

Validity and reliability of the Portuguese version of the American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form

Auristela Duarte de Lima Moser¹, Luiz Alberto Manfre Knaut², Talita Gnoato Zotz³, Karoleen Oswald Scharan⁴

ABSTRACT

Introduction: The American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form (ASES), often used in research in Brazil, although translated and adapted to the Portuguese language, had not had its validity and reliability tested yet. **Objective:** To assess the validity, reliability, and internal consistency of the ASES-PT for shoulder dysfunction. **Materials and methods:** Fifty individuals (26 women; mean age, 39 ± 13 years) participated in the validity assessment, and 38 (19 women; mean age, 37 ± 13 years old) in the reliability assessment, all having shoulder dysfunction. The participants completed the 36-Item Short-Form Health Survey (SF-36), the Disabilities of the Arm, Shoulder and Hand Questionnaire (DASH), and the ASES-PT on two occasions with an interval of seven days. The convergent validity was assessed by use of the Spearman's rank correlation coefficient (ρ), and the analysis of the intrarater reliability used the intraclass correlation coefficient (ICC). The internal consistency was assessed by using Cronbach's alpha. **Results:** The ASES-PT scores correlated with the DASH scores ($\rho = -0.69$, $P = 0.000$) and with the "physical functioning" ($\rho = 0.50$, $P = 0.000$), "role limitation due to physical health" ($\rho = 0.43$, $P = 0.002$) and "bodily pain" domains ($\rho = 0.60$, $P = 0.000$) of the SF-36. The intrarater reliability of the ASES-PT proved to be adequate (ICC = 0.75, $P = 0.000$). The internal consistency (0.794) was satisfactory. **Conclusion:** The validity and reliability study of the ASES-PT supports its use for assessing shoulder dysfunction.

Keywords: shoulder pain, shoulder joint, reproducibility of results, test validity, questionnaires.

© 2012 Elsevier Editora Ltda. All rights reserved.

INTRODUCTION

Shoulder pain is present in 14%–50% of the population.^{1–3} It is estimated that two out of three individuals will have at least one episode of neck or shoulder pain during their lives.⁴ In addition, the incidence of shoulder pain is of eight new cases per year for every 100 workers of the industry and service sectors.⁵ Frequently, that pain and the reduced shoulder mobility have a negative effect on the patient's functional abilities, occupational activities, and quality of life.⁶

To assess and quantify the impact of the musculoskeletal changes on people's lives, functional assessments by use of questionnaires can be performed during treatment. Several questionnaires have been developed for evaluating the function of the upper limbs.⁷ Of those questionnaires, the American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form (ASES) stands out because it was designed for patients with any shoulder dysfunction and specifically for that joint,⁸ in addition to being widely cited in the literature.^{9,10} According to Brazilian studies, the ASES has been applied

Received on 07/18/2011. Approved on 03/05/2012. The authors declare no conflict of interest. Ethics Committee: 5.257.

Post-Graduation Program of Technology in Health, Pontifícia Universidade Católica do Paraná – PUCPR.

1. PhD in Production Engineering, Universidade Federal de Santa Catarina – UFSC; Professor of the Post-Graduation Program of Technology in Health, Pontifícia Universidade Católica do Paraná – PUCPR

2. Master's degree in Rehabilitation, Université de Montréal – CAN; Professor of the Physical Therapy Course, Centro Universitário Campos de Andrade – UNIANDRADE; Member of the Research Nucleus of the Physical Therapy Course, PUCPR

3. Master's degree in Technology in Health, PUCPR

4. Undergraduate student of Physical Therapy, PUCPR

Correspondence to: Auristela Duarte de Lima Moser. Rua Imaculada Conceição, 1155 – Prado Velho. CEP: 80215-901. Curitiba, PR, Brasil.

E-mail: auristela.lima@pucpr.br

since 1988.¹¹ However, its translation and cultural adaptation for Brazilian Portuguese through the appropriate methodological process was only established in 2010.¹²

The original ASES has been elaborated by the Research Committee of the American Shoulder and Elbow Surgeons⁸ and comprises a physician assessment section and a patient self-report section. However, only the patient self-report section contributes to its final score.¹³ That section is composed by an item concerning pain, whose severity is assessed by use of a visual analogue scale (VAS) that ranges from “no pain at all” to “pain as bad as it can be”, and 10 items concerning function, assessed by use of a four-category Likert scale. The score of each section corresponds to 50% of the final test score, which can range from 0 (no functionality) to 100 (normal function).

However, to ratify its use, the Portuguese version of the ASES (ASES-PT) should be investigated regarding its following characteristics: validity (i.e., the instrument’s capacity to measure that which it is intended to measure); reliability (i.e., the instrument’s capacity to be stable and reproducible); and internal consistency (i.e., the capacity of yielding consistent results in repeated measurements).

Thus, this study aimed at assessing the validity, reliability, and internal consistency of the ASES-PT in individuals with shoulder pain.

MATERIALS AND METHODS

Participants

Fifty individuals (26 women; mean age, 39 ± 13 years) were recruited from the school-clinics of the Pontifícia Universidade Católica do Paraná (PUCPR) for assessing the validity of the ASES-PT. All participants met the following criteria: 1) minimum age of 18 years; 2) clinical diagnosis of shoulder dysfunction; 3) neither neurological nor any musculoskeletal disease other than shoulder dysfunction; 4) no cognitive change; and 5) ability to read, understand, and complete the questionnaire studied. For assessing reliability, the sample comprised 38 individuals (19 women; mean age, 37 ± 13 years), due to sample loss of 12 individuals, who did not return to the second assessment. This study has been approved by the Ethics Committee in Research with Human Beings of the institution, according to the protocol number 5.257. All participants provided written informed consent.

Validity

To test the convergent validity, the scores obtained with the ASES-PT were correlated with those obtained with the Portuguese versions of the Disabilities of the Arm, Shoulder and Hand Questionnaire (DASH)¹⁴ and the 36-Item Short-Form Health Survey (SF-36),¹⁵ whose validity and reliability have already been shown.^{14,15} The DASH is a questionnaire comprising 30 questions aimed at measuring the symptoms and physical disabilities related to the upper limbs.¹⁶ The Portuguese version of the SF-36 is a questionnaire with 36 questions regarding the patient’s general quality of life, and covering eight domains (i.e., physical functioning, role limitation due to physical health, bodily pain, general health perceptions, vitality, social functioning, role limitation due to emotional problems, and mental health).¹⁵

Validity was assessed by use of the Spearman’s rank correlation coefficient (ρ), which can range from +1 to -1. A +1 correlation indicates the existence of a perfect positive linear relationship between the variables, while a -1 correlation indicates the existence of a perfect negative linear relationship between the variables. A correlation is considered strong when ρ is greater than 0.60 or lower than -0.60; the correlation is moderate when that coefficient is between 0.30 and 0.60 or -0.30 and -0.60; and it is weak when that coefficient is between -0.30 and 0.30.¹⁷

Reliability

Reliability was tested by use of the intrarater reliability (i.e., test-retest reliability), with intraclass correlation coefficient 1,1 (ICC) and 95% confidence interval. The ICC measures the agreement between the variables studied, and ranges from 0 (no agreement) to 1 (total agreement). Reliability is adequate if ICC is greater than 0.70.¹⁸ Because it is a self-administered questionnaire, its interrater reliability was not assessed.

Internal consistency

Internal consistency, assessed by use of the Cronbach’s alpha, was used to identify to what extent the different items of the questionnaire were associated between themselves. The total values of each item were calculated, as was the alpha’s variation when some items were eliminated. The statistical analyses were performed with the SPSS software, version 10.0 for Windows (SPSS, Chicago, USA).

Experimental procedure

To assess validity, each participant completed the ASES-PT, DASH, and SF-36 questionnaires, randomly applied. Then, the participant was invited to return to the school-clinics after seven days to conclude the reliability assessment,¹⁹ without undergoing any physical therapeutic intervention during that period. Data from the first and second ASES-PT completion were used to assess internal consistency. During the assessments, the participants were always accompanied by one of the examiners.

RESULTS

The scores of the ASES-PT showed correlation with the scores of the Portuguese version of the DASH ($\rho = -0.69$; $P = 0.000$) and with the following SF-36 domains: “physical functioning” ($\rho = 0.50$; $P = 0.000$), “role limitation due to physical health” ($\rho = 0.43$; $P = 0.002$), and “bodily pain” ($\rho = 0.60$; $P = 0.000$). With the other SF-36 domains, the correlations with the ASES-PT were weak and non-significant ($\rho \leq 0.35$; Table 1).

In the first assessment with the ASES-PT, the mean of the scores was 60.1 ± 21.1 , while, in the second assessment, that mean was 63.4 ± 20.7 . The intrarater reliability of the ASES-PT proved to be adequate, as shown in Table 2 ($ICC = 0.75$; $P = 0.000$).

Table 1
Correlation of the ASES-PT scale with the DASH scale and the SF-36 domains (n = 50)

	ASES-PT	
	ρ	Significance
DASH	-0.69	0.000
PF	0.50	0.000
LPH	0.43	0.002
Pain	0.60	0.000
GHP	0.22	0.134
VIT	0.25	0.077
SF	-0.08	0.601
LEP	0.11	0.450
MH	-0.04	0.765

PF: physical functioning; LPH: role limitation due to physical health; GHP: general health perceptions; VIT: vitality; SF: social functioning; LEP: role limitation due to emotional problems; MH: mental health.

Table 2
Intrarater reliability of the ASES-PT and DASH scales and of the SF-36 domains (n = 38)

Scale/domain	ICC (95% CI)	F _(37,38)	Significance
ASES	0.75 (0.57-0.86)	7.031	0.000
DASH	0.86 (0.75-0.93)	13.671	0.000
SF-36			
PF	0.60 (0.36-0.77)	4.114	0.000
LPH	0.57 (0.32-0.75)	3.744	0.000
Pain	0.57 (0.32-0.75)	3.713	0.000
GHP	0.56 (0.30-0.74)	3.623	0.000
VIT	0.67 (0.46-0.81)	5.208	0.000
SF	0.71 (0.52-0.84)	6.092	0.000
LEP	0.54 (0.27-0.73)	3.357	0.000
MH	0.76 (0.59-0.86)	7.492	0.000

ICC: intraclass correlation coefficient 1,1; CI: confidence interval; PF: physical functioning; LPH: role limitation due to physical health; GHP: general health perceptions; VIT: vitality; SF: social functioning; LEP: role limitation due to emotional problems; MH: mental health.

Table 3
Internal consistency of the ASES-PT scale, values of each item and Cronbach’s alpha (n = 38)

Pain Q1 R	0.795
F1 Q1	0.788
F2 Q1	0.794
F3 Q1	0.792
F4 Q1	0.793
F5 Q1	0.793
F6 Q1	0.792
F7 Q1	0.790
F8 Q1	0.769
F9 Q1	0.797
F10 Q1	0.768
Pain Q2 R	0.786
F1 Q2	0.793
F2 Q2	0.794
F3 Q2	0.793
F4 Q2	0.790
F5 Q2	0.794
F6 Q2	0.790
F7 Q2	0.794
F8 Q2	0.774
F9 Q2	0.792
F10 Q2	0.787
Cronbach’s alpha	0.794

Q: question; F: function; R: correlation between the questionnaire items.

The results of the internal consistency of the instrument have shown good reliability indices for all items, with values greater than 0.70 and Cronbach's alpha of 0.794 (Table 3).

DISCUSSION

Measuring instruments are used during rehabilitation programs aimed at the following: obtaining initial information about the patient; elaborating the patient's treatment; monitoring possible changes in symptoms; and assessing the efficacy of the therapeutic procedures used. Despite the existence of its Portuguese version,¹² the ASES-PT had not undergone validation and reliability investigation.

In the present study, the convergent validity of the ASES-PT was demonstrated by use of its correlation with a specific instrument for the upper limbs (i.e., DASH) and a generic instrument (i.e., the SF-36 physical health domains).

The fact that the DASH specifically assesses the symptoms and physical disabilities of the upper limbs justifies the greater correlation of the ASES-PT with that questionnaire. The results are similar to those obtained with the English version of the ASES, which showed a strong correlation when compared with a questionnaire specific for shoulder pain and function (that is, the University of Pennsylvania Shoulder Score – Penn), ($r = 0.78$; $P = 0.01$).⁹ In addition, strong correlations between questionnaires specific for the shoulders have already been reported several times in the literature.^{9,20–22}

Similarly, the moderate and weak correlations between ASES-PT and the SF-36 domains observed in this study repeat the results obtained with the original version of the ASES, in which moderate correlation of the domains “physical functioning” ($r = 0.41$; $P = 0.001$) and “role limitation due to physical health” ($r = 0.33$; $P = 0.008$) was observed with the mean score of the physical components ($r = 0.40$; $P = 0.001$), in addition to the weak correlation of the “role limitation due to emotional problems” ($r = 0.24$; $P = 0.21$) and “mental health” ($r = 0.05$; $P = 0.70$) observed with the mean score of the mental components ($r = 0.15$; $P = 0.25$).⁹

Moderate correlations between different shoulder-specific questionnaires and the SF-36 physical health domains have already been reported in previous studies.^{6,9,20} This can be explained by the fact that their items did not contemplate exclusively the activities of the upper limbs. The weak and non-significant

correlations of the ASES-PT with the SF-36 emotional and mental components result from the fact that such measures assess different constructs.⁹

In the present study, the reliability of the ASES-PT proved to be adequate, as already observed with its English version.^{9,23} However, the ICC of the Portuguese version was slightly lower than that of the English version (that is, 0.75 and 0.84, respectively). A possible explanation for that difference can be the different time intervals between the first and second application of the questionnaire. While the second assessment was applied one to three days after the first for the English version, for the Portuguese version, that interval was of seven days. Thus, a clinical change is more likely to have occurred in the participants between the assessments of the ASES-PT. However, it is worth noting that there is no consensus in the literature regarding the ideal time interval between the assessments to analyze reliability. While long time intervals can be influenced by changes in the patients' clinical findings, short time intervals can be influenced by memory biases.

Regarding internal consistency, the scale showed good indices for all the domains of the instrument, with values over 0.70, considered satisfactory according to the literature.^{9,24,25} The overall reliability index of the instrument was greater than when calculated with the suppression of any item.

Even having assessed the validity, reliability, and internal consistency of the ASES-PT, its responsiveness should also be assessed, which has already been initiated by the authors of the present study to better support its use.

Because the use of questionnaires has been part of the health care routine, a careful analysis of their applicability in the various settings of the therapeutic practices is increasingly important.²⁶ The process of assessing the psychometric properties, which can promote an increase in the resolution potential of therapeutic practices aimed at with this study, is part of that analysis.

CONCLUSION

The ASES-PT is a questionnaire of rapid application, containing clear, objective questions written in a simple language, related to daily activities. In addition, the results of this study have shown the validity, reliability, and internal consistency of the ASES-PT, indicating its adequacy for assessing shoulder function in the clinical and research practice.

REFERENCES

1. Urwin M, Symmons D, Allison T, Brammah T, Busby H, Roxby M *et al.* Estimating the burden of musculoskeletal disorders in the community: the comparative prevalence of symptoms at different anatomical sites, and the relation to social deprivation. *Ann Rheum Dis* 1998; 57(11):649–55.
2. Van der Heijden GJ. Shoulder disorders: a state-of-the-art review. *Baillieres Clin Rheumatol* 1999; 13(2):287–309.
3. Bongers PM. The cost of shoulder pain at work. *BMJ* 2001; 322(7278):64–5.
4. Mäkelä M, Heliövaara M, Sievers K, Impivaara O, Knekt P, Aromaa A. Prevalence, determinants, and consequences of chronic neck pain in Finland. *Am J Epidemiol* 1991; 134(11):1356–67.
5. Leclerc A, Chastang JF, Niedhammer I, Landre MF, Roquelaure Y; Study Group on Repetitive Work. Incidence of shoulder pain in repetitive work. *Occup Environ Med* 2004; 61(1):39–44.
6. Mousavi SJ, Parnianpour M, Abedi M, Askary-Ashtiani A, Karimi A, Khorsandi A *et al.* Cultural adaptation and validation of the Persian version of the Disabilities of the Arm, Shoulder and Hand (DASH) outcome measure. *Clin Rehabil* 2008; 22(8):749–57.
7. Bot SD, Terwee CB, van der Windt DA, Bouter LM, Dekker J, de Vet HC. Clinimetric evaluation of shoulder disability questionnaires: a systematic review of the literature. *Ann Rheum Dis* 2004; 63(4):335–41.
8. Richards RR, An K-N, Bigliani LU, Friedman RJ, Gartsman GM, Gristina AG *et al.* A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg* 1994; 3:347–52.
9. Michener LA, McClure PW, Sennett BJ. American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form, patient self-report section: reliability, validity, and responsiveness. *J Shoulder Elbow Surg* 2002; 11(6):587–94.
10. Angst F, Goldhahn J, Drerup S, Aeschlimann A, Schwyzer HK, Simmen BR. Responsiveness of six outcome assessment instruments in total shoulder arthroplasty. *Arthritis Rheum* 2008; 59(3):391–8.
11. Lino Junior W, Belangero WD. Resultado funcional do tratamento artroscópico da instabilidade do ombro sem a lesão do lábio da glenoide. *Rev Bras Ortop* 2007; 42(7):195–200.
12. Knaut LA, Moser AD, Melo SA, Richards RR. Translation and cultural adaptation to the Portuguese language of the American Shoulder and Elbow Surgeons Standardized Shoulder assessment form (ASES) for evaluation of shoulder function. *Rev Bras Reumatol* 2010; 50(2):176–89.
13. Kirkley A, Griffin S, Dainty K. Scoring systems for the functional assessment of the shoulder. *Arthroscopy* 2003; 19(10):1109–20.
14. Orfale AG, Araújo PM, Ferraz MB, Natour J. Translation into Brazilian Portuguese, cultural adaptation and evaluation of the reliability of the Disabilities of the Arm, Shoulder and Hand Questionnaire. *Braz J Med Biol Res* 2005; 38(2):293–302.
15. Ciconelli RM, Ferraz MB, Santos W, Meinão I, Quaresma MR. Brazilian-Portuguese version of the SF-36: a reliable and valid quality of life outcome measure. *Rev Bras Reumatol* 1999; 39(3):143–50.
16. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). *Am J Ind Med* 1996; 29(6):602–8.
17. Härdén M, Nyström B, Kulich K, Carlsson J, Bengtson A, Edvardsson N. Validity and reliability of a new, short symptom rating scale in patients with persistent atrial fibrillation. *Health Qual Life Outcomes* 2009; 7:65.
18. Godfrey J, Hamman R, Lowenstein S, Briggs K, Kocher M. Reliability, validity, and responsiveness of the simple shoulder test: psychometric properties by age and injury type. *J Shoulder Elbow Surg* 2007; 16(3):260–7.
19. Beaton D, Richards RR. Assessing the reliability and responsiveness of 5 shoulder questionnaires. *J Shoulder Elbow Surg* 1998; 7(6):565–72.
20. Angst F, Goldhahn J, Pap G, Mannion AF, Roach KE, Siebertz D *et al.* Cross-cultural adaptation, reliability and validity of the German Shoulder Pain and Disability Index (SPADI). *Rheumatology (Oxford)* 2007; 46(1):87–92.
21. Lopes AD, Ciconelli RM, Carrera EF, Griffin S, Faloppa F, Dos Reis FB. Validity and reliability of the Western Ontario Rotator Cuff Index (WORC) for use in Brazil. *Clin J Sport Med* 2008; 18(3):266–72.
22. Romeo AA, Mazzocca A, Hang DW, Shott S, Bach BR Jr. Shoulder Scoring scales for the evaluation of rotator cuff repair. *Clin Orthop Relat Res* 2004; (427):107–14.
23. Cook KF, Roddey TS, Olson SL, Gartsman GM, Valenzuela FF, Hanten WP. Reliability by surgical status of self-reported outcomes in patients who have shoulder pathologies. *J Orthop Sports Phys Ther* 2002; 32(7):336–46.
24. Leech NL, Barrett KC, Morgan GA. SPSS for intermediate statistics: use and interpretation. 2.ed. New Jersey: LEA; 2005.
25. Almeida MHM, de Pinho Spinola AW, Iwamizu PS, Okura RI, Barroso LP, Lima ACP. Confiabilidade do Instrumento para Classificação de Idosos quanto à Capacidade para o Autocuidado. *Rev Saude Publica* 2008; 42(2):317–23.
26. Martinez JE, Grassi DC, Marques LG. Análise da aplicabilidade de três instrumentos de avaliação de dor em distintas unidades de atendimento: ambulatório, enfermaria e urgência. *Rev Bras Reumatol* 2011; 51(4):304–8.