

# 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.

Rochelle Furtado (✉ [rfurtad5@uwo.ca](mailto:rfurtad5@uwo.ca))

Western University Faculty of Health Sciences <https://orcid.org/0000-0002-9186-9785>

Joy C MacDermid

Western University

Dianne Bryant

Western University

Kenneth J Faber

Western University

George S Athwal

Western University

---

## Research

**Keywords:** patient reported outcomes, rotator cuff disorders, content validity, quality of life

**Posted Date:** March 2nd, 2020

**DOI:** <https://doi.org/10.21203/rs.2.19542/v3>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Health and Quality of Life Outcomes on March 30th, 2020. See the published version at <https://doi.org/10.1186/s12955-020-01339-7>.

## Abstract

**Background:** The shortened version of the Western Ontario Rotator Cuff Index (Short-WORC) is a patient reported outcome measure that evaluates quality of life (QoL) of patients with rotator cuff pathology. However, formal content validation of the full or Short-WORC has not been reported. This study aims to understand how 1) people interpret and calibrate responses to items on the Short-WORC and 2) compensatory strategies that might enhance function and thereby affect responses. **Methods:** This study uses cognitive interviewing, a qualitative methodology that focuses on the interpretation of questionnaire items. Patients with rotator cuff disorders (n=10), clinicians (n=6) and measurement researchers (n=10) were interviewed using a talk aloud structured interview that evaluated each of the 7 items of the Short-WORC. All interviews were recorded and transcribed verbatim by one researcher (R.F). Analysis was done through an open coding scheme using a previously established framework. **Results:** Overall, the items on the Short-WORC were well received by participants. Through the interviews, the 6 themes of: Comprehension, Inadequate response definition, Reference Point, Relevance, Perspective Modifiers and Calibration Across Items emerged. The items of working above the shoulder (90%), compensating with the unaffected arm (88%) and lifting heavy objects (92%) were the most relevant to participants. Participants calibrated their scores on the items of sleeping and styling (19%) the most. Perspective modifiers of gender, influenced the calibrations of items of styling your hair (30%) and dressing or undressing (19%). Compensatory strategies of task-re allocation and using assistive devices/resources were frequently mentioned by participants. Overall, participants had minor comprehension issues, but found the 7- items of the Short-WORC to be relevant to QoL. **Conclusions:** Therefore, the findings demonstrate that the Short-WORC is not cognitively complex, but varies with patient perspectives. Overall, the Short-WORC provides evidence of demonstrating strong content validity when used for rotator cuff disorder patients.

## Introduction

Rotator cuff disorders (RCDs) include a spectrum of pathologies that can lead to shoulder pain, impairment and activity limitation.<sup>1</sup> While the spectrum of disorders vary, rotator cuff tears are a common problem in the current patient population. Rotator cuff tears are commonly associated with exposure to repetitive movements or strain.<sup>1,2</sup> The prevalence of tears increases with age, affecting more than 60% of patients who are over the age of 60<sup>3</sup> and results in a reduced quality of life (QoL).<sup>1</sup>

The construct of QoL is critical for defining optimal treatments, as the goal of surgery and rehabilitation is to improve the QoL in patients.<sup>3,4</sup> The previous version of the Western Ontario Rotator Cuff Index (WORC) developed by Kirkley et al, is one of the most validated disease-specific questionnaire to measure QoL in RCDs.<sup>5</sup> The WORC focuses on 5 domains; 1) pain and physical symptoms, 2) sports and recreation, 3) work, 4) lifestyle, and 5) emotions.<sup>5</sup> While it has been translated and validated in a variety of different languages<sup>1,2</sup>, the WORC created challenges of patient response burden (time spent to answer questionnaire) and complexity identified through patient interviews and statistical methods<sup>1,2</sup> A shortened version of the Western Ontario Rotator Cuff Index (Short-WORC) was created to address these concerns.<sup>7,27</sup> The Short-WORC by Razmjou et al. contains seven items from the domains of work and lifestyle, focusing on the activity limitations that arise from RCDs.<sup>1,4,6,7</sup> Previously, participants completed a visual analogue scale (VAS) to score their response,<sup>7</sup> however, in the present study we have modified the responsiveness scale to a 0-10-point numeric scale. According to previously literature, the use of a numeric scale reduces patient response burden and increases patient satisfaction.<sup>8</sup> As previously mentioned in prior work validating the Short-WORC,<sup>1,6</sup> the construct of QoL may not be fully retained in this abbreviated questionnaire. In fact, it seems the Short-WORC is assessing the construct of activity limitation and function. Nevertheless, in our preliminary studies and those of others, the Short-WORC has demonstrated measurement properties that are similar to the original WORC.<sup>1,6</sup> While the Short-WORC demonstrates equally strong psychometric performance when items are extracted from the full WORC,<sup>1,6</sup> it has yet to be validated as an independent outcome measure in clinic.<sup>27</sup>

A fundamental aspect of validation is understanding the content validity of a questionnaire.<sup>9</sup> Content validity refers to the extent to which a measure represents all facets of a given construct.<sup>10</sup> When evaluating a shortened PRO, it is advised that while some properties can be obtained from the original study, the property of comprehensiveness should be evaluated from a new study of the shortened PRO.<sup>15</sup> Therefore, it is important that researchers re-evaluate the content validity of a shortened PRO, to verify that it measures the intended construct of the original. According to the Food and Drug Administration's (FDA) guidelines<sup>11</sup>, content validity can be assessed through conducting interviews that seek to evaluate 1) the clarity of the instructions, 2) the content of each item and 3) that the intended meaning of each item is easily interpreted by participants. In this process, the recall and response scales of the PRO are also evaluated.<sup>9,10</sup> Therefore, this study primarily aims to evaluate content validity by exploring how people interpret and calibrate responses to items on the Short-WORC. A secondary aim of this study is to understand how compensatory strategies may influence the way participants interpret and determine responses to the items.

## Methods

### *Study Design*

This study used a descriptive qualitative approach based on the principles of cognitive interviewing to explore participants' interpretations of specific words, constructs (variables that cannot be measured directly but are informed through other variables that are measurable) and phrases on the Short-WORC. This enables an understanding of how participants calibrate options when responding to the measure.<sup>12</sup> Cognitive interviewing uses semi-structured interviews, a talk out loud approach, and probes to understand how patients interpret and respond to items on a self-report questionnaire.<sup>12</sup> These items measure quality of life on the questionnaire. This allows a combination of concurrent (while answering the question) or retrospective (immediately after answering the item) answering, which gathers optimal data quality.<sup>12</sup> Participants were provided with a version of the Short-WORC, that had a numerical scale from 0 -10.

### *Setting and sample*

Interviews were conducted in a small private room at the Hand and Upper Limb Clinical (HULC) Research Laboratory, London, Canada. Patient and healthcare provider participants were recruited from St. Joseph's Health Care London and researchers were recruited from Western University (London, Canada). Participants who met the inclusion criteria were invited to participate in the study; greater than 18 years of age, can speak and read English and did not have another mental or physical ailment that could contraindicate the shoulder injury or not allow them to be able to participate in the interview process. This questionnaire was introduced into standard routine care of core outcome measures for shoulders at HULC. Patients only answered questions about the Short-WORC and no other PRO.

Through purposeful sampling, we aimed to include perspectives of healthcare providers and recipients.<sup>13</sup> Therefore, patients (n=10), healthcare providers (n=6), and measurement researchers (n=10) were recruited. Patients who had received some treatment (i.e. surgery, physiotherapy) for their rotator cuff disorder at HULC (n=6) and patients who had been diagnosed at HULC but not received treatment for their shoulder (n=4) were recruited to achieve a diversity of participant experiences. Patients had a variety of rotator cuff disorders. Diagnoses ranged from under a year (n=4) for patients who have not received treatment, and from 1 year to 5 years of diagnosis (n=6) for patients who had received treatment. Healthcare providers were a mix of both sexes and occupations included: nurses, surgeons, physiotherapists and occupational therapists practicing in London, Ontario. Measurement researchers included trainees of both sexes in the faculty of Health and Rehabilitation Sciences at Western University. Trainees have had exposure to questionnaire design and evaluating psychometric properties. Participation of both men and women of varying age groups allowed for a diversity of experiences. Recruitment for interviews stopped when saturation of the responses was achieved.<sup>12</sup> The study protocol was reviewed and approved by Lawson Healthcare and Western Research Ethics Board (WREB).

#### *Data collection*

Participants provided written informed consent prior to the interview. Interviews were conducted in English by one researcher (RF) and lasted between 40 and 60 minutes. All interviews were recorded on an encrypted tape recorder, and then transcribed verbatim.

The interview structure was informed by previously published work<sup>12,19</sup> and multiple discussions with the research team. Interviews focused on participants' interpretation of each individual item on the Short-WORC. Through the think out loud approach, participants were encouraged to express all their thoughts when responding to each item. Probes such as, "Can you define this word?" or "Can you provide me with an example?", were asked to further explore the rationale of participants' specific responses to each item. Participants described how they determined (calibrated) their responses.

#### *Analysis*

Descriptive statistics (age, sex, occupational status and diagnosis) were collected and are presented in Table 1. The original audio recordings were analyzed by the research team. Analysis of the recordings were done through a descriptive thematic analysis, consisting of open coding.<sup>12</sup> This allowed the scripts to be characterized by fragments, in order for relevant themes to be extracted, categorized and classified first. Next, themes were identified from the responses to each item. Findings were summarized with quotes and percentages as appropriate. A previously established coding system<sup>17</sup> was used to classify issues that affect interpretation. The categories include: Comprehension/clarity (C), Perspective modifiers (PM), Reference point (RP), Calibration across items (CAI), Inadequate response definition (IR) and Relevance (R).

## **Results**

Data analysis resulted in the categorization of 6 themes that describe the issues that participants had when interpreting items. This is further described and illustrated by participants' quotes, as demonstrated in Table 2.

Overall findings demonstrate that individuals interpreted items based on their personal situation, personality traits, biology, work roles and/or environmental factors,<sup>19</sup> which was represented through the theme of perspective modifiers.<sup>20</sup> Certain items such as, styling your hair or dressing, were more influenced by gender. While the genders found this item relevant to QoL, males assigned less importance to this item when compared to items such as, working above shoulder level. Additionally, perspective modifiers also influenced the relevance of doing work outside the house for participants who specified living in an apartment or having designated help prior to the injury. Therefore, item relevancy to the construct of QoL is influenced by a person's biological, environmental or social context.

#### Short-WORC items

##### Instructions on Short- WORC

The Short-WORC focuses on two different domains which have unique sets of instructions. When participants (n=26) were asked to read the instructions out loud, some comprehension (C) issues arose. Specifically, there was misinterpretation around some of the important words in the instructions. For example, when asked to define the word "past week", some participants confused this with "average week".

*"Okay so here I have to answer and think about my lifestyle and what I do on an average day in the week, and how much my shoulder has been kinda affecting or altering those activities." – Measurement Researcher #1, female*

Some participants also struggled with the interpretation of the word "lifestyle". When asked to define, the majority of participants associated lifestyle with only activities of daily living.

*"Lifestyle is my activities during the day. Would the term activities of daily living be better suited instead of lifestyle, is this what you are asking?" - Clinician #4, female*

### 1. How much difficulty do you have sleeping because of your shoulder?

Overall, this item was well understood by most participants and did not demonstrate comprehension issues. When asked to define the term "sleeping", frequent words such as: "at rest, relaxed at night and lying down" arose. The majority of participants (all groups - 73%) considered this item to be relevant to their lifestyle, however, some participants in the patient group suggested that it was only relevant if they slept on the injured shoulder.

*"I always sleep on my left side so my right shoulder (injured) is fine at night." - Patient #8, female*

Additionally, some participants in the patient group (19%) described compensator strategies that allowed them to sleep better at night. Participants in the patient groups further discussed themes of intense shoulder pain, which translated to modifications of their sleeping position to comfortably rest.

*"Since the surgery was on my left shoulder, my left shoulder was out, and so I mostly slept on my back." - patient #4, male.*

### 2. How much difficulty have you experienced with styling your hair because of your shoulder?

Comprehension of this item was generally clear to participants in all three groups. When asked to define the term "styling your hair", phrases of "grooming, blow drying, combing, brushing and using styling products" arose amongst all three groups. Findings demonstrate this item to be relevant to QoL with the majority of participants (all groups- 79%), however, the theme of perspective modifiers heavily influenced the interpretation.

Relevance was a gendered issue with this item. Men in the patient group more often found this item to be irrelevant to their QoL as they had short hair or were bald (30%).

*"I'm bald, I don't need to style my hair!" - Patient #5, male.*

Furthermore, some participants in the group of patients who had not received treatment (19%), used the item of sleeping, as a reference point to calibrate their response to the item of styling your hair. Additionally, female participants in the patient group and measurement researcher group identified that styling their hair was critical to their QoL, and so needed to compensate with the uninjured arm, seek assistance or allot more time in their day for styling.

*"If I were to injure my shoulder, I would still style my hair...I would get someone (roommate) to just help me out if I needed a specific style." - Measurement Researcher #1, female.*

### 3. How much difficulty do you have dressing or undressing?

In general, definitions of "dressing or undressing" were interpreted as "putting on clothes, removing clothes, and getting ready", indicating no comprehension issues with the participant group of patients. Findings further indicated that the patient participants (70%) strongly endorsed the relevance of this item to their QoL. Additionally, some participants in the patient group and measurement researcher group (30%) calibrated their response to this item, based on their scores for the items of sleeping and styling their hair.

*"I would say my answer would be the same as styling my hair...if I chose 5 or 6 in question 2, then I would choose the same answer for question 3." - Measurement Researcher #2, male*

Furthermore, participants in the patient group identified the importance of completing this task and the need to compensate to complete it. Strategies for compensation included: requiring assistance from a device or family member, increasing the allotted time for changing of clothes, or changing the types of clothing worn in order to decrease shoulder movement.

*"I can't reach my back to put on my bra...that's why my husband helps me out." - Patient #6, female.*

### 4. How much difficulty do you experience in daily activities about the house or yard?

Findings indicated some comprehension issues with this item, as participants would interchange the words "lifestyle" and "daily activities" often. When asked to define "daily activities", terms such as: "chores, work, school, and living style" frequently arose amongst all participant groups.

*"Yeah my daily activities are defined by my work and hobbies. My life is my job, family and other activities I do." - Patient #5, female.*

In contrast, definitions of "about the house or the yard" resulted in phrases of: "chores, eating, cleaning, cooking, gardening and yard /outdoors work". Only one female participant in the clinician group initially misinterpreted the meaning of about the house or yard, and defined it as occupational labour that involves working outdoors.

*"This means work outside of the house like employment that you get paid for or yard work. This is both inside or outside the house and external jobs...that's what I think"- Clinician #6, Female*

Overall, all participants in all groups (80%) identified this item to be relevant to their overall quality of life, but some (12%) were concerned with the phrasing of "work in the yard". Due to patient participants' living conditions, some did not require the need to do yard work, i.e. living in an apartment or having designated help prior to the injury.

*"No, I do not do any yard work, my husband always does that." – Patient #8, female.*

Additionally, some compensatory strategies were mentioned from patients such as: seeking assistance from someone else to do their daily activities or modifying the time period or frequency of activities they participated in.

5. How much difficulty do you experience with working above the shoulder?

Overall, this item received positive feedback from all participants in all groups (90%), as many identified this item to be a critical component of recovery. Some female participants in the patient group (10%) identified that they did not need to do much overhead reaching and therefore, found this item less important to their quality of life.

*"I have an office job, I don't need to raise my arms much." – Patient #8, female.*

Definitions of "working above the shoulder" included phrases such as: "overhead reaching, lifting above my head and raising my arms", indicating comprehension was generally good for this item. Participants in the patient group frequently mentioned compensating strategies in order to continue to work above shoulder level, such as: modifying the placement of items for easier access or seeking assistance when needing to reach above shoulder level.

*"I try to use my left hand a lot more to help out and then I keep things within reach. The shelves are much lower in my house and if something is too high for me I use a step ladder." – Patient #10, female.*

6. How much do you use your unaffected arm to compensate for the injured arm?

Definitions of "compensate" led to phrases such as: "using my not injured shoulder and using my healthy shoulder more", indicating no comprehension issues amongst participants. Furthermore, this item was identified as a critical component for QoL by participants in all groups (88%).

*"I use my left hand a lot, which is much harder since I am very right-hand dominant." – patient #6, female.*

In contrast, some female participants in the patient group (12%) indicated that compensating was less relevant, as their injury was on their non-dominant arm.

*"I am right handed; my injury was on my left shoulder...do I compensate? Not frequently".*

*– Patient #3, female.*

7. How much difficulty do you experience lifting heavy objects at or below shoulder level? When asked to define "heavy objects", participants all groups stated words such as: "weight, large and using force", indicating the item was well understood. Overall, this item resulted in a mix of responses depending on what stage of recovery the participant was in. Participants in the patient group who were further along their recovery scored this item lower, while participants who were in the early stages of the injury scored it higher. Nevertheless, the majority of participants (92%) identified this item to be important to QoL. While evident that participants in the patient group understood this item, some (12%) participate in a variety of tasks below shoulder level and therefore, were unsure which tasks to calibrate their score to.

*"I do some yard work and cleaning that can be difficult to bend and pick up things from time to time...I think something in the middle?"- Patient, #3, female*

Additionally, some participants in the patient group discussed compensatory strategies such as re-allocating the task to someone else in order to feel less discomfort.

*"Now with the snow coming, I will have to shovel myself since the weather is bad and I will have to find help."- Patient #7, female.*

## Discussion

Overall, the content validity of the Short-WORC was supported, as most respondents found the items on the Short-WORC to be clear and relevant to their functioning. However, the item of styling your hair was not relevant to a minority of the study sample and had a gender-bias being less relevant to men. Furthermore, it was evident that many patients had developed compensatory strategies as this was mediated in the difficulty reported. Overall, the items received positive feedback, there was no struggle with the recall period, and most of the items were correctly interpreted.

According to the **CO**nsensus-based **S**tandards for the selection of health **M**easurement **I**nstruments (COSMIN), content validity is one of the most important criterion for evaluating a PROs, and should be assessed by both patients and professionals.<sup>25,26</sup> In 2018<sup>28</sup>, COSMIN defined the standards for adequate content validity to be a measure that is comprehensive, comprehensible and relevant.<sup>15,16,28</sup> Comprehension of the items is also an important component of content validation according to the International Society for Pharmacoeconomics and Outcomes Research (ISPOR).<sup>24</sup> Findings indicate that the comprehension level of the Short-WORC is adequate for the intended population. Furthermore, it is recommended to researchers to use appropriate language that does not diverge from the intended meaning.<sup>13,19</sup>

Another component of cognitive interviewing is the evaluation of participants' recall period. The recall period, assesses participants' responses based on the strategies they use when responding to an item.<sup>12</sup> The results confirm the lack of difficulty participants had in the recall phase when responding to the items, and there was no further indication of any unclear reference boundaries that could have impaired their responses.

In addition, another facet of construct validation is item relevancy. Items must be relevant to the intended population and construct being assessed.<sup>12</sup> While results demonstrated a high percentage of relevancy to participants, none of the items were relevant to every individual participant. As anticipated, individuals are unique and have different opinions of what they calibrate as relevant to their QoL or recovery. While researchers try to anticipate this issue in the development of PROs, variables such as: gender, age, lifestyle, or social status will always hinder the relevancy of an item, as they are not generalizable.<sup>13</sup> In the original iteration of the WORC, Kirkley et al<sup>5</sup> used factor analysis and semi-structured interviews to rank the relevancy of items when measuring QoL. Therefore, researchers should use a variety of analytical methods during item selection to enhance the relevancy of items to the construct and the participants.

We conclude that the Short-WORC achieved content validity, based on the evidence presented in this study. Findings demonstrate that all items were relevant to the majority of participants when evaluating their QoL, and the issues were relatively minor. A minority of participants found the item of *styling your hair* to be not relevant to their lifestyle, the majority of participants did. Changing existing measures is a major undertaking since it creates confusion and makes it less possible to compare data across time. Therefore, major issues should be present to warrant these changes. According to COSMIN guidelines<sup>15,26,28</sup>, there is no reason to remove the item, but certain words could be replaced to improve clarity. Overall, it is evident that the comprehension levels of the Short-WORC were easy enough for all participants, and no major comprehension issues were identified that would result in the removal of items. Finally, it is evident that the recall period was accurately evaluated and participants found no difficulty in that process. Therefore, it is evident that the Short-WORC is a valid tool in regard to its content, and qualitatively validates our prior work.<sup>1,6,7</sup>

A secondary purpose of this study was to explore the compensatory strategies that influence the participants calibrate their responses. Findings demonstrated strategies of modifying activity levels, lowering personal expectations or re-allocating tasks in order to avoid stress to the shoulder. Understanding the compensatory strategies for RCDs, provide further insight into why participants with similar impairments might report different levels of functional ability.<sup>18,19</sup> Further, understanding compensatory mechanisms can provide insight into potential for other injuries.<sup>24</sup> For example, overcompensation with the uninjured arm, may increase the risk for an injury in the uninjured arm. Therefore, it is important for researchers to understand the compensatory strategies that might enhance function and affect responses during item development.<sup>22,24</sup>

Overall, limitations in this study included the use of a population that was predominately middle-class Caucasian. However, the demographic region from where this study was conducted is predominately of Caucasian descent.<sup>20,23</sup> Therefore, futures studies should gather information from other ethnic groups in order to compare and contrast the QoL. Additionally, future directions could explore validating the Short-WORC quantitatively through the use of the content validity index<sup>29</sup> to account for this limitation. Furthermore, while the WORC was designed to measure quality of life, the Short-WORC focuses on activity limitation. Therefore, this study cannot be taken as supporting that the WORC and Short-WORC have concurrent content validity.

## Conclusion

In conclusion, the evidence in this study demonstrates that there is no need for change to the items of the Short-WORC, as they tend to be understood by patients with rotator cuff disorders. The Short-WORC reflects the principles of comprehension, relevance and recall, achieving content validity. Overall, the items on the Short-WORC are able to capture aspects of activity limitation, and should be used with a numerical scale from 0 -10. Future studies should assess other psychometric properties such as reliability, validity and responsiveness prospectively.

## List Of Abbreviations

(C) Comprehension

(CAI) Calibration across items

(COSMIN) **C**onsensus-based **S**tandards for the selection of health **M**easurement **I**nstruments

(FDA) Food and Drug Administration

(HULC) Hand and Upper Limb Clinical Research Laboratory

(IR) Inadequate response definition

(ISPOR) International Society for Pharmacoeconomics and Outcomes Research

(PM) Perspective modifier

(PRO) Patient reported outcome

(QoL) Quality of Life

(R) Relevance

(RCD) Rotator cuff disorder

(RP) Reference point

(Short-WORC) Shortened version of the Western Ontario Rotator Cuff Index

(VAS) Visual analogue scale

## Declarations

### Ethics approval and consent to participate

Western Ethics Research Board # R-08-073

### Consent for publication

Not applicable

### Availability of data and material

All data generated or analyzed during this study are included in this published article.

### Competing interests

The authors declare that they have no competing interests

### Funding

The authors received no specific funding for this work.

### Authors' contributions

JM provided overall leadership and contributed to the conception, design and critically revised the manuscript for intellectual content. RF contributed to the conception and design, participated in the data collection and critique process, contributed to analysis and interpretation and drafted sections of the manuscript. DB, KF and GA critically reviewed the manuscript for its intellectual content and contributed in drafting sections. All authors read and approved the final manuscript. All named authors meet the International Committee of Medical Journal Editors (ICMJE) criteria for authorship for this manuscript, take responsibility for the integrity of the work as a whole, and have given final approval to the version to be published.

### Acknowledgement

Joy C. MacDermid was supported by a CIHR Chair in Gender, Work and Health and the Dr. James Roth Research Chair in Musculoskeletal Measurement and Knowledge Translation. Rochelle Furtado is supported in part by a Transdisciplinary Bone & Joint Training Award from the Collaborative Specialization in Musculoskeletal Health Research at Western University, Canada

## References

1. 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 Jul-Sep;29(3):281-91.
2. Puga, Vanessa O. O., Lopes, Alexandre D., & Costa, Leonardo O. P. (2012). Assessment of cross-cultural adaptations and measurement properties of self-report outcome measures relevant to shoulder disability in Portuguese: a systematic review. *Brazilian Journal of Physical Therapy, 16(2)*, 85-93.
3. Johanson, M. E. (2016). Rehabilitation after surgical reconstruction to restore function to the upper limb in tetraplegia: a changing landscape. *Archives of physical medicine and rehabilitation, 97(6)*, S71-S74.
4. Michener LA, Synder, AR. Evaluation of health-related quality of life in patients with shoulder pain: are we doing the best we can? *Clin Sports Med.* 2008; Jul;27(3):491-505
5. Kirkley A, Griffin S. Development of disease-specific quality of life measurement tools. *Arthroscopy.* 2003;19(10):1121-8.
6. Dewan N, MacDermid JC, MacIntyre N. (2018). 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.* 48(5):409-418. doi:10.2519/jospt.2018.7928
7. 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):135-144. doi:10.3138/ptc.2010-51
8. Lee, J. J., Lee, M. K., Kim, J. E., Kim, H. Z., Park, S. H., Tae, J. H., & Choi, S. S. (2015). Pain relief scale is more highly correlated with numerical rating scale than with visual analogue scale in chronic pain patients. *Pain physician, 18(2)*, 195-200.
9. Rentz AM, Skalicky AM, Burslem K, Becker K, Kaschinski D, Esser D, Revicki DA. (2017) The content validity of the PSS in patients with plaque psoriasis. *J Patient Rep Outcomes.*1(1):4. doi: 10.1186/s41687-017-0004-7
10. Lawshe, C.H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*, 563–575.
11. Food Drug Administration (FDA). (2009). Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. *Federal Register, 74(235)*, 65132–65133
12. Willis, G. B. (2004). *Cognitive interviewing: A tool for improving questionnaire design.* Sage Publications.

13. Damman, O. C., Hendriks, M., Rademakers, J., Delnoij, D. M., & Groenewegen, P. P. (2009). How do healthcare consumers process and evaluate comparative healthcare information? A qualitative study using cognitive interviews. *BMC Public Health*, *9*(1), 423.
14. Magasi, S., Ryan, G., Revicki, D., Lenderking, W., Hays, R., Brod, M., . . . Cella, D. (2012). Content validity of patient-reported outcome measures: Perspectives from a PROMIS meeting. *Quality of Life Research*, *21*(5), 739-746.
15. yrvik, M. P., Drendel, A. L., Brandow, A. M., Yan, K., Hoffmann, R. G., & Panepinto, J. A. (2015). A comparison of pain assessment measures in pediatric sickle cell disease: visual analog scale versus numeric rating scale. *Journal of pediatric hematology/oncology*, *37*(3), 190-194.
16. Murtagh, F. E., Addington-Hall, J. M., & Higginson, I. J. (2007). The value of cognitive interviewing techniques in palliative care research. *Palliative Medicine*, *21*(2), 87-93.
17. MacDermid, J. (2018). Cognitive interviewing (CI) - to identify sources of interpretation dissonance in in patient-reported outcome measures (PRO).1-3
18. Mercier, C., Peladeau, N., & Tempier, R. (1998). Age, gender and quality of life. *Community mental health journal*, *34*(5), 487-500.
19. Arumugam, V., MacDermid, J. C., & Grewal, R. (2013). Content analysis of work limitation, Stanford presenteeism, and work instability questionnaires using international classification of functioning, disability, and health and item perspective framework. *Rehabilitation research and practice*.
20. Statistics Canada (2016) Census Profile, 2006 Census. [website] Retrieved from: <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/details/page.cfm?Lang=E&Geo1=CMACA&Code1=555&Geo2=PR&Code2=35&Data=Count&SearchText=London&SearchType=Begins&SearchPR=01&B1=All&GeoLevel=PR&>
21. Thome, S. (2016). *Interpretive description*. Routledge.
22. Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value in health*, *8*(2), 94-104.
23. Packham, T. L., & MacDermid, J. (2014). Content Validity of the Patient-Report Hamilton Inventory for Complex Regional Pain Syndrome. *Journal of Hand Therapy*, *27*(3), e3.
24. Bialocerkowski, A. E. (2002). Difficulties associated with wrist disorders—a qualitative study. *Clinical rehabilitation*, *16*(4), 429-440.
25. Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... & de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of clinical epidemiology*, *63*(7), 737-745.
26. Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., ... & De Vet, H. C. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of life research*, *19*(4), 539-549.
27. Furtado, R., & MacDermid, J. C. (2019). Clinimetrics: Short Western Ontario Rotator Cuff Index. *Journal of physiotherapy*, *65*(1), 56.
28. Terwee, C. B., Prinsen, C. A., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., ... & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Quality of Life Research*, *27*(5), 1159-1170.
29. Shi, J., Mo, X., & Sun, Z. (2012). Content validity index in scale development. *Journal of Central South University. Medical sciences*, *37*(2), 152-155.

## Tables

**Table 1:** Demographic results of study population (n=26)

	Patients who have received treatment (n=6)	Patients yet to be exposed to treatment (n=4)	Clinicians (n=6)	Measurement Researchers (n=10)
<b>Groups</b>				
<b>Mean age (years)</b>	76	44	34	31
<b>Female (%)</b>	67	50	50	50
<b>Male (%)</b>	33	50	50	50
<b>Professional status (%)</b>				
Student	0	25	33	100
Employed	33	50	67	0
Unemployed	0	0	0	0
Sick-leave	0	0	0	0
Retired	67	25	0	0
<b>Affected shoulder (%)</b>				
Left	33	25	N/A	N/A
Right	50	25	N/A	N/A
Bilateral	17	50	N/A	N/A
<b>Diagnosis (%)</b>				
Shoulder pain	67	75	N/A	N/A
Rotator cuff tear	67	25	N/A	N/A
TSA	67	0	N/A	N/A
OA	67	0	N/A	N/A

Table 2: Common issues that arose with the Short-WORC. (n=26)

Item	Comprehension	Perspective Modifiers	Reference Point	Calibration Across Items	Inadequate Response definition	Relevance
sleeping	Understood well					Overall, found to be relevant (73%)  Few participants felt it was only relevant if the injured arm was slept on.
styling your hair	Understood well	Gender issues arose (30%). E.x. men who were bald found this inapplicable		Some participants calibrated this answer with the item of sleeping (19%)		Overall, relevant to 79% of participants
dressing or undressing	Understood well	Some gender issues arose, i.e. men who did not rate this item as important (19%)		Some participants calibrated this answer with the item of sleeping and styling hair (19%)		Overall, relevant to 70% of participants
daily activities about the house or yard	Some comprehension issues with defining daily activities and yard work					Overall was relevant for 80% of participants  12% found work in the yard irrelevant to their QoL
working above the shoulder	Understood well					Overall, 90% of participants found it was relevant to their QoL
compensate with unaffected arm	Understood well					88% of participants found this relevant to their QoL.  12% found if the injury was on the non-dominant hand it was not relevant.
lifting heavy objects at or below shoulder level	Understood well					92% of participants found this to be relevant to QoL