

Translation and cultural adaptation of an Arabic version of the patient scar assessment scale for thyroidectomy patients

Turki Aldrees, MD, Ahmad Albosaily, MD, Latifa Alanazi, MD, Sami Alharethy, MD, Ghada AlHindi, MD, Ahmad Altuwaijri, MD, Saleh Aldhabri, MD.

ABSTRACT

الأهداف: ترجمة وتوثيق النسخة العربية لمقياس تقييم الندبات المرضية (PSAS) للمتحدثين باللغة العربية. الشكل الجمالي لندبة الجرح بعد عمليات إزالة الغدة الدرقية أمر مهم للمريض ونظرة الذاتية لنفسه. أدوات تقييم الندبات عامل مهم في علاجها بناء على أسس علمية مبنية على البراهين.

الطريقة: هذه دراسة كمية استقصائية عملت بطريقة إعطاء المرضى استطلاع من النسخة المعربة من معيار تقييم الندبات للمرضى. الترجمة تمت بإعطاء ثلاثة من أطباء الأنف والأذن والحنجرة المتحدثين بطلاقة باللغة العربية ومن ثم إعادة ترجمة النسخة العربية إلى اللغة الإنجليزية عن طريق متخصصين في اللغة ومقارنتها بالنسخة الإنجليزية الأصلية. تم توزيع استطلاع على المرضى الذين تم إجراء عملية إزالة الغدة الدرقية لهم في مستشفين مختلفين ذات المرجعية الثلاثية في مدينة الرياض. المرضى الذين تم إدخالهم في الدراسة هم من تم إجراء العملية لهم لمدة تزيد عن شهرين ومن تمت موافقتهم للمشاركة ولديهم القدرة على القراءة والفهم باللغة العربية.

النتائج: تم إدراج 50 مريض في الدراسة. وكانت المقاييس الإحصائية جميعها ممتازة (معامل الاتساق الداخلي 0.89) ومعامل الثقة (بين 0.88-0.90). إن أداة التقييم المعربة (PSAS) أظهرت نتائج جيدة في موثوقية الاختبار وبعد إعادة الاختبار. معامل بارسون (Pearson correlation coefficient) بين الاختبار وإعادة الاختبار كان 0.84 ($p > 0.001$) والنطاق المحتمل من 6-60 نقطة مع متوسط معامل الخطأ القياسي 5.14 وأقل نقطة للتغير الطفيف كانت 14.2.

الخلاصة: النسخة العربية من أداة تقييم الندبات للمرضى (PSAS) تعتبر موثوقة للاستخدام بهذه الترجمة في المجتمعات المتحدثة باللغة العربية والتي تسمح بمقارنة النتائج بين الدراسات المختلفة حتى لو اختلفت الدول والتي تساعد على تبادل المعلومات والخبرات على مستوى علمي ودولي.

Objectives: To translate and validate an Arabic version of the patient scar assessment scale (PSAS). The cosmetic appearance of a thyroidectomy scar can critically influence a patient's self-esteem. Moreover, scar evaluation tools are necessary for an evidence-based approach to scar management.

Methods: This quantitative, observational, cross-sectional study was conducted by administering an Arabic-translated version of the PSAS. The translation process included a forward translation into Arabic by 3 fluently bilingual otolaryngologists, a back-translation into English, and a comparison with the original items. The questionnaires were distributed to patients who underwent thyroidectomies. We included patients who underwent surgery at least 2 months previously.

Results: A total of 50 patients were included in this research. The internal consistency was 0.89, with a 95% confidence interval (CI) of 0.88-0.90. The score distributions showed high correlations for all items. The Arabic-translated PSAS showed good test-retest reliability, and the Pearson correlation coefficient between the test and retest administrations was 0.84 ($p < 0.001$). With a possible range of 6-60 points, the standard error of the mean was 5.14, and the minimal detectable change was 14.2.

Conclusion: This Arabic version of the PSAS was reliable for use in Arabic-speaking communities. It will allow for comparisons between the results of investigations conducted in different countries, which aids in the exchange of information within the international scientific community.

*Saudi Med J 2019; Vol. 40 (6): 590-594
doi:10.15537/smj.2019.6.24197*

From the Department of Surgery (Aldrees), College of Medicine, Prince Sattam bin Abdulaziz University, Alkharj, from the Department of Otolaryngology (Albosaily), Head and Neck Surgery, Prince Mohammed bin Abdulaziz Hospital, from the Department of Otolaryngology (Alanazi), College of Medicine; from the Department of Otolaryngology (Alharethy, AlHindi, Aldhabri), Head and Neck Surgery, College of Medicine, King Saud University; Department of Otolaryngology (Aldhabri), King Fahad Medical City; and from the Department of Otolaryngology (Altuwaijri), Head and Neck Surgery, Security Force Hospital, Riyadh, Kingdom of Saudi Arabia.

Received 30th December 2018. Accepted 20th April 2019.

*Address correspondence and reprint request to: Dr. Turki Aldrees, Department of Surgery, College of Medicine, Prince Sattam bin Abdulaziz University, Al-Kharj, Kingdom of Saudi Arabia. E-mail: pt.turki@hotmail.com
ORCID ID: orcid.org/0000-0002-8955-1778*

Scars are the inevitable outcome of most wounds. Depending on multiple factors, including the etiology, site, appearance, and physical symptoms, a scar can impact an individual's perception of their appearance and affect his or her emotional and psychosocial status.¹ Multiple scar therapies, skin closure techniques, and minimally invasive surgical approaches have been designed to limit scars and their associated morbidities. Therefore, it is important to possess a comprehensive tool that can be used to assess scars and their impacts, evaluate the outcomes of the various scar treatments, and determine the benefits of the different surgical techniques. In contrast to objective scar assessment tools, such as colorimeters, adherometers, and digital scar analyses, subjective scar assessment scales are less costly and easier to use, and they can help minimize the time consumption required for objective scar assessments. In addition, subjective scales may capture more information than a scar's physical aspects. Therefore, they are considered to be more comprehensive and clinically useful.^{2,3} Currently, there are many validated subjective scar assessment scales, including the Vancouver scar scale (VSS), Manchester scar scale, Patient and Observer scar assessment scale (POSAS), Stony Brook scar evaluation scale, patient-reported impact of scars measure, and patient scar assessment questionnaire.⁴ Among these, the POSAS is the scale that is used most frequently, followed by the earlier-developed VSS.⁵ Recently, Vercelli et al,⁴ conducted a systematic review of all of these scales, and they reported that the POSAS was the most robust. The POSAS was first developed and validated for the assessment of burn scars by Draaijers et al,⁶ in 2004. Later, it was validated for the evaluation of linear surgical scars.^{7,8} The POSAS is comprised of 2 subscales: a PSAS and an observer scar assessment scale (OSAS). The PSAS includes 6 items (pain, itching, color, stiffness, thickness, and irregularity), and the OSAS includes 5 items (vascularity, pigmentation, thickness, relief, and pliability). In one modification, a sixth item (surface area) was added to the OSAS after a linear regression analysis showed that the input of an observer was most influenced by the scar's surface area.⁸ Each item is rated using a scale ranging from 1-10, with one representing

normal skin and 10 indicating the worst imaginable scar or sensation. The item ratings are summed up in order to obtain a total score of 6-60 for each subscale. Additionally, both the patient and observer rate their overall opinions of the scar's appearance on a scale ranging from 1-10.

To date, the PSAS has been translated and culturally adapted to the French and Italian languages.^{9,10} Therefore, the purpose of this study was to translate and culturally adapt an Arabic version of the PSAS.

Methods. A multistep process was used to produce an Arabic version of the PSAS. This process included a forward translation by 3 fluently bilingual otolaryngologists who translated the original English version of the PSAS into Arabic. Then, this version was back-translated by professionals into English, followed by a comparison to the original PSAS items. The 2 forms of the questionnaire (original and translated) were compared, and the differences were resolved by consensus in a meeting of a panel of experts composed of healthcare providers and language professionals. The committee eventually developed a pre-final version of the questionnaire for field testing.

Cognitive and pilot testing. The Arabic-translated version of the PSAS was tested on 10 patients in order to verify the clarity and detect any potential conceptual or theoretical issues. Then, the patients were asked about their experiences and thoughts regarding that version of the PSAS. No specific constructive feedback was received. Therefore, at that point, the committee met and approved the pre-final version as the final version.

The study protocol was approved by the Research Center, King Fahad Medical City Hospital, Riyadh, Kingdom of Saudi Arabia, and its Ethical Committee, and all of the participants provided verbal informed consent.

Participants and data collection. We conducted an observational, quantitative, cross-sectional study by administering an Arabic-translated version of the PSAS. The questionnaire was distributed between July 2017 and March 2018. We included patients who underwent thyroidectomies (either total or hemithyroidectomies) with or without concurrent neck dissection at 2 different tertiary care hospitals in Riyadh, Kingdom of Saudi Arabia. These patients agreed to participate in the study, and they were able to read and understand Arabic. Those patients who declined to participate in the study and those who were unable to read Arabic were excluded from this research.

The questionnaire was filled out twice by the same patients at least 2 months after surgery and one week later to test the reliability.

Disclosure. Authors have no conflict of interests, and the work was not supported or funded by any drug company. This study was supported by College of Medicine Research Center, Deanship of Scientific Research, King Saud University, Riyadh, Kingdom of Saudi Arabia.

Validation and statistical analysis. The data was analyzed using Statistical Package for Social Science (SPSS), version 19.0 (IBM Corp., Armonk, NY, USA). The descriptive statistical data for the sociodemographic variables were presented as mean values, standard deviations, and percentages. Pearson's correlation coefficient values were used to assess the relationships between the variables. We used Cronbach's alpha coefficients, corrected item-total correlations, and interitem correlation (IIC) matrix analyses to assess the internal consistency and reliability of this measure.

The internal consistency was examined using Cronbach's alpha, which ranges from 0 (no internal consistency; none of the items are correlated with one another) to 1 (perfect internal consistency; all of the items are perfectly correlated with one another). The alpha values were computed using all items.

The test-retest reliability was assessed by administering the Arabic version of the PSAS a second time one week after the first administration. The stability of the responses was estimated using Pearson's correlation coefficients (r) and intraclass correlation coefficients (ICCs) for the responses from the 2 administrations. The test-retest reliability was considered to be weak if the r value was <0.3, moderate if the r value was ≥0.3 and <0.5, and strong if the r value was ≥0.5. An ICC value of ≥0.70 was considered to indicate good test-retest reliability.

It has been suggested previously that 30-40 participants are adequate for translation and cultural validation studies.¹¹

Results. Sample characteristics. A total of 50 participants completed the questionnaire (13 males and 37 females, with a ratio of approximately 1:4). The participants' mean age was 43±13 years old. The largest number of participants had bachelor's degrees (40%), and the other patient characteristics are described in Table 1.

Validity and reliability analyses. The overall internal consistency of the scale items, as measured using Cronbach's alpha coefficients, was good at 0.89 (95% CI: 0.88-0.90). The Pearson's correlation coefficient value between the test and retest administrations was 0.84 (p<0.001), as shown in Figure 1. The test-retest reliability was excellent (ICC=0.9). Table 2 shows the IIC between the first and second administrations for each item. There were no statistically significant differences in the mean differences in the global test scores between the test and retest administrations. A Bland-Altman plot was constructed, which showed no bias in the test scores (Figure 2).

Table 1 - Patient demographic information (N=50).

Characteristic	n (%)
Age in years, (mean±SD)	42.9±12.7
Age in years, range	18-63
Gender	
Male	13
Female	37
M:F ratio	1:4
Time since surgery in years, (mean±SD)	2.6±3.3
Weight in kg, (mean±SD)	81.9±20.9
Height in cm, (mean±SD)	162.2±8.1
BMI, (mean±SD)	30.9±6.7
Educational level	
Primary school	8 (16.0)
Elementary school	8 (16.0)
High school	14 (28.0)
Bachelor's degree	20 (40.0)
Diabetes mellitus	9 (18.0)
Hypertension	14 (28.0)
Asthma	4 (8.0)

BMI - body mass index, SD - standard deviation.

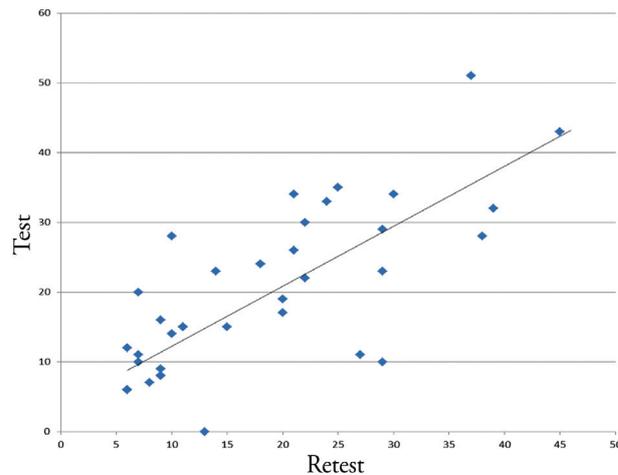


Figure 1 - Correlation of the patient scar assessment scale global score (test and retest). The slope corresponds with the correlation coefficient (0.84), p<0.001.

The median and range of each subscale for the test and retest administrations are shown in Table 3. The standard error of the mean for the total score (with a possible score ranging from 6-60) was 5.14, and the minimal detectable difference with a 95% CI was 14.2.

Discussion. The POSAS developers incorporated the patient's scar assessment, making it both an observer and patient-reported outcome measure, a feature that was lacking in previous assessment scales.⁶ Moreover,

Table 2 - Correlation table for each specific question.

PSAS patient scale questions	IIC matrix	Cronbach α
Has the scar been painful for the past few weeks?	0.840	0.859
Has the scar been itching for the past few weeks?	0.649	0.787
Is the scar color different from the color of your normal skin at present?	0.722	0.839
Is the stiffness of the scar different from your normal skin at present?	0.826	0.904
Is the thickness of the scar different from your normal skin at present?	0.735	0.847
Is the scar more irregular than your normal skin at present?	0.575	0.730
Overall opinion of the scar compared to normal skin	0.846	0.916
Total PSAS score	0.795	0.886

PSAS - patient scar assessment scale, IIC - interitem correlation.

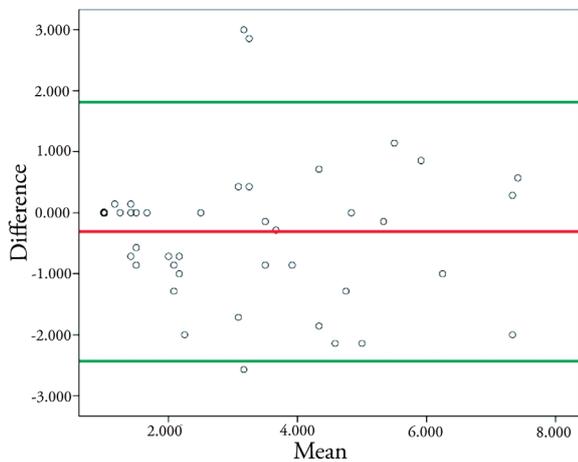


Figure 2 - Bland-Altman plot of total score distribution for test and retest.

multiple studies have proven that the PSAS is a quick, feasible, reliable, and valid scar assessment scale for use in research and busy clinical settings.^{4,7-10} The demonstrated reliability of the PSAS makes it practical for evaluating scars without the need for attending a clinic.¹²

Arabic is the native language of approximately 422 million people in 22 countries, and it is the world's fifth most commonly spoken language (www.UNESCO.org). In this study, we carried out a cross-cultural adaptation and validation of a new Arabic version of the PSAS (Ar-PSAS) for clinical and research use in Arabic-speaking populations. This helps to build a common quantitative tool that can be utilized in cross-cultural or multicenter multinational research. We found no need to translate the OSAS because all of the observers in our medical practice population were capable of understanding and filling out the English form.

In order to ensure that the Ar-PSAS used the same concepts and semantics as the original PSAS (in other

words, to ensure content validity), we followed the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), task force guidelines for the translation and cross-cultural adaptation of patient-reported measures.¹³ No further changes were required after the pilot testing, which demonstrated that this measure was clearly interpreted by the patients.

In order for a scar assessment scale to be suitable, it should be both reliable and valid. Reliability is the degree to which a scale produces consistent and reproducible results. We evaluated the Ar-PSAS by determining its internal consistency and test-retest reliability. The internal consistency indicates the grade to which the items proposed to rationalize a certain construct correlate with each other, and this is measured using the Cronbach's alpha index. The Ar-PSAS showed good internal consistency ($\alpha=0.88$), which was comparable to that of the original English version ($\alpha=0.90$), French version ($\alpha=0.98$), and Italian version ($\alpha=0.80$).⁸⁻¹⁰ The test-retest reliability reflects a scale's ability to produce stable results over time. In this present study, the interval between the test and retest administrations was one week. We believe that this period was sufficiently long to limit the carryover effect, which would cause an overestimation of the scale's reliability, while it was short enough to minimize the changes in the scar's appearance, which would cause an underestimation of the scale's reliability. The Ar-PSAS exhibited high reproducibility, as evidenced by an ICC value of 0.80 and a Pearson's correlation coefficient value of 0.89, which were similar to those of the original and French versions and better than that of the Italian version, which had poor test-retest reliability (ICC <0.7).⁸⁻¹⁰

Study limitations. We did not carry out a concomitant observer assessment of the scars using a validated OSAS. Doing so may have shown similarities or differences between observers' and patients' opinions of the scars in our population. Additionally, this may have helped determine the construct validity of the Ar-PSAS. Moreover, there are no other Arabic scar

Table 3 - Median and range for each item for both test and retest.

PSAS	Test		Retest	
	Median (IQ range)	Range min/max	Median (IQ range)	Range min/max
Pain	1 (1/9)	1/10	1 (1/2)	1/8
Pruritus	1 (1/2)	1/9	1 (1/2)	1/10
Color	2 (1/4)	1/10	3 (1/4)	1/10
Stiffness	1 (1/4)	1/10	2 (1/4)	1/10
Thickness	2 (1/4)	1/10	2 (1/5)	1/10
Irregularity	2 (1/4)	1/10	2 (1/5)	1/10
Total	11 (1/17)	6/46	15 (1/21)	1/51

PSAS - patient scar assessment scale, IQ - intelligence quotient

assessment scales available, which made it difficult to evaluate the concurrent validity of the Ar-PSAS. Finally, we did not assess the ability of the Ar-PSAS to detect the scar changes over time (responsiveness) due to the cross-sectional design of this study.

In conclusion, we adapted a consistent and reliable Ar-PSAS that was validated for use in clinical applications and cross-sectional studies. However, we recommend further future validations of its application in longitudinal research.

Acknowledgment. *This study was supported by the College of Medicine Research Center, Deanship of Scientific Research, King Saud University, Riyadh, Kingdom of Saudi Arabia. The authors gratefully acknowledge Scribendi (www.scribendi.com) for English language editing.*

References

- Brown BC, McKenna SP, Siddhi K, McGrouther DA, Bayat A. The hidden cost of skin scars: quality of life after skin scarring. *J Plast Reconstr Aesthet Surg* 2008; 61: 1049-1058.
- Vercelli S, Ferriero G, Sartorio F, Stissi V, Franchignoni F. How to assess postsurgical scars: a review of outcome measures. *Disabil Rehabil* 2009; 31: 2055-2063.
- Powers PS, Sarkar S, Goldgof DB, Cruse CW, Tsap LV. Scar assessment: current problems and future solutions. *J Burn Care Rehabil* 1999; 20: 54-60; discussion 53.
- Vercelli S, Ferriero G, Sartorio F, Cisari C, Bravini E. Clinimetric properties and clinical utility in rehabilitation of postsurgical scar rating scales: a systematic review. *Int J Rehabil Res* 2015; 38: 279-286.
- Bae SH, Bae YC. Analysis of frequency of use of different scar assessment scales based on the scar condition and treatment method. *Arch Plast Surg* 2014; 41: 111-115.
- Draaijers LJ, Tempelman FR, Botman YA, Tuinebreijer WE, Middelkoop E, Kreis RW, et al. The patient and observer scar assessment scale: a reliable and feasible tool for scar evaluation. *Plast Reconstr Surg* 2004; 113: 1960-1965; discussion 1966-1967.
- Truong PT, Lee JC, Soer B, Gaul CA, Olivotto IA. Reliability and validity testing of the Patient and Observer Scar Assessment Scale in evaluating linear scars after breast cancer surgery. *Plast Reconstr Surg* 2007; 119: 487-494.
- van de Kar AL, Corion LU, Smeulders MJ, Draaijers LJ, van der Horst CM, van Zuijlen PP. Reliable and feasible evaluation of linear scars by the Patient and Observer Scar Assessment Scale. *Plast Reconstr Surg* 2005; 116: 514-522.
- Deslauriers V, Rouleau DM, Alami G, MacDermid JC. Translation of the Patient Scar Assessment Scale (PSAS) to French with cross-cultural adaptation, reliability evaluation and validation. *Can J Surg* 2009; 52: E259-E263.
- Vercelli S, Ferriero G, Bravini E, Stissi V, Ciceri M, Rossetti S, et al. Cross-cultural adaptation, reproducibility and validation of the Italian version of the Patient and Observer Scar Assessment Scale (POSAS). *Int Wound J* 2017; 14: 1262-1268.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)* 2000; 25: 3186-3191.
- Stavrou D, Haik J, Weissman O, Goldan O, Tessone A, Winkler E. Patient and observer scar assessment scale: how good is it? *J Wound Care* 2009; 18: 171-176.
- Wild D, Grove A, Martin M, Eremenco S, McElroy S, Verjee-Lorenz A, et al. 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* 2005; 8: 94-104.