Clin Orthop Relat Res (2024) 00:1-9
DOI 10.1097/CORR.0000000000003083

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

Clinical Research

The Forgotten Joint Score Is a Valid Outcome Measure for Total Ankle Arthroplasty: A Prospective Study

Antonio Mazzotti MD, PhD[1,2], Simone Ottavio Zielli MD[1,2], Elena Artioli MD[1,2], Michele Astolfi PhD[3], Laura Langone MD[1,2], Cesare Faldini MD, PhD[1,2]

Received: 14 December 2023 / Accepted: 21 March 2024 / Published online: 26 April 2024
Copyright © 2024 by the Association of Bone and Joint Surgeons

Downloaded from http://journals.lww.com/clinorthop by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 04/30/2024

Abstract

Background The success of total ankle arthroplasty (TAA) involves objective and subjective parameters such as joint motion and implant survival. Patient-reported outcome measures provide key context from the patient's perspective so that we can evaluate the results of these procedures. The Forgotten Joint Score (FJS) has demonstrated validity and reliability in other arthroplasties, but the existing evidence on the FJS in TAA is very limited.

Question/purposes We aimed to assess the reliability, validity, responsiveness to change, and floor and ceiling effects, as well as to perform a gender analysis, of the FJS in patients who had undergone TAA.

Methods A prospective, observational study enrolled patients who underwent TAA at our institution between June 2021 and May 2023. The three-component TAA was performed through an anterior approach. A total of 120 patients (mean age 62 ± 10 years) participated, which was 83% (120 of 145) of those eligible, and 77% (92 of 120) of those patients underwent follow-up assessments at 12 ± 1 months after surgery. A cross-culturally adapted and validated version of the FJS in our language (Italian) was used. Additionally, three other commonly used patient-reported outcome measures for TAA were administered: the American Orthopaedic Foot and Ankle Society (AOFAS) ankle-hindfoot score, the Manchester-Oxford Foot Questionnaire (MOXFQ), and the VAS.

Results Reliability was demonstrated by excellent internal consistency (Cronbach α coefficient = 0.95) and excellent test-retest reliability (intraclass correlation coefficient = 0.99; standard error of measurement = 2.7). Robust validity was observed, in that the FJS had a strong correlation with the AOFAS ankle-hindfoot score, the MOXFQ, and the VAS (Pearson and Spearman values consistently above 0.7 or below -0.7). Responsiveness to change was observed between 6 and 12 months (Cohen d = 0.37). Low ceiling and floor effects at both 6 months (4% and 3% reached top and bottom scores, respectively) and 12 months (10% and 0% reached top and bottom scores, respectively) were demonstrated, below the recommended 15% interpretability threshold. No association between outcomes and patient gender was observed.

Conclusion Test properties and validity of the FJS were strong in patients who underwent TAA. It seems suitable for clinical use, although future studies should seek to replicate or refute our findings in other patient populations.

Level of Evidence Level II, therapeutic study.

Each author certifies that there are no funding or commercial associations (consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted article related to the author or any immediate family members.
All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research®* editors and board members are on file with the publication and can be viewed on request.
Ethical approval for this study was obtained from the local committee AVEC (reference number CE AVEC: 576/2021/Oss/IOR.)
This work was performed at Istituto Ortopedico Rizzoli, 1st Orthopedic and Traumatology Clinic, Bologna. Italy.

[1]IRCCS Istituto Ortopedico Rizzoli, 1st Orthopaedics and Traumatologic Clinic, Bologna, Italy

[2]Department of Biomedical and Neuromotor Sciences, Alma Mater Studiorum University of Bologna, Bologna, Italy

[3]Department of Neuroscience and Rehabilitation, University of Ferrara, Ferrara, Italy

S. O. Zielli ✉, IRCCS Istituto Ortopedico Rizzoli, Via Pupilli 1, 40136, Bologna, Italy, Email: simoneottavio.zielli@ior.it

Wolters Kluwer

Copyright © 2024 by the Association of Bone and Joint Surgeons. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/clinorthop by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 04/30/2024

Introduction

The Forgotten Joint Score (FJS) [4] is a widely used patient-reported outcomes tool in hip and knee arthroplasty [3, 21, 34], and it is acknowledged for its robust internal consistency, responsiveness, and validity in the mid to long term. The FJS has several advantages over other established scores [3, 21, 26, 35], including superior ability to differentiate between "very good" and "excellent" outcomes (that is, it does not suffer from a ceiling effect) [4, 21]. Given the increasing prevalence of total ankle arthroplasty (TAA) and the importance of patient-reported outcome measures (PROMs) in assessing the results of this intervention [6], applying the FJS to TAA may be worthwhile. This approach may offer a more thorough evaluation of patients' TAA experiences and may enable comparisons with outcomes from other joint arthroplasties.

However, to date, only one study of which we are aware has assessed the validity of the FJS in patients undergoing TAA [29]. This initial investigation suggested that the FJS can be a valuable instrument for postoperative assessments in patients with endstage ankle arthritis undergoing ankle arthrodesis or TAA. However, the sample size used in that study [29] was below the recommended threshold of 10 patients for each questionnaire item, as advocated by some authors [28, 34]. Additionally, the study's cross-sectional design constrained its ability to explore the responsiveness of the aforementioned questionnaire. Moreover, the study's Japanese population, with an average age of 78 years, may not be broadly generalizable to other regions, such as North America or Western Europe.

Given these limitations, we sought to perform a prospective observational investigation to assess the reliability, validity, responsiveness to change, and floor and ceiling effects, as well as to perform a gender analysis, of the FJS in patients who had undergone TAA.

Patients and Methods

Study Design and Setting

We performed a prospective, observational study to evaluate the reliability, validity, responsiveness to change, and floor and ceiling effects, and also, we performed a gender analysis of the FJS for patients who underwent TAA at our institution between June 2021 and May 2023. We used a validated cross-cultural version of the FJS in our language (Italian) according to the translation-retrotranslation method [34].

This study complied with the comprehensive COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist protocol [28].

Patients

Between June 2021 and May 2024, we treated 145 patients with TAA for endstage ankle arthritis. Of those, we considered all patients who were native Italian-language speakers and who had an active email address. We excluded those with other lower-limb joint prostheses, those with a history of major cardiac events, patients who did not wish to participate, those who were unable to complete the questionnaires independently, those with dementia, and minors. Based on these criteria, 94% (136 of 145) were eligible; a further 5% (7 of 145) of patients were excluded because they had other lower-limb prostheses, because they did not wish to participate, or because they were unable to complete the questionnaires independently. Another 6% (9 of 145) had incomplete datasets, which we defined as a questionnaire with fewer than 50% of the items completed, leaving 83% (120) for analysis. Of those, 77% (92 of 120) responded to both questionnaires at 6 and 12 months postoperatively. Each patient gave their informed consent before they were enrolled in the study.

Descriptive Data

Of the included patients, 60% (72 of 120) were men. Their mean age was 62 ± 10 years, and follow-up was conducted between 12 and 13 months (mean 12 ± 1 months) after surgery. A three-component TAA featuring a nitride-coated titanium surface and a polyethylene mobile insert was used. The surgical procedures, all performed by two surgeons (AM, CF), employed an anterior approach using the interval between the extensor hallucis longus and the anterior tibialis tendon.

Description of Experiment

An online questionnaire was sent via email to all patients who underwent TAA between June 2021 and May 2023, with responses collected at 6 and 12 months postsurgery. The minimum surveillance period was 6 months. Patients who did not return the questionnaire within 2 weeks received a reminder telephone call, and those who did not respond within another month were definitively excluded from the study (with recorded reasons if provided). An orthopaedic surgeon (AM) was available to respond to any questions from patients.

Data were independently collected by two investigators (SOZ, EA). Each patient was deidentified and assigned a unique study number, and the results were recorded in an Excel document.

Copyright © 2024 by the Association of Bone and Joint Surgeons. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/clinorthop by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 04/30/2024

Outcome Measures

In addition to the FJS questionnaire, three of the most commonly used PROMs for TAA were administered to the patients: the American Orthopaedic Foot and Ankle Society (AOFAS) ankle-hindfoot scale, the Manchester-Oxford Foot Questionnaire (MOXFQ), and the VAS [32].

FJS

The FJS was first described by Behrend et al [4]. The original FJS has 12 items that measure patients' ability to forget the presence of an artificial joint in their daily life, graded on a 5-point Likert scale. The raw results (0 to 60 points) are then converted to a 0-point to 100-point scale. The higher the score, the better the outcome, corresponding to a lower perception of the prosthesis by patients.

AOFAS Ankle-Hindfoot Scale

The AOFAS ankle-hindfoot scale is a semiobjective scale that evaluates pain, function, and alignment of the ankle and hindfoot. Scores range from 0 to 100, with painless, perfectly functioning ankles receiving 100 points. The version we used is the one translated to and adapted for the Italian language by Leigheb et al. [23].

MOXFQ

The MOXFQ consists of a patient self-assessment questionnaire with 16 items, evaluating the results of foot and ankle operations in three areas: pain, ability to stand and walk, and limitations to social interaction. Raw scores are convertible to a 0 to 100 metric, where 100 indicates the most severe impairment. The version we used was translated to and adapted for the Italian language by Marinozzi et al. [25].

VAS

The VAS in medicine is a subjective assessment tool that is used to measure the intensity or magnitude of the patient's perception of various factors, such as pain [9]. It typically consists of a straight line, usually 10 cm, with anchor points at the ends, representing the absence or minimum intensity and the maximum intensity of the measured experience, respectively. Patients are invited to draw a point along the VAS line corresponding to their perception of the experience, and the resulting distance between the minimum intensity anchor point and the point marked by the patient is then measured in mm or cm; these lengths represent the intensity felt by the patients.

Scores Achieved

At the 6-month follow-up, the FJS score was 53 ± 31, the AOFAS ankle-hindfoot scale was 65 ± 22, the MOXFQ was 36 ± 29, and the VAS was 4.1 ± 2.5. These results are comparable to those of other studies at the 6-month time-point after TAA [18, 31].

Study Size

Based on the guidelines for health status questionnaires, which recommend at least 10 patients per questionnaire item [38], a minimum of 120 patients was established for our study. To evaluate test-retest reliability, only the first 60 patients were involved, as performed in a similar study [34]. To minimize the risk of changes related to clinical evolution, only patients with follow-up of at least 12 months were included for the test-retest reliability analysis. These patients completed the FJS questionnaire twice, with 2 to 6 weeks between the two questionnaires.

Ethical Approval

The study protocol was approved by our local ethics committee.

Statistical Analysis

We evaluated four parameters: reliability, validity, responsiveness to change, and floor and ceiling effects. Furthermore, a gender analysis was conducted to identify any differences between the two genders.

Reliability

Reliability reflects the extent to which measurements are free from error [38]. To assess reliability, we calculated internal consistency, test-retest reliability, and measurement error.

Internal consistency assesses the interrelatedness among questionnaire items [28]. A Cronbach α higher than 0.7 indicate a good internal consistency [38].

Test-retest reliability measures the repeatability of results over time in patients with a static health status [28]. According to the COSMIN checklist recommendations, retests were performed 2 weeks after the 12-month

Copyright © 2024 by the Association of Bone and Joint Surgeons. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/clinorthop by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 04/30/2024

postoperative test to ensure that no important changes in the patient's health had occurred. Test-retest reliability was determined using intraclass correlation coefficients (ICC), where an ICC greater than 0.7 indicates good reliability [38]. To ensure test-retest reliability and responsiveness, patients received the questionnaire 2 weeks after the initial test, with reminders every 2 weeks if needed.

Measurement error quantifies systematic and random errors in a patient's score that are unrelated to true changes in the construct being measured [28]. The standard error of measurement (SEM) was calculated using the formula $SD \cdot \sqrt{(1\text{-ICC})}$, where SD is the standard deviation [38]. The smallest detectable change denotes the smallest individual score change interpretable as a real change and was computed as: $SEM \cdot 1.96 \cdot \sqrt{2}$ [38].

Validity

Testing validity assesses whether a test accurately measures what it purports to measure. Given the absence of a gold standard for measuring patient-reported outcomes, we assessed the validity of the FJS through convergent construct and clinical validity. The convergent construct validity measures the comparability of FJS scores against other functional questionnaires: the AOFAS, MOXFQ, and VAS. We evaluated convergent validity with the Spearmen rank and Pearson correlation coefficient r; an r value greater than 0.7 indicates a strong correlation, an r value between 0.3 and 0.7 indicates a moderate correlation, and an r value below 0.3 means poor correlation [28].

Responsiveness to Change

Responsiveness to change can be defined as the capability of a test or questionnaire to detect clinically relevant variations over time [28]. In particular, it represents a PROM's ability to respond according to changes in the patient's clinical status. It is calculated by measuring the change between the 6-month and 12-month follow-up FJS score. This change is reported as an effect size for the mean change in terms of the Cohen d. Furthermore, the smallest detectable change is used as a secondary measure of responsiveness, that is, if the smallest detectable change is smaller than the minimum important change (MIC), a positive rating for responsiveness is attributed [38]. The MIC is computed as half of the SD [30].

Floor and Ceiling Effects

The floor and ceiling effect pertains to the ability of a questionnaire to effectively capture and differentiate responses at the extremes of its scale, particularly when scores reach the minimum (0) or maximum (100) values [19, 28]. When FJS questionnaires produce scores at the extremes of the scale, it may not accurately depict the unique conditions of patients, leading to the perception of seemingly identical situations. According to the COSMIN checklist, a satisfactory level of interpretability and content validity is attained when the combined percentage of patients scoring the lowest (0) or the highest (100) on the scale is less than 15% [19, 28].

Gender Analysis

The study cohort contained 72 men and 48 women. A subgroup gender analysis was conducted. An unpaired two-sample t-test was performed to identify any potential differences in outcomes between men and women undergoing TAA.

Results

Reliability

In terms of internal consistency, the Cronbach α coefficient was 0.95, which exceeded the 0.7 threshold, suggesting good data consistency and reliability. Test-retest reliability was excellent based on the ICC (two-way random effect) for the 12-month scores and the scores reported 2 weeks later (Fig. 1). Interrater reliability (at 0.99) was also excellent. The SEM was equal to 2.7, and the smallest detectable change was determined to be 7.5.

Validity

Convergent validity analysis consistently yielded values above 0.7 or below -0.7, demonstrating a good positive or negative correlation, respectively (Fig. 2). In particular, we observed a positive correlation between the FJS and the AOFAS (Pearson 0.88, Spearman 0.90), whereas the correlation between the FJS and the MOXFQ (Pearson -0.86, Spearman -0.86), and FJS and VAS (Pearson -0.79, Spearman -0.80) was negative. This negative correlation arises from the MOXFQ's design, where better clinical outcomes are indicated by lower scores [25].

Responsiveness to Change

We found a mild-to-moderate responsiveness of the FJS to changes in the patient's health status (Cohen d = 0.37; $p < 0.001$).

Copyright © 2024 by the Association of Bone and Joint Surgeons. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/clinorthop by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 04/30/2024

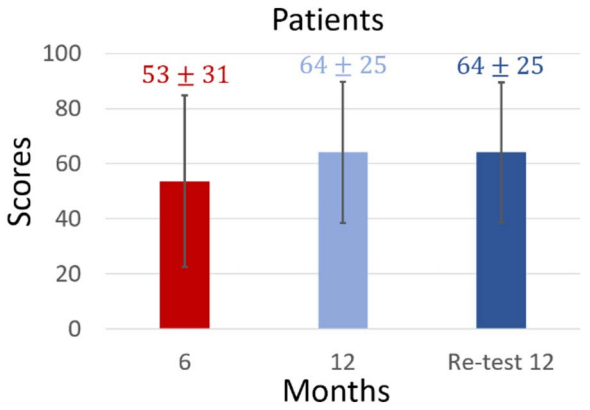


Fig. 1 The results of the FJS at 6 and 12 months after surgery are shown here, along with retest findings at 12 months.

Floor and Ceiling Effects

We observed low ceiling and floor effects for both the 6-month (4% and 3%, respectively) and 12-month (10% and 0%, respectively) follow-up questionnaires (Fig. 3). Although the ceiling effect percentage increased at the 12-month assessment, it remained below the recommended 15% interpretability threshold, whereas the floor effect decreased as expected.

Gender Analysis

We found no differences between the mean scores of men and those of women, suggesting no relationship between gender and distribution of FJS scores (the mean ± FJS for men was 64 ± 26, compared with 65 ± 23 for women; p = 0.91).

Discussion

After recovery from arthroplasty, patients should not perceive the joint as artificial during activities of daily living [15]; that is, we hope that our patients will forget they have a TAA. However, foot and ankle surgeons acknowledge that this often is not the case, because some patients may experience joint awareness, including pain, functional limitations, reduced sports activities, and additional

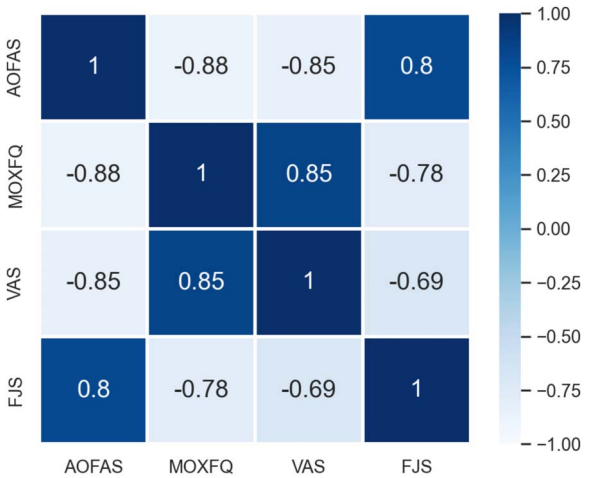


Fig. 2 This heatmap shows the r-value (correlation coefficient) of the FJS compared with the AOFAS, MOXFQ, and VAS.

Copyright © 2024 by the Association of Bone and Joint Surgeons. Unauthorized reproduction of this article is prohibited.

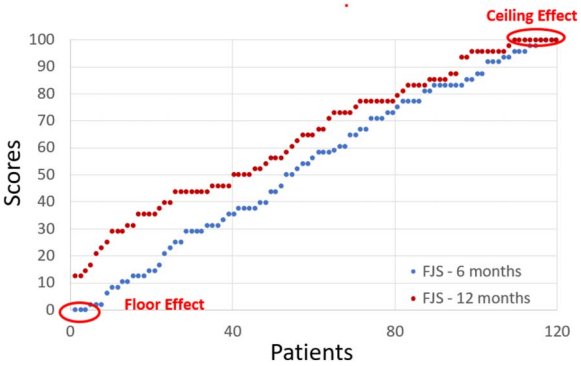


Fig. 3 This graph demonstrates the distribution of the FJS results at 6 months (blue dots) and 12 months (red dots) postsurgery, encompassing both minimum (floor) and maximum (ceiling) scores.

sensations of discomfort [24]. The FJS, introduced in 2012, is designed to measure a patient's ability to forget about an implanted joint [4]. Widely accepted in clinical practice and research for hip, knee, and shoulder arthroplasties, where the FJS has been validated, the FJS offers value to patients in those clinical contexts because of its low ceiling effect. That is, it effectively distinguishes among patients with good, very good, and excellent outcomes [7]. Despite these strengths, only one study we know of has assessed the validity of the FJS in patients undergoing TAA, and it had a number of limitations [29]. We therefore sought to prospectively evaluate the reliability, validity, responsiveness to change, and floor and ceiling effects, and performed a gender analysis of the FJS to subjectively assess how much patients perceive their TAA.

Limitations

This study relied on questionnaires, and the absence of objective measurements such as ROM or strength assessments limited the comprehensiveness of our evaluation of convergent validity. However, because so-called objective measurements are often not as important to patients as their perceptions of the results of surgery, we do not believe this is a disqualifying problem. The sample was relatively small, rendering it unsuitable for principal component factor analysis and other statistical methods that could have offered a more detailed assessment of dimensionality. Nonetheless, the case series described in this study is more than twice as large as the only comparable study of which we are aware [29], which also was not prospective. Despite a mean follow-up duration of 12 months, not all patients were present for the 12-month visit. Specifically, only 77% of the enrolled patients completed the 12-month questionnaire. We did not intend to conduct a clinical evaluation, which is typically recommended at a minimum interval of 24 months. Instead, the focus was to validate a new questionnaire in an anatomic area investigated solely by another study [29] and to compare the questionnaire to other widely used scores. Nine patients could not be analyzed because of incomplete datasets. An incomplete questionnaire might indicate an ineffective test; however, in 5 of 9 patients excluded for this reason, the FJS questionnaire had been completed, while the other questionnaires, which were administered subsequently, were missing. Given the inability to concretely validate the test's validity, these patients were excluded. As for the other four excluded patients, they did not answer any questions on the test and simply returned the received form, and thus were not considered. Only a small percentage of patients did not respond to question 12 of the FJS, which asks, "Are you aware of your ankle joint when you are doing your favorite sport or hobby?" As noted in a recent systematic review [1], TAA typically does not result in patients returning to high-level sports, especially those involving contact and repeated mechanical stresses. The introduction of the term "hobby" prompted patients to contemplate a wider array of recreational activities, and these pursuits should be given greater consideration in evaluating patients' overall quality of life. In the only other study of which we are aware that has used the FJS in patients who have undergone TAA [29], this specific question focused exclusively on sports activities, and it registered the highest percentage of missing responses (40%).

In addition, we used a validated cross-cultural version in our language instead of administering the questionnaire in the original language. The aim of a questionnaire should be to assess intervention outcomes independently of the geographic and cultural environment in which the patient is

Copyright © 2024 by the Association of Bone and Joint Surgeons. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/clinorthop by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 04/30/2024

Downloaded from http://journals.lww.com/clinorthop by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 04/30/2024

situated. Nevertheless, achieving scores that are intersubjectively valid, irrespective of the social and cultural nuances of each country, proves challenging. Therefore, researchers continuously strive toward the ongoing goal of establishing these scores as universally valid. SCORES clash with reality; thus, making these scores universally valid should be a constant goal for researchers. Questions such as "Are you aware of your ankle joint when you are standing up from a low-sitting position" can be interpreted very differently. Patients in Western cultures may understand this as getting up from a chair, armchair, or sofa. By contrast, in some Asian societies, sitting on the ground with crossed legs plays an important role in everyday life. Therefore, the same question might be interpreted as "getting up from the ground," and so this question may be interpreted very differently between patients in Asia and those in Europe, the United States, or Australia. In this way, our study is an extension of the only previous study of the FJS in TAA [29], which was performed in Japan. We believe our study may align more with the original intent of the authors who first developed the FJS, but evaluations of the FJS both in Western and Asian cultures of course are equally important.

One of the steadily growing trends in the field of PROMs is computerized adaptive testing [12, 13]. Despite the potential these systems hold, we believe that their limitations outweigh their benefits, particularly in the context of current clinical practice. A crucial challenge stems from the lack of standardized questions, introducing complexities when comparing results across diverse studies or analyzing trends over time [13]. Furthermore, the substantial demand for continuous validation necessitates an ongoing process to ensure the reliability and validity of measurements [11]. Additionally, we have concerns about potential variability in patient responses to the computerized format, especially among older patients who often undergo TAA and who may not be facile with computer-based tools.

Reliability

The FJS exhibited comparable values not only to the only previous study of TAA [29] but also in other anatomic areas [34]. The results of the 12-month test-retest further confirmed there was excellent interrater reliability, aligning with what has been demonstrated by others [34, 35].

Validity

The FJS displayed the strongest correlation coefficient with the AOFAS ankle-hindfoot scale, which is widely used in assessing results after TAA [2, 5, 7, 8, 17, 20], but which has a number of shortcomings that have been documented elsewhere [10, 14, 37]. Furthermore, the FJS's validity was reaffirmed by a high correlation with the MOXFQ, another frequently used questionnaire in evaluating TAA [36, 39]. Conversely, the weakest correlation was observed with the VAS. Although lower, this coefficient remains substantially above the 0.7 threshold, confirming there was a robust correlation between the two scores. Nevertheless, this finding can be perceived as an advantageous aspect of the FJS. The VAS is a unidimensional score that exclusively evaluates patient-reported pain, usually with one activity or at rest, potentially failing to encompass the impact of TAA on daily activities and restrictions more broadly [2]. Consequently, it is not as comprehensive a reflection of the prosthesis' influence on the patient's overall pain and function.

Responsiveness to Change

We found that the FJS showed only a mild-to-moderate degree of responsiveness to change in a patient's clinical condition between 6 and 12 months postoperatively. This observation may be attributed to the relatively extended recovery period typically associated with TAA, often resulting in postoperative ankle immobilization for many weeks [27]. This postoperative protocol can often lead to a delay in complete functional recovery, resulting in the small observed difference between the 6-month and 12-month postoperative assessments. Notably, the FJS was responsive to change when applied to TAA, with the observed MIC being larger than the smallest detectable change in our study. Although this is the first study we know of investigating the responsiveness to change in the FJS in TAA, this outcome aligns with comparable findings observed when applying the FJS to TKA [16].

Floor and Ceiling Effects

Consistent with prior studies, the FJS demonstrated low floor and ceiling effects, with the latter consistently maintained below 10% [4, 21]. The other available study in TAA likewise highlighted a total ceiling and floor effect of 8% [29], and we have confidence in our main finding on this point. Moreover, when we compared the ceiling and floor effect prevalence across the various scores used in this study, the FJS consistently displayed the lowest percentages. This noteworthy characteristic highlights the FJS's ability to effectively differentiate between patients' results [33].

Gender Analysis

Men's and women's scores on the FJS did not differ in this study, consistent with findings reported by others [22].

Copyright © 2024 by the Association of Bone and Joint Surgeons. Unauthorized reproduction of this article is prohibited.

Currently, there is no apparent justification for using distinct criteria to assess ankle prostheses between patients of different genders; however, more-detailed studies on the needs and goals of patients of different genders ought to be performed to provide better information on this important topic.

Downloaded from http://journals.lww.com/clinorthop by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 04/30/2024

Conclusion

The FJS demonstrated strong internal consistency and a low floor and ceiling effect, effectively differentiating between patients with good outcomes and those with excellent outcomes. Its robust validity, indicated by strong correlations with established measures such as the AOFAS and MOXFQ, supports its utility in patients undergoing TAA. No differences between men and women were observed in the FJS. Our findings suggest that the FJS will enrich surgeons' understandings about patients' experiences after TAA. Further studies should prospectively evaluate medium- to long-term TAA clinical outcomes to confirm the responsiveness of the FJS. We believe the FJS would be most helpful alongside other well-established scores in evaluating patients undergoing TAA in ways that leverage its distinctive traits, in particular low floor and ceiling effects and adaptability across diverse cultures and populations.

Acknowledgment We thank Dr. Johannes Giesinger and all licensors of the Forgotten Joint Score-12.

References

1. Arceri A, Mazzotti A, Zielli S, et al. Return to sports after total ankle arthroplasty: a systematic review and meta-analysis. *J Orthop*. 2023;44:57-65.
2. Bagheri K, Anastasio AT, Poehlein E, et al. Outcomes after total ankle arthroplasty with an average follow-up of 10 years: a systematic review and meta-analysis. *Foot Ankle Surg*. 2023;30:64-73.
3. Baumann F, Ernstberger T, Loibl M, Zeman F, Nerlich M, Tibesku C. Validation of the German Forgotten Joint Score (G-FJS) according to the COSMIN checklist: does a reduction in joint awareness indicate clinical improvement after arthroplasty of the knee? *Arch Orthop Trauma Surg*. 2016;136:257-264.
4. Behrend H, Giesinger K, Giesinger JM, Kuster MS. The "forgotten joint" as the ultimate goal in joint arthroplasty: validation of a new patient-reported outcome measure. *J Arthroplasty*. 2012;27:430-436.e1.
5. Bianchi A, Martinelli N, Caboni E, Raggi G, Manfroni F, Sansone V. Long-term follow-up of Bologna-Oxford (BOX) total ankle arthroplasty. *Int Orthop*. 2021;45:1223-1231.
6. Brown TD, Johnston RC, Saltzman CL, Marsh JL, Buckwalter JA. Posttraumatic osteoarthritis: a first estimate of incidence, prevalence, and burden of disease. *J Orthop Trauma*. 2006;20:739-744.
7. Brunner S, Barg A, Knupp M, et al. The Scandinavian total ankle replacement: long-term, eleven to fifteen-year, survivorship analysis of the prosthesis in seventy-two consecutive patients. *J Bone Joint Surg Am*. 2013;95:711-718.
8. Clough T, Bodo K, Majeed H, Davenport J, Karski M. Survivorship and long-term outcome of a consecutive series of 200 Scandinavian Total Ankle Replacement (STAR) implants. *Bone Joint J*. 2019;101:47-54.
9. Delgado DA, Lambert BS, Boutris N, et al. Validation of digital visual analog scale pain scoring with a traditional paper-based visual analog scale in adults. *J Am Acad Orthop Surg Glob Res Rev*. 2018;2:e088.
10. D'souza JJ. American Orthopedic Foot-and-Ankle Society score, where are we now? – a narrative review of quality of life measures in foot-and-ankle surgery. *Open Access Maced J Med Sci*. 2020;8:133-136.
11. Gamper E-M, Martini C, Petersen MA, Virgolini I, Holzner B, Giesinger JM. Do patients consider computer-adaptive measures more appropriate than static questionnaires? *J Patient Rep Outcomes*. 2019;3:7.
12. Gausden EB, Levack A, Nwachukwu BU, Sin D, Wellman DS, Lorich DG. Computerized adaptive testing for patient reported outcomes in ankle fracture surgery. *Foot Ankle Int*. 2018;39:1192-1198.
13. Gershon RC, Rothrock N, Hanrahan R, Bass M, Cella D. The use of PROMIS and assessment center to deliver patient-reported outcome measures in clinical research. *J Appl Meas.* 2010;11:304-314.
14. Guyton GP. Theoretical limitations of the AOFAS scoring systems: an analysis using Monte Carlo modeling. *Foot Ankle Int*. 2001;22:779-787.
15. Hudak PL, McKeever PD, Wright JG. Understanding the meaning of satisfaction with treatment outcome. *Med Care*. 2004;42:718-725.
16. Ingelsrud LH, Roos EM, Terluin B, Gromov K, Husted H, Troelsen A. Minimal important change values for the Oxford Knee Score and the Forgotten Joint Score at 1 year after total knee replacement. *Acta Orthop*. 2018;89:541-547.
17. Jastifer JR, Coughlin MJ. Long-term follow-up of mobile bearing total ankle arthroplasty in the United States. *Foot Ankle Int*. 2015;36:143-150.
18. Jeyaseelan L, Si-Hyeong Park S, Al-Rumaih H, et al. Outcomes following total ankle arthroplasty: a review of the registry data and current literature. *Orthop Clin North Am.* 2019;50:539-548.
19. Kulthanan K, Chularojanamontri L, Tuchinda P, Rujitharanawong C, Maurer M, Weller K. Validity, reliability and interpretability of the Thai version of the urticaria control test (UCT). *Health Qual Life Outcomes*. 2016;14:61.
20. Kvarda P, Peterhans U-S, Susdorf R, Barg A, Ruiz R, Hintermann B. Long-term survival of HINTEGRA total ankle replacement in 683 patients: a concise 20-year follow-up of a previous report. *J Bone Joint Surg Am*. 2022;104:881-888.
21. Larsson A, Rolfson O, Kärrholm J. Evaluation of Forgotten Joint Score in total hip arthroplasty with Oxford Hip Score as reference standard. *Acta Orthop*. 2019;90:253-257.
22. Lee G-W, Lee K-B. Does gender influence the outcomes of total ankle arthroplasty in patients with ankle osteoarthritis? *J Orthop Surg*. 2020;15:207.
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;87:38-45.
24. Lewis S, Price M, Dwyer KA, et al. Development of a scale to assess performance following primary total knee arthroplasty. *Value Health*. 2014;17:350-359.

Copyright © 2024 by the Association of Bone and Joint Surgeons. Unauthorized reproduction of this article is prohibited.

Downloaded from http://journals.lww.com/clinorthop by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3i3D0OdRyi7TvSFl4Cf3VC4/OAVpDDa8KKGKV0Ymy+78= on 04/30/2024

25. Marinozzi A, Martinelli N, Panascì M, et al. Italian translation of the Manchester-Oxford Foot Questionnaire, with re-assessment of reliability and validity. *Qual Life Res*. 2009;18:923-927.
26. Matsumoto M, Baba T, Homma Y, et al. Validation study of the Forgotten Joint Score-12 as a universal patient-reported outcome measure. *Eur J Orthop Surg Traumatol*. 2015;25:1141-1145.
27. Mazzotti A, Viglione V, Gerardi S, et al. Post-operative management after total ankle arthroplasty: a systematic review of the literature. *Foot Ankle Surg*. 2022;28:535-542.
28. Mokkink LB, Prinsen CAC, Bouter LM, de Vet HCW, Terwee CB. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Braz J Phys Ther*. 2016;20:105-113.
29. Noguchi K, Yamaguchi S, Teramoto A, et al. Reliability and validity of the Forgotten Joint Score-12 for total ankle replacement and ankle arthrodesis. *PLoS One*. 2023;18:e0286762.
30. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care*. 2003;41:582-592.
31. Onggo JR, Nambiar M, Phan K, Hickey B, Galvin M, Bedi H. Outcome after total ankle arthroplasty with a minimum of five years follow-up: a systematic review and meta-analysis. *Foot Ankle Surg*. 2020;26:556-563.
32. Parker J, Nester CJ, Long AF, Barrie J. The problem with measuring patient perceptions of outcome with existing outcome measures in foot and ankle surgery. *Foot Ankle Int*. 2003;24:56-60.
33. Pena F, Agel J, Coetzee JC. Comparison of the MFA to the AOFAS outcome tool in a population undergoing total ankle replacement. *Foot Ankle Int*. 2007;28:788-793.
34. Sansone V, Fennema P, Applefield RC, et al. Translation, cross-cultural adaptation, and validation of the Italian language Forgotten Joint Score-12 (FJS-12) as an outcome measure for total knee arthroplasty in an Italian population. *BMC Musculoskelet Disord*. 2020;21:23.
35. Shadid MB, Vinken NS, Marting LN, Wolterbeek N. The Dutch version of the Forgotten Joint Score: test-retesting reliability and validation. *Acta Orthop Belg*. 2016;82:112-118.
36. Smith TO, Dainty J, Loveday DT, et al. Outcomes in Ankle Replacement Study (OARS). *Bone Joint J*. 2023;105:895-904.
37. SooHoo NF, Vyas R, Samimi D. Responsiveness of the foot function index, AOFAS clinical rating systems, and SF-36 after foot and ankle surgery. *Foot Ankle Int*. 2006;27:930-934.
38. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34-42.
39. Yau J, Emmerson B, Kakwani R, Murty AN, Townshend DN. Patient-reported outcomes in total ankle arthroplasty: patient specific versus standard instrumentation. *Foot Ankle Spec*. 2024;17(1 suppl):30S-37S.

Copyright © 2024 by the Association of Bone and Joint Surgeons. Unauthorized reproduction of this article is prohibited.