Evidencia en Fisioterapia

Homepage: www.evidenciaenfisioterapia.com ISSN: 2603-7440

[Review]

**Fernández, Rubén.**

*FT. Contacto: @RubenFMat*

# IS THE SHOULDER SYMPTOM MODIFICATION PROCEDURE A RELIABLE TOOL?

**Fernández, Rubén. Is the shoulder symptom modification procedure a reliable tool? Evidencia en Fisioterapia. Septiembre, 2018**

## 1. Introduction

The rotator cuff tendinopathy, subacromial impingement syndrome or rotator cuff related shoulder pain (RCRSP) is the most common disorder of the shoulder, accounting for 44-65% of all complaints of shoulder pain (1). In addition, it is often persistent, with 54% of sufferers reporting ongoing symptoms after 3 years (2). The RCRSP is characterized by the presence of pain located in the anterior-lateral shoulder region which is reproduced with some orthopedic tests (Neer, Hawkins-Kennedy, Jobe…) and/or with active resisted movements of the shoulder (2).

The Shoulder Symptom Modification Procedure (SSMP) is a method of assessment and a guide for treatment of the shoulder proposed by Jeremy Lewis in 2009 (2). It was created because of the inability of traditional orthopedic tests to discriminate the tissue causing the symptoms in people with RCRSP (2).

The SSMP consist of the identification of the movement, posture or activity which most

reproduces the symptoms of the patient and the application of some modification procedures to see if the symptoms respond (2). These procedures can be clustered into 4 subgroup categories:

1. Thoracic kyphosis.

2a. Scapular position.

2b. Winging scapula.

3. Humeral head procedures.

4. Additional procedures.

The purpose of this review is to analyze the published literature upon the reliability of the SSMP.

## 2. Methods

An electronic search of MEDLINE was carried out from inception to May 2018. The terms “SSMP” OR “Shoulder Symptom Modification Procedure” were used. In addition to this, a manual search was also conducted.

Quality appraisal was carried out with the Quality Appraisal of Diagnostic Reliability (QAREL) Checklist (Table 1) (4). The QAREL has shown to have a good inter-rater reliability with just poor reliability in the items 1 and 10, possibly due to the heterogeneity of the tests included in the study (5).

## 3. Results

The searching of the electronic database produced 32 records in total, a further one was identified through a Google search. The characteristics and results of each study are shown in table 2.

| Table 1. Quality Appraisal of Diagnostic Reliability (QAREL) Checklist (4). | |
|---|---|
| Item | **Description** |
| **1** | Was the test evaluated in a sample of subjects who were representative of those to whom the authors intended the results to be applied? |
| **2** | Was the test performed by raters who were representative of those to whom the authors intended the results to be applied? |
| **3** | Were raters blinded to the findings of other raters during the study? |
| **4** | Were raters blinded to their own prior findings of the test under evaluation? |
| **5** | Were raters blinded to the results of the accepted reference standard or disease status for the target disorder (or variable) being evaluated? |
| **6** | Were raters blinded to clinical information that was not intended to be provided as part of the testing procedure or study design? |
| **7** | Were raters blinded to additional cues that were not part of the test? |
| **8** | Was the order of examination varied? |
| **9** | Was the stability (or theoretical stability) of the variable being measured taken into account when determining the suitability of the time-interval between repeated measures? |
| **10** | Was the test applied correctly and interpreted appropriately? |
| **11** | Were appropriate the statistical measures of agreement used? |

### Methodological Quality Appraisal

Methodological quality was assessed with the Quality Appraisal of Reliability Studies (QAREL) checklist (Table 3) (4). Of the three articles included in this review, only one of them had a good methodological quality according to QAREL (9) with 9/11 points. The study of Lewis et al. (3) had 6/11 points and the study of Meakins et al. (6) had 5/11 points. Although item number 8 was satisfied for all the three studies, only Bahat & Kerner (9) used randomization. The item number 9 was scored as "unclear" in the study of Meakins et al. (6) because of the short time between the two raters, so we cannot ensure independency of assessments. In the study of Bahat & Kerner (9) this item was scored as "yes" because a 20-40 minutes period was considered enough to ensure independency of assessments.

## 4. Discussion

### Reliability of the SSMP

In 2016, Lewis et al. (3) published the study "Inter-rater reliability of the Shoulder Symptom Modification Procedure in people with shoulder pain". The researchers recruited a sample of 11 patients with shoulder pain who were assessed by Jeremy Lewis with the SSMP. After each procedure, the patient was asked if his/her symptoms had worsened, improved completely, partially improved or had not undergone changes. The assessments were recorded on videotape obtaining a total of 167 videos. A sample of 36 physiotherapists and 1 osteopath analyzed the videos individually. They had to determine, based upon the response of the patient to the procedure in the video, if the patients symptomatology had worsened, improved completely, partially improved or had not undergone changes. That is, they evaluated the reliability of the clinicians' listening skills to determine the meaning of the message transmitted by the patient. They obtained a good reliability for all procedures (Krippendorff's □ > 0.762). Based on their results, we can conclude that clinicians can reliably understand the message transmitted by the patient after each SSMP procedure. However, we cannot conclude that the SSMP itself has a good reliability.

## Table 2. Characteristics of included studies.

| Author | Study type | Sample and raters | Statistic parameter | Results |
|---|---|---|---|---|
| **Lewis et al. 2016 (3)** | Inter-rater reliability study | 11 patients with unilateral shoulder pain. There were no specifications upon inclusion and/or exclusion criteria.<br><br>Assessments were conducted by 36 physiotherapists and 1 osteopath. | Krippendorff's □ | Substantial to almost perfect reliability was obtained for all procedures (□ □> 0.762) |
| **Bahat & Kerner HS, 2016 (9)** | Intra-rater and inter-rater reliability study | 115 patients >18 year old with shoulder pain not referred below the elbow. Exclusion criteria: Systemic problems or neurologic conditions; unstable fracture/dislocation; and pregnancy<br><br>Assessments were conducted by 3 physiotherapists. | Cohen's Kappa Coefficient | Poor to moderate (intra-rater and inter-rater) reliability for almost all procedures (k<0.60). |
| **Meakins, May & Littlewood (6)** | Inter-rater reliability study | 26 patients >18 years old with shoulder and/or upper arm pain reproduced on elevation and isometric resisted testing into external rotation and/or abduction. Exclusion criteria: more than 25% loss of passive movement into external rotation and/or elevation when compared to the contralateral side, signs of cervical spine involvement, any positive distal neurological signs or symptoms on examination and a history of shoulder surgery or gross orthopaedic trauma within the past 12 months.<br><br>Assessments were conducted by 9 physiotherapists. | Cohen's Kappa Coefficient | Moderate reliability (k=0.47; 95% CI, 0.20-0.71). |

| Table 3. Methodological quality appraisal of included studies according to QAREL Checklist (4). | | | |
|---|---|---|---|
| **Ítem** | **Lewis et al. 2016 (3)** | **Bahat & Kerner HS, 2016 (9)** | **Meakins, May & Littlewood A, 2018 (6)** |
| **1** | N | Y | Y |
| **2** | Y | Y | Y |
| **3** | Y | Y | Y |
| **4** | NA | N | NA |
| **5** | NA | NA | NA |
| **6** | Y | Y | NC |
| **7** | NC | Y | NC |
| **8** | Y | Y | Y |
| **9** | NA | Y | NC |
| **10** | Y | Y | NC |
| **11** | Y | Y | Y |
| **Total** | **6/11** | **9/11** | **5/11** |
| **N = No; Y = Yes; NC = Unclear; NA = Not applicable** | | | |

In 2018, Meakins et al. (6) published the study "Reliability of the Shoulder Symptom Modification Procedure and association of within-session and between-session changes with functional outcomes". In addition to the inter-rater reliability of the SSMP, they also evaluated the association of within-session and between-session changes with functional outcomes; however, this section will not be analyzed in this paper. They recruited a sample of 26 patients with shoulder pain who were assessed by 9 physiotherapists. They recorded the patient response (total relieve or partial relieve (4 or more points reduction in the Numeric Pain Rating Scale) of pain) or lack of response to the SSMP in a standardized form:

- 0 = No response to SSMP.
- 1 = Response to Thoracic Repositioning Procedures.
- 2 = Response to Scapula Repositioning Procedures.
- 2a = Response to Scapula Winging Procedures.
- 3 = Response to Humeral Head Procedures.
- 4 = Response to Neuromodulation Procedures.

They obtained a Cohen's Kappa Coefficient value of 0.47 (95 CI, 0.20-0.71) and a total percentage of agreement value of 57.7%. This study deserves some considerations. Firstly, the authors used a 4 point cut-off in the NPRS as criteria for partial relieve without providing any justification for this value. That is relevant because the Minimally Clinically Important Difference in patients with shoulder pain has been estimated in 2 points (7, 8) or in 30% of pain relieve (8). In second place, the authors just recorded the global response for each category of the SSMP without providing any information of the response for each procedure. For example, the category "Humeral Head Procedures" contains at least 10 different procedures (3). It would be interesting to record the individual response to each procedure because if the SSMP is a reliable tool, each of its procedures should be too. Finally, the authors established that the different categories were incompatible, so the assessors couldn't choose more than one category for each patient; however, they didn't provide any criteria to know which category choose in a patient that has a partial relieve in two of them.

The latest study is, perhaps, the least known of the three, both by its authors and by the journal in which it is published. It is the study of Bahat & Kerner, published in the

Journal of Novel Phytiotherapies in 2016 (9). They analyzed inter-rater and intra-rater reliability of the SSMP. They recruited a sample of 115 patients with shoulder pain (25 for intra-rater reliability study and 90 for inter-rater reliability study) who were assessed by 3 physiotherapists. They defined a positive test result as greater than 30% improved symptoms as a result of the modification. They used Cohen's Kappa Coefficient to analyze intra-rater and inter-rater (3 pairs of clinicians). They studied the reliability of each procedure individually and also the reliability of each category of procedures. In the majority of the procedures they obtained a kappa value of less than 0.60. In the inter-reliability study they also found a kappa of less than 0.60 for the majority of procedures, further there was a lot of variability in the kappa value between procedures and also between pairs of raters. The kappa values of less than 0.60 are below the recommendations of Cadogan (10) and Scholtes (11) for the minimum reliability needed for orthopaedic tests for their use in clinical practice, so based upon these recommendations and the results of this study, the SSMP doesn't seem to have sufficient reliability to recommend its use in clinical practice.

**Considerations for future research on the reliability of the SSMP**

The SSMP has, as it is described in the literature, some aspects that deserve consideration when planning the design of reliability studies. When Jeremy Lewis proposed it in 2009, he established that (2):

- The initial procedure and the order of the procedures can vary.
- If a procedure results in the complete cessation of symptoms, then the testing process is over.
- If there is only a partial reduction in symptoms, then other component of the SSMP is tested to determine whether a better response is achieved.
- Regarding to the response to the SSMP, we should take into account:
    - For a positive response in the symptom criteria, there should be an improvement equal or better than 30%.
    - There should be consistency in the patients response to the SSMP procedure selected, with at least two positive responses in the same session.
- With the SSMP we look for the procedure which produces the greatest relieve in

symptomatology, so if one produces a 100% relieve, the testing process is over.

- In addition to the standard procedures described, it is stated that it can also be used a combination of procedures to see if the combination of two procedures that partially alleviates symptomatology produce a total relieve in combination.

Firstly, given that the SSMP is composed of a series of procedures, the studies that aimed to evaluate the reliability of the SSMP should analyze each procedure individually, as Bahat & Kerner did (9).

In second place, it is stated that the initial procedure and the order of them can vary. However, if one of them produces a 100% relieve of symptomatology the SSMP testing process is over, as in the study of Bahat & Kerner. This should be avoided in reliability studies because of the loss of data that could alter the possible results of the reliability of each procedure, making also difficult to control the sample size used to assess each of them, varying the statistical power. It would be more appropriate in research to perform all the procedures in all the patients included in the study despite their response to them because they are not incompatible and although there is a 100% reduction in symptoms with one procedure, it doesn't mean that the other ones cannot relieve them too.

Thirdly, although a positive response to the SSMP is described in terms of improved symptomatology or range of motion, only the first one has been used in reliability studies (6, 9). It would be more appropriate if clinicians and investigators agreed in the criteria used for a positive response to the SSMP.

Another important point is the elevated number of procedures included in the SSMP. Although in clinical practice all of them are performed continuously, as Meakins et al. (6) and Bahat & Kerner (9) did, it could be that, in order to ensure independency between procedures, there should be a washout period between each procedure assessment and not only between raters. What's more, as Meakins et al. (6) and Bahat & Kerner (9) says, it would be more appropriate to have longer washouts periods between raters, up to one or several days. Given that the SSMP is a tool for guiding treatment, the response to the procedures should be consistent between different days, if not, it just could serve as a within-session decision making tool with poor clinical utility.

Finally, the clinical recommendations of combining different procedures that produce a partial relieve of symptomatology to see if their combination produces a greater relieve

should be avoided in reliability studies until reliability of each procedure on its own is proved.

## 5. Conclusion

Based on published research, we cannot ensure the reliability of the SSMP and recommend its use in clinical practice. However, due to the lack of studies and their poor methodology, we cannot make definite statements until more and better studies are published.

## 6. References


1) Michener LA, McClure PW, Karduna AR. Anatomial and biomechanical mechanisms of subacromial impingement syndrome. Clin Biomech (Bristol, Avon). 2003; 18(5): 369-79.
2) Lewis JS. Rotator cuff tendinopathy/subacromial impingement síndrome: is it time for a new method of assessment? Br J Sports Med. 2009; 43(4): 259-64.
3) Lewis JS, McCreesh K, Barratt E, Hegedus EJ, Sim J. Inter-rater reliability of the Shoulder Symptom Modification Procedure in people with shoulder pain. BMJ Open Sport Exerc Med. 2016; 2(1): e000181.
4) Lucas NP, Macaskill P, Irwig L, Bogduk N. The development of a quality appraisal tool for studies of diagnostic reliability (QAREL). J Clin Epidemiol. 2010; 63(8): 854-61.
5) Lucas N, Macaskill P, Irwig L, Moran R, Rickards L, Turner R, et al. The reliability of a quality appraisal tool for studies of diagnostic reliability (QAREL). BMC Med Res Methodol. 2013; 13: 111.
6) Meakins, May & Littlewood A, May S, Littlewood C. Reliability of the Shoulder Symptom Modification Procedure and association of within-session and between-session changes with functional outcomes. BMJ Open Sport Exerc Med. 2018; 4(1): e000342.
7) Michener LA, Snyder AR, Leggin BG. Responsiveness of the numeric pain rating scale in patients with shoulder pain and the effect of surgical status. J Sport Rehabil. 2011; 20(1): 115-28.

8) Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain. 2001; 94(2): 149-58.
9) Bahat & Kerner HS, Kerner O. The Shoulder Symptom Modification Procedure (SSMP): A reliability study. J Nov Physiother. 2016; S3: 011.
10) Cadogan A, Laslett M, Hing W, McNair P, Williams M. Interexaminer reliability of orthopaedic special tests used in the assessment of the shoulder pain. Man Ther. 2011; 16(2): 131-5.
11) Scholtes VA, Terwee CB, Poolman RW. What makes a measurement instrument valid and reliable? Injury. 2011; 42(3): 236-40.

**Infographic**