

RESEARCH

Open Access



# Translation, cross-cultural adaptation and validation of psychometric properties of foot function index in Urdu-speaking population with ankle and foot disorders

Aqsa Anjum<sup>1</sup>, Sana Tauqeer<sup>1\*</sup>, Ayesha Arooj<sup>1</sup>, Hassan Javed<sup>1,3</sup>, Hammad Shakeel<sup>1</sup> and Ammarah Ikram<sup>2</sup>

## Abstract

**Background** Foot and ankle problems are frequently prevalent, especially among the elderly, ranging from 70 to 80%. In primary care, foot, and ankle complaints stand out as one of the most frequent reasons for orthopedic consultations. Patient-reported outcome measures are significant in the assessment burden of any condition on the effects of intervention as well as research. The Foot Function Index (FFI) is a region-specific tool that was identified as one of the most commonly used evaluation tools for foot complaints. This study aimed to translate, cross-cultural adapt, and test the psychometric properties of FFI in the Urdu language.

**Methodology** The FFI was translated into Urdu language (FFI-U) following Beaton et al. translation guidelines. The data were collected from 230 Urdu-speaking participants with different foot and ankle pathologies. Data collection started after the written informed consent from the participants. All participants completed the FFI-U, Visual Analogue Scale (VAS), SF-36, and the Foot and Ankle Outcome Score (FAOS) at baseline while only 30 participants completed ULFI-U after one week for test-retest reliability. The psychometric properties involved reliability and validity testing. Reliability was assessed where internal consistency was measured using Cronbach's alpha and test-retest reliability through the intra-class correlation coefficient (ICC). FFI-U was tested for face validity and construct validity (convergent and discriminant). Psychometric criteria were examined against priori hypotheses, and alpha level (p-value < 0.05) was considered statistically significant.

**Results** FFI-U demonstrated good reliability with internal consistency ( $\alpha=0.86$ ) and test-retest reliability with intra-class correlation coefficient = 0.845 (0.78–0.89). A moderate correlation was found using Pearson correlation between FFI-U total score and physical components of SF-36, VAS (pain, disability), and FAOS ( $\gamma=-0.65, 0.72, 0.71, -0.68$ ) respectively, indicating convergent validity however, a weak correlation was found with mental components of SF-36 ( $\gamma=-0.25$ ) demonstrating discriminant validity. Face validity was assessed at the pre-final testing stage by interviewing patients. There were no floor and ceiling effects found for FFI-U.

\*Correspondence:

Sana Tauqeer  
sana.tauqeer@uipt.uol.edu.pk

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

**Conclusion** The FFI-U has been found reliable, valid, and feasible tool to be used as a patient-reported outcome measure to assess functional levels with different foot and ankle disorders in Urdu speaking population.

**Keywords** Foot function index, Reliability, Validity

## Introduction

The foot represents a complex musculoskeletal structure, with common patient complaints revolving around related disorders such as pain and deformities [1]. The prevalence of foot pain rises with age and obesity (Body Mass Index > 30), with higher occurrence among females [1, 2]. The higher occurrence of foot disorders in females may include several factors such as improper footwear [3], biomechanical variations such as wider pelvis and lower muscle mass may cause over-pronation of foot [4], hormonal changes, and lifestyle or occupational demands [5]. Even after adjusting for gender, age, and BMI, individuals experiencing foot pain tend to score lower on health-related quality-of-life measures [2]. Moreover, depending on the activity and degree of competition, the foot accounted for up to 20% of all sports injuries [6]. Disorders linked to the foot substantially curtail patients' activities and exacerbate their quality of life, it is underscored that evaluating foot functions is pivotal for gauging progress in both conservative and surgical treatments [7]. Foot disorders are widespread, affecting around 15–25% of the adult population with resultant foot pain impacting gait, balance, and daily activities [8].

The foot pain usually limits locomotion, impairs balance, and compromises functional activities of daily living. Researchers and medical practitioners use self-reported outcome instruments to assess the conditions of the ankle and foot as an assessment of treatment [9]. These tools enable the use of reliable patient perception measurements, and certain instruments have been standardized to monitor, evaluate, and assess the results of a particular intervention [10]. The self-reported questionnaires have proven to be an effective means to monitor outcome measures across various medical fields [11]. Direct information on the patient's perceived health can be obtained through the use of patient-reported outcome measures (PROMs) [10]. Healthcare systems that prioritize the needs of their patients are gaining popularity, and PROMs are crucial instruments for gathering patient feedback regarding the burden and influence of their health conditions [12]. PROMs, and standardized and validated questionnaires, are crucial in clinical practice and research [12, 13]. These tools help determine patient needs, clinical status, and treatment goals, enhance clinician-patient communication, and evaluate a patient's functional status and well-being [9].

Foot health is measured using a variety of PROMs and the Foot Function Index (FFI) is one example of a region-specific PROM; other PROMs are disease-specific [14].

According to a review conducted in 2020 to identify different outcome measurement tools for foot and ankle, 9% of described studies utilized FFI [15]. According to another systematic review in 2008, the FFI stands out as one of the most commonly employed assessment tools for foot-related problems [16]. This questionnaire is known for its feasibility, clarity, participant understanding, and short completion time [17]. The FFI was formulated by Budiman-Mak, Conrad, and Roach in 1991 [14]. In 2006, it was revised into a long and short version (FFI-RL & FFI-RS) by adding psychosocial components [18]. However, all three versions are under use [19]. A review conducted by developers of FFI in 2013 to evaluate between original FFI and FFI revised version where it was documented that 51 studies out of 78 studies have employed original FFI with three subscales [19]. The FFI was formulated to measure the impact of foot health in the fields of podiatry, rheumatology, and orthopedic medicine [19] on activities of daily living and functions in terms of three components of the questionnaire i.e. pain, disability, and activity restriction. These three components act as subscales of questionnaires in which pain (9 items), activity limitation (9 items), and disability (5 items) [14]. According to the guidelines of the International Classification of Function (ICF), these three areas correspond well with patients' responses to foot problems [20].

The original FFI has been translated into fifteen languages i.e. Saudi Arabic [21], Italian [22], Thai [23], Egyptian Arabic [24], French [25], German [26], Spanish [27], Brazilian Portuguese [28], Chinese [29], Danish [30], Korean [31], Persian [32], Dutch [33], Taiwan Chinese [34], and Gujrati [35], whereas FFI-RL has been translated into two languages i.e. Brazilian Portuguese [36] and Turkish [37] and FFI-RS has been translated to Polish [38] and Norwegian [39] languages. All the studies have shown good reliability and validity for FFI [21–39]. Since standardization is crucial when employing assessment tools, questionnaires developed in other languages are required to have their psychometric qualities assessed and translated to establish study equivalency [40]. The present study aims to translate and cross-culturally adapt the original FFI into Urdu language. There is no translation of the original FFI into the Urdu language to date, which is the foremost reason to translate the original version in this study. FFI is an easy-to-use PROM that would help the patients to understand and respond accordingly; hence, an Urdu version for the Urdu-speaking population would be beneficial. Henceforth, it was required to translate and cross-culturally adapt the FFI into Urdu

language to improve comprehension for the Urdu-speaking population.

## Methods

### Study design

The study was designed as a cross-sectional clinimetric study. The formal permission to translate and cross-culturally adapt the original FFI into the Urdu language was obtained from the original developers of the scale. The STROBE guidelines were used to report the study [41].

### Ethical approval

The study protocol was approved by the ethical review board of the University of Lahore called Research Ethics Committee (REC) with reference number REC-UOL-408-05-2023 dated 27/03/2023. The study was completed in one year in which data were collected from May 2023 to December 2023. A written informed consent was obtained from every participant before data collection.

### Setting

The data were collected from two-hundred-and-thirty Urdu-speaking participants with different foot problems. The recruitment of participants was made through The Primary Podiatric Clinic, Lahore, and the musculoskeletal outpatient department of the physiotherapy department at The University of Lahore Teaching Hospital, Lahore.

### Sample size

According to standards for the respondent-to-item ratio [42], which ranges from 10:1 (230 respondents for a 23-item questionnaire), this study recruited 230 participants who had been diagnosed with foot problems discussed below in the eligibility criteria.

### Eligibility criteria

Male and female participants aged  $\geq 18$  years with diagnosis of foot problems from distal to the talocrural joint such as planter medial neuropathia causing heel pain, heel pain due to an increase in weight-bearing activities, plantar fasciitis, positive windlass test, metatarsalgia such as pain in metatarsals or positive Mulder sign [43], presenting complaint of ankle sprain or history of minimum one ankle sprain [44], inflammatory symptoms such as pain or swelling, and finally, the interruption in physical activity due to sprain for a minimum of one day [44]. Only Urdu-speaking participants were included who were able to read and understand the Urdu language. Participants were excluded from any neurological disorders, foot pain as a result of neurological/neuromuscular disorder, or foot deformities associated with any lower limb injuries.

### Instruments

Three assessment tools in addition to FFI were used i.e. Short-form health survey (SF-36), Foot and Ankle Outcome Score (FAOS), and Visual Analogue Scale (VAS) were utilized for assessing reliability and validity.

### Foot Function Index

The Foot Function Index (FFI) is a self-reporting tool that evaluates various aspects of foot function based on patient-centered values [14]. Good reliability has been found for FFI with test-retest reliability as 0.87–0.69 while internal consistency ranged from 0.96–0.73 and an acceptable structural validity [45]. The FFI has three sub-scales i.e. pain (9 items), disability (9 items), and activity limitation (5 items), with a total of 23 items. The scoring for FFI is to obtain the score for each sub-scale first, the sum of all the items that the patient has responded to divided by the possible score for the sub-scale and multiplied by 100 [14]. The item is deemed not applicable (NA) if the patient does not carry out the activity specified by one of the subscale items (such as not employing orthotic devices). The final score for the questionnaire will be calculated as the sum of the final % of all sub-scales/ 3(total number of sub-scales). The patient will score from 0 to 9 with 0 as good and 9 as the worst. Hence, the results may differ from 0 to 100% as these are directly proportional to impairment that is higher the percentage the higher be functional alteration [14].

### Short-Form Health Survey (SF-36)

The SF-36 is a general health and multi-function outcome measure that is comprised of 36 components [46]. This instrument has eight further subscales which include Physical Functioning (PF), Bodily Pain (BP), Role Physical (RP), General Health (GH), Vitality (VT), Social Functioning (SF), Role Emotional (RM), and Mental Health (MH). These are designated into two main scores i.e. physical component score (PCS) and the mental component score (MCS). The health would be perceived as better with a higher score. The score is usually calculated through two summary scores in terms of PCS and MCS [47]. However, online software [48] has been used to calculate its scoring for the present study. The reason to utilize SF-36 in the study is its good psychometric properties in the context of foot disorders [49].

### The Foot and Ankle Outcome Score (FAOS)

The FAOS is a modification of the Knee Injury and Osteoarthritis Outcome Score designed to assess foot and ankle symptoms and functional limitations [50]. The 42 items that make up the FAOS evaluate five distinct patient-relevant parameters: nine questions related to pain; seven items related to other symptoms such as stiffness, swelling, and range of motion; seventeen items

related to activities of daily living; five items related to sports and recreational activities; and four items related to foot and ankle-related quality of life [51]. Five Likert boxes (no, mild, moderate, severe, and extreme) were utilized for each item. Every item had a score ranging from zero to four, and the total of the items in each of the five sub-scales was used to determine its score [52]. The score for each subscale was calculated using the formula:  $100 - (\text{Scored items} / \text{possible score in the sub-scale}) \times 100$ . Scored ranged from 0 to 100, 0 indicates severe symptoms while 100 reflects no symptoms [51]. The original FFI has been attached in the supplementary files.

### **Visual Analogue Scale**

The visual analog scale (VAS) is one of the pain rating scales, Hayes and Patterson were the first ones to use this in the year 1921 [53]. The scale has two ends: “no pain” (0 cm on the left) and “worst pain” (10 cm on the right). A single handwritten mark is placed at one position along a 10-cm line, symbolizing a continuum between the two ends of the scale. The scores are based on self-reported measures of symptoms. The patient’s pain is measured in centimeters, starting at the left end of the scale and going all the way to their marks [54]. Greater pain and disability are indicated with a higher score while no pain and disability with a score close to zero. The cut-points for VAS of 10 cm scale are recommended as (0–4 mm), (5–44 mm), (45–74 mm), and (75–100 mm) denoted as no pain, mild pain, moderate pain, and severe pain, respectively, providing score range from 0 to 100 [55].

### **Phases of study**

The present study consists of two following phases:

- 1) Translation and cross-cultural adaptation of FFI.
- 2) Psychometric properties of FFI.

### **Translation and cross-cultural adaptation of FFI**

The translation and cross-cultural adaptation were made by following standard guidelines proposed by Beaton et al. [56]. The steps for translation include forward translation, synthesis, backward translation, expert committee review, pre-testing, and expert committee approval. 1-The first phase of adaptation is the forward translation. The tool translated from the original target language, English, to the new target language i.e., Urdu. The two native bilingual Urdu translators have translated the original FFI into Urdu language. One of the translators was an informed translator with a background in physiotherapy, called T1. The second translator was blinded to the study with no medical background and called T2. Both translators were fluent and proficient in English and Urdu languages. 2-The synthesis of forward translations (T12) was extracted by translated versions from T1 and T2. This

step was completed by the authors of the study and discussed with a moderator from the field of physiotherapy. 3-Two back-translators who were proficient in both Urdu and English (bilingual), with English being the primary spoken language, independently translated the T12 version back to its original version (English) (BT1, BT2). The original FFI was kept blinded from the BT1 and BT2 translators to reduce the possibility of bias arising during the translation process. A brief report was provided to each translator upon completion of their translation process. 4-The expert committee members comprised of all the translators (forward and backward), one language professional, one senior physiotherapist, and principal investigators. The task of this committee was to discuss all translated versions of FFI and to compose a pre-final version of FFI-Urdu (FFI-U) for field testing. 5- The pre-final version was then pilot-tested on 20 participants with a history of different foot problems to check if they were able to understand FFI-U and to collect their views and feedback for the questionnaire. The pilot testing stage has been discussed in detail in the [results](#) Sect. 6- After pilot testing, expert members discussed participants’ comments and demonstrated the final version of FFI-U. Figure 1 demonstrates the process of translation for FFI-U.

### **Psychometric Properties**

The psychometric properties i.e. reliability and validity were tested for FFI-U and have been discussed under statistical analysis as follows:

#### **Statistical analysis**

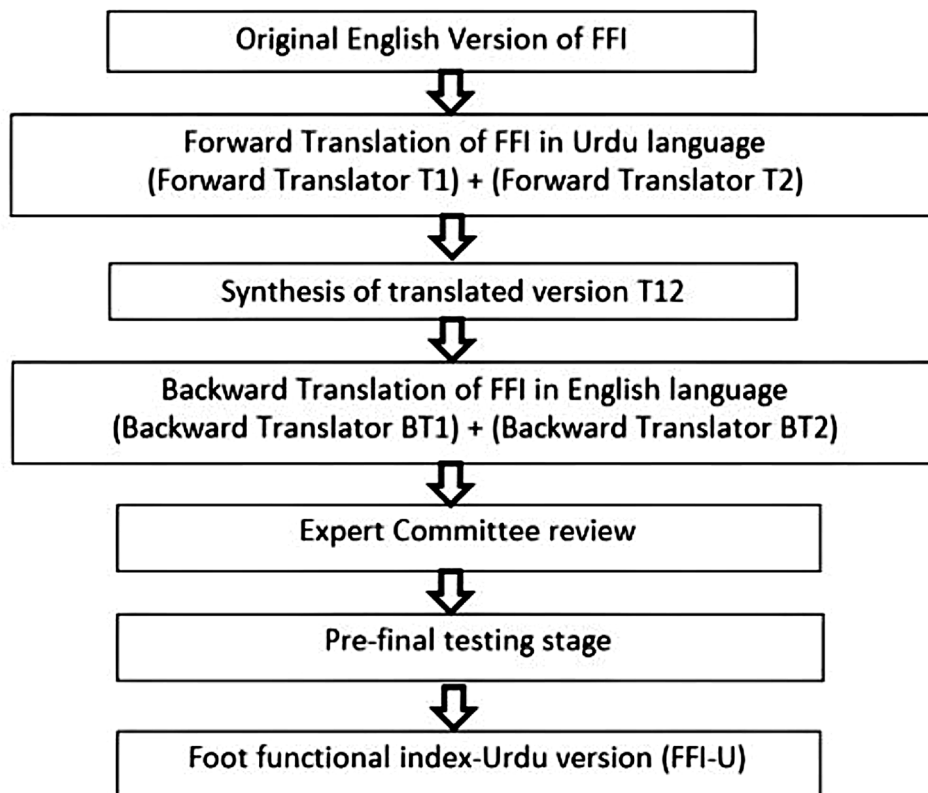
The Statistical Package of Social Science (SPSS) version 25.0 was used for the analysis of data. A statistically significant P-value was defined as less than 0.05. A priori hypothesis was used to validate the values of the psychometric characteristics. The participant’s attributes were examined using descriptive statistics.

#### **Reliability testing**

The reliability of FFI-U has been examined through internal consistency and test-retest reliability.

The Cronbach’s alpha was used to measure internal consistency for FFI-U. The value for Cronbach’s alpha ranged from 0 to 1 where zero indicates no internal consistency while 1 is for perfect internal consistency [57]. The grading system for internal consistency as proposed by George et al. [58] where Excellent internal consistency  $\geq 0.90$ , Good internal consistency  $\geq 0.80$ , acceptable  $\geq 0.70$ , questionable  $\geq 0.60$ , poor  $\geq 0.50$ , and unacceptable if  $< 0.50$ .

The test-retest reliability was measured through the intra-class coefficient (ICC) [59]. The reliability would be considered weak if ICC values  $< 0.5$ , moderate between 0.50 and 0.70, good reliability when 0.75–0.90, and any



**Fig. 1** Translation process of FFI-U

value  $\geq 0.90$  would be considered excellent test-retest reliability [59].

The Standard Error of Measurement (SEM) was used to calculate the measurement error and calculated the Minimal Detectable Change (MDC95) with a 95% confidence interval [60, 61]. The formulas  $SEM = SD \times \sqrt{1 - ICC}$  [62] and  $MDC = 1.96 \times \sqrt{2} \times SEM$  [63], were used for the calculation.

#### Validity testing

The validity of FFI-U was measured through face and construct validity. The face validity was obtained during cognitive debriefing while conducting interviews with participants at the pilot testing stage. The construct validity has been assessed using convergent validity by correlating FFI-U with three instruments i.e. FAOS, VAS and SF-36. The Spearman rank ( $r$ ) correlation was used to obtain convergent validity. A correlation ranging between  $-1$  to  $+1$  indicates a negative perfect to positive perfect correlation. The correlation's cut-off values are as follows:  $0.0-0.19$  is considered very weak,  $0.2-0.39$  is considered weak,  $0.4-0.59$  is considered moderate,  $0.6-0.89$  is considered substantial, and  $0.9-1.0$  is considered very strong [64]. VAS is an easy-to-understand scale that measures pain and disability that has also been used in the previous translation studies of FFI [17, 25, 65] whereas SF-36 is a generic outcome measure examining health in two

dimensions physical and mental [66]. A priori hypothesis has been made based on the previously available literature [25, 65, 67]. For convergent validity, we assumed that there would be a moderate to strong positive correlation of FFI-U subscales with physical components of SF-36 i.e. (RP and PF) and VAS (pain and disability). For divergent validity, we assumed weak correlations for subscales of FFI-U with mental components of SF-36. A 75% agreement between the results and the hypothesis indicated good validity [68].

#### Floor and ceiling effects

Floor and ceiling effects explain the acceptability or feasibility of an instrument [69]. The acceptability of any scale is measured by calculating missing responses/items from the questionnaire [69]. It was predicted that no floor and ceiling effects would be observed however there was an estimation of  $<5\%$  missing questions from the scale. According to McHorney and Tarlov, if more than 15% of individuals received the highest or lowest score, floor and ceiling effects were considered to be present [70]. The Floor and ceiling effects for FFI-U have been calculated from the number of participants who achieved the highest and lowest scores.

**Table 1** Demographic and clinical characteristics

Variables	Mean ± SD / Frequency (%)
Age	39 ± 11.2
BMI	27.41 ± 6.7
Onset of pain (days)	19.3 ± 5.0
<b>Gender</b>	
Males	129 (56)
Females	101 (44)
<b>Foot/ankle complaints</b>	
Ankle sprain	44 (19)
Metatarsalgia	25 (11)
Planter fasciitis	47 (20)
Achillies tendinitis	31 (13)
Pes planus	02 (0.8)
Hallux valgus	15 (6.5)
Rheumatoid arthritis of foot/ankle	42 (18)
Traumatic injury at foot/ankle	27 (12)

**Table 2** Descriptive analysis

Descriptive statistics	Pain	Disability	Activity limitation	Total FFI-U score
Mean	55.08	57.98	31.8	48.3
Standard deviation	23.76	21.21	11.65	15.87
Minimum	3	0	0	2.81
Maximum	100	95	90	75.56

## Results

### Translation and cross-cultural adaptation

The translation of the original FFI to the Urdu language comprehended no difficulty in terms of item comprehension and language. However, cultural adaptations are necessary, with the concern on measurement of gait parameters. The notion “blocks” was used in Item 12 to evaluate distance which is not a commonly used term in the Urdu-speaking population. Hence, after research and consulting previous literature, one block was found equal to 200 m and we employed 800 m for four blocks [17, 25].

### Pre-testing of the final version

Twenty participants with ages ranging from 32 ± 4.7 years from a podiatry clinic were included. These participants were seen with foot problems such as ankle sprain, planter fasciitis, metatarsalgia, and others. During the pretest stage, participants were interviewed and no one raised concern for any item. Hence, the scale was found feasible to comprehend and easy to understand, giving an acceptable face validity.

### Demographic characteristics

The study comprised of 56% males and 44% females with an average of 39 ± 11.2 years. Further, the demographic characteristics along with descriptive statistics for the 230 participants included in the study have been presented in Table 1.

**Table 3** Results for floor and ceiling effects, missing items and not applicable

FFI-U Subscales	Ceiling Score N (%)	Floor Score N (%)	Not applicable N (%)
Pain	7 (3.04)	8 (3.47)	17 (7.39)
Disability	4 (1.73)	5 (2.17)	11 (4.78)
Activity limitation	10 (4.34)	7 (3.04)	21 (9.13)

### FFI-U scoring characteristics

The scores were normally distributed for FFI-U. Table 2 illustrates the descriptive statistics for each sub-scale i.e. pain, disability and activity limitation, and total scores.

### Readability Index

The readability of FFI-U was measured by Flesch-Kincaid readability test formula [71]:  $206.835 - 1.015 \times (\text{words/sentences}) - 84.6 \times (\text{syllables/words})$ . This method was used manually as readability measurement methods are not available for Non-English language especially Urdu language. But manual adaptations can be made, though these are subjective [72]. We found out a score of 80 which shows standard to good readability. The readability scores as 0–30 being very difficult to read, 60–70 standard readability while 90–100 as good readability [71].

### Floor and ceiling effects

No participant reached the maximum or minimum score for FFI-U; hence no floor and ceiling effects were recorded. Table 3 demonstrates the effects of floor and ceiling for each sub-scale as well as the overall FFI-U score.

### Missing items

Missing items were excluded and rated as ‘Not applicable.’ The calculation of each item FFI-U score was ensured by a 10 cm horizontal line as originally explained by developers of FFI, also described in the methods section. At the end of each item ‘NA’ was given for patients, if that item did not match the condition of patients. Similarly, if an item was missed, it was rated similarly to items rated NA. For instance, out of 9 items in the pain subscale, 5 items have been marked 6 while 3 items were reported as NA and one item was missing, the sum of the rated items would be 30. The sum is divided by 45 (attended items (5) \* total items (9) = 45) and eventually multiplied by 100. Hence, the score for pain subscale in this example would be 66.67. The missing items and items marked as NA are also presented in Table 3.

### Reliability analysis

The Cronbach’s alpha ( $\alpha$ ) was measured for internal consistency where the ‘excellent’  $\alpha$  value was found for disability sub-scale i.e.  $\alpha=0.93$ . The values for pain and activity limitation were also found ‘good’ with  $\alpha=0.89$

and 0.87, respectively. The total score of FFI-U was found 'good' with  $\alpha=0.86$  for internal consistency.

The ICC (1, 2) was used to measure test-retest reliability analysis—a two-way random effects model which states that every rater measures every subject, and raters are thought to be typical of a wider group of comparable raters [73]. Out of 230 participants, data were collected from 30 participants on the 7th day of assessment for test-retest reliability. These participants were those who had scored moderate pain and disability levels on the VAS scale, upon their first round of assessment. These participants were well informed that they could not take any treatment between their first and second assessments. Only those participants were included in test-retest reliability analyses, who agreed by giving verbal and written consent. A good to excellent test-retest reliability was found for pain, disability, and activity limitation subscales with  $ICC_{2,1} = 0.87, 0.90,$  and  $0.82$  respectively. The  $ICC_{2,1}$  for the total score of FFI-U was found good i.e.  $0.845 (0.78–0.89)$ . The SEM was computed to be 3.19, while the MDC95 was determined to be 9.8.

The reliability analysis for FFI-U is presented in Table 4.

#### Validity analysis

The construct validity was measured by calculating correlation using the Pearson correlation coefficient between FFI-U and other questionnaires (SF-36, FAOS, and VAS). There were negative findings for the Pearson's correlation coefficient ( $\gamma$ ) between the SF-36 and the FFI-U. This can be demonstrated with the certitude that a higher score of SF-36 denotes better health status and a higher score of FFI denotes deteriorated health status [17]. A moderate correlation was found for the physical components (PE, GH, RP, and BP) with all three subscales of FFI-U. The mental components of SF-36 (SE, RE, VT, and MH) had a weak correlation with all the subscales of FF-U. A moderate negative correlation was found for the physical component score (PCS) stating ( $\gamma = -0.65, p\text{-value} < 0.05$ ) with the total score of FFI-U. A weak negative correlation was found for the mental components score (MCS) stating ( $\gamma = -0.25, p\text{-value} > 0.05$ ).

The correlation coefficient was found positive between FFI-U with VAS (pain and disability) as these scales indicate a low score for worse health outcomes and a higher score for better health outcomes. A substantial positive correlation was found with VAS-pain, and VAS-disability, while a moderate negative correlation FAOS with ( $\gamma = 0.72, 0.71, -0.68$ ) respectively with the total score of FFI-U. However, a weak correlation was found with the activity limitation sub-scale of FFI-U with VAS (pain & disability).

The correlation analysis for FFI-U supports the hypothesis about discriminant and convergent validity. As, it was hypothesized earlier referring to the [methods](#) section,

**Table 4** Reliability analysis of FFI-U

Sub-scales	Internal consistency ( $\alpha$ )	Test-retest (30 participants)		Intra-class correlation coefficient
		Test score $\pm$ SD	Retest score $\pm$ SD	
Pain	0.89	60.87 $\pm$ 13.76	54.19 $\pm$ 22.14	0.82 (0.73–0.85)
Disability	0.93	65.23 $\pm$ 27.87	67.65 $\pm$ 19.43	0.90 (0.72–0.95)
Activity limitation	0.81	31.98 $\pm$ 9.56	30.61 $\pm$ 11.98	0.87 (0.79–0.88)
Total FFI-U score	0.86	52.69 $\pm$ 14.11	50.81 $\pm$ 12.64	0.845 (0.78–0.89)

**Table 5** Validity analysis of FFI-U

Variables	Pain	Disability	Activity Limitation	FFI-U Total score
Physical functioning	-0.57	-0.76	-0.54	-0.62
Role physical	-0.41	-0.43	-0.37	-0.48
Bodily pain	-0.51	-0.32	-0.43	-0.44
General health	-0.14	-0.23	-0.21	-0.27
Vitality	-0.27	-0.42	-0.38	-0.41
Social functioning	-0.35	-0.37	-0.32	-0.37
Role emotional	-0.24	-0.18	-0.21	-0.27
Mental health	-0.17	-0.21	-0.25	-0.23
PCS	-0.64	-0.71	-0.60	-0.65
MCS	-0.32	-0.21	-0.34	-0.25
VAS-pain	0.88	0.74	0.68	0.72
VAS-disability	0.75	0.76	0.71	0.71
FAOS	-0.45	-0.77	-0.76	-0.68

FFI-U would show a good correlation with physical components of SF-36 and VAS, while mental components of SF-36 were found in weak correlation. The strong correlation of FFI-U with physical components SF-36 and VAS shows convergent validity whereas discriminant validity is the weak correlation of mental components of SF-36.

The validity analysis has been presented in Table 5.

#### Discussion

The present study aimed to translate and culturally adapt the FFI into Urdu language and measure its psychometric properties. The results of the study demonstrated FFI-U as a reliable and valid tool to be used for the Urdu-speaking population with foot and ankle complaints.

FFI is known as one of the most commonly used patient-reported outcome measures [19, 74]. However, some problems were incorporated with the original version due to the component of activity limitation with concerns raised about its internal consistency and low reproducibility [14, 75, 76]. Agel et al. documented a significant number of drop-outs, non-responses, and ceiling effects due to activity limitation subscale [76]. For this reason, the cross-cultural studies for languages such as Italian [67], German [26], Taiwan Chinese [34], and

Dutch [33] versions of FFI have completely deleted this sub-scale and concluded strong reliability and validity. Moreover, many studies reported increased drop-outs or missed responses for items concerning orthotic devices as these were not relevant to the populations being studied [33, 75, 76]. Similarly, we have faced the same circumstances during this study and found missed responses for items relating to orthotic devices. However, it could be argued that the original FFI tested on rheumatoid arthritis patients only [14], and hence, results may be affected while applying the scale to patients with other foot disorders. Nevertheless, keeping in view all these concerns, the developers of FFI re-evaluated the scale and revised the outcome measure using Rasch analysis [18], which is a powerful tool for analyzing the content of the questionnaire [77]. This resulted in an outcome measure consisting of 68 items, which was contracted to a short form with 34 questions which was still with an excess of 11 items from the original FFI. Furthermore, the original FFI is still widely used in clinics and for research purposes due to its easy scoring method [14]. Therefore, these two were primary reasons to utilize an original version of FFI for translation and cultural adaptation into the Urdu language, which also helped us in appropriate comparisons with the already reported clinical studies.

Cultural adaptation of a questionnaire is as important as translating it into the target language to avoid any systematic errors [78]. Therefore, while formulating FFI-U all discrepancies and systematic considerations were carefully addressed, proofread, and cross-checked. The only cultural adaptation made for FFI-U was the use of the term 'meters' instead of 'blocks'. A similar cross-cultural adaptation was previously made in the French [25], Persian [32], Arabic [17], and Thai [65] versions of FFI.

The present study has enrolled participants with acute as well as chronic conditions of the foot and ankle to endorse a wider aspect of foot conditions. Similarly, enrolling multiple conditions can also be seen in different language versions of FFI such as Italian [67], Brazilian Portuguese [28], Saudi Arabic [17], and Chinese [29] such as ankle sprain, hallux valgus and varus, planter fasciitis/fasciopathy, metatarsalgia and painful fat feet. However, some studies have only measured one disorder of the foot using FFI [21, 22]. In the present study, the rationale behind including a variety of foot conditions is to increase the use of the tool relentlessly by the clinicians for their patients with most of the foot conditions.

FFI is renowned for its brevity, ease of use, and short completion time [76]. The mean completion time recorded for FFI-U was 7.2 min which is suggestive of the feasibility of this questionnaire. However, the mean time was calculated informally by the therapists while collecting data and no formal scale such as a stop-watch was used. Nonetheless, the questionnaire was completed by

all the participants without any complaints. A completion time of 5–10 min was recorded for the original FFI as reported by Budiman et al. (1991) [14], whereas the Arabic version took up to 5.2 min [17] and the French version reported 10 min [25].

This study has demonstrated good internal consistency i.e. 0.86 for a total score of FFI-U and good to excellent Cronbach alpha was found for three sub-scales of FFI-U indicating the measurement of the same construct by all the items. The internal consistency results of FFI-U can be found comparable with other versions such as French version ( $\alpha=0.85-0.97$ ) [25], Turkish version ( $\alpha=0.82-0.94$ ), Arabic version ( $\alpha=0.88-0.93$ ) [21], Italian version ( $\alpha=0.95$ ), Spanish version ( $\alpha=0.69-0.95$ ) [27] and Persian version ( $\alpha=0.88-0.95$ ) [32]. It can be observed from the results of other translated versions of FFI, that the internal consistency for activity limitation was fair-moderate, which ended up as moderate internal consistency for the whole scale. For instance, Cronbach's alpha value for activity limitation was ( $\alpha=0.69$ ) for the Spanish version [27], and ( $\alpha=0.61$ ) for the Brazilian Portuguese version resulting in a fair  $\alpha$ -value for the total score [28]. Similarly, the original version showed a fair value ( $\alpha=0.71$ ) for the original version of FFI with  $\alpha=0.87$  for the total score [14]. However, the present study showed a good  $\alpha$ -value for FFI-U with activity limitation as  $\alpha=0.81$  which has been found similar to French [25], Arabic [21], and Chinese [29] versions of FFI. Hence, FFI-U demonstrated measurement of the same construct within all sub-scales.

The test-retest reliability of FFI-U demonstrated good to excellent results for all three subscales i.e. (ICC=0.82, 0.90, and 0.87) and good test-retest for the total score (ICC<sub>2, 1</sub>=0.845). One of the reasons could be the stability of symptoms as activity limitation and disability are those symptoms that do not fluctuate as pain. The results of the present study are comparable to ICC values of the original FFI i.e. 0.84 for disability and 0.81 for activity limitation while 0.69 for pain [14]. The ICC values of FFI-U were comparable to other versions of FFI [21, 25–27].

The construct validity of FFI-U was confirmed through convergent and discriminant validity when correlated with different outcome measures. The SF-36 has been considered a gold standard for measuring the correlation of patient-reported outcome measures. This is also evident in several previous validation studies including different versions of FFI as well [17, 26, 28, 37, 67]. In the present study, a moderate correlation was found for physical components ( $\gamma=-0.65$ ) while a weak correlation with mental components of SF-36 ( $\gamma=-0.25$ ) with FFI-U, which can also be seen in different versions of FFI such as Taiwan Chinese found moderate correlation [34] while German [26] and Arabic versions [17] found moderate to high correlation with SF-36. A high correlation can be seen in the Italian [67] and Korean [31] versions of FFI



whereas a weak correlation has been demonstrated by the Turkish version [37] and a weak to fair correlation was observed in the Brazilian Portuguese version of FFI [28]. However, it could be argued while considering the higher correlation between German and Korean versions of FFI that has deleted the activity limitation component of the FFI scale [26, 31].

Similarly, FFI showed a high correlation with VAS (pain, disability) and FAOS with Pearson correlation values ( $\gamma=0.72, 0.71, -0.68$ ), respectively. Caution is required while applying results for the correlation of FAOS and FFI-U as we did not use sub-scales for this scale and only correlated with the total score. However, the same can be seen in the Brazilian Portuguese version of FFI where a total score of FFI was correlated with subscales of FAOS [28]. Nevertheless, the results showed a higher correlation as the scale measures the pain and disability levels in the foot as seen in FFI.

From the correlation analysis, convergent validity has been confirmed i.e. FFI showed a higher correlation with activity limitation, disability, and pain components of all outcome measures used, however, a weak correlation was observed for mental components of SF-36 confirming discriminant validity.

The results from the reliability and validity analysis demonstrated comparable measurements with the studies conducted using the original version of FFI in the English language [14, 19, 79], which supports the view that the Urdu version of FFI is clinically applicable to foot disorders.

There are some limitations observed in the present study that should be considered such as the lack of longitudinal psychometric variables analysis, for instance, reproducibility, error score, sensitivity to change, or responsiveness that shows clinically important differences. Furthermore, structural validity was not measured. Therefore, future studies are recommended to evaluate these psychometric variables analysis. In addition to these, the influence of covariates such as age, gender, employment, and diagnosis could affect the outcomes. It has been informally observed that an increased age population had more FFI scores which would have impacted the results. The findings may be unique for participants with increasing age. Henceforth, demographics should be considered while performing analysis in future studies and the association of these variables should be measured with changed FFI-U scores to allow evaluation of clinical improvements in each group after treatment. Finally, the sample size for test-retest reliability was small as only 30 participants were those who maintained no treatment between the first and second assessments. Therefore, it is recommended to imply these results with

caution, and further studies are encouraged to demonstrate test-retest reliability for FFI-U on a larger scale of population.

On the other hand, the strengths of the present study are that it has been formulated with guidelines from the literature regarding the minimal number of subjects needed for each item to guarantee the psychometric analysis of the questionnaire [80]. Since there are 23 questions on the FFI-U, a minimum of 230 subjects would be required [42]; nonetheless, this study had 230 participants. Moreover, biases were avoided at the maximum by presenting all the outcome measures in the same order. One of the strengths of the present study was the inclusion of diverse foot problems which increases the external validity of results by enhancing the generalizability.

The FFI-U has considerable clinical implications, particularly in the context of improving patient care in Urdu-speaking regions. It improves the quality of communication between healthcare providers and patients, resulting in accurate evaluations and tailored treatments. This translation guarantees standardised and culturally relevant data collection, which in turn enhances the accuracy of diagnosis, monitoring, and treatment outcomes. Furthermore, it enables a wider population to access foot health assessments, empowers patients, and supports public health initiatives, thereby promoting health equity. The FFI-U also contributes to global research by facilitating cross-cultural comparisons and extending the comprehension of foot health in a variety of contexts.

Taken together, the present study confirms that FFI-U is a valid and reliable tool to be administered in clinical settings for patients with different foot conditions.

## Conclusion

The FFI-U has been found reliable, valid, and feasible tool to be used as a patient-reported outcome measure to assess different foot disorders in the Urdu-speaking population. Thus, the clinicians and researchers might use FFI-U in their respective settings for the assessment of multiple foot disorders.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12891-024-07857-5>.

Supplementary Material 1

Supplementary Material 2

## Acknowledgements

We would like to thank Ms Maria Sohail for getting us permission to translate from the developers of the Foot Functional Index.

## Author contributions

AqA and ST: Study conception and design, AA and HS: Acquisition of data, HJ, HS and AI: Analysis and interpretation of data, AA and AI: Drafting

of manuscript, HS and ST: Critical revision AqA and HJ: proofreading of manuscript. All authors: final approval of the manuscript.

#### Funding

No funding source.

#### Data availability

Data is provided within the manuscript or supplementary information files.

#### Declarations

##### Ethics approval and consent to participate

The Research ethics committee (REC) of the University of Lahore provided the ethical approval for this study with reference number (REC-UOL-408-05-2023). A written informed consent was taken from the participants and all ethical rules were kept in consideration during the study. The relevant guidelines by the declaration of Helsinki were followed within the study.

##### Consent for publication

Not applicable.

##### Financial disclosure

None.

##### Competing interests

The authors declare no competing interests.

##### Author details

<sup>1</sup>University Institute of Physical Therapy, University of Lahore, Lahore, Pakistan

<sup>2</sup>University of the West of Scotland, Glasgow, Scotland

<sup>3</sup>Pakistan Society for Rehabilitation of Differently Abled, Lahore, Pakistan

Received: 5 May 2024 / Accepted: 6 September 2024

Published online: 19 September 2024

#### References

- Rao S, Riskowski JL, Hannan MT. Musculoskeletal conditions of the foot and ankle: assessments and treatment options. *Best Pract Res Clin Rheumatol*. 2012;26(3):345–68.
- Hill CL, Gill TK, Menz HB, Taylor AW. Prevalence and correlates of foot pain in a population-based study: the North West Adelaide health study. *J Foot Ankle Res*. 2008;1:1–7.
- Mika A, Oleksy L, Mika P, Marchewka A, Clark BC. The effect of walking in high- and low-heeled shoes on erector spinae activity and pelvis kinematics during gait. *Am J Phys Med Rehabil*. 2012;91(5):425–34.
- Hornestam JF, Arantes PMM, Souza TR, Resende RA, Aquino CF, Fonseca ST, et al. Foot pronation affects pelvic motion during the loading response phase of gait. *Braz J Phys Ther*. 2021;25(6):727–34.
- Simonsen EB. Contributions to the understanding of gait control. *Dan Med J*. 2014;61(4):B4823.
- Lievers WB, Goggins KA, Adamic P. Epidemiology of Foot Injuries Using National Collegiate Athletic Association Data from the 2009–2010 through 2014–2015 Seasons. *J Athl Train*. 2020;55(2):181–7.
- Hajebrahimi F, Tarakci D, Arman N, Emir A, Bursali A, Tarakci E. Cross-cultural adaptation, validity and reliability of Turkish version of Oxford Ankle Foot Questionnaire for children with congenital talipes equinovarus. *Foot Ankle Surg*. 2021;27(4):439–42.
- Thomas MJ, Roddy E, Zhang W, Menz HB, Hannan MT, Peat GM. The population prevalence of foot and ankle pain in middle and old age: a systematic review. *Pain*. 2011;152(12):2870–80.
- Yagci G, Erel S, Okunakol V. Validation of the Turkish version of the revised foot function index for patients with foot and ankle disorders. *Foot Ankle Surg*. 2020;26(6):624–9.
- Churruca K, Pomare C. Patient-reported outcome measures (PROMs): a review of generic and condition-specific measures and a discussion of trends and issues. 2021;24(4):1015–24.
- Krüger L, Rolvien T, Seller M, Hubert J, Beil FT, Strahl A. Reliability, validity and clinical applicability of the German version of the European Foot and Ankle Society score. *Foot Ankle Surg*. 2022;28(8):1307–13.
- Howell D, Molloy S, Wilkinson K, Green E, Orchard K, Wang K, et al. Patient-reported outcomes in routine cancer clinical practice: a scoping review of use, impact on health outcomes, and implementation factors. *Ann Oncol*. 2015;26(9):1846–58.
- Basch E, Barbera L, Kerrigan CL, Velikova G. Implementation of patient-reported outcomes in routine medical care. *Am Soc Clin Oncol Educational Book*. 2018;38:122–34.
- Budiman-Mak E, Conrad KJ, Roach KE. The foot function index: a measure of foot pain and disability. *J Clin Epidemiol*. 1991;44(6):561–70.
- Hijji FY, Schneider AD, Pyper M, Laughlin RT. The popularity of outcome measures used in the foot and ankle literature. *Foot Ankle Spec*. 2020;13(1):58–68.
- Van Der Leeden M, Steultjens MP, Terwee CB, Rosenbaum D, Turner D, Woodburn J, et al. A systematic review of instruments measuring foot function, foot pain, and foot-related disability in patients with rheumatoid arthritis. *Arthritis Care Research: Official J Am Coll Rheumatol*. 2008;59(9):1257–69.
- Khan S, Faulkner S, Algarni FS, Almalki A, Almansour A, Altowajiri AM. Foot function index for arabic-speaking patients (FFI-Ar): translation, cross-cultural adaptation and validation study. *J Orthop Surg Res*. 2022;17(1):212.
- Budiman-Mak E, Conrad K, Stuck R, Matters M. Theoretical model and Rasch analysis to develop a revised foot function index. *Foot Ankle Int*. 2006;27(7):519–27.
- Budiman-Mak E, Conrad KJ, Mazza J, Stuck RM. A review of the foot function index and the foot function index – revised. *J Foot Ankle Res*. 2013;6(1):5.
- Stucki G, Cieza A, Ewert T, Kostanjsek N, Chatterji S, Üstün TB. Application of the International Classification of Functioning, disability and health (ICF) in clinical practice. *Disabil Rehabil*. 2002;24(5):281–2.
- Amri MI, Alzhrani MM, Alanazi AD, Alqahtani MM, Kashoo FZ. Cross-cultural adaptation and validation of the arabic version of the foot function index in patients with chronic lateral ankle instability. *J Foot Ankle Res*. 2022;15(1):21.
- Vetrano M, Vulpiani MC, Erroi D, Vadala A, Ferretti A, Saraceni VM. Cross-cultural adaptation and reliability of the Italian version of the foot function index (FFI-I) for patients with plantar fasciitis. *J Sports Med Phys Fit*. 2014;54(5):636–43.
- Bovonsunthonchai S, Thong-On S, Vachalathiti R, Ritiravoranont W, Suwanarat S, Smith R. Thai version of the foot function index: a cross-cultural adaptation with reliability and validity evaluation. *BMC Sports Sci Med Rehabilitation*. 2020;12(1):1–11.
- Abdelhamed SS, Ibrahim AHM, Draz A, Abdel-Aal NM. Cross-Cultural Adaptation. Validity and reliability of Egyptian Arabic version of foot function index for patients with Foot and Ankle disorders. *J Pharm Negat Results*. 2022;4:226–33.
- Pourtier-Piotte C, Pereira B, Soubrier M, Thomas E, Gerbaud L, Coudeyre E. French validation of the foot function index (FFI). *Annals Phys Rehabilitation Med*. 2015;58(5):276–82.
- Naal FD, Impellizzeri FM, Huber M, Rippstein PF. Cross-cultural adaptation and validation of the foot function index for use in german-speaking patients with Foot complaints. *Foot Ankle Int*. 2008;29(12):1222–31.
- Paez-Moguer J, Budiman-Mak E, Cuesta-Vargas AI. Cross-cultural adaptation and validation of the foot function index to Spanish. *Foot Ankle Surg*. 2014;20(1):34–9.
- Martinez BR, Staboli IM, Kamonseki DH, Budiman-Mak E, Yi LC. Validity and reliability of the foot function index (FFI) questionnaire brazilian-portuguese version. *Springerplus*. 2016;5:1–7.
- Gonzalez-Sanchez M, Ruiz-Munoz M, Li GZ, Cuesta-Vargas AI. Chinese cross-cultural adaptation and validation of the foot function index as tool to measure patients with foot and ankle functional limitations. *Disabil Rehabil*. 2018;40(17):2056–61.
- Jorgensen J, Andreasen J, Rathleff M. Translation and validation of the Danish Foot Function Index (FFI-DK). *Scand J Med Sci Sports*. 2015;25(4):e408–13.
- Huh J-W, Eun I-S, Ko Y-C, Park M-J, Hwang K-M, Park S-H, et al. Reliability and validity of the Korean version of the foot function index. *J Foot Ankle Surg*. 2016;55(4):759–61.
- Mousavian A, Mohammadi A, Seyed-Hosseini S-H, Shahpari O, Elahpour N, Orooji A, et al. Reliability and validity of the Persian version of the foot function index in patients with foot disorders. *Archives Bone Joint Surg*. 2019;7(3):291.
- Kuyvenhoven MM, Gorter KJ, Zuihoff P, Budiman-Mak E, Conrad KJ, Post MW. The foot function index with verbal rating scales (FFI-5pt): a clinimetric evaluation and comparison with the original FFI. *J Rheumatol*. 2002;29(5):1023–8.

34. Wu S-H, Liang H-W, Hou W-H. Reliability and validity of the Taiwan Chinese version of the foot function index. *J Formos Med Assoc.* 2008;107(2):111–22.
35. Chauhan M, Translation. Cross Cultural Adaptation and Validation of the foot function index questionnaire for Gujarati Speaking People with Foot Complaints.
36. Yi LC, Cabral ACC, Kamonseki DH, Budiman-Mak E, Vidotto MC. Translation and cultural adaptation of the revised foot function index for the Portuguese language: FFI-R Brazil. *Sao Paulo Med J.* 2017;135:573–7.
37. KÜLÜNKOĞLU B, Firat N, Yildiz NT, Alkan A. Reliability and validity of the Turkish version of the Foot function index in patients with foot disorders. *Turk J Med Sci.* 2018;48(3):476–83.
38. Rutkowski R, Gałczyńska-Rusin M, Gizińska M, Straburzyński-Lupa M, Zdanowska A, Romanowski MW et al. Adaptation and validation of the foot function index-revised short form into Polish. *BioMed research international.* 2017;2017.
39. Mørk M, Hoksrud AF, Soberg HL, Zucknick M, Heide M, Groven KS, et al. Psychometric properties of the Norwegian foot function index revised short form. *BMC Musculoskelet Disord.* 2022;23(1):1–10.
40. Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J Anaesth.* 2017;11(Suppl 1):S80–9.
41. Cuschieri S. The STROBE guidelines. *Saudi J Anaesth.* 2019;13(Suppl 1):S31–4.
42. Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi J Anaesth.* 2017;11(Suppl 1):S80.
43. Espinosa N, Maceira E, Myerson MS. Current concept review: metatarsalgia. *Foot Ankle Int.* 2008;29(8):871–9.
44. Hale SA, Hertel J. Reliability and sensitivity of the Foot and Ankle Disability Index in subjects with chronic ankle instability. *J Athl Train.* 2005;40(1):35.
45. Budiman-Mak E, Conrad KJ, Mazza J, Stuck RM. A review of the foot function index and the foot function index-revised. *J Foot Ankle Res.* 2013;6(1):1–37.
46. Ware J, Kosinski M, Bjorner J, Turner-Bowker D, Gandek B, Maruish M. Development. User's Manual for the SF-36v2® Health Survey 2nd edn Lincoln (RI): QualityMetric Incorporated. 2007;200(7).
47. org S-. FAQ: is there one summary score that is a combined score for the various subscales in the SF-36® so that a single score could be used for each patient?.
48. Saris-Baglama RN, Dewey CJ, Chisholm GB, Plumb E, King J, Kosinski M et al. QualityMetric health outcomes™ scoring software 4.0: installation guide. Lincoln (RI): QualityMetric Incorporated. 2010.
49. SooHoo NF, Vyas R, Samini D. Responsiveness of the foot function index, AOFAS clinical rating systems, and SF-36 after foot and ankle surgery. *Foot Ankle Int.* 2006;27(11):930–4.
50. Tapaninaho K, Uimonen MM, Saarinen AJ, Repo JP. Minimal important change for foot and ankle outcome score (FAOS). *Foot Ankle Surg.* 2022;28(1):44–8.
51. Imoto AM, Peccin MS, Rodrigues R, Mizusaki JM. Translation, cultural adaptation and validation of foot and ankle outcome score (FAOS) questionnaire into Portuguese. *Acta Ortopédica Brasileira.* 2009;17:232–5.
52. Roos EM, Brandsson S, Karlsson J. Validation of the foot and ankle outcome score for ankle ligament reconstruction. *Foot Ankle Int.* 2001;22(10):788–94.
53. Couper MP, Tourangeau R, Conrad FG, Singer E. Evaluating the effectiveness of visual analog scales: a web experiment. *Social Sci Comput Rev.* 2006;24(2):227–45.
54. Delgado P, Vargas C, Ackerman R, Salmerón L. Don't throw away your printed books: a meta-analysis on the effects of reading media on reading comprehension. *Educational Res Rev.* 2018;25:23–38.
55. Aun C, Lam YM, Collett B. Evaluation of the use of visual analogue scale in Chinese patients. *Pain.* 1986;25(2):215–21.
56. 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.
57. Nunnally J. *Psychometric theory.* (No Title). 1994.
58. George D, Mallery P. *IBM SPSS statistics 26 step by step: a simple guide and reference.* Routledge; 2019.
59. Lee KM, Lee J, Chung CY, Ahn S, Sung KH, Kim TW, et al. Pitfalls and important issues in testing reliability using intraclass correlation coefficients in orthopaedic research. *Clin Orthop Surg.* 2012;4(2):149–55.
60. Gabel CP, Michener LA, Burkett B, Neller A. The Upper Limb Functional Index: development and determination of reliability, validity, and responsiveness. *J Hand Ther.* 2006;19(3):328–49.
61. Cuesta-Vargas AI, Gabel CP, Bennett P. Cross cultural adaptation and validation of a Spanish version of the lower limb functional index. *Health Qual Life Outcomes.* 2014;12(1):1–9.
62. Portney LG, Watkins MP. *Foundations of clinical research: applications to practice.* Pearson/Prentice Hall Upper Saddle River, NJ; 2009.
63. Koo T, Li M. Cracking the code: providing insight into the fundamentals of research and evidence-based practice a guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15(2):155–63.
64. Zaiğ A, Berteu P. Methods for testing discriminant validity. *Manage Mark J.* 2011;9(2):217–24.
65. Bovonsunthonchai S, Thong-On S, Vachalathiti R, Intiravoranon W, Suwanarat S, Smith R. Thai version of the foot function index: a cross-cultural adaptation with reliability and validity evaluation. *BMC Sports Sci Med Rehabilitation.* 2020;12:1–11.
66. Lins L, Carvalho FM. SF-36 total score as a single measure of health-related quality of life: scoping review. *SAGE open Med.* 2016;4:2050312116671725.
67. Martinelli N, Scotto GM, Sartorelli E, Bonifacini C, Bianchi A, Malerba F. Reliability, validity and responsiveness of the Italian version of the foot function index in patients with foot and ankle diseases. *Qual Life Res.* 2014;23:277–84.
68. Aragon V, Oyama S, Oliaro S, Padua D, Myers J, Brindle T. Quality criteria were proposed for measurement properties of health status questionnaires. *J Athl Train.* 2012;47(5):507–15.
69. Šimkovic M, Träuble B. Robustness of statistical methods when measure is affected by ceiling and/or floor effect. *PLoS ONE.* 2019;14(8):e0220889.
70. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res.* 1995;4(4):293–307.
71. Flesch R. Flesch-Kincaid readability test. Retrieved Oct. 2007;26(3):2007.
72. Wang L-W, Miller MJ, Schmitt MR, Wen FK. Assessing readability formula differences with written health information materials: application, results, and recommendations. *Res Social Administrative Pharm.* 2013;9(5):503–16.
73. Chenani KT, Madadzadeh F. Guideline for selecting types of reliability and suitable intra-class correlation coefficients in clinical research. *J Biostatistics Epidemiol.* 2021;7(3):305–9.
74. Lau JT, Mahomed NM, Schon LC. Results of an internet survey determining the most frequently used ankle scores by AOFAS members. *Foot Ankle Int.* 2005;26(6):479–82.
75. Saag KG, Saltzman CL, Brown CK, Budiman-Mak E. The foot function index for measuring rheumatoid arthritis pain: evaluating side-to-side reliability. *Foot Ankle Int.* 1996;17(8):506–10.
76. Agel J, Beskin JL, Brage M, Guyton GP, Kadel NJ, Saltzman CL, et al. Reliability of the foot function index: a report of the AOFAS outcomes Committee. *Foot Ankle Int.* 2005;26(11):962–7.
77. Boone WJ. Rasch Analysis for Instrument Development: Why, When, and How? CBE life sciences education. 2016;15(4).
78. Gjersing L, Caplehorn JR, Clausen T. Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. *BMC Med Res Methodol.* 2010;10:13.
79. Goldstein CL, Schemitsch E, Bhandari M, Mathew G, Petrisor BA. Comparison of different outcome instruments following foot and ankle trauma. *Foot Ankle Int.* 2010;31(12):1075–80.
80. Stratford PW. Getting more from the literature: estimating the Standard Error of Measurement from Reliability studies. *Physiotherapy Can.* 2004;56:027.

## Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.