ORIGINAL ARTICLE

Classical Test Theory and Rasch Analysis Validation of the Upper Limb Functional Index in Subjects With Upper Limb Musculoskeletal Disorders

Elisabetta Bravini, PT,[a] Franco Franchignoni, MD,[b] Andrea Giordano, PhD,[c] Francesco Sartorio, MSc, PT,[b] Giorgio Ferriero, MD, PhD,[b] Stefano Vercelli, PhD, PT,[b] Calogero Foti, MD[d]

From the [a]School in Advanced Sciences and Technology in Rehabilitation Medicine and Sports, Tor Vergata University, Rome; [b]Unit of Occupational Rehabilitation and Ergonomics, Salvatore Maugeri Foundation, Scientific Institute of Veruno, Veruno; [c]Unit of Bioengineering, Salvatore Maugeri Foundation, Scientific Institute of Veruno, Veruno; and [d]Department of Physical and Rehabilitation Medicine, Tor Vergata University Hospital, Rome, Italy.

Abstract

Objective: To perform a comprehensive analysis of the psychometric properties and dimensionality of the Upper Limb Functional Index (ULFI) using both classical test theory and Rasch analysis (RA).

Design: Prospective, single-group observational design.

Setting: Freestanding rehabilitation center.

Participants: Convenience sample of Italian-speaking subjects with upper limb musculoskeletal disorders (N=174).

Interventions: Not applicable.

Main Outcome Measures: The Italian version of the ULFI. Data were analyzed using parallel analysis, exploratory factor analysis, and RA for evaluating dimensionality, functioning of rating scale categories, item fit, hierarchy of item difficulties, and reliability indices.

Results: Parallel analysis revealed 2 factors explaining 32.5% and 10.7% of the response variance. RA confirmed the failure of the unidimensionality assumption, and 6 items out of the 25 misfitted the Rasch model. When the analysis was rerun excluding the misfitting items, the scale showed acceptable fit values, loading meaningfully to a single factor. Item separation reliability and person separation reliability were .98 and .89, respectively. Cronbach alpha was .92.

Conclusions: RA revealed weakness of the scale concerning dimensionality and internal construct validity. However, a set of 19 ULFI items defined through the statistical process demonstrated a unidimensional structure, good psychometric properties, and clinical meaningfulness. These findings represent a useful starting point for further analyses of the tool (based on modern psychometric approaches and confirmatory factor analysis) in larger samples, including different patient populations and nationalities.

Archives of Physical Medicine and Rehabilitation 2015;96:98-104

© 2015 by the American Congress of Rehabilitation Medicine

Region-specific patient-reported outcome measures of the upper limb assess the affected arm as a functional unit. This solves the problem of having to use multiple measures in patients with multiple impairments, enables comparison across different local disorders, and allows for greater uniformity in research.[1] The selection of a patient-reported outcome measure greatly depends on its content (the construct being measured) and measurement properties. To properly detect functional improvements and change over time, an outcome measure should be endowed with good psychometric properties, including evidence of reliability and validity (content validity, construct validity, responsiveness), interpretability of scores, quality of translation, and acceptable patient and investigator burden.[2]

The Upper Limb Functional Index (ULFI)[3] has been proposed as an alternative to the Disability of the Arm, Shoulder, and Hand (DASH)[4,5] and QuickDASH[6] questionnaires to assess

all types of upper limb musculoskeletal disorders. Recently, the new ULFI version with a 3-point response showed good clinimetric characteristics[7] in English-speaking individuals when analyzed by means of classical test theory (CTT). However, to our knowledge, a Rasch analysis (RA) of this scale has never been performed. Indeed, the advantages of using new approaches (eg, RA) in the development and evaluation of clinical tools for health care have been clearly documented in the literature.[8,9] As an example, RA examines how well each item performs in terms of its relevance or usefulness for measuring the underlying construct, the amount of the construct targeted by each question, the possible redundancy of an item relative to other items in the scale, and the appropriateness of the response categories. Moreover, scales fitting the Rasch model fulfill the requirements for the main mathematical manipulations of the scores.

The aim of this study was to perform a comprehensive analysis of the psychometric properties (including dimensionality) of the ULFI Italian version,[10] using both CTT and RA to gain insights for a possible refinement of the questionnaire.

Methods

Participants

A convenience sample of 174 patients was consecutively recruited between October 2011 and September 2013 at our rehabilitation clinic. Inclusion criteria were age ≥18 years and presence of upper limb musculoskeletal problems requiring physiotherapy in a conservative or after-surgery rehabilitation program. Exclusion criteria were as follows: significant upper limb motor comorbidity (eg, as a result of stroke, Parkinson disease, polyneuropathy, cervical radiculopathy, systemic rheumatic condition) and problems with reading or understanding the Italian language.

This study was approved by the Salvatore Maugeri Foundation Institutional Review Board and the local ethics committee and was undertaken in compliance with the Declaration of Helsinki.

Assessment

Upper Limb Functional Index
The ULFI is an upper limb patient-reported questionnaire assessing health-related quality of life and upper extremity dysfunction.[3] In this study, the most recent version of the ULFI was adopted,[7] its Italian version,[10] which was previously produced through the process of validation and cross-cultural adaptation as suggested by Beaton et al.[11] An initial statement states the following: "Your arm may make it difficult to do some things you normally do. This list contains sentences people use to describe themselves with such problems. Think of yourself over the last few days. If an item describes you, mark the box. If not, leave the box blank. If an item partly describes you, use a half mark ($^1/_2$)." After that, the common stem "Due to my arm…" is followed by

---

List of abbreviations:

CTT	classical test theory
DASH	Disability of the Arm, Shoulder, and Hand
MnSq	mean square
PCA	principal component analysis
RA	Rasch analysis
ULFI	Upper Limb Functional Index

---

25 items, each with the aforementioned 3 response options.[3] For statistical convenience in applying factor analysis and RA, the values of the 3 response options (yes=1, partly=0.5, no=0) were converted to integers, becoming yes (2), partly (1), and no (0). The points were summed (raw score) and multiplied by 2 to give a total score ranging from 0 (no functional limitation) to 100 (maximum functional limitation).

Procedure

The 25-item ULFI Italian version[10] was administered to all patients immediately before a physical therapy program for upper limb musculoskeletal disorders. Patients were required to answer all questions; therefore, there were no missing values. The questionnaire was scored as recommended by its developers.[7]

Statistical analysis

Internal consistency
The items of the ULFI Italian version were assessed for internal consistency by calculating the Cronbach alpha (a minimum of .90 is recommended for clinical application in individuals)[12] and item-to-total correlation, omitting that item from the total (a correlation of .40 was chosen as the minimum value for satisfactory item-to-total correlation).[13] SPSS version 16.0[a] was used to analyze these data.

Dimensionality
Because of the unclear structure of responses of the ULFI Italian version, an estimate of the number of meaningful factors in the response matrix was obtained with a parallel analysis.[14] A single factor structure is critical to ensure that all questions within the scale are reporting on the same underlying variable or construct. If >1 dimension is present, the items considered not pertaining to the latent tract being measured should be identified and then progressively dropped, until a sufficient unidimensional score is obtained (in the current study, this process was carried out using RA, which will be subsequently discussed). Finally, after the interventions on the questionnaire suggested by RA, a new parallel analysis and an exploratory factor analysis appropriate for ordinal data[15] were performed to confirm the unidimensionality of the set of the 19 remaining items. For parallel and exploratory factor analysis, FACTOR 8.1 software[b] was used.[16]

Rasch analysis
The responses to the ULFI Italian version items underwent RA (rating scale model)[17] using WINSTEPS software 3.68.2[c] according to the following steps. The first step was scale diagnostics. We investigated whether the response categories of the scale were being used effectively and consistently, according to the guidelines suggested by Linacre.[18,19] In the second step, internal construct validity was assessed by determining how well the empirical data fit the Rasch model.[20,21] Infit and outfit mean square (MnSq) statistics were calculated, considering (according to our sample size) a MnSq between 0.8 and 1.2 as an indicator of acceptable fit.[22] Items outside this range were considered as underfitting (MnSq>1.2) or overfitting (MnSq<0.8).[23] We used fit statistics to identify problematic items; then, expert judgment was used to decide which ones to omit. Moreover, the Rasch model provides estimates of the level of difficulty achieved by each item (item difficulty measure) and estimates of the global amount of latent traits in each individual (ability). These 2 parameters are expressed on a common interval scale in logit units,