RESEARCH ARTICLE

Validation of the Persian Version of the American Orthopedic Foot and Ankle Society Score (AOFAS) Questionnaire

Sayyed-Hadi Sayyed-Hosseinian, MD; Golnaz Ghayyem Hassankhani, MD; Farshid Bagheri, MD; Niloufar Alavi, BSc; Babak Shojaie, MD, FA; Alireza Mousavian, MD

Research performed at Orthopedic Research Center, Ghaem Medical and Teaching Hospital, Mashhad University of Medical Sciences, Mashhad, Iran

Abstract

Background: American Orthopedic Foot and Ankle Society Score (AOFAS) is a reliable and reproducible measurement tool which is commonly used for the assessment of foot and ankle conditions. In this study we aimed to translate and assess the psychometric properties of the Persian version of AOFAS questionnaire.

Methods: In this study, we enrolled 53 patients with ankle and hindfoot conditions. Our study was conducted according to five staged cross-cultural adaption steps including translation, synthesis, back translation, expert committee review, and pretesting. After that reliability of the subjective parts calculated by Cronbach's alpha and the intraclass correlation coefficient (ICC) and the reliability of the objective items estimated using Cohen's kappa test. Also, construct validity was assessed by testing the Persian AOFAS against the SF-36 questionnaire.

Results: Chronbach's alpha coefficient was 0.696, which was considered acceptable. Furthermore, the test-retest reliability measured by using the ICC for the subjective subscales was 0.853 (*P<0.001*). The reliability of testing the objective subscales was calculated by using Kappa, which indicated acceptable values. Pearson correlation coefficient between AOFAS and SF-36 was 0.415 (*P=0.008*). In addition, floor and ceiling effects were calculated 1.9% and 7.5% respectively.

Conclusion: In our study, Persian translation of AOFAS demonstrated acceptable validity and reliability with no need to be culturally adapted.

Level of evidence: II

Keywords: AOFAS, Persian, Translation, Validation

Introduction

A reliable, valid, and reproducible measurement tool is required to evaluate and compare the treatment outcomes in a qualitative manner (1, 2). Traditional objective instruments are physical examination and radiographic measurements. None of these measurements are representatives of patient outcomes and foot radiographic findings in particular are not reproducible in terms of inter and intra

Corresponding Author: Farshid Bagheri, Shahid Kamyab Emdadi Hospital, Mashhad University of Medical Sciences, Mashhad, Iran
Email: Ebrahimih2@mums.ac.ir

AB&JS

THE ONLINE VERSION OF THIS ARTICLE
ABJS.MUMS.AC.IR

PERSIAN TRANSLATION OF AOFAS

observer reliability. Recent trend is toward more subjective patient reported outcome scores (3-12). American Orthopedic Foot and Ankle Society Score (AOFAS) is commonly used for the assessment of foot and ankle conditions, which is comprised of three domains each assessing a specific feature of the lower extremity (13). AOFAS contains four anatomic scales including ankle-hindfoot, midfoot, hallux metatarsophalangeal-interphalangeal joints, and lesser metatarsophalangeal-interphalangeal joints. Each scale consists of three subscale including pain, function, and alignment with the total value score of 100 in each scale (14). These domains are divided into objective and subjective sub domains in which physicians rate to 5 questions while 4 are rated by the patients.

Due to extensive usage of AOFAS in various studies by foot and ankle surgeons, we aimed to translate and assess the psychometric properties of the Persian version, which can be used in Persian speaking countries including Iran, Tajikistan, Afghanistan, as well as parts of Iraq, and Pakistan.

Materials and Methods
Patients
In this study, we enrolled 53 patients from the foot and ankle clinic from November 2016 to April 2017 after receiving Institutional Review Board (IRB) approval from Mashhad

University of Medical Sciences. All of the patients agreed to complete informed consent form before entering the study. Inclusion criteria were ankle and hindfoot conditions and ability to speak Persian fluently. We excluded patients who were not able to independently complete the form (illiterate or none fluent speakers), with radiating pain because of the spine disorders, or with neuropathic pain such as charcot-marie tooth disease.

Translation and cross-cultural adaptation
Our study was conducted according to the recommended five staged cross-cultural adaption steps which are currently in use (15).These stages consist of translation, synthesis, back translation, expert committee review, and pretesting (15).

Translation: at first, three native Persian speakers including two orthopedists and one non-medical expert translated the original version of the questionnaire into Persian.

Synthesis, back translation, and expert committee review: the translated questionnaires were reconciled to achieve a unified questionnaire. Next, a native English speaker, who was also fluent in Persian language, back translated this version into English, which was followed by comparing that to the original one in a meeting to resolve probable contradictions and ambiguities.

Pretesting: Finally, 20 patients filled the Persian version to assess face validity and comprehensiveness of the questionnaire. Face validity indicates that each

item of a questionnaire is convenient for its overall concept. Face validity assesses if most people of the society would percept the same concept from a test. In other words, it answers to: do the questions really appear to be valid to the respondents(16). Following this step, we addressed difficulties in understanding the description of the items [Figure 1].

Statistical analysis
For statistical analysis, SPSS version 16.0 (SPSS Inc., Chicago IL) was used. Demographic statistics were calculated for all variables including frequency, percentage, mean, and standard deviation (SD). Also validation and psychometric assessment were performed using internal consistency, test–retest reliability, construct validity, and ceiling and floor effects measurement.

Assessment of psychometric properties
During the first visit, 53 patients were asked to fill out the Persian AOFAS and the 36-Item Short Form Survey (SF- 36) questionnaire. Fifty patients filled the AOFAS questionnaire for the second time during the second visit after a mean of four days (range: 2 to 7 days). Patients did not receive any major treatment during this period, including surgery or injection.

Reliability testing
Reliability estimates the consistency or reproducibility of a questionnaire by using various tests (17). In this study, we used Cronbach's alpha and the intraclass correlation coefficient (ICC) for the subjective parts of the questionnaire that was filled by the patients. Cronbach's alpha estimates the degree of relationship and direction of response among the items within a domain (18). Cronbach's alpha ranges between 0 to 1 with stronger relationship when it is closer to 1 (18). We also used ICC to assess the similarity in responses between test and retest after an average of 4 days. This shows reliability and reproducibility of the questionnaire (17). ICC value of > 0.70 is often considered reliable (19).

To test the reliability of the objective items that filled by two different physicians, we used Cohen's kappa test. It is similar to ICC, which is applied to test inter and intra rater reliability (20). Kappa coefficient closer to 1 is indicative of higher reliability between the raters. (20).

Construct validity
Validity estimates the concept that the questionnaire is supposed to measure (17). By applying a formerly validated questionnaire in Persian, the construct validity of a newly translated one can be analyzed. We used SF-36 questionnaire, which has been previously validated and translated into Persian (21). Construct validity was assessed using Pearson's correlation coefficient.

Floor and ceiling effects
These effects are measurable in studies with the

پرسشنامه AOFAS پشت پا

« مشــخــصــات فــردی »

تاریخ روز:	سن:	شماره ملی:	نام و نام خانوادگی:
سمت درگیر: راست ☐ چپ ☐		شماره پرونده:	شغل:

درد: (٤٠ امتیاز)

٤٠ ندارم

٣٠ خفیف (گهگاهی)

٢٠ متوسط (هر روز)

٠ شدید (تقریبا همیشه)

عملکرد: (٥٠ امتیاز)

محدودیت حرکتی و نیاز به وسیله کمکی

١٠ محدودیتی در فعالیت ندارم

٧ محدودیتی در فعالیت های روز مره ندارم , در فعالیت ورزشی و تفریحی محدودیت دارم.

٤ محدودیت در فعالیتهای روزمره و تفریحی دارم. و از عصا استفاده میکنم.

٠ محدودیت شدید در فعالیت های روزمره و تفریحی دارم و از واکر یا ویلچر یا بریس استفاده میکنم.

حد اکثر طول پیاده روی :(بر حسب بلوک)

٥ بیش از ٦ بلوک

٤ ٤–٦ بلوک

٢ ١–٣ بلوک

٠ کمتر از ١ بلوک

سطوح پیاده روی

٥ مشکلی در سطوح نا هموار ندارم

٣ در سطوح ناهموار, پله و سطح شیب دار کمی مشکل دارم.

٠ در سطوح ناهموار, پله و سطح شیب دار مشکل زیادی دارم.

اختلال در راه رفتن

٨ ندارم یا خفیف

٤ واضح (متوسط)

٠ شدید

دامنه حرکت در پلان ساژیتال (فلکشن و اکستنشن)

٨ طبیعی یا محدودیت کم(٣٠ درجه یا بیشتر)

٤ محدودیت متوسط (١٥ تا ٢٩ درجه)

٠ محدودیت شدید (کمتر از ١٥ درجه)

دامنه حرکتی عقب پا (اورژن و اینورژن)

٦ طبیعی یا محدودیت کم (٧٥ تا ١٠٠ ٪ طبیعی)

٣ محدودیت متوسط (٢٥ تا ٧٤٪ طبیعی)

٠ محدودیت شدید (کمتر از ٢٤٪ طبیعی)

پایداری مچ پا و عقب پا

٨ پایدار

٠ کاملا نا پایدار

الاینمنت (امتداد) : ١٠ امتیاز

١٠ خوب , پا پلانتیگراد و میدفوت در الاینمنت طبیعی

٨ متوسط , پا پلانتیگراد و درجاتی از اختلال محور در مید فوت , علامت ندارد

٠ بد , پا پلانتیگراد نیست و اختلال محور شدید در مید فوت دارد و علامتدار است.

Figure 1. Persian version of the AOFAS with acceptable psychometric properties.

sample size of at least 50 participants. These effects are not commonly acceptable if more than 15% of participants' responses show the lowest or highest possible scores. This means that the patients with the highest or lowest scores cannot be differentiated, which indicates limited responsiveness and low content validity when these effects are calculated more than 15% of the respondents (19).

Results

Demographic characteristics of participants are summarized in Table 1. The mean scores of AOFAS taken as test and retest were 75 and 76, respectively. Also the SF-36 questionnaire had the mean scores of 80 [Table 2].

Chronbach's alpha coefficient for measuring internal consistency ranged between 0.631 and 0.783 while the total was 0.696, which was considered acceptable [Table 3]. Furthermore, the test-retest reliability measured by using the ICC for the subjective subscales was 0.853 indicating a very good intra-rater reliability (*P<0.001*). The reliability of testing the objective subscales was calculated by using Kappa, which indicated acceptable values. Kappa coefficient was 0.47 for gait abnormality (*P<0.001*), 0.79 for sagittal ROM (*P<0.001*), 0.66 for hindfoot motion (*P<0.001*), 0.37 for stability (*P=0.008*), and 0.86 for alignment (*P<0.001*).

Pearson's correlation coefficient between AOFAS and SF-36 measured for the validity of the Persian version was 0.415, which was statistically significant (*P=0.008*) [Table 2]. In addition, floor and ceiling effects were calculated 1.9 % and 7.5 % respectively both of which were small and acceptable.

Table 1. Characteristics of patients with the ankle and hind foot problems (N=53)		
Age (Mean(SD))		32.1 (11.9)
Sex (no. (%))	Male	38.0 (70.4)
	Female	15.0 (27.8)
Affected Side (no. (%))	Right	38.0 (62.3)
	Left	19.0 (35.2)
	Both	1.00 (1.90)
Diagnosis	Calcaneus Fracture	8.00 (30.8)
	Ankle Fracture	6.00 (23.1)
	Others (talar osteochondral defect, ankle sprain, Achilles tendon rupture, and ankle ganglion cyst)	12.00 (46.2)
Occupation (no. (%))	Student	5.00 (11.9)
	Self-employed	27.0 (64.3)
	Teacher	2.00 (4.80)
	Housewife	8.00 (19.0)
Educational level (no. (%))	Primary school	8.00 (21.1)
	Junior high school	15.0 (39.5)
	High School diploma	7.00 (18.4)
	Bachelor degree	8.00 (21.1)

Table 2. Average scores of patient with the ankle and hind foot problems and construct validity expressed by Pearson's correlation between AOFAS and SF36				
Questionnaires	Mean	SD	Pearson correlation	*P*
AOFAS Retest	76.3	19.5	–	–
AOFAS Test	75.4	20.8	0.415	0.008
SF36	80.4	13.8		

Table 3. Internal consistency and test-retest reliability of the Persian version of the American Orthopedic Foot and Ankle Society (AOFAS) Ankle-Hindfoot Questionnaire							
Internal Consistency		**Cronbach Alpha**	**Kappa Intraclass Correlation**	**Intraclass correlation**	**Upper Band**	**Lower band**	**Sig**
Subjective domain	Pain	0.783	_	0.853	0.917	0.741	<0.001
	Activities limitations	0.656	_				
	Maximum walking distance	0.678	_				
	Walking surfaces	0.653	_				
Objective domain	Gait abnormality	0.656	0.474	_	_	_	<0.001
	Sagittal ROM	0.668	0.788	_	_	_	<0.001
	Hindfoot motion	0.656	0.662	_	_	_	<0.001
	Stability	0.701	0.369	_	_	_	0.008
	Alignment	0.631	0.856	_	_	_	<0.001
	Total	0.696	_	_	_	_	_

Discussion

Several studies have investigated the reliability and validity of AOFAS questionnaire in patients with ankle and foot disorders (1, 22-24). In this study, we assessed the psychometric properties in patients with any ankle and foot condition (25). Our study revealed an acceptable reliability and validity for Persian version of this questionnaire in both subjective and objective domains.

Like the other studies, our study had some limitations. First, the small number of participants is always a limitation. SF-36 questionnaire was a general measure, which could not show a high coefficient although the correlation was still significant.

The AOFAS questionnaire was translated into Persian with no need to be culturally adapted. The internal coherence between each item and the total score was acceptable. Furthermore, the intra-rater reliability was measured using the subjective subscales of the Persian version of the AOFAS questionnaire, which indicated a very good result. Also, the inter-rater reliability was calculated using Kappa test.

The intra-rater reliability of subjective subscales in other studies showed excellent results, which was very good in our study (22, 24). Our result can possibly arise from the fact that pain item has a great effect on subjective scores. Since pain level cannot change during four days, intra-rater reliability cannot differ during this short period. In addition, it is possible that subjective subscales could be recalled easily after a short period for the retest (24). Comparing kappa and ICC revealed higher reliability in subjective subscales that might also be influenced by the pain as noted before.

Validity of the Persian AOFAS was measured by testing against SF-36, which had been validated into Persian previously. Choosing SF36 as a questionnaire to validate another orthopedic one has been recommended (13). A systematic review revealed no significant construct validity comparing AOFAS with the SF-36, while there was a moderate correlation between AOFAS and the other measures including foot function index (FFI) (1, 26). As such, it has been suggested to use both AOFAS and SF-36 questionnaires in case of an ankle and hindfoot condition as they complete each other due to the fact that SF-36 demonstrates the overall patient's health status whereas AOFAS measures a specific region's function (27).

In our study, Persian translation of AOFAS demonstrated acceptable validity and reliability with no need to be culturally adapted. Due to the fact that this questionnaire is shorter than the other foot and ankle assessment tools such as SF-36 and FFI, it is more appreciable among clinicians as well as the patients with foot and ankle problems both for the treatment assessment and the follow-up.

Patient consent: Informed consent was obtained from the study participant

Disclosure: The authors report no conflict of interest concerning the materials or methods used in this study or the findings specified in this paper.

Sayyed-Hadi Sayyed-Hosseinian MD
Farshid Bagheri MD
Orthopedic Research Center, Shahid Kamyab Hospital, Mashhad University of Medical Sciences, Mashhad, Iran

Golnaz Ghayyem Hassankhani MD
Niloufar Alavi BSc
Alireza Mousavian MD
Orthopedic Research Center, Ghaem Hospital, Mashhad University of Medical Sciences, Mashhad, Iran

Babak Shojaie MD FA
Traumatology, Hand and Orthopedic Surgery Department, st. Marien Medical Campus, Friesoythe, Germany

PERSIAN TRANSLATION OF AOFAS

References

1. Ibrahim T, Beiri A, Azzabi M, Best AJ, Taylor GJ, Menon DK. Reliability and validity of the subjective component of the American Orthopaedic Foot and Ankle Society clinical rating scales. J Foot Ankle Surg. 2007; 46(2):65-74.

2. Kachooei AR, Badiei Z, Zandinezhad ME, Ebrahimzadeh MH, Mazloumi SM, Omidi-Kashani F, et al. Influencing factors on the functional level of haemophilic patients assessed by FISH. Haemophilia. 2014; 20(2):185-9.

3. Mousavian A, Ebrahimzadeh MH, Birjandinejad A, Omidi-Kashani F, Kachooei AR. Translation and cultural adaptation of the Manchester-Oxford Foot Questionnaire (MOXFQ) into Persian language. Foot (Edinb). 2015; 25(4):224-7.

4. Ebrahimzadeh MH, Vahedi E, Baradaran A, Birjandinejad A, Seyyed-Hoseinian SH, Bagheri F, et al. Psychometric properties of the Persian Version of the simple shoulder test (SST) questionnaire. Arch Bone Jt Surg. 2016; 4(4):387-92.

5. Kachooei AR, Ebrahimzadeh MH, Erfani-Sayyar R, Salehi M, Salimi E, Razi S. Short Form-McGill Pain Questionnaire-2 (SF-MPQ-2): a cross-cultural adaptation and validation study of the persian version in patients with knee osteoarthritis. Arch Bone Jt Surg. 2015; 3(1):45-50.

6. Ebrahimzadeh MH, Birjandinejad A, Golhasani F, Moradi A, Vahedi E, Kachooei AR. Cross-cultural adaptation, validation, and reliability testing of the shoulder pain and disability index in the Persian population with shoulder problems. Int J Rehabil Res. 2015; 38(1):84-7.

7. Ebrahimzadeh MH, Birjandinejad A, Kachooei AR. Cross-cultural adaptation, validation, and reliability of the Michigan Hand Outcomes Questionnaire among Persian population. Hand Surg. 2015; 20(1):25-31.

8. Baradaran A, Ebrahimzadeh MH, Birjandinejad A, Kachooei AR. Cross-cultural adaptation, validation, and reliability testing of the modified oswestry disability questionnaire in Persian population with low back pain. Asian Spine J. 2016; 10(2):215-9.

9. Moradi A, Menendez ME, Kachooei AR, Isakov A, Ring D. Update of the quick DASH questionnaire to account for modern technology. Hand (N Y). 2016; 11(4):403-9.

10. Ebrahimzadeh MH, Moradi A, Vahedi E, Kachooei AR, Birjandinejad A. Validity and reliability of the persian version of shortened disabilities of the arm, shoulder and hand questionnaire (Quick-DASH). Int J Prev Med. 2015; 6(1):59.

11. Ebrahimzadeh MH, Kachooei AR, Vahedi E, Moradi A, Mashayekhi Z, Hallaj-Moghaddam M, et al. Validity and cross-cultural adaptation of the persian version of the oxford elbow score. Int J Rheumatol. 2014; 2014(10):381237.

12. Kachooei AR, Moradi A, Janssen SJ, Ring D. The influence of dominant limb involvement on DASH and quick DASH. Hand (N Y). 2015; 10(3):512-5.

13. Coster MC, Rosengren BE, Bremander A, Brudin L, Karlsson MK. Comparison of the self-reported foot and ankle score (SEFAS) and the American Orthopedic foot and ankle society score (AOFAS). Foot Ankle Int. 2014; 35(10):1031-6.

14. Kostuj T, Krummenauer F, Schaper K, Stief F, Zettersten K, Baums MH, et al. Analysis of agreement between the German translation of the American Foot and Ankle Society's Ankle and Hindfoot Scale (AOFAS-AHS) and the Foot Function Index in its validated German translation by Naal et al. (FFI-D). Arch Orthop Trauma Surg. 2014; 134(9):1205-10.

15. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine. 2000; 25(24):3186-91.

16. Bolarinwa OA. Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. Niger Postgrad Med J. 2015; 22(4):195-201.

17. Bravo G, Potvin L. Estimating the reliability of continuous measures with Cronbach's alpha or the intraclass correlation coefficient: toward the integration of two traditions. J Clin Epidemiol. 1991; 44(4-5):381-90.

18. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ. 2011; 2(1):53-5.

19. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007; 60(1):34-42.

20. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med. 2012; 22(3):276-82.

21. Montazeri A, Goshtasebi A, Vahdaninia M, Gandek B. The short form health survey (SF-36): translation and validation study of the Iranian version. Qual Life Res. 2005; 14(3):875-82.

22. Rodrigues RC, Masiero D, Mizusaki JM, Imoto AM, Peccin MS, Cohen M, et al. Translation, cultural adaptation and validity of the" American Orthopaedic Foot and Ankle Society (AOFAS) Ankle-Hindfoot Scale". Acta Ortop Bras. 2008; 16(2):107-11.

23. Leigheb M, Janicka P, Andorno S, Marcuzzi A, Magnani C, Grassi F. Italian translation, cultural adaptation and validation of the "American Orthopaedic Foot and Ankle Society's (AOFAS) ankle-hindfoot scale". Acta Biomed. 2016; 87(1):38-45.

24. Analay Akbaba Y, Celik D, Ogut RT. Translation, cross-cultural adaptation, reliability, and validity of turkish version of the american orthopaedic foot and ankle

PERSIAN TRANSLATION OF AOFAS

society ankle-hindfoot scale. J Foot Ankle Surg. 2016; 55(6):1139-42.

25. Vosoughi AR, Roustaei N, Mahdaviazad H. American orthopaedic foot and ankle society ankle–hindfoot scale: a cross-cultural adaptation and validation study from Iran. Foot Ankle Surg. 2017; 7731(17):30044-9.

26. Button G, Pinney S. A meta-analysis of outcome rating scales in foot and ankle surgery: is there a valid, reliable, and responsive system? Foot Ankle Int. 2004; 25(8):521-5.

27. Ceccarelli F, Calderazzi F, Pedrazzi G. Is there a relation between AOFAS ankle-hindfoot score and SF-36 in evaluation of Achilles ruptures treated by percutaneous technique? J Foot Ankle Surg. 2014; 53(1):16-21.