

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,500

Open access books available

135,000

International authors and editors

165M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



# Patient-Report Outcome Measures for Ankle-Related Functionality

*Tarcísio Santos Moreira*

## Abstract

A patient's subjective perception about his/her own functional status and also about health-related quality of life represents a challenge both for clinicians and researchers, particularly in the field of rehabilitation. Clinicians often overlook the functional limitations and disability experienced by patients. Because functional limitations and disability are most important to the patient, it is essential that clinicians quantify dysfunction at this level. Client-based assessment instruments, like questionnaires, are tools suitable for comprising the domains of activity and social participation and are often the selected instrument for the assessment of health-related quality of life. In this chapter, the main aspects of such outcome measures are discussed in order to help clinicians and researchers in the selection of appropriate assessment tools in their daily practice.

**Keywords:** self-report measures, psychometric analysis, ankle/foot disorders, item response theory, Rasch model

## 1. Introduction

The main purpose of this chapter is to present a clinical and scientific perspective of the applicability of self-report outcome measures on the assessment of rehabilitation intervention designed to treat ankle/foot disorders. Some aspects about the development of these clinical tools will be briefly discussed in order to offer clinicians and scientists some criteria to better select an appropriate measure for a given clinical goal. In addition, a presentation is given of how the item response theory (IRT), by using Rasch analysis, can be used to assess such clinical measurement tools and to enrich rehabilitation data concerning clinical intervention effectiveness.

## 2. Patient-based outcome measures: the importance of quantifying patient's perspectives about health treatments

Throughout the history of orthopedic field, the main focus of clinical relevant outcome measures used to be those related to body structure and function, quantifying movement restrictions, such as range of motion or functional impairments like muscle strength. By this point of view, treatment goals and definitions of successful interventions captured mainly what is direct important for the healthcare professional rather than for the patient. However, what is crucial and naturally understandable for patients is functionality and disability, which brings a clear necessity

to measure dysfunction at this level. Furthermore, the International Classification of Functionality (ICF) proposed by the World Health Organization (WHO) suggests that health issues should be considered by taking into account individual function, activity, and social participation. The impact that injuries, illnesses, and any other harm might have upon health, especially over functionality and quality of life, must be considered in a context of clinical evaluation in the health area.

Usually, ankle dysfunctions require the involvement of a wide variety of health-care professionals in order to achieve excellence in recovery and functionality. This is particularly true when the treatment plan includes surgery, medications, rehabilitation program, and many other interventions carried out by different professionals. Multidisciplinary healthcare thus requires instruments that work across distinct disciplines in a sense of combining them and unifying perspectives.

The health team may have priorities that can diverge from patient-specific needs or beliefs. This may happen often because both look for a health condition from different backgrounds and starting points and this communication noise may lead to inappropriate treatment plan and can decrease patient's compliance.

Focusing on the patient, who is most interested in full recovering of his/her health, quantifying subjective perception of functional status as well as health-related quality of life represents a challenge both for clinicians and researchers, particularly in the field of rehabilitation. A patient-centered or also called client-based assessment tool is needed and should meet the clinical needs of both the patient and the healthcare team, in such a way that it must be practical and accepted by everyone involved in a treatment context [1–17].

Client-based assessment instruments, like questionnaires, are tools suitable to comprise the domains of activity and social participation being commonly the selected instrument for the assessment of health-related quality of life. They have the ultimate goal of transforming subjective measures into objective data that can be quantified and analyzed. Self-report questionnaires are useful both for clinical and scientific research purposes once they combine efficiency, reliability, and low cost and, at the same time, meeting the necessity to quantify patient-centered clinical outcome measures.

### **3. Practical scenario: clinical use of self-report assessment tools**

Every clinical outcome measure may have five goals in order to be useful in a clinical-based scenario. The acronym that exemplifies this feature is known as SMART goals and can be visualized in **Figure 1**.

#### **3.1 Target population and purpose of the measurement tool**

One way to classify questionnaires and functional scales is by their assessment application (see **Figure 2**). In this case, they can be categorized as being generic or specific. Generic questionnaires measure overall health, within biopsychosocial approach, and are intended to be applicable across a wide spectrum of diseases, interventions, demographic, and cultural subgroups. The most famous and used instrument that encompasses this properties is the 36-item short-form health survey (SF-36), which measure health-related quality of life in two main domains of mental and physical health. On the other hand, disease-specific measures aim to assess the most important traits usually affected by a condition of interest and that can be used to determine clinical improvement or deterioration. The foot and ankle outcome score (FAOS) and the American Orthopedic Foot and Ankle Society ankle-hindfoot scale (AOFAS) are both examples of condition-specific measures.



**Figure 1.**  
*SMART goals for a clinical assessment tool.*



**Figure 2.**  
*Questionnaires and functional scale classification schema.*

Another form in which self-report tools can be organized is in relation to the clinical function (see **Figure 2**). Within this context, they can be discriminative or evaluative instruments. The selection of one type over the other depends on the desired use of the instrument. Discriminative instruments, such as the Cumberland ankle instability tool (CAIT), can be used to identify individuals with a particular disorder, in this case, chronic ankle instability. Evaluative instruments are developed to follow up and measure an individual's change, thus assessing the effectiveness and outcome of treatment. The foot and ankle ability measure (FAAM) and lower extremity functional scale (LEFS) are examples of evaluative instruments. Information acquired from an evaluative instrument is useful only if evidence is available to support the interpretation of scores obtained in the specific population in which the instrument is intended to be used.

### **3.2 Practicality and feasibility**

The main factors when considering practicality are the following:

1. The time expended to self-answer or to administer the questionnaire
2. The necessity of formal training or prior experience
3. Special or specific set up
4. Scoring method and use of electronic devices or software

Although patients may appear to have time to complete self-report measures in the waiting room before being seen for therapy, lengthy or numerous self-report forms may interfere with patient care. Some people may fatigue while completing self-report forms, and this fatigue could influence their responses. When selecting a self-report instrument, clinicians should pay attention on the time needed to fulfill the questionnaire. Usually, the authors report this time in scientific publications or in user's manual, when applicable. For a self-administered questionnaire, ideally no more than 10 min should be needed to answer all items.

Another important point to consider is the form of test administration. Some self-report questionnaires, especially translated versions, may require a structured interview for a proper measurement reliability. Although this procedure is likely to increase the time needed to fill in the questionnaire, it is essential to achieve an acceptable level of accuracy in measurement. Instructions for taking a test are sometimes not sufficient, and special training and experience may be necessary. Usually, there is no need for special training, but test familiarization and reading the user's guide, when applicable and suggested by the test developers, may be helpful. A great advantage of self-report measures is that no equipment or specific setups, and no professional or support staff are needed to help with the tests.

A strong point for a clinician is the importance of immediate feedback and interpretation of test result. Consequently, a scoring method that can be done manually without any software or computer assistance is desirable. It is common practice for a test's score to be attained by just summing up the individual items score and then transforming this result into a percentage. It is important to know the correct interpretation of this value, whether 100% means full function or the worst score for functionality.

Some instruments result in a single composite score or in a composite score and subscale scores for components of the item being measured. A single composite score can be desirable for communicating findings to others and for identifying people who are at risk for chronic ankle instability, for example. A composite score can be useful for discriminative instruments once a single cutoff score is an important clinical information for a diagnosis process. However, a single composite score may not represent a comprehensive analysis about physical function or ankle-related functionality for a specific task or domain. Subscale scores for components of physical function, like activities of daily living or sport related tasks, may be more useful for planning intervention and monitoring outcomes.

### **4. Psychometric properties**

In order to ascertain that a questionnaire has proper methodological quality, information about its psychometric properties must be available. Ideally, although potentially time-consuming, information about the development of the

questionnaire can sometimes be useful for a better comprehension about target population and/or medical conditions. Some author's suggest eight criteria that should be taken into account when assessing the quality of such outcome measures. These include (1) conceptual and measurement model; (2) content and construct validity; (3) reliability; (4) responsiveness; (5) floor and ceiling effects; (6) internal consistency; (7) feasibility to answer, administer, and interpret; and (8) cultural and language adaptations (translations). Reference values for each of these variables have also been suggested aiming to help clinicians and researchers in the selection and use of the clinical assessment tool that best suits their necessity.

For the purposes of this chapter, four basic variables will be addressed in detail. They contain the minimum information needed to select and use a self-report questionnaire. They are validity, reliability, internal consistency, and responsiveness.

#### 4.1 Validity

Content validity examines the extent to which the concepts of interest are comprehensively represented by the items in the questionnaire. It is very important to know about the following aspects regarding the development of a questionnaire for an appropriate judgment of content validity:

- **Measurement aim of the questionnaire:** it can be discriminative, evaluative, or predictive. Different items may be valid for different aims.
- **Target population:** it indicates whether the items were at the appropriate level of difficulty for the sample or the population for which the questionnaire was developed. If a questionnaire is intended to measure the functional status of patients with ankle/foot disorders, it is expected that items like *standing on tiptoes* should be much more relevant for such group than it would be for patients with knee problems. Nevertheless item's appropriateness, also the item's difficulty level, is another issue to be considered. Different populations demand different outcome measures. Ankle-related functionality of volleyball professional players, for example, requires items that measure function in a higher level of ability, with more challenging functional tasks, such as *jumping* and *landing*. In sum, a detailed description of the target population is crucial for judging the comprehensiveness and the applicability of the questionnaire for a given population.
- **Concepts:** for what the questionnaire was developed to measure. Clinicians must be aware about the relevant concepts that a questionnaire is able to measure. Quality of life, functionality, and symptoms are examples of different concepts a questionnaire may assess. These different outcome levels should clearly be distinguished and measured by separate subscales. Self-report instruments measures at the level of individual's capacity, that is, what he/she thinks they are able to do. Functional scales usually assess individual's performance, which is what he/she actually can do.
- **Item selection and item reduction:** a thorough list of potential items relating to symptoms, signs, and limitations can be gathered from literature review and input from expert clinicians (in the case of ankle-related functionality, from physicians, surgeons, and physical therapists) who treat individuals with foot and ankle-related disorders. Another important source of information is individuals with musculoskeletal pathologies within this scope. A common procedure is ask for all these people to rate each potential item from  $-2$  (not important) to  $+2$  (very important), and after that, reject all items with a score below  $+1$ .

- **Interpretability of the items:** completing the questionnaire should not require reading skills beyond that of a 12-year-old to avoid missing values and unreliable answers. To meet this recommendation, items should be as short as possible and, written with friendly vocabulary, understandable for a layperson out of health area. Another two points are direct questions, one attribute at a time, and direct reference about the time frame to which the questionnaire refers to.

Evidence for construct validity includes how the scores on the instrument relate to other measures of the construct, in a manner that is consistent with theoretically derived hypotheses concerning the concepts that are being measured. Construct validity should be assessed by testing predefined hypotheses. When testing expected correlations between measures, this can be called convergent validity or divergent validity when dealing with expected differences in scores between “known” groups.

Similar to construct validity is the criterion validity, which refers to the extent to which scores on a particular instrument relate to a gold standard. In a situation where there is no gold standard test or, at least, a well-established measurement tool for a given clinical condition, the analysis of criterion validity can become quite challenging. In these cases, face validity can be achieved by the process of item selection and item reduction. This indicates whether a measure appears to have been designed to measure what it is supposed to measure, in case, ankle-related functionality. Face validity, while contributing to the validity of the data obtained with a measure, is not represented by the outcome of a statistical test but by the judgment of the tester to make sure the measure has been used under similar conditions of measurement.

Evidence of validity is the first step when choosing an instrument to assess and interpret the effect of pathology and subsequent impairment on physical function, as well as to compare clinical intervention effectiveness.

## 4.2 Reliability

Reliability relates to score stability, and it concerns the degree to which patients can be distinguished from each other, despite measurement error. Reliability coefficients such intra-class correlation coefficient (ICC) take into account three sources of variation, that is:

- The variation among individuals, also known as interindividual variation
- The personal variation, which is the same as intraindividual variation
- At last, a variation that combines those previous mentioned, which is the error attributed to the measurement itself (measurement error)

Index like ICC is used for continuous measures and is expressed as a ration between 0 (low reliability) and 1 (high reliability). High reliability is particularly important for discriminative purposes because the difference observed in a measure should be a perfect reflection of a real change and not overlapped or shadowed by any sources of error.

Authors and test developers should provide clear information about which reliability measure they have used; if ICC is the case, the two-way random effect model is the best option for the far majority of cases. Pearson correlation coefficient is inadequate, because systematic differences are not considered. The correspondent of ICC for ordinal measures is the weighted Cohen’s kappa coefficient, which is so

the preferred option for such variables. In groups or samples with 50 subjects or above, the value of 0.70 is the minimum recommendation for both indexes.

### **4.3 Internal consistency**

When adding up items with the purpose to measure, a construct is very important to know if those items are well correlated with each other and with the total score generated by them. In other words, it is highly desirable to know if the instrument is homogenous or unidimensional and so if the questionnaire as a whole measure the same concept or construct. This measure of unidimensionality is the internal consistency and is quite presumable that it should be as high or as good as possible.

There are many ways to measure internal consistency. Usually they complement each other in the process of measuring unidimensionality. The principal component analysis or the factor analysis is both very good ways to determine whether the items form only one overall dimension or not. Also confirmatory or exploratory analysis, when applicable, is useful to determine if a given group of items measure one same construct, and therefore are grouped in one scale, or if it would be better to join the items in two or more subscales. The Rasch model is also a way to measure internal consistency by using the fit statistics, which is used to assess unidimensionality, and can be simply explained as a ratio between the observed response and the response predicted by the model. This analysis is also important for evidence of construct validity, which will be better explained in the proper section ahead in this chapter.

Once the scale(s) is(are) defined, then the Cronbach's alpha is the appropriate measure of choice. Here we have two possible situations:

- A very low Cronbach's alpha indicates that there is no reason to group the items together in a same scale or questionnaire, because there are not well correlated with each other
- On the other hand, a very high Cronbach's alpha suggests that maybe there are redundant items, which means that they measure almost the same attribute of functionality. When this happens it is valuable to judge if one or more items could be removed from the questionnaire.

Cronbach's alpha should be interpreted with caution when applied to questionnaires with too many items, approximately more than 20 items. In these cases, the index is usually very high, because Cronbach's alpha is dependent upon the number of items in a scale. The reference value for adequate internal consistency when using the Cronbach's alpha range between 0.70 and 0.95.

### **4.4 Responsiveness**

A large number of definitions and methods have been proposed for assessing responsiveness. A very good comprehensive definition for responsiveness is the ability of the instrument to detect clinically important changes in an individual's status over time even when these changes are small. This ability is the accuracy of the instrument that must be able to differentiate clinical observed changes from measurement error. Even though an instrument can capture very small changes, what really matters is to know if a change is clinically relevant. The Guyatt's responsiveness ratio (RR) does precisely this comparison by relating the variability found within the subject with between the subjects. The reference value for RR is 1.96 because this happens when the minimal important change equals the smallest detectable change.



Another adequate and common measure of responsiveness is the area under the receiver-operating characteristics (ROC) curve. It is very useful to define cutoff scores for discriminative purposes and to define injury severity. The reference value for the area under the curve is at least 0.70.

One point that impact negatively on responsiveness is the presence of floor or ceiling effects. They are considered to be present when more than 15% of respondents achieved the lowest or highest possible score. Thus, the responsiveness is limited because changes cannot be measured in these patients nor is it possible to distinguish one from another, which compromises reliability.

Limitations in measurements, such as ceiling or floor effects, can usually be avoided by selecting measures that have been demonstrated to provide meaningful information about people who are similar to those being measured. In other words, the target population of each measurement tool must be considered by matching the sample, e.g., patients with the appropriate questionnaire or functional scale.

## 5. “Traditional statistics” and item response theory (Rasch and factor analysis)

The Rasch model and the factor analysis constitute two ways of assessing psychometric properties of an instrument and can be, and frequently are, used in functional scales development. These two statistical procedures have the same theoretical model, which is the item response theory. The basic concept behind IRT is that the probability of choosing a response for each item is a function of both the subject's or patient's ability and the difficulty level of each item.

When applying the concepts of IRT to psychometric properties analysis, it is possible to obtain more detailed information about validity, accuracy, and targeting that helps understanding the clinical meaning of a self-report instrument. It goes beyond just looking at the final score of a questionnaire or at cutoff scores. This closer look at outcome measures like functional scales adds information to those obtained by traditional statistical tests, e.g., Cronbach alpha or ICC. IRT not only improves the methodological quality when elaborating new instruments but gives clearly insights into effects of intervention as well, whether comparing groups or the subject longitudinally.

Rasch analysis can be applied to examine instruments or assessment scales applicable in wide spectrum of disciplines, including studies in health area, education, marketing, economy, and social sciences. In the majority of evaluations, a well-defined group is selected to answer a series of predefined items. The Rasch model offers a mathematical theoretical reference by which researches that elaborate instruments are able to create comparable measures. The main point behind this model is the concept of unidimensionality, which can be summarized by the idea that useful clinical measures involve the analysis of only one human attribute at the time. In other words, it implies that the instrument measures a single latent ability. Taking a self-report questionnaire as an example, this would mean the items are organized according to their difficulty level and are placed in a single linear hierarchic scale.

The Rasch model transforms ordinal scales into interval measures. This process allows us to calibrate item difficulty and subject's ability in a same linear *continuum*, which is divided into equal intervals or *logits*. The *logits* is defined by items and works similarly as a ruler on which individuals are organized accordingly to their level of ability. The probabilistic model of Rasch analysis can be defined by the following formula:

$$P_{ni}(x = 1) = f(B_n - D_i)$$

where  $P$  is the probability of an “ $n$ ” individual to succeed on a given event “ $i$ ” in any trial. This probability equals to the mathematical function  $f$  of the subtraction from the “ $n$ ” individual’s ability “ $B$ ” in relation to the “ $i$ ” item’s difficulty level “ $D$ .” This probability can be extrapolated for multilevel items, i.e., for non-dichotomous responses. As a result of this procedure under IRT concepts, an item characteristic curve can be drawn, which represents the probability of choosing a response for each item based on the subject’s or patient’s ability. A typical item characteristic curve is defined by two properties: item’s difficulty and item’s discrimination power. Taking back the ruler analogy cited above, the difficulty of an item functions as a location index that is where in the *continuum* of ability the item works better. Hard items function with high-ability individuals as well as easy items do the same with low-ability subjects. By discrimination power, it means how well an item can separate individuals whose abilities are below or above the item location. Graphically, this property appears as the steepness of the item characteristic curve and can be interpreted as the steeper the greater the discrimination power. A flat curve means that the probability of a right answer is nearly the same with low or high levels of ability. It is worthwhile to stress out that this two properties only describe the form of item characteristic curve, and consequently how well an item function, but it cannot be used as a proof of item validity. Applying these concepts for multilevel item, Likert-like scale, for example, each answer possibility would have its own curve with distinct peaks. All the curves together should measure the spectrum of ability measured by the item.

If all items meet this probabilistic expectation, it is possible to state that the questionnaire, as a whole, assesses an unidimensional construct. This probabilistic framework constitutes the basis of Rasch model and thereby makes it possible to organize items by their difficulty level as well as by the patients’ ability level, both based on the observed answering pattern.

Questionnaires should be responsive to changes in the status of the patient across the spectrum of ability. Another benefit of IRT is that it provides the amount of information that each item contributes at varying levels of ability. Easy items should provide information among low-ability levels examinees, and conversely, hard items that describe difficult tasks give information among high-ability examinees. The questionnaire final score is, therefore, the sum of all these information collected by each item, and the accuracy of it is directly proportional to the amount of information provided. The target of an evaluative instrument is to provide information across all ability ranges. Therefore, such an appropriate evaluative questionnaire should contain items that assess an individual’s ability to perform activities that span from easy to more challenging ones.

The results of an item characteristic curve are valuable only when the following requirements are met:

1. Unidimensionality

- a. The questionnaire measures a single latent trait.

2. Local independence

- a. The answer for each item is independent from another item.

3. No time constraints

- a. Should be no time limit or restriction when answering the test.

#### 4. No guessing as an answer

- a. A correct answer may not be due to guessing but reflect the person's ability.

This implies that only one latent ability accounts for the individual's response for each of the items contained on the instrument, which is exactly the unidimensionality mentioned throughout this section. Both factor analysis and the Rasch model can ascertain this aspect of construct validity. Those items that did not fit to the model should be revised or eliminated accordingly with scale's goals.

## 6. Questionnaires and functional scales

Measures should be chosen based on whether they have been designed for and have been used with people similar to the people to be measured. For example, to assess an elite athlete's functionality after an ankle sprain, sport subscale of the foot and ankle ability measure questionnaire should provide more clinically useful information than the whole lower extremity functional scale (LEFS). This is not because one instrument is better than the other is, but because it is a better instrument selection to the target population or the right patient.

The next section shows four self-report questionnaires that are currently available for measuring ankle-related functionality. Some systematic reviews report that these instruments have very good psychometric properties and are quite useful for clinical scenarios as well as for research contexts. A brief overview for each of them is provided. For full information, we suggest to read the original studies about their development.

### 6.1 Lower extremity functional scale

The LEFS is a measure of activity limitation developed for musculoskeletal conditions of the lower extremity. On this scale, participants rate the difficulty in performing 20 activities of the lower extremity on a 5-point Likert scale, rating grade 0, meaning "extreme difficulty or unable to perform activity," to grade 4, meaning "no difficulty". The responses are summed to give a score ranging from 0 to 80, with 0 indicating high functional limitation and 80 indicating low functional restriction. LEFS was tested in a heterogeneous population with different lower limb conditions and was found to have high internal consistency (.96) and high test-retest reliability ( $r = .86$ ) and correlated well with the physical function subscale and the physical component summary scores of the medical outcomes study 36-item short-form health survey ( $r = .80$  and  $.64$ , respectively).

### 6.2 Foot and ankle ability measure

The FAAM is composed of two subscales named activities of daily living (ADL) and sports subscale, respectively. ADL subscale has 21 items and the sports subscale 8 items. Each item is scored on a 5-point Likert scale anchored by 4 (no difficulty at all) and 0 (unable to do). Item score totals, which range from 0 to 84 for the ADL subscale and from 0 to 32 for the sports subscale, are transformed to percentage scores. A higher score represents a higher level of function for each subscale.

### 6.3 Foot and ankle outcome score

The FAOS is a 42-item questionnaire divided into 5 subscales: "pain" (9 items), "other symptoms" (7 items), "activities of daily living" (17 items), "sport and

recreation function” (5 items), and “foot- and ankle-related quality of life” (4 items). Each question can be scored on a 5-point Likert scale (from zero to four) and each of the five subscale scores is calculated as the sum of the items included. Raw scores are then transformed to a 0 to 100, worst to best score.

#### 6.4 Cumberland ankle instability tool

The CAIT is a nine-item questionnaire designed to be a discriminative instrument of chronic ankle instability. The questionnaire is structured so that the feeling of instability is reported for different types of activities such as running, walking, hopping, and descending stairs. The nine items generate a total score from 0 to 30 for each foot, in which 0 is the worst possible score, meaning severe instability, and 30 is the best possible score, meaning normal stability. The CAIT is a reliable (ICC 0.96) instrument that can discriminate stable from unstable ankles and measure the severity of functional ankle instability.

**Table 1** summarizes the main psychometric properties of each instruments reported above.

	LEFS	FAAM	FAOS	CAIT
<b>Validity</b>	92% of variance explained at baseline Concurrent validity $r = .80$ and $.87$ , at short and medium term follow up, respectively (correlation with Olerud-Molander ankle score)	Experts and patients were involved in item generation and reduction	Item selection and reduction by patients ( $n = 213$ ) Experts: not involved	CAIT and LEFS ( $\alpha = .50$ , $P < .01$ ) and VAS ( $\alpha .76$ , $P < .01$ ) Construct validity and internal reliability were acceptable ( $\alpha = .83$ ; point measure correlation for all items, $>0.5$ ; item reliability index, $.99$ )
<b>Reliability</b>	No information	ADL subscale: ICC = $.89$ ; SEM = 2,1 points Sport subscale: ICC = $.87$ ; SEM = 4,5 points	Subscale pain, $rs = .96$ ; subscale symptoms, $rs = .89$ ; subscale ADL, $rs = .85$ ; subscale sports, $rs = .92$ ; subscale quality of life, $rs = .92$	ICC, 0,96
<b>Internal consistency</b>	$\alpha = .92$ at base line; $.94$ short term; $.90$ long term.	Cronbach alpha for ADL subscale, $\alpha = .96$ in stable group ( $n = 79$ ); in changed group, $\alpha = .98$ ( $n = 164$ ) Cronbach alpha for sport subscale from a combined sample, $\alpha = .98$	Subscale pain, $\alpha = .94$ ; subscale symptoms, $\alpha = .88$ ; subscale ADL, $\alpha = .97$ ; subscale sports, $\alpha = .94$ ; subscale “quality of life,” $\alpha = .92$	The threshold CAIT score was 27.5 (Youden index, 68.1); sensitivity was 82.9% and specificity was 74.7%.
<b>Responsiveness</b>	Guyatt = 1.99 AUC ROC = 0.79 (95% CI = 0.70-0.88)	MDC ADL subscale, 5.7 MDC Sport subscale, 12.3	No information	No information

**Table 1.**  
 Main psychometric properties of ankle-related self-report measures.

## A. Appendix

### Foot and Ankle Ability Measure (FAAM)

Please answer **every question** with one response that most closely describes to your condition within the past week.

If the activity in question is limited by something other than your foot or ankle mark not applicable (N/A).

	No difficulty	Slight difficulty	Moderate difficulty	Extreme difficulty	Unable to do	N/A
Standing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Walking on even ground	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Walking on even ground without shoes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Walking up hills	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Walking down hills	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Going up stairs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Going down stairs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Walking on uneven ground	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Stepping up and down curbs	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Squatting	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Coming up on your toes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Walking initially	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Walking 5 minutes or less	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Walking approximately 10 minutes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Walking 15 minutes or greater	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

©2000 RobRoy Martin

Because of your **foot and ankle** how much difficulty do you have with:

	No difficulty at all	Slight difficulty	Moderate difficulty	Extreme difficulty	Unable to do	N/A
Home Responsibilities	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Activities of daily living	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Personal care	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Light to moderate work (standing, walking)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Heavy work (push/pulling, climbing, carrying)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Recreational activities	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

How would you rate your current level of function during your usual activities of daily living from 0 to 100 with 100 being your level of function prior to your foot or ankle problem and 0 being the inability to perform any of your usual daily activities?

.0 %

[ ]

FAAM Sports Scale

Because of your **foot and ankle** how much difficulty do you have with:

	No difficulty at all	Slight difficulty	Moderate difficulty	Extreme difficulty	Unable to do	N/A
Running	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Jumping	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Landing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Starting and stopping quickly	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cutting/lateral movements	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Low impact activities	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ability to perform activity with your normal technique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Ability to participate in your desired sport as long as you would like	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

How would you rate your current level of function during your sports related activities from 0 to 100 with 100 being your level of function prior to your foot or ankle problem and 0 being the inability to perform any of your usual daily activities?

.0 %

Overall, how would you rate your current level of function?

- Normal    
  Nearly normal    
  Abnormal    
  Severely abnormal

IntechOpen

IntechOpen

### **Author details**

Tarcísio Santos Moreira  
Master of Rehabilitation Science, Universidade Federal de Minas Gerais,  
Belo Horizonte, Minas Gerais, Brazil

\*Address all correspondence to: [tarcisiosam@gmail.com](mailto:tarcisiosam@gmail.com)

### **IntechOpen**

---

© 2019 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 



## References

- [1] Hale SA, Hertel J. Reliability and sensitivity of the foot and ankle disability index in subjects with chronic ankle instability. *Journal of Athletic Training*. 2005;**40**(1):35-40
- [2] Moreira TS, Sabino GS, Resende MA. Instrumentos clínicos de avaliação funcional do tornozelo: Revisão sistemática. *Fisioter Pesqui*. 2010;**17**:88-93
- [3] Sabino GS, Coelho CM, Sampaio RF. Utilização da Classificação Internacional de Funcionalidade, Incapacidade e Saúde na avaliação fisioterapêutica de indivíduos com problemas musculoesqueléticos nos membros inferiores e região lombar. *Acta Fisiátrica*. 2008;**15**(1):24-30
- [4] Steiner WA, Ryser L, Huber E. Use of ICF model as a clinical problem solving tool in physical therapy and rehabilitation medicine. *Physical Therapy*. 2002;**82**:1098-1107
- [5] Carcia CR, Martin RL, Drouin JM. Validity of the Foot and Ankle Ability Measure in athletes with chronic ankle instability. *Journal of Athletic Training*. 2008;**43**:179-183
- [6] Vanswearingen JM, Brach JS. Making geriatric assessment work: Selecting useful measures. *Physical Therapy*. 2001;**81**(6):1233-1251
- [7] Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*. 2007;**60**:34-42
- [8] Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Medical Research Methodology*. 2010;**10**(22)
- [9] Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health related patient-reported outcomes. *Journal of Clinical Epidemiology*. 2010;**63**:737-745
- [10] Martin RL, Irrgang JJ, Burdett RG, et al. Evidence of validity for the foot and ankle ability measure (FAAM). *Foot & Ankle International*. 2005;**26**:968-983
- [11] Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. Protocol of the COSMIN study: CONsensus-based standards for the selection of health Measurement INSTRUMENTS. *BMC Medical Research Methodology*. 2006;**6**:2
- [12] Baker FB. The item characteristic curve. In: Boston C, Rudner L, editors. *The Basics of Item Response Theory*. 2nd ed. College Park (MD): ERIC Clearinghouse on Assessment and Evaluation; 2001. pp. 5-11
- [13] Bond TG, Fox CM. Rasch modeling applied: Rating scale design. In: Bond TG, Fox CM, editors. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. 2nd ed. Mahwah (NJ): Lawrence Erlbaum Associates; 2007. pp. 219-234
- [14] Binkley JM, Stratford PW, Lott SA, et al. The lower extremity functional scale: Scale development measurement properties and clinical application. *Physical Therapy*. 1999;**79**(4):371-383
- [15] Moreira TS, Magalhães LC, Silva RD, Martin RL, Resende MA. Translation, cross-cultural adaptation and validity of the foot and ankle ability measure questionnaire. *Disability and Rehabilitation*. 2016;**15**:1-12
- [16] Roos EM, Brandsson MD, Karlsson J. Validation of the foot and

ankle outcome score. *Foot & Ankle International*. 2001;**22**(10):788-794

[17] Hiller CE, Refshauge KM, Bundy AC, et al. The Cumberland ankle instability tool: A report of validity and reliability testing. *Archives of Physical Medicine and Rehabilitation*. 2006;**87**:1235-1241

IntechOpen

IntechOpen