

The Dutch Shoulder Pain and Disability Index (SPADI): a reliability and validation study

Marloes Thoomes-de Graaf · Gwendolijne G. M. Scholten-Peeters ·
Edwin Duijn · Yasmayne Karel · Bart W. Koes ·
Arianne P. Verhagen

Accepted: 25 November 2014 / Published online: 4 December 2014
© Springer International Publishing Switzerland 2014

Abstract

Purpose To evaluate the reliability and validity of the Dutch Shoulder Pain and Disability Index (SPADI-D).

Background The SPADI is recommended and frequently used. However, the validity and reliability of the SPADI-D are unknown.

Methods The study population consisted of patients consulting a physical therapist for shoulder pain. We assessed construct validity, using known groups, convergent validity (SDQ) and divergent validity (EQ5D) for which the mean difference or Spearman correlations coefficients were calculated. The factor structure was assessed using principal component factor analysis, and we calculated Cronbach's alpha and the ICC to assess the reliability.

Results A total of 356 patients and a randomly selected group of 74 subjects for the reliability analysis were included. There was a significant difference between extreme groups (a high/low level of pain and work absence/presence) in SPADI score. The correlation between the SPADI and the SDQ was 0.69, with the EQ5D

mobility-item 0.25 and with the depression-item 0.14. The SPADI consisted of one factor according to principal component factor analysis, which showed high internal consistency (Cronbach's alpha = 0.94 for the total score), and the test–retest reliability was good (ICC = 0.89).

Conclusion The Dutch SPADI is a valid and reliable questionnaire for patients in primary care in assessing functional disability.

Keywords SPADI · Validation · Reliability · Shoulder

Introduction

One of the main complaints in patients with shoulder pain is functional disability [1]. Treatment of shoulder pain is usually aimed at pain reduction and improvement of functional disabilities [2]. Consequently, outcome measurements should include an instrument (e.g., questionnaire) for the evaluation of functional disabilities [3].

There are several self-administered shoulder pain and disability questionnaires. Patients ranked the Shoulder Disability Questionnaire (SDQ) and the Shoulder Pain and Disability Index (SPADI) as the most relevant questionnaires [4]. The SPADI was the least time-consuming, both the SDQ and the SPADI appear to be convenient and easy to complete [4].

The SPADI was originally developed in English [5]. It has been translated and validated in several languages and showed excellent reliability and responsiveness [6–9].

The Royal Dutch Society for Physical Therapy has recommended implementation of the Dutch SPADI (SPADI-D) in a clinical guideline for patients with shoulder pain [10]. Nevertheless, the SPADI-D has not been validated and tested for reliability.

M. Thoomes-de Graaf (✉) · G. G. M. Scholten-Peeters ·
E. Duijn · Y. Karel · A. P. Verhagen
Research Group Diagnostics, Avans University of Applied
Science, Breda, The Netherlands
e-mail: m.degraaf.2@erasmusmc.nl

M. Thoomes-de Graaf · G. G. M. Scholten-Peeters · E. Duijn ·
Y. Karel · B. W. Koes · A. P. Verhagen
Department of General Practice, Erasmus Medical Centre,
Rotterdam, The Netherlands

G. G. M. Scholten-Peeters
Faculty of Human Movement Sciences, Research Institute
MOVE, VU University, Amsterdam, The Netherlands

Therefore, the aim of this study is to evaluate the reliability and construct validity of the SPADI-D for patients with shoulder pain in primary care.

Methods

Patients with shoulder pain were recruited from primary care physical therapy clinics and signed informed consent. This study was approved by the Medical Ethics Committee of the Erasmus Medical Center (MEC-2011-414) [11].

Baseline measurement

Patients received an online questionnaire that included the SPADI-D, SDQ and EuroQol five-item quality of life questionnaire (EQ-5D-3L).

The SPADI is designed to measure pain and disability associated with shoulder pain. It consists of 13 items and response options range from 0 to 10, where 0 represents “no pain/no difficulty” and 10 “worst pain imaginable/very difficult.” The total score varies between 0 and 100; a higher score indicates a higher level of pain-related disability [5].

The SDQ is a pain-related disability questionnaire consisting of 16 items. Response options are “yes,” “no” or “not applicable.” The SDQ-score can range from 0 to 100 with a higher score indicating more severe disability [2]. The SDQ was originally designed and validated in Dutch, and internal consistency and responsiveness are good [2, 12].

The Dutch EQ-5D-3L is a quality of life questionnaire covering 5 dimensions of health: mobility, self-care, usual activities, pain/discomfort and anxiety/depression and an official language version [13]. Response options are “no problems,” “some problems” and “extreme problems.”

Test–retest measurement

A randomly selected group of patients received a second SPADI-D after 1 week. The time interval was chosen to minimize recall bias as well as progression bias and is often considered appropriate [14]. A sample size of approximately 80 is considered acceptable [15].

Analysis

Analyses were performed with SPSS22. Handling of missing items was performed as described by the original authors of the SPADI and SDQ [5, 12].

All data were checked on normality, using a stem-and-leaf plot, Q-plot and whisker box. Nonparametric tests were used if data were not normally distributed.

Known groups validity We assumed that patients with high initial pain (>7 on the Numeric Rating Scale in the

preceding 24 h) and work absence would have a higher level of perceived disability. Both groups have been chosen a priori [7, 12]. The independent *t* test was used to test the difference between known groups.

Convergent validity High correlations ($r \geq 0.60$) were expected [15] between the scores on SPADI-D and the SDQ, as both aim to measure the same construct.

Divergent validity Low correlations ($r < 0.30$) were expected [15] between the items “mobility” (as patients with shoulder pain and healthy subjects do not differ significantly in the amount of time spent walking [16]) and “anxiety/depression” (as low correlations were found between anxiety/depression and activity limitations for patients with shoulder pain [17]).

Factor structure We conducted a principal component factor analysis with and without varimax rotation. Data were checked for suitability. We used the scree test and parallel analysis [18–20] to extract the number of factors. Items loading higher than 0.40 on one factor and lower than 0.30 on any other were acceptable [21]. Ultimately, the stability of our model was assessed using two random splitting halves (subsamples) [22], and we performed this five times to assess if our findings were consistent.

Internal consistency Internal consistency was calculated using Cronbach’s alpha and only for the scale(s) that was extracted from our factor analysis. A Cronbach’s alpha between 0.70 and 0.95 is considered “good” [23].

Test–retest The intraclass correlation coefficient (ICC) using a two way mixed model was used to calculate the test–retest reliability. The ICC can range from 0.00 (no stability/agreement) to 1.00 (perfect stability/agreement). An ICC of 0.70 is considered to be acceptable [23]. We checked the test–retest data for extreme values and assessed whether this influenced the ICC.

Results

Patient characteristics

Due to missing variables out of 389 patients, 356 patients were included in this analysis and 74 in the test–retest reliability analysis. The mean age was 49.5 (SD 13) years, and 47 % was male. Demographic characteristics are reported in Table 1.

The data of the SPADI-D at baseline and at re-test were considered as normally distributed, in contrast to the data of the SDQ and EQ-5D-3L.

Validity

Differences between “known groups” were statistically significant and considered clinically relevant (Table 2).

Table 1 Demographic characteristics of the included patients

	Total (356)	Test–retest (74)
Gender (male)	166 (47 %)	29 (39 %)
Age (SD)	49.5 (13.1)	51.4 (12.7)
Duration of shoulder pain in weeks	12 (6–26)	16 (8–40)
Median (IQR)		
SPADI baseline score	46.7 (21.3)	50.8 (22.6)
Mean (SD)		
Use of medication	171 (49 %)	37 (51 %)
Initial pain (NRS)	6 (4–7)	6 (4–8)
Median (IQR)		
Work absence	38 (out of the 255 working patients) (15 %) ^a	5 (out of the 48 working patients) (11 %) ^b

SD standard deviation, NRS Numeric Rating Scale, IQR inter-quartile range

^a Based on data of 252 working patients

^b Based on 47 working patients

This means that the SPADI-D is able to differentiate between different groups.

The Spearman correlation between the SPADI-D and SDQ was high ($r = 0.69$), meaning that the convergent validity of SPADI-D with SDQ is good. The Spearman correlation between the SPADI-D and EQ-5D-3L_mobil-ity-item ($r = 0.25$) and the EQ-5D-3L_depression-item ($r = 0.14$) was low. This means the SPADI-D and EQ-5D-3L measure a different construct.

Factor structure

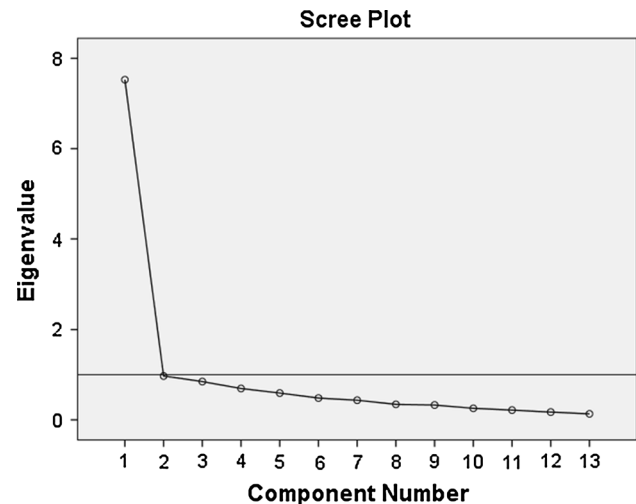
Parallel analysis revealed that the eigenvalue of the first factor should be above 1.44 and of the second factor above 1.33 to be extracted. Only one factor was extracted (see Fig. 1), the eigenvalue of the second factor was 0.97. A one-factor solution explained 57.9 % of the variance and the second factor added only 7 %. Findings were consistent with all five analyses based on two random subsamples. This means that we consider the SPADI-D to have one factor.

Table 2 Extreme groups correlation coefficients

	Group	SPADI-D (mean, SD)	Mean difference
Pain-SPADI-D	High initial pain >7	59.4 (18.0)	-21.4* $p = 0.00$ (95 % CI -25.4 to -17.5)
	Low initial pain <7	37.9 (18.9)	
Work absence-SPADI-D	Work absence +	50.5 (20.6)	-7.6* $p = 0.04$ (95 % CI -14.8 to -0.3)
	Work absence -	43.0 (21.1)	

SD standard deviation, CI confidence interval

* Significant

**Fig. 1** Scree plot. A scree plot of eigenvalues, the demarcation point indicates one factor. The results are based on 298 patients

Reliability

The internal consistency and test–retest reliability were good [Cronbach’s alpha = 0.94; ICC = 0.89 (95 % CI 0.83–0.93)]. After exclusion of two patients with extreme values, the ICC was 0.90 (95 % CI 0.85–0.94). Both indicate a high level of agreement.

Discussion

This study shows that the SPADI-D consists of one factor only and can be considered as a valid and reliable questionnaire. It discriminates well between known groups and correlates well with the SDQ, and internal consistency and test–retest reliability are high.

One SPADI validation study used similar “known groups,” showing a higher mean difference for work absence compared to ours [7]. Differences with this study were that their population was smaller and had a higher baseline SPADI score, and they did not present the percentage of people that could not work due to their shoulder pain.

Correlation coefficients found in other studies for convergent validity varied between moderate and high

(0.33–0.85) depending on the comparator [4, 24–26]. Few studies evaluated divergent validity of the SPADI and none used the EQ-5D-3L [25, 26].

Only one study reported a factor structure as originally described by Roach [26], the majority of studies could not confirm this loading pattern or reported a one-factor structure [6, 27–29]. One study concluded that people do not distinguish between pain and disability and a possible explanation for this finding could be the wording of the SPADI items. The disability items ask respondents to indicate the amount of difficulty they have with specified functions. It is possible that when people report their difficulty in performing an activity, they consider pain to be part of what makes the activity difficult [29].

The Cronbach's alpha found in other studies ranged between 0.90 and 0.95 [5, 26–28, 30], and ICC values ranged between 0.88 and 0.94 [6, 7, 30–32], both consistent with ours.

Our study has some limitations. First, the translation process of the SPADI-D was not published, and it is unknown if it is performed as recommended [33]. Nevertheless, the SPADI-D is commonly used in clinical practice and research and is also integrated in multiple patient-management software programs in the Netherlands. Second, we did not use the general perceived effect scale to check if patients were indeed stable between the test and the re-test. However, it is unlikely that patients would have been recovered within 1 week, due to the duration of complaints and the mean number of weeks patients usually need to recover [34]. The extreme value analysis showed that differences after exclusion were minimal.

On the other hand, we used an adequate sample size to perform factor analysis [22]. There is increasing consensus among statisticians that parallel analysis is superior to other procedures and typically yields optimal solutions to the number of components problem [18].

Acknowledgments This study is financed by the SIA-RAAK Grant serving exclusively for lectureships and knowledge networks at Universities of Applied Sciences.

Conflict of interest I affirm that I have no financial affiliation (including research funding) or involvement with any commercial organization that has direct financial interest in any matter included in this manuscript.

Ethical standard The Medical Ethical Committee of the Erasmus University approved this study, nr mec-2011-414.

References

1. Feleus, A., Bierma-Zeinstra, S. M., Miedema, H. S., Verhaar, J. A., & Koes, B. W. (2008). Management in non-traumatic arm, neck and shoulder complaints: Differences between diagnostic groups. *European Spine Journal*, *17*(9), 1218–1229.
2. van der Windt, D. A., van der Heijden, G. J., de Winter, A. F., Koes, B. W., Deville, W., & Bouter, L. M. (1998). The responsiveness of the Shoulder Disability Questionnaire. *Annals of the Rheumatic Diseases*, *57*(2), 82–87.
3. Mintken, P. E., Glynn, P., & Cleland, J. A. (2009). Psychometric properties of the shortened disabilities of the Arm, Shoulder, and Hand Questionnaire (QuickDASH) and Numeric Pain Rating Scale in patients with shoulder pain. *Journal of Shoulder and Elbow Surgery*, *18*(6), 920–926.
4. Paul, A., Lewis, M., Shadforth, M. F., Croft, P. R., Van Der Windt, D. A., & Hay, E. M. (2004). A comparison of four shoulder-specific questionnaires in primary care. *Annals of the Rheumatic Diseases*, *63*(10), 1293–1299.
5. Roach, K. E., Budiman-Mak, E., Songsiridej, N., & Lertratanakul, Y. (1991). Development of a Shoulder Pain and Disability Index. *Arthritis Care & Research*, *4*(4), 143–149.
6. Jamnik, H., & Spevak, M. K. (2008). Shoulder Pain and Disability Index: Validation of slovene version. *International Journal of Rehabilitation Research*, *31*(4), 337–341.
7. Christiansen, D. H., Andersen, J. H., & Haahr, J. P. (2013). Cross-cultural adaption and measurement properties of the Danish version of the Shoulder Pain and Disability Index. *Clinical Rehabilitation*, *27*(4), 355–360.
8. Ekeberg, O. M., Bautz-Holter, E., Tveita, E. K., Keller, A., Juel, N. G., & Brox, J. I. (2008). Agreement, reliability and validity in 3 shoulder questionnaires in patients with rotator cuff disease. *BMC Musculoskeletal Disorders*, *9*, 68.
9. Bot, S. D., Terwee, C. B., van der Windt, D. A., Bouter, L. M., Dekker, J., & de Vet, H. C. (2004). Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. *Annals of the Rheumatic Diseases*, *63*(4), 335–341.
10. Jansen, M. J., Brooijmans, F., Geraets, J. J. X. R., Lenssen, A. F., Ottenheijm, R. P. G., Penning, L. I. F., et al. (2011). KNGF Evidence Statement Subacromiale klachten. *Nederlands Tijdschrift voor Fysiotherapie*, *121*(1).
11. Karel, Y. H., Scholten-Peeters, W. G., Thoomes-de Graaf, M., Duijn, E., Ottenheijm, R. P., Koes, B. W., et al. (2013). Current management and prognostic factors in physiotherapy practice for patients with shoulder pain: Design of a prospective cohort study. *BMC Musculoskeletal Disorders*, *14*(1), 62.
12. de Winter, A. F., van der Heijden, G. J., Scholten, R. J., van der Windt, D. A., & Bouter, L. M. (2007). The Shoulder Disability Questionnaire differentiated well between high and low disability levels in patients in primary care, in a cross-sectional study. *Journal of Clinical Epidemiology*, *60*(11), 1156–1163.
13. Lamers, L. M., McDonnell, J., Stalmeier, P. F., Krabbe, P. F., & Busschbach, J. J. (2006). The Dutch tariff: Results and arguments for an effective design for national EQ-5D valuation studies. *Health Economics*, *15*(10), 1121–1132.
14. Streiner, D. L., & Norman, G. R. (2008). *Health measurement scales a practical guide to the development and use*. Oxford: Oxford university press.
15. Mokkink, L. B., Terwee, C. B., Knol, D. L., Stratford, P. W., Alonso, J., Patrick, D. L., et al. (2010). The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Medical Research Methodology*, *10*, 22.
16. Hallman, D. M., Ekman, A. H., & Lyskov, E. (2014). Changes in physical activity and heart rate variability in chronic neck-shoulder pain: Monitoring during work and leisure time. *International archives of occupational and environmental health*, *87*(7), 735–744.
17. Roh, Y. H., Noh, J. H., Oh, J. H., Baek, G. H., & Gong, H. S. (2012). To what degree do shoulder outcome instruments reflect patients' psychologic distress? *Clinical Orthopaedics and Related Research*, *470*(12), 3470–3477.

18. O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, 32(3), 396–402.
19. Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179–185.
20. Patil, V. H., Surendra, N. S., Sanjay, M., & Donovan, T. (2008). "Parallel analysis engine to aid determining number of factors to retain computer software. <http://smishra.faculty.ku.edu/parallelengine.htm>; utility developed as part of Patil, Vivek H., Surendra N. Singh, Sanjay Mishra, and Todd Donovan. (2008). "Efficient theory development and factor retention criteria: A case for abandoning the 'eigenvalue greater than one' criterion". *Journal of Business Research*, 61 (2), 162–170.
21. Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment*, 7, 286–299.
22. de Vet, H. C., Ader, H. J., Terwee, C. B., & Pouwer, F. (2005). Are factor analytical techniques used appropriately in the validation of health status questionnaires? A systematic review on the quality of factor analysis of the SF-36. *Quality of Life Research*, 14(5), 1203–1218. (discussion 1219–1221, 1223–1204).
23. Terwee, C. B., Bot, S. D., de Boer, M. R., van der Windt, D. A., Knol, D. L., Dekker, J., et al. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60(1), 34–42.
24. Staples, M. P., Forbes, A., Green, S., & Buchbinder, R. (2010). Shoulder-specific disability measures showed acceptable construct validity and responsiveness. *Journal of Clinical Epidemiology*, 63(2), 163–170.
25. Cloke, D. J., Lynn, S. E., Watson, H., Steen, I. N., Purdy, S., & Williams, J. R. (2005). A comparison of functional, patient-based scores in subacromial impingement. *Journal of Shoulder and Elbow Surgery*, 14(4), 380–384.
26. Hill, C. L., Lester, S., Taylor, A. W., Shanahan, M. E., & Gill, T. K. (2011). Factor structure and validity of the Shoulder Pain and Disability Index in a population-based study of people with shoulder symptoms. *BMC Musculoskeletal Disorders*, 12, 8.
27. Tveita, E. K., Sandvik, L., Ekeberg, O. M., Juel, N. G., & Bautz-Holter, E. (2008). Factor structure of the Shoulder Pain and Disability Index in patients with adhesive capsulitis. *BMC Musculoskeletal Disorders*, 9, 103.
28. MacDermid, J. C., Solomon, P., & Prkachin, K. (2006). The Shoulder Pain and Disability Index demonstrates factor, construct and longitudinal validity. *BMC Musculoskeletal Disorders*, 7, 12.
29. Roddey, T. S., Olson, S. L., Cook, K. F., Gartsman, G. M., & Hanten, W. (2000). Comparison of the University of California—Los Angeles Shoulder Scale and the simple shoulder test with the Shoulder Pain and Disability Index: Single-administration reliability and validity. *Physical Therapy*, 80(8), 759–768.
30. Bicer, A., & Ankarali, H. (2010). Shoulder Pain and Disability Index: A validation study in Turkish women. *Singapore Medical Journal*, 51(11), 865–870.
31. Tveita, E. K., Ekeberg, O. M., Juel, N. G., & Bautz-Holter, E. (2008). Responsiveness of the Shoulder Pain and Disability Index in patients with adhesive capsulitis. *BMC Musculoskeletal Disorders*, 9, 161.
32. Jeldi, A. J., Aseer, A. L., Dhandapani, A. G., & Roach, K. E. (2012). Cross-cultural adaptation, reliability and validity of the Indian (Tamil) version of the Shoulder Pain and Disability Index. *Hong Kong Physiotherapy Journal*, 30(2), 99–104.
33. Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., et al. (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 Health*, 8(2), 94–104.
34. Kooijman, M., Barten, J., Swinkels, I., & Veenhof, C. J. (2010). *En trendcijfers 2006–2010 fysiotherapie*. Landelijke Informatievoorziening Paramedische Zorg. Utrecht: NIVEL. <http://www.nivel.nl/lipz>