The Journal of Foot & Ankle Surgery 54 (2015) 46–50

Contents lists available at ScienceDirect

The Journal of Foot & Ankle Surgery

journal homepage: www.jfas.org

Validity, Reliability, and Responsiveness of the Korean Version of American Academy of Orthopedic Surgeons Foot and Ankle Questionnaire

Ji-Beom Kim, MD [1], Je Kyun Kim, MD [1], Sang Gyo Seo, MD [2], Dong Yeon Lee, MD [3]

[1] Orthopedic Surgeon, Department of Orthopedic Surgery, Seoul National University Hospital, Yeongeon-dong, Jongno-gu, Seoul, Korea
[2] Instructor, Department of Orthopedic Surgery, Seoul National University Hospital, Yeongeon-dong, Jongno-gu, Seoul, Korea
[3] Associate Professor (Clinical), Department of Orthopedic Surgery, Seoul National University Hospital, Yeongeon-dong, Jongno-gu, Seoul, Korea

ARTICLE INFO

Level of Clinical Evidence: 3

Keywords:
AAOS
discriminant validity
outcome assessment
test-retest reliability
transcultural adaptation
translation
VAS pain

ABSTRACT

The American Academy of Orthopaedic Surgeons Foot and Ankle Questionnaire (AFAQ) reflects patients' subjective disorder due to foot and ankle conditions. We evaluated the validity, reliability, and responsiveness of the Korean version of the AFAQ, after translation and transcultural adaptation of the original AFAQ into the Korean language. A total of 206 patients were enrolled, including 152 with chronic problems (experimental group) and 54 with acute problems (control group). We used the intraclass correlation coefficient to assess the test–retest reliability and Cronbach's α to assess internal reliability. Pearson's correlation coefficient was used to assess the criterion validity by correlating the Korean AFAQ scores with those from other validated scales (American Orthopaedic Foot and Ankle Society Hallux-Metatarsophalangeal-Interphalangeal scale, American Orthopaedic Foot and Ankle Society Ankle-Hindfoot scale, and visual analog scale for pain). To analyze discriminant validity, we evaluated the difference between the experimental and control groups using the Student t test. Of the 152 patients in the experimental group, 29 revisited our clinic postoperatively and repeated the Korean AFAQ. To analyze responsiveness, we used paired t tests to evaluate postoperative changes. In terms of test–retest reliability, the intraclass correlation coefficient ranged from 0.979 to 0.999. In terms of internal reliability, Cronbach's α was 0.528 for the stiffness and swelling subscale and greater than 0.7 for all other subscales. In terms of criterion validity, Pearson's correlation coefficient ranged from 0.492 to 0.699. The probability of the null hypothesis for discriminant validity and responsiveness was statistically significant ($p < .001$ and $p = .021$, respectively). These results showed that the Korean version of the AFAQ had the same concept and intention as the original version and is reliable, valid, and responsive.

Disorders involving the foot and ankle increase with age owing to degenerative changes involving bone, joint, neuromuscular tissues, and tendon (1). Active lifestyles, shoe gear, and an increased body mass can influence the prevalence of foot and ankle disorders (2). In fact, foot and ankle disorders are usually not correlated directly to age; rather, they correlate with quality of life. However, assessing the effect of foot and ankle disorders on each individual can be quite difficult, because, even in cases of similar severity, they could have different effects, depending on the lifestyle of the patient (3).

To assess foot and ankle problems, determining the subjective malfunction and severity of pain is as important as finding the anatomic malformation (4). Various questionnaires have been used as an instrument to measure the subjective discomfort due to the disorders and/or the improvement after treatment (5–10). However, no reference standard instrument has yet been determined for assessment of the foot and ankle. For example, the RAND Medical Outcomes Study Short Form-36 questionnaire was supposed to represent the subjective discomfort of patients well, but it does not focus on foot and ankle problems (11) but, instead, is a general health measurement. The American Orthopaedic Foot and Ankle Society (AOFAS) scales focus on foot and ankle problems and seem to be the most widely used scales in foot and ankle research in recent years (6). The AOFAS scale studies reported the scales could evaluate the outcomes of surgery (12), and their subjective component had acceptable validity (13). However, other studies have reported that the AOFAS scales are difficult to perform and lack reliability (14), validity, and precision (15–17).

Financial Disclosure: This study was supported by Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Science, ICT and Future Planning (grant NRF-2012R1A1A2044067).

Conflict of Interest: None reported.

Address correspondence to: Dong Yeon Lee, MD, Associate Professor (Clinical), Department of Orthopedic Surgery, Seoul National University Hospital, Yeongeon-dong, Jongno-gu, Seoul, Korea, 110-744.

E-mail address: leedy@snu.ac.kr (D.Y. Lee).

1067-2516/$ - see front matter © 2015 by the American College of Foot and Ankle Surgeons. All rights reserved.
http://dx.doi.org/10.1053/j.jfas.2014.08.011

The American Academy of Orthopaedic Surgeons (AAOS) (9) developed an instrument for subjective patient assessment (i.e., the AAOS Lower Limb Instrument), and the AAOS Foot and Ankle Questionnaire (AFAQ) is one of the major parts of the AAOS Lower Limb Instrument, focusing on foot and ankle problems. The AFAQ is a self-administered questionnaire designed for efficient collection of patient-oriented data. It consists of 25 items and combines the Global Foot and Ankle scale (GFAS) and the Shoe Comfort scale (SCS). The GFAS is formatted into 4 subscales that address pain, function, stiffness and swelling, and "giving way" (9). "Giving way" of an ankle means the ankle is unable to support the person's weight unexpectedly and suddenly. The SCS is formatted into 5 questions about the type of shoes that are comfortable to the patient. An assessment instrument should be examined for reliability, validity, and responsiveness. After an assessed instrument has been determined to be reliable, valid, and responsive, physicians and surgeons can trust the instrument and use it. The AFAQ was expected to reflect a patient's subjective discomfort well and to produce meaningful data. Although AAOS group insisted that the AFAQ showed high reliability and validity (9), similar support for the AFAQ in the published data, to our knowledge, is nonexistent.

The original AFAQ was developed in the English language and was intended for use in English-speaking countries (9). To validate the AFAQ in Korea, the original AFAQ had to be translated into the Korean language. Although it is not easy to translate an English-language questionnaire into another language without losing some of the meaning in translation, we created the Korean version of the AFAQ with adaptation of the cultural nuances in accordance with internationally published recommendations (18,19). The purpose of the present report was to demonstrate the reliability, validity, and responsiveness of the Korean version of the AFAQ.

Patients and Methods

Our hospital's institutional review board approved the present study, and all enrolled patients provided written informed consent. Before beginning the present study, we ensured that the original AFAQ had not already been translated into the Korean language. In addition, we contacted Johanson, the author of the AAOS Lower Limb Instrument (9), and he granted us permission to produce the Korean version of the AFAQ. We performed 2 steps in the present study. The first step was to create the Korean version of the AFAQ by translation and transcultural adaptation. The second step was to examine the reliability, validity, and responsiveness of the Korean version of the AFAQ.

The original AFAQ was translated into Korean by 2 Koreans, an orthopedic surgeon and a nonmedical person. Both translators were native Koreans and fluent in English. They created the Korean version of the AFAQ separately. To reconcile these 2 Korean versions of the AFAQ, the 2 translators and 1 Korean nurse met and discussed the translations. After producing a reconciled Korean version of the AFAQ, 2 bilingual speakers (Korean–English) translated the reconciled Korean version of the AFAQ into English together, without any background information pertaining to the original AFAQ. A committee consisting of 2 back translators and 2 forward translators met to compare the back-translated version of the AFAQ with the original AFAQ. The committee determined that some expressions in the back-translated version of the AFAQ were not equivalent to those of the original AFAQ. To improve these expressions, the reconciled Korean version of the AFAQ was revised. Finally, 20 Korean outpatients from our clinic reviewed the revised Korean version of the AFAQ and were asked whether they had any problems in understanding and answering the revised Korean version. After the pilot study, several ambiguous expressions were modified further. Errors in the revised Korean version of the AFAQ were then corrected, and the Korean version was finalized.

To conduct the second step of our investigation, we recruited patients who had undergone an operation at our foot and ankle clinic from January 2012 to March 2013. The patients who had had a chronic foot and ankle problem for more than 1 year were enrolled in the experimental group, and those with an acute foot and ankle problem that had been present for less than 1 month were enrolled in the control group. After admission, the patients enrolled in our study completed the Korean version of the AFAQ for the first time and the visual analog scale for pain (VAS pain). Twenty-four hours later, they completed the Korean version of the AFAQ for the second time. An orthopedic surgeon investigator then visited the patients in the experimental group and completed the other, previously validated, foot-related quality-of-life scales. If patients had a forefoot problem, the AOFAS Hallux-Metatarsophalangeal-Interphalangeal (Hallux-MTP-IP) scale was completed. If patients had a hindfoot or ankle problem, the AOFAS Ankle-Hindfoot scale was completed. After leaving the hospital, the patients in the experimental group visited our clinics and again completed the Korean version of the AFAQ. The patients in the control group completed the Korean AFAQ 1 time at admission. The patients in the experimental group completed the Korean AFAQ 3 times, at admission, 24 hours after completing the Korean AFAQ at admission, and after the operation.

We assessed the test–retest reliability and internal reliability to understand the reliability of the Korean version of the AFAQ. Test–retest reliability was evaluated for 2 of the questionnaires, which were completed at a 24-hour interval by the patients in the experimental group before surgery. We calculated the intraclass correlation coefficient as a measure of test–retest reliability. Internal reliability was evaluated for the different items in each of the subscales. We calculated Cronbach's α coefficient to understand the internal reliability of the Korean version of the AFAQ.

The criterion validity and discriminant validity were also assessed to understand the validity of the Korean version of the AFAQ. Criterion validity was analyzed using the correlation between the mean standardized score of the Korean version of the GFAS and other validated scales, including the AOFAS Hallux-MTP-IP scale, AOFAS Ankle-Hindfoot scale, and VAS pain scale. We calculated the Pearson correlation coefficients to assess criterion validity. To assess discriminant validity, we used Student's t test to determine the difference in the mean standardized score of the Korean version of the AFAQ between the control and experimental groups.

Responsiveness indicates the sensitiveness of an instrument to detect clinically meaningful changes in the state of patients. Postoperatively, the patients visited our clinic again to check their status. They also completed the Korean version of the AFAQ a second time. We used paired t tests to compare the differences between the preoperative mean standardized score of the Korean version of the GFAS and the postoperative mean standardized score of the Korean version of the GFAS. The Statistical Package for Social Sciences, version 18.0 (SPSS, Chicago, IL), was used for statistical analysis.

Results

A total of 206 patients were included in our study, of whom 152 (73.79%) were enrolled in the experimental group and 54 (26.21%) in the control group. In the experimental group, 152 patients (73.79%) provided us with useable questionnaire information from the first test after enrollment. After a 24-hour interval, of the 152 patients in the experimental group, 144 (94.73% of the experimental group) provided useable questionnaire information from the second test. All 54 patients (26.21%) in the control group provided useable information. All 152 patients in the experimental group provided VAS pain scale data, 38 (25.00% of the experimental group) provided chronic forefoot information by completing the AOFAS Hallux-MTP-IP scale, and 46 (30.26% of the experimental group) provided chronic ankle and/or hindfoot information by completing the AOFAS Ankle-Hindfoot scale. In the experimental group, 29 patients (19.07% of experimental group) visited our clinic again after the operation and completed the Korean version of the AFAQ. The mean average interval from surgery to the follow-up visit when the Korean AFAQ was completed again was 8.2 (range 5.1 to 13.9) months.

Of the 152 patients in the experimental group, 61 were male (40.13% of the experimental group) and 91 were female (59.87% of the experimental group), and their mean age was 49 (range 14 to 81) years. Of the 54 patients in the control group, 31 were male (57.41% of the control group) and 23 were female (42.59% of the control group), and their mean age was 46 (range 17 to 76) years. All the patients enrolled in our study were literate and appeared to understand the questionnaires. A statistical description of the prevalence of the diagnoses by patient group is given in Table 1.

Test–retest reliability was calculated for the 144 patients (69.9%) in the experimental group who had provided us useable information after completing the questionnaire twice. Internal reliability was calculated for the 152 patients (73.79%) in the experimental group who had provided us with useable questionnaire results after completing the questionnaire the first time after enrollment into the study. The mean standardized scores of the subscales and statistical reliability coefficients are presented in Table 2. The test–retest reliability was considered excellent, because the intraclass correlation coefficient ranged from 0.979 to 0.999. Internal reliability was considered moderate to high, because Cronbach's α ranged from 0.528