Cross-cultural adaptation and validation of the Ankle Osteoarthritis Scale for use in French-speaking populations

Magalie Angers, MD
Amy Svotelis, PhD
Frederic Balg, MD
Jean-Pascal Allard, MD

Accepted for publication
Dec. 22, 2015

Correspondence to:
M. Angers
Department of Surgery
Orthopaedic Division
Centre intégré universitaire de santé et de services sociaux de l'Estrie — Centre hospitalier universitaire de Sherbrooke
300 King St E.
Sherbrooke QC J1G 1B1
magalie.angers@usherbrooke.ca

DOI: 10.1503/cjs.010415

Background: The Ankle Osteoarthritis Scale (AOS) is a self-administered score specific for ankle osteoarthritis (OA) with excellent reliability and strong construct and criterion validity. Many recent randomized multicentre trials have used the AOS, and the involvement of the French-speaking population is limited by the absence of a French version. Our goal was to develop a French version and validate the psychometric properties to assure equivalence to the original English version.

Methods: Translation was performed according to American Association of Orthopaedic Surgeons (AAOS) 2000 guidelines for cross-cultural adaptation. Similar to the validation process of the English AOS, we evaluated the psychometric properties of the French version (AOS-Fr): criterion validity (AOS-Fr v. Western Ontario and McMaster Universities Arthritis Index [WOMAC] and SF-36 scores), construct validity (AOS-Fr correlation to single heel-lift test), and reliability (AOS-Fr test–retest). Sixty healthy individuals tested a prefinal version of the AOS-Fr for comprehension, leading to modifications and a final version that was approved by C. Saltzman, author of the AOS. We then recruited patients with ankle OA for evaluation of the AOS-Fr psychometric properties.

Results: Twenty-eight patients with ankle OA participated in the evaluation. The AOS-Fr showed strong criterion validity (AOS:WOMAC r = 0.709 and AOS:SF-36 r = –0.654) and construct validity (r = 0.664) and proved to be reliable (test–retest intraclass correlation coefficient = 0.922).

Conclusion: The AOS-Fr is a reliable and valid score equivalent to the English version in terms of psychometric properties, thus is available for use in multicentre trials.

Contexte : L'Ankle Osteoarthritis Scale (AOS) est une échelle d'auto-évaluation de l'arthrose de la cheville très fiable, et dont la validité conceptuelle et critérielle est élevée. De nombreux essais multicentriques randomisés récents ont utilisé l'AOS, mais faute d'une version française, la participation de la population francophone est limitée. Notre objectif était donc de créer une version française et d'en valider les propriétés psychométriques pour veiller à ce qu'elle soit équivalente à la version anglaise originale.

Méthodes : La traduction a été effectuée conformément aux lignes directrices de 2000 de l'American Association of Orthopaedic Surgeons (AAOS) en matière d'adaptation interculturelle. Comme ce fut le cas pour le processus de validation de l'échelle anglaise, nous avons évalué les propriétés psychométriques de la version française (AOS-Fr) : validité critérielle (AOS-Fr contre le Western Ontario and McMaster Universities Arthritis Index [WOMAC] et les scores du questionnaire SF-36), validité conceptuelle (corrélation de l'AOS-Fr au test d'élévation sur la pointe d'un seul pied) et fiabilité (test–retest de l'AOS-Fr). Soixante personnes en santé ont fait l'essai d'une version préfinale de l'AOS-Fr pour en évaluer l'intelligibilité, ce qui a entraîné des modifications, et la version définitive a été approuvée par C. Saltzman, auteur de l'AOS. Nous avons ensuite recruté des patients atteints d'arthrose de la cheville pour évaluer les propriétés psychométriques de l'AOS-Fr.

Résultats : Vingt-huit patients atteints d'arthrose à la cheville ont participé à l'évaluation. Une forte validité critérielle (AOS:WOMAC : r = 0,709 et AOS:SF-36 : r = –0,654) et conceptuelle (r = 0,664) a été mise en évidence, et l'échelle s'est avérée fiable (coefficient de corrélation intraclasse = 0,922 pour le test–retest).

Conclusion : L'AOS-Fr est une échelle fiable et valide équivalente à la version anglaise sur le plan des propriétés psychométriques; elle peut donc être utilisée pour les essais multicentriques.

Ankle osteoarthritis (OA) is defined as degenerative changes of the tibiotalar joint. It is less common than hip and knee OA, and as opposed to those larger joints, less than 10% of cases of OA are primary cases of ankle OA.[1] As the most common cause of ankle OA is post-traumatic, many young patients are affected by this condition, and in this otherwise active population, ankle OA treatment can be challenging.[2] The progression of ankle OA leads to invalidity and remains difficult to treat without causing functional limitation.[3] A study by Glazebrook and colleagues[4] in 2008 demonstrated that disability associated with ankle OA was at least as severe as that associated with hip OA.

One of the main treatment options for ankle OA is ankle arthrodesis. However, long-term studies have reported secondary subtalar arthrosis in as many as 50% of patients at 10 years.[5] While early total ankle arthroplasty (TAA) designs were associated with high failure rates and complications, biomechanical progress brought forward a new generation of TAA designs with good to excellent results in the intermediate term.[6] This renewal in interest for TAA led surgeons to perform this procedure more often. Long-term multicentre studies are required to help surgeons decide whether fusion or TAA is the best treatment option for their patients. Good randomized controlled trials rely on the use of good functional outcome evaluation tools to provide high-quality and valid results.

The Ankle Osteoarthritis Scale (AOS) is an adaptation of the Foot Function Index, modified specifically for ankle OA, and is often used for research purposes.[7,8] The AOS is a self-administered score divided in 2 subscales of 9 items. The first section evaluates pain, while the other is designed to evaluate functional limitations. Each item is answered using a visual analogue scale. The score is known for excellent reliability and strong construct and criterion validity.[8] The AOS has been recommended for use in conjunction with the SF-36 for the evaluation of end-stage ankle OA owing to its level of responsiveness and lack of constraint of patient responses.[9] Notably, many recent randomized multicentre trials have used the AOS as a measure of ankle OA.[6,10]

The usability and validity of any functional scale is only as good as its ability to evaluate patients cross-culturally. In order to evaluate a non-English–speaking population, functional score questionnaires in the patients' native language should be created and validated.[8,11,12] At this time, the participation of French-speaking orthopedic patients in multicentre studies is limited by the absence of a French version of the AOS. Our goal was to translate the AOS into French and validate the French version by evaluating its psychometric properties.

Methods

Translation

The study was approved by the ethics committee of the Centre intégré universitaire de santé et de services sociaux de l'Estrie — Centre hospitalier universitaire de Sherbrooke (CIUSSS de l'Estrie CHUS; protocol #09–130). Following the guidelines for the cross-cultural adaptation process written by the American Association of Orthopaedic Surgeons (AAOS) in 2000, the translation was performed using a 6-step process:[8] initial translation, synthesis, back translation, expert committee, test of the prefinal version and submission of the document to the developer. For the initial translation, 2 independent translators whose mother tongue was French translated the scale from English to French. The synthesis step required these 2 translators to meet, discuss and compose a synthesized version of the translated AOS. Two independent translators, blind to the original scale, whose native language was English then translated the synthesized version back to English. An expert committee composed of a linguist, 2 orthopedic surgeons and the 4 translators revised the whole process and consolidated the prefinal version. Sixty individuals without ankle OA then tested the prefinal version by answering the questionnaire and evaluating their comprehension of each item. From the data collected in this step, some sections of the test were changed to improve comprehension and readability. For the final step, the methods for obtaining the corrected French version (AOS-Fr) were submitted to Dr. Charles Saltzman, the developer of the English version, who approved the methodology and use (Appendix 1, available at canjsurg.ca).

Evaluation of psychometric properties

In order to validate the use of the AOS-Fr, we set out to evaluate the psychometric properties of the test among patients with ankle OA using the same process as the original study validating the English version of the AOS.[7] We recruited patients with degenerative changes isolated to the ankle at the outpatient orthopedic clinic of CIUSSS de l'Estrie CHUS, Sherbrooke, Que. Most patients were seen for a follow-up of their ankle OA, but a few of them were new consultations. The diagnosis of ankle OA was based on the presence of degenerative changes evident on weight-bearing radiographs. To be included, patients had to consider French as their mother tongue and be able to read and write in French. Patients younger than 18 years and those with additional foot and/or ankle pathologies were excluded.

At the first visit, we collected sociodemographic data (age, sex, body mass index [BMI], occupation) and pertinent medical history (diabetes, neuropathy, ankle or foot fractures), and patients were asked to perform single heel-lift tests of both the affected and unaffected sides. The participants completed 3 questionnaires: the AOS-Fr, SF-36 and Western Ontario and McMaster Universities Arthritis Index (WOMAC). We assessed criterion validity by comparing the AOS-Fr to the WOMAC and the SF-36 scores, as was done in the original validation of the AOS.[7] A research assistant gave participants instructions for

completing the questionnaires, and the patients were left to complete the questionnaires alone. Construct validity was established by examining the correlation between the AOS-Fr scores and the single heel-lift test. Finally, to measure test–retest reliability, patients were asked to complete the AOS-Fr a second time and to complete a small questionnaire detailing any modification to the treatment (shoes, orthotics, medication, surgery, injection) 1 week after the first visit and return the questionnaire by mail. Any modification to their treatment between the completion of the 2 AOS-Fr questionnaires would nullify the test–retest reliability, so these patients would be excluded from the analysis.

Power and sample size calculation

The required number of patients was estimated based on properties to assess. The reproducibility (test–retest) of the questionnaire is an important property to assess. Its lower acceptable limit is 0.85. Because the reproducibility of the English version would require at least 0.97, we estimated that the French version would be at least 0.90. A sample of 10 patients was therefore sufficient to obtain a power of 99% with a reproducibility of 0.90.[13,14] Concerning validity, our goal was to obtain a statistically significant correlation with the 3 selected tests. We hypothesized that we would obtain the lowest correlation between the AOS-Fr and the SF-36 questionnaire, as the SF-36 is not specific to ankle OA. By estimating an average correlation of 0.50, a sample of 30 patients was required to obtain a significant correlation with a power of 80% and $p < 0.05$.[13,14]

Results

Population

A total of 28 patients were included in the study: 18 men and 10 women. Thirteen patients had OA of the right ankle, 12 patients had OA of the left ankle, and 3 patients had bilateral ankle involvement. The mean age was 61 (range 25–83) years, and the mean BMI was 31.2 (range 22.7–46.9; Table 1).

Test–retest reliability

To determine reliability of the AOS-Fr, we compared the scores of the 2 AOS tests completed by the participants, excluding those whose treatments were modified between the tests, as modification may have led to a difference in functional outcome. Between the first visit and the second time answering the AOS-Fr 1 week later, treatment was modified in 9 patients. One patient had an ankle surgery, 2 received a cortisone ankle injection, 3 began wearing orthotics and 3 modified their shoe wear. All of these patients were excluded from the test–retest analysis, leaving 19 patients (Fig. 1). The intraclass correlation coefficient (ICC) between the AOS-Fr at first visit and after 1 week was excellent. The ICC for the entire score was 0.922 (95% confidence interval [CI] 0.8–0.97), with 0.895 (95% CI 0.729–0.96) for the pain subscale and 0.915 (95% CI 0.779–0.967) for the disability subscale (Table 2).

Criterion validity

We used the examination of the WOMAC correlation to the AOS-Fr completed at the first visit to measure criterion validity. Two patients didn't complete the WOMAC properly and were excluded, leaving 26 patients for the analysis (Fig. 1). The Pearson correlation coefficient between the global AOS-Fr and the WOMAC scores was $r = 0.709$ ($p < 0.001$). When comparing the WOMAC scores with the AOS-Fr pain and disability subscale scores, the correlation coefficients were $r = 0.677$ ($p < 0.001$) and $r = 0.698$ ($p < 0.001$), respectively (Table 2).

We then proceeded to calculate the correlation between the AOS-Fr and SF-36 scores. As described in the original publication of the AOS, we expected a negative correlation between the AOS-Fr and SF-36 scores since an increase in the SF-36 score represents a better function, while a high AOS score suggests poor function.[7] Indeed, the correlation coefficient obtained between the AOS-Fr and SF-36 scores was $r = –0.654$ ($p = 0.001$). When evaluated separately, the SF-36 strongly correlated with the pain section ($r = –0.620$, $p = 0.003$) and with the disability section ($r = –0.618$, $p = 0.003$) of the AOS-Fr (Table 2).

Construct validity

Since the single heel-lift ratio is influenced by the results of both ankles, we excluded 3 patients who had bilateral ankle OA, leaving 25 patients for the analysis (Fig. 1). Similar to the SF-36, the single heel-lift is reciprocal to the AOS score, giving a negative correlation. The ICCs between the AOS-Fr and the single heel-lift were $r = –0.664$ ($p < 0.001$) overall, $r = –0.664$ ($p < 0.001$) for the global score, $r = –0.542$ ($p < 0.001$) for the pain score and $r = –0.628$ ($p < 0.001$) for the disability score (Table 2).

Table 1. Baseline demographic and clinical data for participants with ankle OA ($n = 28$)

Characteristic	Participants
Sex, male:female	18:10
Age, mean ± SD, yr	61 ± 16.5
Age, range, yr	25–83
BMI, mean ± SD	31.2 ± 6.35
BMI, range	22.7–46.9
Affected ankle, right:left:bilateral	13:12:3

BMI = body mass index; OA = osteoarthritis; SD = standard deviation.

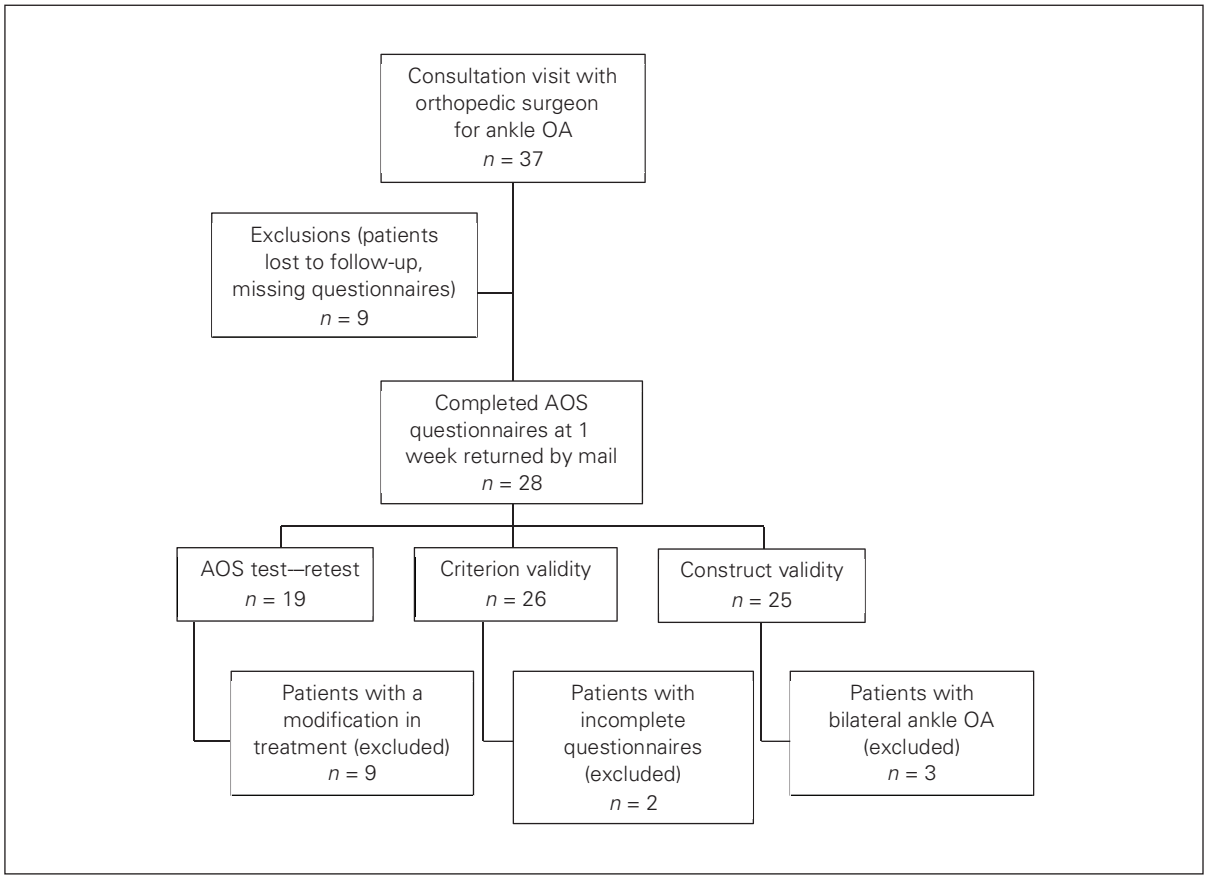

Fig. 1: Breakdown of the 37 participants recruited for the study. Owing to different factors (loss to follow-up, incomplete questionnaires, and treatment modifications between visits), some participants were not included in our analyses. AOS = Ankle Osteoarthritis Scale; OA = osteoarthritis.

Table 2. Validity of the AOS-Fr

Comparison	ICC
Test–retest validity*	
AOS-Fr v. AOS-Fr 1-wk	0.922 (95% CI 0.800–0.970)
AOS-Fr v. AOS-Fr 1-wk, pain subscale	0.895 (95% CI 0.729–0.960)
AOS-Fr v. AOS-Fr 1-wk, disability subscale	0.915 (95% CI 0.779–0.967)
Criterion validity†	
AOS-Fr v. WOMAC	0.709§
AOS-Fr/PAIN vs WOMAC	0.677§
AOS-Fr, disability subscale v. WOMAC	0.698§
AOS-Fr v. SF-36	–0.654§
AOS-Fr, pain subscale v. SF-36	–0.620§
AOS-Fr, disability subscale v. SF-36	–0.618§
Construct validity‡	
AOS-Fr v. single heel-lift ratio	–0.664§
AOS-Fr, pain subscale v. single heel-lift ratio	–0.542§
AOS-Fr, disability subscale v. single heel-lift ratio	–0.628§

AOS-Fr = Ankle Osteoarthritis Scale, French version; CI = confidence interval; ICC = intraclass correlation coefficient; WOMAC = Western Ontario and McMaster Universities Arthritis Index.

*Calculated using ICC.

†Measured using the Pearson correlation coefficient comparing the WOMAC and SF-36 to the AOS-Fr at the same visit.

‡Calculated using the Pearson correlation coefficient between the AOS-Fr and the single heel-lift ratio at the same visit.

§Significant at $p < 0.01$, 2-tailed.

Discussion

The purpose of this study was to produce a valid French version of the AOS that is equivalent to the English version. The final AOS-Fr was validated for reliability, construct validity and criterion validity. The result of the test–retest reliability showed an excellent ICC of 0.922, which was almost equal to the original version (ICC = 0.97).[7] Similarly, evaluation of the criterion validity gave a moderate correlation between the SF-36 and AOS-Fr disability scale (r = –0.618, p = 0.003), as found in the English version (r = –0.66, p = 0.001). However, we detected a stronger correlation with the pain subscale (r = –0.620, p = 0.003) than that calculated in the original version (r = –0.34, p < 0.20).[7] When studying the correlation between WOMAC and the AOS-Fr pain and disability subscale scores we obtained correlations of 0.667 and 0.698, respectively, which is in line with the results of the English version (0.65 and 0.79, respectively). Examination of the construct validity by calculation of the correlation between the single heel-lift ratio and the AOS-Fr demonstrated less correlation than the original version but still a strong and significant correlation (AOS-Fr: r = –0.664, p < 0.001 v. original: r = 0.88 for the global score; AOS-Fr: r = –0.542, p = 0.003 v. original: r = 0.90 for the pain score; AOS-Fr: r = –0.628, p < 0.001 v. original: r = 0.63 for the pain score).[7] Although our study included slightly fewer patients than required in our sample size calculation, we obtained strong and significant correlations with all of our tests.

Conclusion

The AOS-FR is a reliable and valid score. It is equivalent to the original English version in terms of psychometric properties.

Acknowledgements: The authors thank the Université de Sherbrooke medical students for their help with data collection and the Fondation de recherché et enseignement en orthopédie de Sherbrooke for funding this study.

Affiliations: From the Department of Surgery, Orthopaedic division, Centre intégré universitaire de santé et de services sociaux de l'Estrie — Centre hospitalier universitaire de Sherbrooke (CIUSSS de l'Estrie CHUS), Sherbrooke, QC (Angers, Svotelis, Balg); and the Department of Surgery, Orthopaedic and traumatology division, Centre de Santé et des Services Sociaux de Chicoutimi, Chicoutimi, QC (Allard).

Competing interests: None declared.

Contributors: M. Angers, F. Balg and J.-P Allard designed the study. M. Angers, A. Svotelis and J.-P. Allard acquired the data, which M. Angers and A. Svotelis analyzed. M. Angers, A. Svotelis and F. Balg wrote the article, which all authors reviewed and approved for publication.

References

1. Valderrabano V, Horisberger M, Russell I, et al., Etiology of ankle osteoarthritis. *Clin Orthop Relat Res* 2009;467:1800-6.
2. Thomas RH, Daniels TR. Ankle arthritis. *J Bone Joint Surg Am* 2003;85-A:923-36.
3. Coester LM, Saltzman CL, Leupold J, et al. Long-term results following ankle arthrodesis for post-traumatic arthritis. *J Bone Joint Surg Am* 2001;83-A:219-28.
4. Glazebrook M, Daniels T, Younger A, et al. Comparison of health-related quality of life between patients with end-stage ankle and hip arthrosis. *J Bone Joint Surg Am* 2008;90:499-505.
5. Ling JS, Smyth NA, Fraser EJ, et al. Investigating the relationship between ankle arthrodesis and adjacent-joint arthritis in the hindfoot: a systematic review. *J Bone Joint Surg Am* 2015;97: 513-20.
6. Nguyen MP, Pedersen DR, Gao Y, et al. Intermediate-term follow-up after ankle distraction for treatment of end-stage osteoarthritis. *J Bone Joint Surg Am* 2015;97:590-6.
7. Domsic RT, Saltzman CL. Ankle osteoarthritis scale. *Foot Ankle Int* 1998; 19:466-71.
8. Beaton, DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976)* 2000. 25:3186-91.
9. Madeley NJ, Wing KJ, Topliss C, et al. Responsiveness and validity of the SF-36, Ankle Osteoarthritis Scale, AOFAS Ankle Hindfoot Score, and Foot Function Index in end stage ankle arthritis. *Foot Ankle Int* 2012;33:57-63.
10. Daniels TR, Mayich DJ, Penner MJ. Intermediate to long-term outcomes of total ankle replacement with the Scandinavian Total Ankle Replacement (STAR). *J Bone Joint Surg Am* 2015;97:895-903.
11. Bullinger M, Alonso J, Apolone G, et al. Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. International Quality of Life Assessment. *J Clin Epidemiol* 1998;51:913-23.
12. Perneger TV, Leplege A, Etter JF. Cross-cultural adaptation of a psychometric instrument: two methods compared. *J Clin Epidemiol* 1999;52:1037-46.
13. Streiner DL, Norman GR *Health measurement scales. A practical guide to their development and use. 4th edition*. New York: Oxford University Press; 2008.
14. McDowell I. *Measuring health: a guide to rating scales and questionnaires. 3rd edition*. New York: Oxford University Press; 2006.