

## RESEARCH ARTICLE

# Psychometric Properties of Full and Shortened Persian-version of Western Ontario Rotator Cuff Index Questionnaires in Persian-speaking Patients with Shoulder Pain

Maryam Daghighi, MSc, PT<sup>1</sup>; Hossein Negahban, PhD, PT<sup>1,2</sup>; Neda Mostafaei, PhD, PT<sup>1</sup>; Mohammad H. Ebrahimzadeh, MD<sup>2</sup>; Ali Moradi, MD<sup>2</sup>; Amir R. Kachooei MD<sup>2</sup>; Aref Saidi, PhD, PT<sup>3</sup>

Research performed at the Mashhad University of Medical Sciences, Mashhad, Iran

Received: 15 March 2022

Accepted: 17 May 2022

## Abstract

**Background:** This study aimed to translate the shortened Western Ontario Rotator Cuff (Short-WORC) questionnaire into Persian and determine the psychometric features of WORC and Persian-Short-WORC in patients with shoulder pain.

**Methods:** A total of 130 patients completed Persian-WORC and -Short-WORC, Shoulder Pain And Disability Index (SPADI), shortened Disability of Arm, Shoulder, and Hand (Quick-DASH), WORC, as well as Short-Form health survey (SF-36), in the evaluation and re-evaluation sessions with an interval of 5-7 days to assess reliability and validity. To determine responsiveness, all patients completed questionnaires and a global rating of change scale before and after the 4-week physiotherapy. Intra-class correlation coefficient (ICC) was used for assessing reliability, two-tailed Pearson ( $r$ ) for validity, as well as longitudinal validity, and receiver operating characteristics (ROC) curve analysis for responsiveness.

**Results:** The ICC was 0.95 (confidence interval: 0.93-0.96) for Short-WORC. A strong correlation was found between Short-WORC, SPADI ( $r=-0.82$ ), Quick-DASH ( $r=-0.79$ ), WORC ( $r=0.92$ ), SF-36 physical ( $r=0.76$ ), and SF-36 mental ( $r=0.71$ ). Floor and ceiling effects were not detected. The responsiveness of Short-WORC and WORC was proven with an area under the curve of  $>0.90$ , and their minimal important change was 28.56 and 26.28 points, respectively.

**Conclusion:** The Persian version of WORC has good psychometric properties to measure disability and health-related quality of life in patients with shoulder pain.

**Level of evidence:** IV

**Keywords:** Persian, Psychometrics, Short-WORC, WORC

## Introduction

Shoulder pain is one of the most prevalent musculoskeletal conditions in primary settings (1), which accounts for a high recurrence rate (1), long recovery period (2), and poor health-related quality of life (HRQOL) (2). This condition includes a wide range of impairments, including frozen shoulder, subacromial pain syndrome, complete/partial rotator cuff (RC)

tear, RC tendonitis, and bursitis (3). One of the most common tools for assessing HRQOL is the Western Ontario Rotator Cuff (WORC) questionnaire. Kirkley et al. (2003) designed WORC as an HRQOL questionnaire for the RC disease (4). It was used to assess HRQOL in partial and full thickness tearing of the RC, tendonitis with calcification, tendon inflammatory conditions, RC

**Corresponding Author:** Hossein Negahban, Department of Physical Therapy, School of Paramedical Sciences, Mashhad University of Medical Science, Mashhad, Iran  
Email: honegahban@yahoo.com



THE ONLINE VERSION OF THIS ARTICLE  
ABJS.MUMS.AC.IR

complaints that need surgery, and shoulder pain (5). The WORC has been translated and validated in many languages and approved as a responsive questionnaire (5-7). Psychometric properties of the Persian-WORC have been proved to evaluate Iranian patients complaining of shoulder pain (8); however, its responsiveness has not been explored yet.

Although the principal value of the WORC is its capability to evaluate the effect of shoulder pain on HRQOL dimensions, there is no evidence to address the factor validation of each subscale (9, 10). The time and accuracy required to answer all questions have also become a challenge in collecting information from crowded clinical settings (8, 9). Therefore, in 2012, Razmjou et al. designed and evaluated the shortened Western Ontario Rotator Cuff (Short-WORC) to address response burden concerns and the absence of factor validity (9). The findings showed that the Short-WORC questionnaire is highly valid, reliable, responsive, has a high level of factor validity, and requires less time and effort to respond (9). Recent studies revealed the Short-WORC had good reliability, validity, as well as responsiveness, and thus, can be used for patients

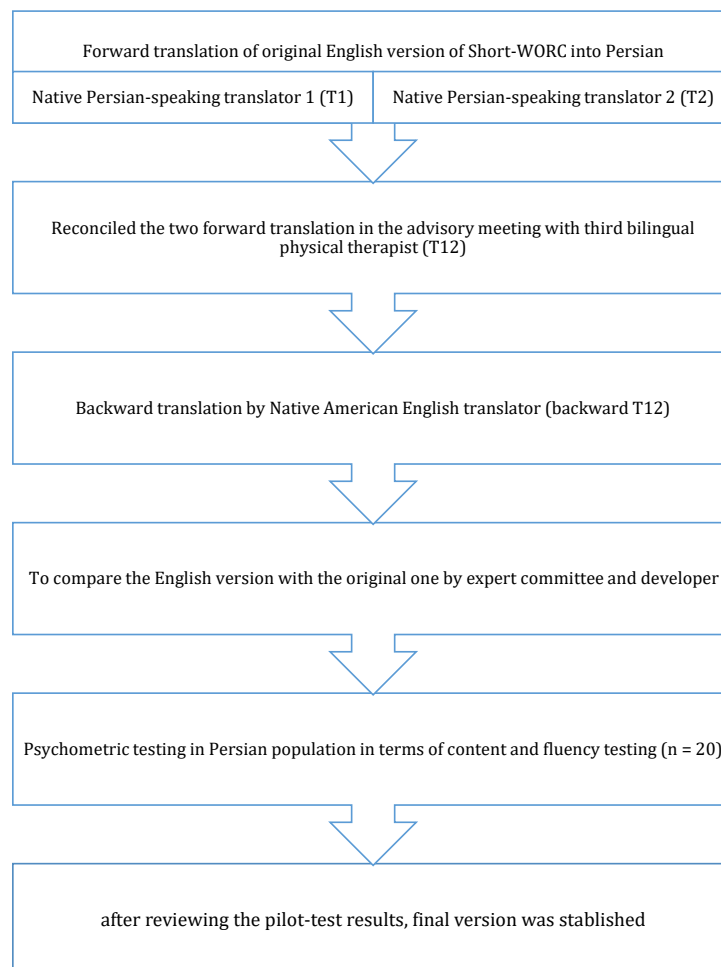
with RC disorders, including shoulder pain, RC tear, total shoulder arthroplasty, osteoarthritis, and RC repair (11-13). However, Short-WORC has not been translated into Persian, and its psychometric properties have not been explored yet.

The racial and linguistic differences between communities living in different countries can affect how Patient-Reported Outcome Measures (PROMs) are completed. Therefore, to use the Short-WORC questionnaire for Iranian patients, it is necessary to translate it and conceptually equate its text. This study thus aimed to cross-culturally evaluate the psychometric properties of the Short-WORC translated into Persian and investigate the responsiveness of the Persian-Short-WORC and the WORC

## Materials and Methods

### Translation process

The Short-WORC questionnaire was translated into Persian using internationally recognized cross-cultural translation and adaptation guidelines (14). The schematic trend of translation is shown in figure 1 [Figure 1]. The



**Figure 1. Translation process.**

developer gave permission to translate the original questionnaire.

First, the original Short-WORC was translated into Persian independently by two native speakers of Persian unfamiliar with the questionnaire (forward translation, T1 and T2). Translators were supposed to emphasize conceptual rather than literal equivalence. In the second step, a bilingual physical therapist evaluated the translations in terms of the quality of the Persian translation, made modifications if needed (in terms of clarity, everyday language use, as well as acceptability), solved and reconciled any disagreements in the advisory meeting with translators, and corrected any conceptual errors or discrepancies. Therefore, the two forward translations came into one (T12). In the third step, the T12 version was translated to English (backward-T12) by one bilingual translator (Persian and English), who was blind to the primary version of the questionnaire and the purpose of the study. In the fourth step, to compare the backward-T12 with the original version, the questionnaire was evaluated by an expert committee (two translators, three investigators, three expert physiotherapists, three expert Shoulder orthopedists, and one linguistics Ph.D. holder), as well as the developer. It was also reviewed for the grammatical structure of the sentences and similarities in meanings. After a minor revision, the pre-final version was established. In the fifth step, 20 patients with shoulder pain completed the Persian-Short-WORC to detect any problematic, uncomfortable, or confusing items. After completing the questionnaire, the authors asked them, "if any question is hard to understand or not?" the possible answers were: 1. Yes or 2. No. Therefore, after the expert committee analyzed the pilot results, the Persian-Short-WORC was eventually prepared for psychometric analysis.

### Study design and the sample

From April 2019 to February 2021, the authors conducted a prospective multicenter study using convenient non-probability sampling, with a sample size of 130 patients. Inclusion criteria were: literate Persian-speaking patients between 18 to 85 years old who had shoulder pain, including frozen shoulder, impingement syndrome, RC tear, tendinitis, and bursitis. Patients were excluded if they had structural shoulder problems, a shoulder joint surgery, and dislocation history of any of the shoulder girdle joints during the previous three months, pain referred to upper extremities past from the neck and chest during the preceding three months, any systematic disease, and inability to complete the questionnaires. Patients included in this study were from seven facilities (four hospital outpatient physiotherapy clinics and three physiotherapy clinics) spread across three populated cities. Three physiotherapists, who had three years of experience in shoulder rehabilitation and using the PROMs, assessed patients. Ethical approval was obtained from the Research Ethics Committee (Approval number: 1398.246), and before participating in the research program, all participants provided their written informed consent. Table 1 shows the characteristics of the patients under study [Table 1].

To assess the reliability of Short-WORC, all patients completed the Persian-Short-WORC in two stages, with an interval of 5-7 days. To ensure that patients' symptoms and general condition had not changed, they were asked, 'Has your shoulder pain-related health status changed since the first visit?' Possible answers were: 1. Yes or 2. No. Only patients with no change in their function and HRQOL were included on the retest day.

To evaluate the construct validity of the Short-WORC questionnaire, the Persian-version of shortened Disability of the Arm Shoulder and Hand (Quick-DASH), Shoulder Pain And Disability Index (SPADI), WORC, and 36-Item Short-Form Health Survey (SF-36) questionnaires were used for a sample of 130 patients. In addition, the

**Table 1. Sociodemographic and clinical characteristics of participants completing the Short-WORC and WORC questionnaires**

Sociodemographic data	N (%) unless stated
Age (yr), Mean (SD)	46 (12.79)
Height (m), Mean (SD)	1.66 (0.11)
BMI (kg/m <sup>2</sup> ), Mean (SD)	27.18 (4.30)
Year of education (yr), number (%)	6-8 68 (39.7)
9-12	74 (49.3)
>12	8 (11.4)
Job	Employed 64 (49.2)
	Unemployed 66 (50.8)
Marital	Single 16 (12.3)
	Married 114 (87.7)
Sex	Men 56 (43.1)
	Women 74 (56.9)
Type of shoulder pain	SAPS 46 (36.9)
	Frozen shoulder 24 (16.2)
	Tendinitis 24 (18.5)
	Bursitis 9 (6.9)
	Osteoarthritis 10 (7.7)
	Partial RC tear 5 (3.8)
	Full RC tear 6 (4.6)
	GH instability 6 (4.6)
Time since disease (yr)	< 6 month 61 (46.9)
	6 month <x<1 year 37 (28.5)
	< 1 year 32 (24.6)

SD: Standard deviation, WORC: Western Ontario rotator cuff, Short-WORC: Shortened version of Western Ontario Rotator Cuff, BMI: Body mass index, SAPS: Subacromial pain syndrome, RC: Rotator cuff, GH: Glenohumeral instability

authors explored the interpretability of Short-WORC by evaluating the floor and ceiling effects (13).

A sample of 130 patients, who wished to attend physiotherapy sessions, was included to evaluate responsiveness. They completed Short-WORC and WORC at inclusion and after the 4-week physiotherapy. In the last physiotherapy session, they also completed the 7-item Global Rating of Change (GRC) scale as an external standard (15). Based on the COSMIN checklist, responsiveness depends only on the selected external standard and not on the type of therapy (16); therefore, the authors did not precisely control the type of intervention. However, for the consistency of intervention among clinics, the overall physiotherapy program was the same. It included RC strengthening, as well as sensory-motor and stretching exercises.

### Instruments

The WORC consists of 21 items presented under five domains, and each item has a score range of 0-100 mm, with the total score ranging from 0 (the best HRQOL) to 2,100 (the worst HRQOL) (4). The WORC scores can be reported as percentages, with the total score represented from 0% (the lowest score) to 100% (the highest score) by subtracting the total from 2,100, dividing it by 2,100, and multiplying it by 100 (17). The Persian-WORC has an acceptable level of reproducibility and validity (8).

The Short-WORC consists of all work and lifestyle domain items, except one lifestyle item related to roughhousing (question 17) (9). Its score ranges from 0 (the best possible score) to 700 (the worst possible score). Additionally, the total Short-WORC score can range from 0% (the lowest score) to 100% (the highest score) by subtracting the score from 700, dividing it by 700, and multiplying it by 100 (9).

The Quick-DASH has 11 items that measure the upper limb's physical disability and symptoms. The scores of all completed responses are produced out of five. The score could be represented in percentages by subtracting one from it and multiplying it by 25. Final scores range from 0 (the best functional status) to 100 (the worst functional status). The Persian Quick-DASH has been validated to use in Iran and has adequate ability to determine changes in disability among patients with shoulder pain (18, 19).

The SPADI is a self-administered, 13-item questionnaire. The score is obtained by dividing the total score by 130 and multiplying it by 100. Total scores range from 0 (the best functional status) to 100 (the worst functional status). The SPADI is a valid and responsive questionnaire to detect patients (20). Ebrahimzadeh et al. (2014) evaluated the psychometric properties of the Persian-SPADI and reported its adequate validity (21). Negahban et al. (2014) explored the responsiveness of the Persian-SPADI and reported its excellent responsiveness (22). Additionally, the findings of the study by Jerosch-Herold et al. (2018) presented that SPADI has a good structural validity in assessing patients with shoulder pain (23).

The SF-36 is primarily used in health outcome surveys mostly to evaluate the health status in general (24). The

SF-36 has good psychometric properties to measure HRQOL among the Iranian population (24). Possible scores range from 0 (the poorest HRQOL) to 100 (the best HRQOL) (24).

The 7-item GRC scale represents, as the external anchor, the patient's view on the amount of change that occurred (15). Each item of the scale has a different meaning: 1=very much worse, 2=much worse, 3=slightly worse, 4=no change, 5=slightly better, 6=much better, and 7=very much better. The authors defined the cut-off point for GRC as follows: scores less than or equal to 5 as "unimportant change" and scores higher than 5 as "important change."

### Statistical analysis

The authors used the SPSS software (version 24) to analyze data and considered  $P < 0.05$  as statistically significant. Based on the COSMIN recommendation, 50 to 100 participants were enough for evaluating psychometric properties (25). Therefore, with a possible 15% of attrition, the estimated desired sample size was considered 130 participants. Data were assessed for completeness, the percentage of missing data, and the presence of outliers. The statistical significance was set at an alpha less than 0.05 at a 95% confidence interval (CI). Additionally, descriptive statistics were examined to determine data distributions. The normal distribution of variables was evaluated by the Kolmogorov-Smirnov test, the results of which indicated that all variables had normal distribution, except for the mean changes of all variables (26). Therefore, Pearson ( $r$ ) and paired t-tests were used for normally distributed data and Spearman's rank correlation coefficient ( $r_s$ ) for evaluating the longitudinal validity of data with non-normal distribution (22).

To assess the reliability of the Short-WORC, the authors investigated the related reliability by intra-class correlation coefficient (ICC). Therefore, to evaluate the amount of variation over time in stable patients, the 2-way random-effects model of ICC (ICC<sub>2,1</sub>) was used, as proposed by Shrout and Fleiss (27). The ICC was classified in terms of the agreement; an ICC of  $> 0.70$  indicates good reliability, whereas an ICC of  $< 0.70$  represents poor to moderate reliability (25). Cronbach's Alpha was assessed to investigate the internal consistency (25). The obtained scores from the first administration of Short-WORC were used to calculate this value. Cronbach's Alpha value of  $> 0.70$  was considered a good internal consistency (25). The standard measurement error (SEM) and minimal detectable change (MDC) were utilized to assess absolute reliability. The SEM is estimated by  $SD \times \sqrt{(1 - R)}$ , with  $R = ICC$  and  $SD = \sqrt{\text{total variance}}$ . It was used to calculate the MDC by  $1.96 \times \sqrt{2} \times SEM$ . Normally, smaller SEM and MDC values indicate the clinical usefulness of the scale (11).

The authors assessed the construct validity of Short-WORC, performing hypothesis testing using two-tailed Pearson ( $r$ ) (28). All hypotheses were defined before data collection and based on previous similar literature (12, 29). It was hypothesized that Short-WORC would have 1) a strong positive correlation with WORC ( $r > 0.70$ ),

2) a strong negative correlation with SPADI ( $r > -0.70$ ), 3) a strong negative correlation with Quick-DASH ( $r > -0.70$ ), and 4) a moderate positive correlation with SF-36 ( $r > 0.60$ ). In addition, it was hypothesized that Short-WORC would have a higher positive correlation with SF-36 physical score than SF-36 mental score. According to the previous literature, the correlation coefficient can be considered weak ( $< 0.30$ ), moderate ( $0.30-0.70$ ), or strong ( $> 0.70$ ) (30).

Floor and ceiling effects represent the percentage of patients with the minimum and maximum scores on the Short-WORC questionnaire from the first testing session. Floor and ceiling effects were considered if more than 15% scored at either end of the scale (31).

In a longitudinal context, the authors followed the COSMIN framework, which represented responsiveness as an approach to validity (25). Additionally, De vet et al. (2011) introduced two approaches for assessing responsiveness, a criterion and a construct approach (31). They believed the criterion approach was used for situations with a gold standard, and that there is only one acceptable gold standard for a shortened version of a questionnaire. In the present study, WORC is the gold standard for Short-WORC. However, to assess WORC responsiveness, the GRC was used as the gold standard. The Spearman's rank correlation coefficient ( $r_s$ ) was utilized to examine the correlation with the external anchor (GRC) (16). The responsiveness of the Short-WORC and WORC was considered acceptable if  $r_s \geq 0.70$  (25).

The receiver operating characteristics (ROC) curve analysis (95%CI) was used to assess the ability of the two questionnaires (Short-WORC and WORC) to correctly classify patients as improved and not-improved. Sensitivity and 1-specificity are plotted at several cut-off points, and Area under the curve (AUC) can be estimated in ROC curve analysis. The AUC was used as an indicator of responsiveness (32) and measures the instrument's ability to discriminate between two groups, according to an external anchor. The authors considered adequate responsiveness if a  $ROC_{AUC}$  of  $> 0.70$  was represented (25). The minimal important change (MIC) was considered the optimal cut-off point on the ROC curve (i.e., the value for which the sum of the proportions of misclassification ([1-sensitivity] + [1-specificity]) is small) (25).

## Results

A total of 228 patients met the inclusion criteria, while

95 were unwilling or unable to participate, and three were excluded because of systematic problems. Finally, 130 patients were included, whose characteristics are shown in Table 1 [Table 1]. The authors found no missing value, dropout, or outlier in the sample. Participants did not receive any additional care during the test-retest intervals or the study period, and no adverse effects were reported. The authors checked the presence of missing responses at the item level in Short-WORC and all other assessment tools for each patient and asked the patients to complete the questionnaire if any missing responses existed. Moreover, during completion, when patients did not want to answer a question, they could easily withdraw from the study. However, none of them declined to answer any questions on the questionnaires.

The authors found the Persian-Short-WORC a satisfactory equivalent to the original version, and the developer also accepted the questionnaire. Only a few adjustments were made in wording during the translation process due to semantic differences. As a case in point, there was a disagreement about questions 3 and 7, and after the meeting, the translators reached a conclusion and consensus. The pilot test results showed that the instrument could be applied in clinical settings. All patients answered 'no' to the question we asked, completed the questionnaire without any difficulties, and claimed that the questions were transparent. Therefore, there were no difficulties or modifications.

All 130 patients had stable health conditions during the test-retest interval and enrolled in the test-retest reliability process. Table 2 shows the excellent level of ICC ( $= 0.95$ ) for Short-WORC [Table 2]. There was a slight mean difference ( $= 1.69$ ) between the test-retest measures of the Short-WORC. With Cronbach's alpha of  $= 0.97$ , the internal consistency was also excellent. On the other hand, the SEM equaled 5.31 scale points while the MDC was 10.40.

All 130 patients agreed to participate in the validity analysis. As shown in Table 3, there was a significantly strong correlation between the Short-WORC and Quick-DASH scores ( $r = -0.79$ ,  $P < 0.001$ ), SPADI scores ( $r = -0.82$ ,  $P < 0.001$ ), and WORC scores ( $r = 0.92$ ,  $P < 0.001$ ) [Table 3]. As hypothesized, higher correlations were observed for SF-36 physical ( $r = 0.76$ ,  $P < 0.001$ ) than SF-36 mental ( $r = 0.71$ ,  $P < 0.001$ ). No floor or ceiling effects were observed on the Short-WORC. The analyses of validity

**Table 2. Test-retest reliability and internal consistency of the Persian Short-WORC (n = 130)**

Questionnaire	Test Mean (SD)	Re-test Mean (SD)	ICC (95% CI) CA (n = 130)	CA (n = 130)	SEM	MDC
Short-WORC	31.70 (24.34)	40.85 (24.58)	<b>0.95</b> (0.93 - 0.96)	<b>0.97</b>	2.12	10.40

ICC and Cronbach's alpha greater than 0.70 are in bold.

ICC: Intra-class correlation coefficient, CA: Cronbach's alpha;

SEM: Standard error of the mean, MDC: Minimal detectable

change, Short-WORC: Shortened version of Western Ontario

Rotator Cuff.

**Table 3. Construct and content validity analysis of Short-WORC and longitudinal validity analysis of Short-WORC and WORC (n = 130)**

Construct validity (Pearson r)					
Questionnaire	WORC (r)	SPADI (r)	Quick-DASH (r)	Physical SF-36 (r)	Mental SF-36 (r)
Short-WORC	<b>0.92</b>	<b>-0.82</b>	<b>-0.79</b>	<b>0.76</b>	<b>0.71</b>
Interpretability					
Questionnaire	Ceiling effects (%)			Floor effects (%)	
Short-WORC	0			1 (1/100)	
Longitudinal validity (Spearman $r_s$ )					
Questionnaire	Short-WORC ( $r_s$ )	WORC ( $r_s$ )	Likert scale ( $r_s$ )	-----	-----
Short-WORC	-----	<b>0.93</b> <b>(&lt;0.001)</b>	<b>0.72</b> <b>(&lt;0.001)</b>	-----	-----
WORC	<b>0.93</b> <b>(&lt;0.001)</b>	-----	<b>0.74</b> <b>(&lt;0.001)</b>	-----	-----

Significant correlation coefficients at  $p < 0.001$  are in bold.

Short-WORC: Shortened version of Western Ontario Rotator Cuff, WORC: Western Ontario Rotator Cuff, Quick-DASH: shortened version of Disability of Arm, Shoulder, and Hand, SPADI: Shoulder Pain And Disability Index, SF-36: Short form health survey

**Table 4. Area under the receiver operating characteristic (AUC) curve for Short-WORC and WORC according to external, dichotomized measure of global rating scale (improved versus unimproved) (n=100)**

Questionnaire	AUC (95% CI)	MIC (Optimal cutoff value)	Sensitivity (95% CI)	Specificity (95% CI)
Short-WORC	<b>0.89 (0.83-0.95)</b>	28.56	0.83 (0.70-0.91)	0.83 (0.67-0.92)
WORC	<b>0.90 (0.85-0.96)</b>	26.28	0.86 (0.74-0.93)	0.81 (0.64-0.90)

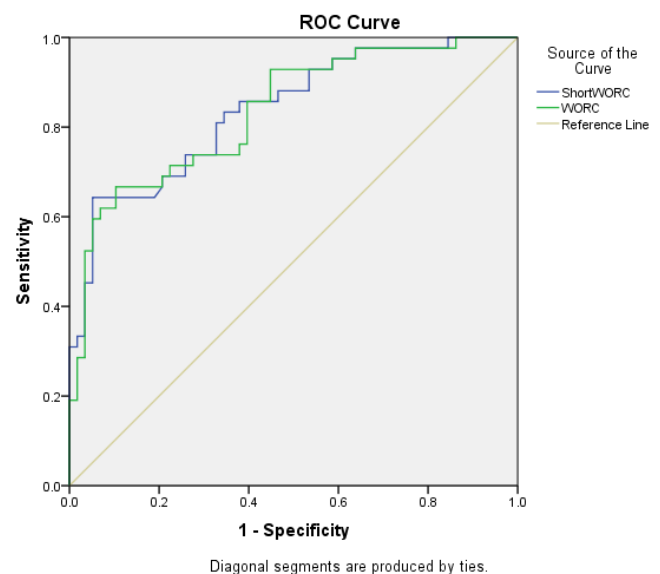
AUC (area under curve) equal or greater than 0.70 are in bold.

Short-WORC: Shortened version of Western Ontario Rotator Cuff, WORC: Western Ontario Rotator Cuff, MIC: Minimal Important Change.

hypotheses confirmed all the pre-defined hypotheses and indicated an acceptable validity (4/4, 100%).

Out of 130 patients, 100 individuals enrolled in the responsiveness process. Table 3 demonstrates a high longitudinal validity for Short-WORC and WORC scores. The mean change of the Short-WORC questionnaire was highly correlated with the mean change scores of WORC scores ( $r=0.93$ ) and GRC ( $r=0.72$ ) after four weeks of physiotherapy. In addition, the mean change of the WORC scale was highly correlated with GRC scores ( $r=0.74$ ).

Table 4 shows good and excellent responsiveness with an AUC of 0.89 (95%CI: 0.83-0.95) for the Short-WORC and an AUC of 0.90 (95%CI: 0.80-0.96) for the WORC scales [Table 4]. The MIC values were equal to 26.28 and 28.56 points for Short-WORC and WORC, respectively. Based on the GRC scale, 59 patients were classified as improved and 41 as not-improved. The number of patients identified as very much better, much better, slightly better, no change, slightly worse, much worse, and very much worse were 42, 17, 26, 15, 0, 0, and 0, respectively. The schematic curve of Short-WORC and WORC responsiveness are represented in Figure 2 [Figure 2].

**Figure 2. Responsiveness of Short-WORC and WORC questionnaire in receiver operating curve (ROC).**

## Discussion

The findings showed an excellent level of test-retest reliability and internal consistency of the Short-WORC. The Short-WORC and WORC were highly responsive to detect changes following the physiotherapy intervention.

The authors followed recently-published guidelines and recommendations for the translation and adaptation processes (14). There was only one deviation from the translation approach due to financial constraints. In the present study, the authors collaborated with just one translator for the backward translation resulting in a potential bias (14). The guideline standard trend confirmed the Short-WORC was highly analogous to the original version. In addition, validity was confirmed by checking the translation process in an expert committee and testing with a pilot sample in the clinical setting to check for understandability and acceptability. The involvement of the original author throughout the process was a strong positive point.

Only a few studies explored the reliability of the Short-WORC questionnaire (9, 11, 33). In studies by Razmjou et al. (2012), Dewan et al. (2016), and Furtado et al. (2020), the findings confirmed good internal consistency, as well as absolute reliability, and the results were slightly less powerful than what the present study reported (9, 11, 33). These small variations could be due to the difference in populations (RC disorders which need operation) with more disability levels, longer time intervals (a week before surgery and three months postoperatively) and different statistical methods (using the 2-way mixed model), as well as the small sample size (n=51) and the longer time interval between test-retest sessions (9, 11, 33). As all estimates of the related reliability exceeded the accepted score, this questionnaire had acceptable reliability across multiple contexts.

Here the authors tried to discuss studies exploring the validation of Short-WORC. In the study by Razmjou et al. (2012), high correlation scores were observed between two versions of WORC, similar to the present study (9). Dewan et al. (2018) also explored its validity in patients who had RC repair, revealing that Short-WORC had an excellent, fair, and good correlation with WORC, SPADI, and Quick-DASH, respectively (12). Contrary to the present study findings, the floor effect of their study at baseline was 18.8% on the Short-WORC, which was slightly more than the standard threshold of 15%. While their findings suggest caution in using the Short-WORC to evaluate worsening, the present study findings could strongly recommend Persian Short-WORC to evaluate improving/worsening conditions. It was hypothesized that the difference between their results and those in the present study is due to different samples and sample sizes. This might be because short versions are always more compromised than the original version, and a larger sample size is needed to detect results (34).

The findings revealed that both Persian-Short-WORC and WORC questionnaires were highly responsive to determining patient condition changes, which supports

previous results (12, 35-37). Correlation analysis also confirmed a strong correlation between the mean change scores of the Short-WORC and WORC. Both Short-WORC and WORC showed acceptable responsiveness (22). In addition, Terwee et al. (2007) recommended that when patients' MDC exceeds SEM, the change is an essential and real one, which can be considered MIC; therefore, the instrument can detect clinically important changes over time (MIC>SEM) (38). Similarly, Short-WORC MIC exceeded the SEM in this study, so the authors could strongly recommend that when Short-WORC and WORC scores reach the MIC value, real and clinical change happens.

Three studies explored the WORC responsiveness, yet only Braun et al. (2018) demonstrated the responsiveness of WORC using anchored-based methods (35-37). They concluded that the MIC for WORC was estimated at 14.28 points for 'improved' patients with a shoulder-related disability, which was slightly lower than what the present study reported (=26.28). It was hypothesized that the difference was because of different statistical methods (logistic regression), different sample sizes (=64), and more disability levels. The other two studies explored internal responsiveness using distribution-based methods (standard response mean). The present findings may be comparable with their findings, which showed that WORC has acceptable responsiveness, as did the present findings (35, 36). Although no study has ever been conducted on the responsiveness of Short-WORC through anchor-based methods in patients with shoulder pain, one original validation study explored the internal responsiveness (12). The present results may be comparable to the findings by Dewan et al. (2018), exploring the internal aspect of responsiveness and showing the high responsiveness of Short-WORC. The present study explored the external element and showed the same result.

Both Persian Short-WORC and WORC questionnaires have good psychometric properties in test-retest reliability, validity, and responsiveness. Therefore, Short-WORC and WORC questionnaires had a high level of psychometric properties to measure functional limitations and HRQOF in patients with shoulder pain in clinical settings.

## Limitation

This study suffered from some limitations. According to the guidelines, a 2-week interval is often appropriate to reduce recall bias (14); however, the time interval between repeated questionnaire completions in this study (5 to 7 days with a mean of 6±64 days) was not controlled for ethical reasons. The other limitation was collaborating with only one backward translator due to financial constraints.

## Acknowledgment

This study is research project of Mrs. Daghiani Special thanks to Mashhad University of Medical Science for the financial support.

**Source of Support:** This work was supported by a research grant in Mashhad University of Medical Sciences, Mashhad, Iran

**Declaration of interest:** None of the authors have any financial or other interests relating to the manuscript to be submitted for publication in Archive of Bone & Joint Surgery.

**Declaration of informed consent:** There is no information (names, initials, hospital identification numbers or photographs) in the submitted manuscript that can be used to identify patients.

Maryam Daghighi MSc PT<sup>1</sup>

Hossein Negahban PhD PT<sup>1,2</sup>

Neda Mostafaei PhD PT<sup>1</sup>

Mohammad Hosein Ebrahimzadeh MD<sup>2</sup>

Ali Moradi MD<sup>2</sup>

Amir R. Kachooei MD<sup>2</sup>

Aref Saidi PhD PT<sup>3</sup>

<sup>1</sup> Department of Physical Therapy, School of Paramedical Sciences, Mashhad University of Medical Sciences, Mashhad, Iran

<sup>2</sup> Orthopedic Research Center, Mashhad University of Medical Sciences, Mashhad, Iran

<sup>3</sup> Faculty Member, University of Lahore, Pakistan

## References

1. Greving K, Dorrestijn O, Winters JC, Groenhof F, van der Meer K, Stevens M, et al. Incidence, prevalence, and consultation rates of shoulder complaints in general practice. *Scand J Rheumatol.* 2012;41(2):150-5.
2. Östör AJK, Richards CA, Prevost AT, Speed CA, Hazleman BL. Diagnosis and relation to general health of shoulder disorders presenting to primary care. *Rheumatology.* 2005;44(6):800-5.
3. Lewis J. Rotator cuff related shoulder pain: Assessment, management and uncertainties. *Manual Therapy.* 2016;23:57-68.
4. Kirkley A, Alvarez C, Griffin S. The development and evaluation of a disease-specific quality-of-life questionnaire for disorders of the rotator cuff: The Western Ontario Rotator Cuff Index. *Clin J Sport Med.* 2003;13(2):84-92.
5. Furtado R, MacDermid JC, Nazari G, Bryant DM, Faber KJ, Athwal GS. Cross-cultural adaptations and measurement properties of the WORC (Western Ontario rotator cuff index): a systematic review. *Health Qual Life Outcomes.* 2020;18(1):17.
6. Wiertsema SH, Rietberg MB, Hekman KM, Schothorst M, Steultjens MP, Dekker J. Reproducibility of the Dutch version of the Western Ontario rotator cuff Index. *J Shoulder Elbow Surg.* 2013;22(2):165-70.
7. Brix LD, Bjørnholdt KT, Nikolajsen L, Kallestrup K, Thillemann TM. The Danish version of the Western Ontario Rotator Cuff Index. *Dan Med J.* 2020;67(2).
8. Mousavi SJ, Hadian MR, Abedi M, Montazeri A. Translation and validation study of the Persian version of the Western Ontario Rotator Cuff Index. *Clin Rheumatol.* 2009;28(3):293-9.
9. Razmjou H, Stratford P, Holtby R. A Shortened Version of the Western Ontario Rotator Cuff Disability Index: Development and Measurement Properties. *Physiother Can.* 2012;64(2):135-44.
10. Wessel J, Razmjou H, Mewa Y, Holtby R. The factor validity of the western ontario rotator cuff index. *BMC Musculoskelet Disord.* 2005;6(1):1-7.
11. Dewan N, MacDermid JC, MacIntyre N, Grewal R. Reproducibility: reliability and agreement of short version of Western Ontario rotator cuff index (short-WORC) in patients with rotator cuff disorders. *J Hand Ther.* 2016;29(3):281-91.
12. Dewan N, MacDermid JC, MacIntyre N. Validity and Responsiveness of the Short Version of the Western Ontario Rotator Cuff Index (Short-WORC) in Patients With Rotator Cuff Repair. *J Orthop Sports Phys Ther.* 2018;48(5):409-18.
13. Furtado R, MacDermid JC, Bryant DM, Faber KJ, Athwal GS. Interpretation and content validity of the items of the numeric rating version short-WORC to evaluate outcomes in management of rotator cuff pathology: a cognitive interview approach. *Health Qual Life Outcomes.* 2020;18(1):88.
14. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine.* 2000;25(24):3186-91.
15. Houweling TA. Reporting improvement from patient-reported outcome measures: A review. *Clinical Chiropractic.* 2010;13(1):15-22.
16. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol.* 2000;53(5):459-68.
17. Holtby R, Razmjou H. Measurement properties of the Western Ontario rotator cuff outcome measure: a preliminary report. *J Shoulder Elbow Surg.* 2005;14(5):506-10.
18. Ebrahimzadeh MH, Moradi A, Vahedi E, Kachooei AR, Birjandinejad A. Validity and Reliability of the Persian Version of Shortened Disabilities of the Arm, Shoulder and Hand Questionnaire (Quick-DASH). *Int J Prev Med.* 2015;6:59.
19. Budtz CR, Andersen JH, de Vos Andersen N-B, Christiansen DH. Responsiveness and minimal important change for the quick-DASH in patients with shoulder disorders. *Health Qual Life Outcomes.* 2018;16(1):226.
20. Williams JW, Jr, Holleman DR, Jr, Simel DL. Measuring shoulder function with the Shoulder Pain and Disability Index. *J Rheumatol.* 1995;22(4):727-32.
21. Ebrahimzadeh MH, Birjandinejad A, Golhasani F,



- Moradi A, Vahedi E, Kachooei AR. Cross-cultural adaptation, validation, and reliability testing of the Shoulder Pain and Disability Index in the Persian population with shoulder problems. *Int J Rehabil Res*. 2015;38(1):84-7.
22. Negahban H, Behtash Z, Sohani SM, Salehi R. Responsiveness of two Persian-versions of shoulder outcome measures following physiotherapy intervention in patients with shoulder disorders. *Disabil Rehabil*. 2015;37(24):2300-4.
23. Jerosch-Herold C, Chester R, Shepstone L, Vincent JJ, MacDermid JC. An evaluation of the structural validity of the shoulder pain and disability index (SPADI) using the Rasch model. *Qual Life Res*. 2018;27(2):389-400.
24. Motamed N, Ayatollahi AR, Zare N, Sadeghi-Hassanabadi A. Validity and reliability of the Persian translation of the SF-36 version 2 questionnaire. *East Mediterr Health J*. 2005;11(3):349-57.
25. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539-49.
26. Negahban H, Pouretzad M, Yazdi MJS, Sohani SM, Mazaheri M, Salavati M, et al. Persian translation and validation of the Kujala Patellofemoral Scale in patients with patellofemoral pain syndrome. *Disabil Rehabil*. 2012;34(26):2259-63.
27. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420.
28. Zhaeentan S, Legeby M, Ahlstrom S, Stark A, Salomonsson B. A validation of the Swedish version of the WORC index in the assessment of patients treated by surgery for subacromial disease including rotator cuff syndrome. *BMC Musculoskelet Disord*. 2016;17:165.
29. Razmjou H, Bean A, van Osnabrugge V, MacDermid JC, Holtby R. Cross-sectional and longitudinal construct validity of two rotator cuff disease-specific outcome measures. *BMC Musculoskelet Disord*. 2006;7(1):26.
30. Andresen EM. Criteria for assessing the tools of disability outcomes research. *Arch Phys Med Rehabil*. 2000;81:S15-20.
31. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in Medicine: A Practical Guide*. Cambridge: Cambridge University Press; 2011.
32. Miller, J. Philip, Rao, C. Radhakrishna, Rao, D. C. *Essential Statistical Methods for Medical Statistics*. Burlington: North Holland; 2011.
33. Furtado R, MacDermid JC, Bryant DM, Faber KJ, Athwal GS. Evaluating the reproducibility of the short version of the Western Ontario Rotator Cuff Index (Short-WORC) prospectively. *JSES Int*. 2020;4(1):197-201.
34. Locker D, Jokovic A, Clarke M. Assessing the responsiveness of measures of oral health-related quality of life. *Community Dent Oral Epidemiol*. 2004;32(1):10-8.
35. St-Pierre C, Dionne CE, Desmeules F, Roy JS. Reliability, validity, and responsiveness of a Canadian French adaptation of the Western Ontario Rotator Cuff (WORC) index. *J Hand Ther*. 2015;28(3):292-9.
36. de Witte PB, Henseler JF, Nagels J, Vliet Vlieland TP, Nelissen RG. The Western Ontario rotator cuff index in rotator cuff disease patients: a comprehensive reliability and responsiveness validation study. *Am J Sports Med*. 2012;40(7):1611-9.
37. Braun C, Handoll HH. Estimating the Minimal Important Difference for the Western Ontario Rotator Cuff Index (WORC) in adults with shoulder pain associated with partial-thickness rotator cuff tears. *Musculoskelet Sci Pract*. 2018;35:30-3.
38. 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(1):34-42.