Original article

Validity, reliability and responsiveness of the French language translation of the Western Ontario Shoulder Instability Index (WOSI)

C. Gaudelli a, F. Balg b, V. Godbout c, S. Pelet d, A. Djahangiri e, S. Griffin f, D.M. Rouleau a,∗

a Hôpital de Sacré Coeur de Montréal, Montréal, Québec, Canada
b Centre Hospitalier Universitaire de Sherbrooke, Sherbrooke, Québec, Canada
c Centre Hospitalier de l’Université de Montréal, Montréal, Québec, Canada
d Centre Hospitalier Affilé Universitaire de Québec, Pavillon Enfant-Jésus, Québec, Canada
e Centre Hospitalier Universitaire Vaudois, Lausanne, Switzerland
f University of Western Ontario, London, Ontario, Canada

A R T I C L E  I N F O

Article history:
Accepted 11 September 2013

Keywords:
Shoulder instability
Functional outcome measures
Translation and cultural adaptation
Outcome measure validation
WOSI

A B S T R A C T

Background: The WOSI (Western Ontario Shoulder Instability Index) is a self-administered quality of life questionnaire designed to be used as a primary outcome measure in clinical trials on shoulder instability, as well as to measure the effect of an intervention on any particular patient. It is validated and is reliable and sensitive. As it is designed to measure subjective outcome, it is important that translation should be methodologically rigorous, as it is subject to both linguistic and cultural interpretation.

Objective: To produce a French language version of the WOSI that is culturally adapted to both European and North American French-speaking populations.

Materials and methods: A validated protocol was used to create a French language WOSI questionnaire (WOSI-Fr) that would be culturally acceptable for both European and North American French-speaking populations. Reliability and responsiveness analyses were carried out, and the WOSI-Fr was compared to the F-QuickDASH-D/S (Disability of the Arm, Shoulder and Hand–French translation), and Walch-Duplay scores.

Results: A French language version of the WOSI (WOSI-Fr) was accepted by a multinational committee. The WOSI-Fr was then validated using a total of 144 native French-speaking subjects from Canada and Switzerland. Comparison of results on two WOSI-Fr questionnaires completed at a mean interval of 16 days showed that the WOSI-Fr had strong reliability, with a Pearson and interclass correlation of r = 0.85 (P < 0.01) and ICC = 0.84 [95% CI = 0.78–0.88]. Responsiveness, at a mean 378.9 days after surgical intervention, showed strong correlation with that of the F-QuickDASH-D/S, with r = 0.67 (P < 0.01). Moreover, a standardized response means analysis to calculate effect size for both the WOSI-Fr and the F-QuickDASH-D/S showed that the WOSI-Fr had a significantly greater ability to detect change (SRM 1.55 versus 0.87 for the WOSI-Fr and F-QuickDASH-D/S respectively, P < 0.01). The WOSI-Fr showed fair correlation with the Walch-Duplay.

Discussion: A French-language translation of the WOSI questionnaire was created and validated for use in both Canadian and Swiss French-speaking populations. This questionnaire will facilitate outcome assessment in French-speaking settings, collaboration in multinational studies and comparison between studies performed in different countries.

Type of Study: Multicenter cohort study.

Level of evidence: II.

© 2013 Published by Elsevier Masson SAS.

1. Introduction

The incidence of traumatic shoulder instability is 1.7% in the general population [1,2]. In addition to traumatic etiology, there are many other categories of symptomatic shoulder instability that significantly increase its overall prevalence [3]. Clinical presentation is very different from that of other shoulder pathologies. After joint reduction, pain is rarely a major complaint, whereas apprehension and loss of confidence in the shoulder lead to progressive reduction in sports activities and impaired quality of life (QOL) [4]. It is therefore of interest to measure the QOL impact of this pathology. One way of measuring QOL is through general QOL questionnaires such as the Index of Well-Being [5], Sickness Impact Profile [6] and SF 36.

∗ Corresponding author. 5400, Boulevard Gouin Ouest, Montréal, Québec, Canada. Tel.: +14 338 2222, 3427.
E-mail address: dominique.rouleau@umontreal.ca (D.M. Rouleau).

1877-0568/S – see front matter © 2013 Published by Elsevier Masson SAS.
http://dx.doi.org/10.1016/j.otsr.2013.09.007

However, they are poor at detecting small but clinically important changes in patients with specific medical conditions [8,9]. Disease-specific questionnaires such as the Western Ontario Shoulder Instability Index (WOSI) were created to remedy the situation. The WOSI is a self-administered QOL questionnaire designed as primary outcome measure in clinical trials. It can also be used in the clinical setting, to monitor an individual’s progress. It comprises 4 domains (physical symptoms, sports/recreation/work, lifestyle and emotions), with scores ranging between 0 and 2100. A percentage score can also be calculated to make it easier to compare the WOSI with other outcome scores. It is filled out by the patient with minimal instructions required and has high reliability, validity and responsiveness [3,10–14]. It has had extensive psychometric testing and is a recommended tool for the study of shoulder instability [15–19].

However, the subjective nature of patient-administered questionnaires makes them particularly sensitive to cultural and linguistic interpretation. Literal translation does not necessarily produce a questionnaire that is appropriate for the study country, as it may encompass activities either not practiced there or not accorded the same importance as in the original country. Thus guidelines for producing a translated outcome measure that is culturally appropriate for the study country have been developed by the American Academy of Orthopaedic Surgeons (AAOS) [20–22]. The process includes producing 2 independent translations, back-translating these versions back into the original language, and having a committee review them to produce a pre-final version. The pre-final version must then be validated to ensure the translation has lost neither reliability nor responsiveness. The reliability of a questionnaire measures whether there is a change in the score when the questionnaire is taken by the same patient at 2 different time-points despite no clinically significant changes having occurred. The responsiveness of a questionnaire is its ability to detect and quantify the change in clinical status that occurs after a given intervention or event.

Because the WOSI is one of many QOL questionnaires, we sought to correlate it with other widely-used functional assessment scores. The F-QuickDASH-D/S, a validated French-language translation of the QuickDASH (Disability of the Arm, Shoulder and Hand) outcome measure, was used as gold standard [23,24]. It consists of 11 questions and looks at the degree of difficulty in performing various physical activities as a result of pain, stiffness, tingling or weakness of the hand, arm or shoulder. The impact of this difficulty on social activities, work, sleep and self-image is also assessed. It is a disability index, so the higher the score the greater the limitations the patient experiences. The second measure, the Walch-Duplay score, is a non-validated outcome measure commonly used in Europe [25]. It has both an objective and a subjective component, and is filled out by an examiner but largely consists of patient-reported outcomes such as apprehension and return to sport. The examiner also records range of motion for the affected shoulder.

The hypothesis was that we would be able to produce a validated French translation of the WOSI that would remain highly reliable and responsive and would be acceptable by both Canadian and European French-speaking populations. We also sought to compare this translation to other scores (F-QuickDASH-D/S, Walch-Duplay).

2. Methods

2.1. Translation process

We followed the methodology recommended by several publications regarding translation and cultural validation of research outcome measures [20–22]. This comprised initial translation performed by 2 native French-speaking translators (T) to produce 2 independent initial French versions (T1 and T2). These translations were then synthesized to make a consensus version (T1 + T2). Two independent native English translators (ET) then performed 2 back-translations of this consensus version (ET1 and ET2). These back-translations were then compared to the original English-language WOSI questionnaire by the research team and by the WOSI copyright holder (SG). An expert committee of orthopedic surgeons, 2 Québécois, 2 Swiss and 1 French (LN), then agreed on a pre-final French-language version. This version was then validated, as outlined below, to produce the final version (WOSI-Fr), shown in Appendix 1. The committee also ensured that the questionnaire could be understood by subjects with a 6th-grade reading level.

2.2. Cultural adaptation

The translators were instructed to translate the questionnaire so as to establish linguistic equivalence, rather than literally. The goal was to obtain a questionnaire that would be acceptable to both Canadian and European French-speaking populations. The committee therefore included experts from both North America and Europe, and the questionnaire likewise was validated on patients in both geographical locations.

2.3. WOSI-Fr validation

2.3.1. Patient selection

Patients were recruited from 5 centers located in Québec, Canada and Lausanne, Switzerland and were selected from the surgical waiting lists of participating surgeons. Patients were included if they were awaiting shoulder stabilization surgery and met the following criteria: symptomatic instability, whether anterior, posterior or multidirectional; traumatic or non-traumatic; native French speakers, able to read French; agreeing to fill in questionnaires both in the pre-operative period and at their 1-year follow-up visit. Patients had to be willing to provide informed consent, and in the case of minors, parental consent to inclusion was required. Exclusion criteria were significant psychiatric or psychological disorder and inability or unwillingness to give informed consent for the study. This study was approved by the ethics board at each institution.

2.3.2. Reliability

Subjects were contacted via telephone. If they agreed to participate, demographic data was collected and they were asked to fill out two WOSI-Fr questionnaires at 1 week’s interval (WF-1 and WF-2). Subjects who experienced a traumatic event, a significant instability event or a treatment modality other that physical therapy within that week were excluded. Scores at the 2 time-points were compared on Pearson’s and interclass correlation coefficients.

2.3.3. Responsiveness

Subjects were asked to complete a WOSI-Fr questionnaire 1 year after surgical stabilization (WF-3). Results on this questionnaire were compared with those of the first questionnaire (WF-1) to determine the effect of stabilization surgery. A 1-year interval was used to ensure that potential instability events after return to regular activities were included: most patients were not allowed to return to high-risk activities until 6 months after surgery. Pearson’s correlation and standardized mean analysis were used to compare the magnitude of difference between WOSI-Fr and the previously validated F-QuickDASH-D/S scores.

2.3.4. Correlation with other scores

Walch-Duplay scores were also calculated pre-operatively, and compared with the WF-1 score, again using Pearson’s correlation.
3. Results

3.1. Patient Demographics

A total of 178 subjects were eligible for the study; 144 were included in the reliability analysis and 49 in the responsiveness analysis (patients awaiting surgery (n = 32) or not having completed 1 year of follow-up (n = 23) were excluded from the responsiveness analysis). Demographic and clinical data are shown in Table 1 and a subject flowchart is shown in Fig. 1.

3.2. Translation and cultural adaptation

The 2 initial translations (T1 and T2) contained several minor differences. The translators were asked to justify their translations, and a consensus version (T1+2) was produced in collaboration with the research team. The 2 back-translations were performed and sent to the copyright holder, who approved them after requesting that 1 word be modified. The expert committee then met to review the T1+2 version. This team included members from Canada, Switzerland and France, to ensure cultural equivalence was achieved. A pre-final version was created and accepted by all members. No further changes were required by the results of the questionnaire validation process. The final version is presented in Appendix 1.

3.3. Reliability analysis

The mean interval WF-1 and WF-2 was 16 days (range, 3–357 days). Mean WF-1 and WF-2 scores were 1146 ± 392 and 1195 ± 415 respectively (Fig. 2) or, in percentage terms, 54.6% ± 19 and 56.9% ± 20 respectively. Pearson’s correlation and interclass correlation coefficient comparing WF-1 and WF-2 both showed strong correlations: r = 0.85 (P < 0.01) and ICC = 0.84 [95% CI 0.78–0.88]. Re-analysis using only subjects with an interval of 6 to 14 days between WF-1 and WF-2 (116 subjects) showed a similar strong correlation: r = 0.88 (P < 0.01) and ICC = 0.87 [95% CI 0.82–0.91]. Independent sample t-tests showed that the difference between the entire cohort and the 6–14 day interval subgroup was non-significant (P = 0.587).

The reliability of the F-QuickDASH-D/S was also calculated: Pearson’s correlation was r = 0.75 (P < 0.01) and ICC was 0.75 [95% CI 0.66–0.81]. Comparing this ICC to that of the WOSI-Fr, the confidence intervals for the F-QuickDASH-D/S and that of the entire WOSI-Fr cohort overlap, indicating similar reliability for the 2 measures.

3.4. Correlation with other outcome measures

The pre-operative WOSI-F (WF-1) was compared on Pearson’s correlation with the F-QuickDASH-D/S and Walch-Duplay scores, also taken prior to surgical intervention. WOSI-F showed good correlation with F-QuickDASH-D/S (r = 0.65, P < 0.01) and fair correlation with Walch-Duplay scores (r = −0.31, P < 0.01).

3.5. Responsiveness analysis

The post-operative WOSI-Fr questionnaire (WF-3) was completed at a mean 453 days (range, 302–707 days) after WF-1, or 379 days (189–634) days after surgery. Four subjects had posterior instability, diagnosed at the time of surgery, and the remainder had anterior instability. Thirty-two subjects underwent arthroscopic Bankart surgery, 13 open Bankart stabilization, 3 Latarjet procedure and 2 glenoid bone grafting from another source. WF-3 scores at final follow-up are presented in Table 1. There was a significant improvement between pre- and post-operative WOSI-Fr scores. There was a concomitant improvement in F-QuickDASH-D/S scores at the same time-points (Table 1). The respective improvement in WOSI-Fr and F-QuickDASH-D/S scores showed good correlation on Pearson’s correlation coefficient: r = 0.67, P < 0.01. Standardized means analysis (standardized response mean: SRM) was also undertaken. This is an effect size index, obtained by dividing the mean change in scores by the standard deviation of the score, and is one of the best measures to estimate responsiveness. SRM was 1.55 for the WOSI-Fr and 0.87 for the F-QuickDASH-D/S, both corresponding to strong correlation (SRM > 0.8 indicates strong correlation, 0.5–0.8 moderate correlation and < 0.5 weak correlation). Comparing SRM for the WOSI-Fr and the F-QuickDASH-D/S on paired samples t-test found a significant difference (P < 0.01). Thus, although both the WOSI-Fr and F-QuickDASH-D/S are highly responsive, the WOSI-Fr proved significantly more responsive than the F-QuickDASH-D/S.

4. Discussion

The WOSI questionnaire was successfully translated into the French language (WOSI-Fr) and was adapted for both European and North American French-speaking populations, using a widely accepted translation methodology.

The translated version showed good reliability: i.e., scores did not change significantly between different time-points without clinically relevant change in the underlying condition. Responsiveness, or the ability of the WOSI-Fr to detect clinically significant change in a patient’s condition, was also good. In fact, the WOSI-Fr proved better able to detect change in patients with shoulder instability than did the F-QuickDASH-D/S. This has been previously documented for the original, English versions of the questionnaires. The underlying reason is that the DASH was designed to measure...
pain and dysfunction in the entire upper limb; since shoulder instability patients do not typically experience much pain outside of dislocation episodes, the WOSI focuses more on apprehension and function, and is thus better suited to measure what is clinically relevant for shoulder instability patients.

A fair negative correlation was found with the Walch-Duplay score (negative, because a lower WOSI-Fr score represents a better functional outcome whereas a higher Walch-Duplay score indicates better function). Correlation between the 2 scores was made in a previous study [26], and is merely fair because the 2 scores are designed to measure different patient factors: the WOSI is solely subjective and patient-administered, whereas the Walch-Duplay score has both a subjective and objective component and is administered by the examiner. Thus, a low-demand patient with poor range of motion and some persistent instability may score poorly on Walch-Duplay, being relatively inactive, but better on the WOSI because they are performing at their desired level; conversely, a highly active patient may score well on Walch-Duplay but, because they demand so much from their shoulder, their relative dissatisfaction is demonstrated by a poorer WOSI score.

A major strength of this project was that an accepted translation and cultural adaptation protocol was followed. French-speaking members of the committee were recruited from different sites in different countries in order to see whether a single version of the WOSI could be made which would be adapted to the different populations. Another strength was that a larger number of subjects than recommended in the translation protocol were recruited for both the reliability and responsiveness analyses, adding statistical
power to our conclusions. One limitation was that the 1-year follow-up rate, and thus numbers for the responsiveness analysis, was lower than that for the reliability analysis. There are 2 major reasons for this. One is that patients were enrolled in this cohort on an ongoing basis and were included in the reliability analysis even if they had not had surgery, to maximize numbers. Another reason was that shoulder instability subjects are typically younger and more geographically mobile and usually show good short to medium term evolution: this makes them more difficult to contact and bring back for follow-up. Measures were taken to increase the follow-up rates but, even in the best-case scenario, there would be a high attrition rate, although the WOSI itself is fairly simple to answer, and can easily be filled out while waiting for a clinical appointment.

As shoulder surgery continues to evolve, it is important to measure accurately the results of any surgical or non-surgical intervention, to optimize treatment for any given condition. It is also important to measure outcomes in a way that is relevant to the patient, which the WOSI has been shown to accomplish. With this validated translation, there is now a new tool to facilitate multinational multicenter studies.

5. Conclusion

A French-language translation of the WOSI questionnaire was created and validated for use in both Canadian and Swiss French-speaking populations. The validity and responsiveness of this French version supports its use as equivalent to the English version.

Disclosure of interest

The authors declare that they have no conflicts of interest concerning this article.

Acknowledgment

The authors would like to thank Dr Lionel Neyton (Centre Orthopédique Sancy, Lyon, France) for his participation in the translation committee and in the generation of the WOSI-Fr questionnaire.

Appendix A. Supplementary data

Supplementary data (French version of the WOSI score) associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.otrsr.2013.09.007.

References
