japan

T. Aizawa
Department of Orthopaedic Surgery, Iwaki Kyoritsu General Hospital, Iwaki, Japan

S. Kuriyama
Division of Epidemiology,
Department of Public Health and Forensic Medicine,
Tohoku University School of Medicine, Sendai, Japan

Introduction

Shoulder instability is a common injury, especially among the young and active population. The incidence of shoulder instability has been reported to be approximately 2% of the population [1]. A large number of patients commonly experience recurrent instability after a first anterior shoulder dislocation. According to the reports about high-risk populations, for example young athletes and some subpopulations, this number approaches 100% [2, 3].

The common complaints of shoulder instability are quite different from those of other shoulder pathologies. After joint reduction, pain is rarely a major complaint, whereas apprehension and loss of confidence in the shoulder are characteristic symptoms arising in association with some activities. With increased severity, patients may reduce

their sports activities and experience a decrease in their quality of life [4].

Several instruments have been developed to assess the clinical outcome of the management of shoulder conditions including instability problems. However, most of these were derived from clinical data and depend on the judgment of an examiner. Especially for shoulder instability, assessment tools that focus on the patient's subjective viewpoint should be necessary to evaluate the true disability in their activities of daily living.

Kirkley et al. [5] were the first to develop a questionnaire as a disease-specific quality of life scoring system for shoulder instability; the Western Ontario Shoulder Instability Index (WOSI). The WOSI questionnaire is a patient-administered questionnaire for evaluating symptoms and their effects on physical activity and quality of life. It can be used to compare different populations with shoulder instability, and facilitate comparisons between studies. The WOSI score has already been used to evaluate the outcome of treatment in several studies [6–9]. In clinical practice, the WOSI questionnaire is widely used in English-speaking countries to determine patients' severity of impairment and their functional improvement, and for monitoring the outcome of treatment [10]. In order to use the WOSI questionnaire for non-English speaking patients, it must be translated into other languages; it should also be culturally adapted and properly evaluated. The reliability and validity of the Swedish and German versions of the WOSI have been reported [11, 12]. The purposes of this study were:

1 to translate and culturally adapt the English WOSI questionnaire to Japanese; and
2 to evaluate reliability and validity for the Japanese population.

Materials and methods

Translation of the WOSI questionnaire into Japanese was performed in several steps. The original English version was translated into Japanese by one of the co-authors (E.I.), a senior shoulder surgeon who had learned English as a second language. Back translation of the Japanese version into English was performed independently by another co-author (N.Y.), who is also a shoulder surgeon. Subsequently, Sharon H. Griffin, the author and owner of the WOSI questionnaire, checked for any discrepancies. After several revisions of the translation and careful corrections, the final WOSI Japanese version was certified as Official Japanese Translation by Dr Griffin.

In this study, the Japanese version of the WOSI was tested for patients with shoulder instability. All subjects were given written information about the purpose and procedure of the study, and informed consent to participation was obtained. Ethics approval was obtained from the Ethics Committee at the Medical Faculty in our institution. The protocol and publication of the study were approved by our institutional review board.

Eighty-five patients, mean age 27 years (range 16–62), with shoulder instability were included in the reliability and validity evaluation. The patients were recruited in the three clinics in Japan. The inclusion criteria were the same as in the original study by Kirkley [5]. All subjects had been diagnosed with either traumatic or non-traumatic instability of the shoulder on the basis of their history and clinical findings. There were 76 patients with traumatic instability and 9 patients with non-traumatic instability.

Reliability

The reliability of assessment tools can be divided into two major classes, internal consistency and reproducibility (test–retest). Internal consistency of the WOSI Japanese version questionnaire was assessed by use of the WOSI score from 85 patients, which were obtained at first visit. To assess the reproducibility of the WOSI Japanese version, 59 of the 85 patients also participated in the test–retest evaluation and completed the WOSI questionnaire again 2 weeks after the first visit. Correlation between the total results from both tests was analyzed statistically to assess the test–retest reproducibility of the Japanese version.

Validity

Validity is an index of how well a tool measures what it is supposed to measure. If there is a "gold standard" with which the results can be compared, validation may be clearly demonstrated. In the absence of a "gold standard", for example for quality of life with shoulder instability, it is necessary to compare the tool with other measures to determine validity. In this study, the validity of the WOSI Japanese version was evaluated by comparing the WOSI score with results from the Rowe rating scale [13] from the 85 patients, as was performed in previous validation of the Swedish and German versions of the WOSI [11, 12].

At the first visit, all 85 patients underwent evaluation on the Rowe rating scale and using the WOSI Japanese version. They were also required to complete the Quick DASH (the Japanese Society for Surgery of the Hand version of the Quick Disability of Arm, Shoulder, and Hand questionnaire [14]) and the SF-36 (the official Japanese version of the 36-Item Short Form Health Survey, version 1.2 [15, 16]) at the same time. The Quick DASH and the SF-36 scores were compared with the WOSI score. Similar to the WOSI, both measures are patient self-administered

Table 1 A summary of the reliability tests of the WOSI

	Pearson's r	ICC	Cronbach's alpha
Test–retest ($n = 59$)			
WOSI total score	0.91	0.91	
Domain			
Physical symptoms	0.86	0.86	
Sport/recreation/work	0.65	0.64	
Lifestyle	0.83	0.82	
Emotions	0.80	0.79	
Internal consistency ($n = 85$)			0.84

questionnaires. The Quick DASH score is a general upper extremity score used for patients with any condition of any joint of the upper extremity. We hypothesized that the WOSI score would be more consistent with the Quick DASH score than with the overall health measurement by use of the SF-36.

Statistical analyses

Internal consistency was assessed by calculation of Cronbach's alpha. The test–retest data were analyzed by use of Pearson's product–moment correlation coefficient and the intraclass correlation coefficient (ICC). Comparison of the WOSI with the other measurements to assess validity was performed by calculating the Pearson's correlation coefficient. The level of significance was set at $p < 0.05$.

Results

Reliability

Table 1 summarizes the results from evaluation of the reliability of the WOSI Japanese version. Internal consistency for the 21 questions in the WOSI was 0.84, as measured by use of Cronbach's alpha. Strong test–retest correlation (Pearson's $r = 0.91$, ICC = 0.91) was indicative of good reproducibility of the total score of the WOSI. In analysis of each domain separately, the values obtained were indicative of adequate reliability. Among the four domains, reliability was lowest for the sports/recreation/work domain (Pearson's $r = 0.65$, ICC = 0.64).

Validity

Validity, expressed as Pearson's r, was 0.63 between the WOSI Japanese version score and the Quick DASH; this was the highest correlation in this study (Table 2). As expected, the WOSI score correlated significantly with the Rowe rating scale (Pearson's $r = 0.42$, $p < 0.0001$).

Table 2 Correlation between the WOSI and the Quick DASH, the Rowe rating scale, and the SF-36

	Correlation with the WOSI Pearson's r
Quick DASH***	0.63
Rowe rating scale***	0.42
SF-36	
Physical functioning***	0.36
Role physical*	0.26
Body pain**	0.34
General health*	0.27
Vitality*	0.28
Social functioning	0.10
Role emotional	0.16
Mental health*	0.26

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$; Pearson's correlations

Regarding the SF-36, the physical subscales such as physical functioning (PF) and bodily pain (BP) seemed to be highly correlated with the WOSI score (Pearson's $r = 0.36$, 0.34, respectively). However, values were lower for the social functioning (SF) and role-emotional (RE) subscales of the SF-36 (Pearson's $r = 0.10$, 0.16, respectively).

Discussion

A standardized assessment tool which focuses on shoulder instability from the patient's perspective can be used for comparison of clinical outcomes among different institutions or treatment methods. Because the original WOSI is known to be an easy and reliable self-administered questionnaire, it would be beneficial for Japanese researchers if the Japanese version of the WOSI could be used for their clinical studies with the same level of reliability as the original version.

This study clearly demonstrated the high reliability of the Japanese version of the WOSI questionnaire. Both internal consistency and reproducibility were regarded as good. Especially, we found that the test–retest reliability of the WOSI Japanese version was excellent. In the literature, ICC values for the original English, Swedish, and German versions were reported as 0.95, 0.94, and 0.87, respectively [5, 11, 12]. This study revealed that the ICC for the WOSI Japanese version was 0.91, which seemed comparable with those of WOSIs in other languages.

When we analyzed the four domains of the Japanese version separately, the domain for sports/recreation/work had the lowest value. It is interesting to note that the lowest value among the four domains was found for this domain in

the other authors' studies [11, 12]. However, the value of 0.64 for this domain can still be regarded as acceptable for test–retest reliability according to the criteria of the outcome measures [17].

As for the validity, we found that the scores of the WOSI Japanese version correlated significantly with the Rowe's rating scale (Pearson's $r = 0.42$, $p < 0.0001$). However, the correlation coefficient was slightly lower than that for other languages (0.59–0.62) [5, 11, 12]. It is possible that differences between patients' treatment status caused these discrepancies among the studies; the data were collected from untreated patients in our study, whereas they were collected from those under treatment [5] or from those who had undergone primary treatment more than 8 years previously [11], or data for subjects both before and after treatment were pooled [12].

In this study, the highest correlation coefficient was between the WOSI Japanese version scores and the Quick DASH scores, which was similar to results previously reported by Kirkley et al. [5]. A number of similar items in both instruments might cause such correlation, as Kirkley et al. [5] pointed out in their study. On the other hand, the correlation coefficient between the SF-36 and the WOSI or DASH scores was lower than that between the WOSI and the Quick DASH. We assumed that the WOSI would reflect the function of the upper limb better than the general health in any language.

A number of physician-based results, for example range of motion and manual muscle testing, have weighed heavily in outcome assessment in the past. However, in an ideal means of direct evaluation of shoulder function, the patient's point of view must have priority [18]. The WOSI score may enable assessment of symptoms important to the patients other than the instability itself [11].

This Japanese version of the WOSI questionnaire can be filled out easily and quickly, and requires minimum assistance during follow-up for the Japanese population. The validated WOSI has successfully been used to measure outcome in several case studies including randomized controlled trials [6, 8, 19]. Translations into other languages will enable us to perform an international comparison of results from different population groups with shoulder instability.

A major limitation of this study must be considered when evaluating the results. Responsiveness or sensitivity to changes between measurements before and after treatment (over time) was not analyzed. Although most of the patients underwent surgical stabilization, we did not investigate the consequence of such treatment. Further studies focusing on responsiveness over time would be necessary for use of the WOSI Japanese version score in the clinical setting.

Conclusion

In this study we culturally adapted and validated the Japanese version of the WOSI questionnaire for the Japanese population. The results obtained using this new Japanese version of the WOSI were comparable with those obtained by use of the original version.

The content of the Japanese version of the WOSI questionnaire is available (as supplementary material) in the on-line version of this publication in the Journal of Orthopaedic Science or in the Japanese abstract in the Journal of the Japanese Orthopaedic Association.

Acknowledgments We sincerely thank Drs. Norimasa Takahashi, Nobuaki Kawai, and Hisahiro Tonotsuka for their help.

Conflict of interest The authors did not receive and will not receive any benefits or funding from any commercial party related directly or indirectly to the subject of this article.

References

1. Owens BD, Duffey ML, Nelson BJ, DeBerardino TM, Taylor DC, Mountcastle SB. The incidence and characteristics of shoulder instability at the United States Military Academy. Am J Sports Med. 2007;35:1168–73.
2. Larrain MV, Botto GJ, Montenegro HJ, Mauas DM. Arthroscopic repair of acute traumatic anterior shoulder dislocation in young athletes. Arthroscopy. 2001;17:373–7.
3. te Slaa RL, Wijffels MP, Brand R, Marti RK. The prognosis following acute primary glenohumeral dislocation. J Bone Joint Surg Br. 2004;86:58–64.
4. Dawson J, Fitzpatrick R, Carr A. The assessment of shoulder instability. The development and validation of a questionnaire. J Bone Joint Surg Br. 1999;81:420–6.
5. Kirkley A, Griffin S, McLintock H, Ng L. The development and evaluation of a disease-specific quality of life measurement tool for shoulder instability. The Western Ontario Shoulder Instability Index (WOSI). Am J Sports Med. 1998;26:764–72.
6. Kirkley A, Griffin S, Richards C, Miniaci A, Mohtadi N. Prospective randomized clinical trial comparing the effectiveness of immediate arthroscopic stabilization versus immobilization and rehabilitation in first traumatic anterior dislocations of the shoulder. Arthroscopy. 1999;15:507–14.
7. Westerheide KJ, Dopirak RM, Snyder SJ. Arthroscopic anterior stabilization and posterior capsular plication for anterior glenohumeral instability: a report of 71 cases. Arthroscopy. 2006;22:539–47.
8. Salomonsson B, Abbaszadegan H, Revay S, Lillkrona U. The Bankart repair versus the Putti-Platt procedure: a randomized study with WOSI score at 10-year follow-up in 62 patients. Acta Orthop. 2009;80:351–6.
9. Krueger D, Kraus N, Pauly S, Chen J, Scheibel M. Subjective and objective outcome after revision arthroscopic stabilization for recurrent anterior instability versus initial shoulder stabilization. Am J Sports Med. 2011;39:71–7.
10. Plancher KD, Lipnick SL. Analysis of evidence-based medicine for shoulder instability. Arthroscopy. 2009;25:897–908.

11. Salomonsson B, Ahlström S, Dalén N, Lillkrona U. The Western Ontario Shoulder Instability Index (WOSI): validity, reliability, and responsiveness retested with a Swedish translation. Acta Orthop. 2009;80:233–8.
12. Hofstaetter JG, Hanslik-Schnabel B, Hofstaetter SG, Wurnig C, Huber W. Cross-cultural adaptation and validation of the German version of the Western Ontario Shoulder Instability index. Arch Orthop Trauma Surg. 2010;130:787–96.
13. Rowe CR, Patel D, Southmayd WW. The Bankart procedure: a long-term end-result study. J Bone Joint Surg Am. 1978;60:1–16.
14. Imaeda T, Toh S, Wada T, Uchiyama S, Okinaga S, Kusunose K, Sawaizumi T. Impairment Evaluation Committee, Japanese Society for Surgery of the Hand. Validation of the Japanese Society for Surgery of the Hand Version of the Quick Disability of the Arm, Shoulder, and Hand (Quick DASH-JSSH) questionnaire. J Orthop Sci. 2006;11:248–53.
15. Fukuhara S, Bito S, Green J, Hsiao A, Kurokawa K. Translation, adaptation, and validation of the SF-36 Health Survey for use in Japan. J Clin Epidemiol. 1998;51:1037–44.
16. Fukuhara S, Ware JE Jr, Kosinski M, Wada S, Gandek B. Psychometric and clinical tests of validity of the Japanese SF-36 Health Survey. J Clin Epidemiol. 1998;51:1045–53.
17. Salter K, Jutai J, Foley N, Teasell R. Outcome measures in stroke rehabilitation. 2005. Accessed cited 2006 October 5, 2006). Available from: http://www.ebrsr.com.
18. Romeo AA, Bach BR Jr, O'Halloran KL. Scoring systems for shoulder conditions. Am J Sports Med. 1996;24:472–6.
19. Skare O, Schrøder CP, Reikerås O, Mowinckel P, Brox JI. Efficacy of labral repair, biceps tenodesis, and diagnostic arthroscopy for SLAP Lesions of the shoulder: a randomised controlled trial. BMC Musculoskelet Disord. 2010;11:228.