Translation and cultural validation of clinical observational scales – the Fugl-Meyer assessment for post stroke sensorimotor function in Colombian Spanish

Nubia E. Barbosa, Sandra M. Forero, Claudia P. Galeano, Edgar D. Hernández, Nancy S. Landinez, Katharina S. Sunnerhagen & Margit Alt Murphy

To cite this article: Nubia E. Barbosa, Sandra M. Forero, Claudia P. Galeano, Edgar D. Hernández, Nancy S. Landinez, Katharina S. Sunnerhagen & Margit Alt Murphy (2018): Translation and cultural validation of clinical observational scales – the Fugl-Meyer assessment for post stroke sensorimotor function in Colombian Spanish, Disability and Rehabilitation, DOI: 10.1080/09638288.2018.1464604

To link to this article: https://doi.org/10.1080/09638288.2018.1464604
Translation and cultural validation of clinical observational scales – the Fugl-Meyer assessment for post stroke sensorimotor function in Colombian Spanish

Nubia E. Barbosa, Sandra M. Forero, Claudia P. Galeano, Edgar D. Hernández, Nancy S. Landinez, Katharina S. Sunnerhagen and Margit Alt Murphy

ABSTRACT

Purpose: Fugl-Meyer Assessment (FMA) is the most widely used and recommended clinical scale for evaluation of sensorimotor impairment post stroke, but an official Spanish version is not available today. This study aimed to establish methodological structure for translation and cultural adaptation process and perform a transcultural validation of the upper and lower extremity FMA to Colombian Spanish.

Methods: Procedures included forward and backward translation, step-wise reviewing by bilingual and professional experts to ensure conceptual and semantic equivalence. Validation included a pilot evaluation of item-level agreement on 10 individuals with stroke at the Central Military Hospital of Colombia.

Results: Comprehensive step-wise procedure for transcultural validation was established. Low agreement (less than 70%) was detected for items assessing arm movements within synergies and for coordination/speed subscale. All points of disagreement were systematically reviewed and agreed upon when drafting the final version of the Spanish FMA.

Conclusions: Use of FMA will allow unified description of stroke severity and motor recovery in Spanish speaking countries. This will open up possibility to compare stroke and rehabilitation outcomes with other countries and regions world-wide. Comprehensive methodological procedures provided can facilitate introduction of well-established clinical scales in other languages.

IMPLICATIONS FOR REHABILITATION

• The Fugl-Meyer Assessment (FMA) of upper and lower extremity is the most used and recommended clinical scale for evaluation of sensorimotor impairment after stroke.
• The Spanish version of FMA, validated in this study, is now first time available for use in research and clinical practice.
• Use of FMA will allow unified description of stroke severity and motor recovery in Spanish speaking countries, which in turn opens up possibility to compare stroke and rehabilitation outcomes with other countries and regions world-wide.

Introduction

Deficits in motor functions are the most widely recognized impairments and the main cause of disability post stroke [1]. The most frequent motor sequel after stroke is hemiparesis, which hampers the ability to undertake, maintain, modify, and control movements of the contralesional side. Upper and lower extremity motor function influence activities of daily living, balance, gait and mobility, which all are identified as top 10 research priorities by the stroke survivors, caregivers, and health professionals [2].

Evaluation of motor function is essential for understanding the mechanisms of motor control and motor learning. Multiple clinical scales are available to determine the functional ability and motor function in individuals with hemiparesis [3,4]. Use of standardized and validated scales allows more precise assessment of the degree of impairment and will allow more optimal prediction and evaluation of rehabilitation outcome [5,6]. These scales should be robust, easily administered, valid, reliable, and able to capture change over time [3,4].

The Fugl-Meyer Assessment (FMA) developed and introduced in 1975 by Fugl-Meyer et al. [7], was originally published both in English and Swedish. It was the first quantitative instrument for evaluation of hemiparetic patients. FMA is valid, reliable, responsive, and most widely used standardized clinical scale for evaluation of sensorimotor function after stroke [3,4,8]. FMA is increasingly used in research in many countries but most in Canada, Italy, Japan, the Netherlands, and USA [8]. The original FMA consists of several parts including evaluation of motor function, balance, sensation, passive range of motion, and pain [7], but the motor evaluation of upper and lower extremity have been most used [8]. The motor domains of the FMA are used...
extensively [9–12] and show excellent reliability, validity, and good responsiveness [13–16]. The motor domains of the FMA are recommended for assessment of the level of sensorimotor function and for evaluation of the progress during motor recovery [6,11].

Currently, there is no official translation of the FMA into Spanish available. Although there are recommendations for translation and cultural adaptation process for patient-reported outcome measures [17,18], standardized procedures for observational clinical ordinal scales have only been described sparsely [19]. The Spanish translation and cultural validation of the FMA would significantly improve the sensorimotor assessment procedures for the patients with stroke in Spanish speaking countries and in this way make comparisons of patient populations with other countries possible. The motor domains of the FMA have shown the highest validity, reliability, and responsiveness for persons with stroke, and therefore only these parts will be considered in the current study.

The aim of this study was to perform a transcultural translation of the motor function domains of the FMA for upper extremity and lower extremity into Colombian Spanish, and perform a cultural validation of this translation in a sample of stroke patients admitted to the Central Military Hospital of Colombia. The study aimed to establish the methodological structure and process of the translation process applicable for clinical observational rating scales.

Methods

Study design

The translation and cultural validation process followed a standardized forward and backward translation protocol to determine the conceptual, semantic, and operational equivalence [17–20]. A rigorous multistep structure was applied for reviewing, revising, and drafting of the final translation. This process included official translators, bilingual clinical experts, experts with clinical and research expertise within the Central Military Hospital of Colombia and National University of Colombia as well as experts familiar with the original scale at University of Gothenburg. The translated version of the FMA was subsequently tested in a pilot trial by three experienced physiotherapists at the Central Military Hospital of Colombia in order to identify inconsistencies in wording that might influence the scoring, understanding, interpretation, and cultural application.

The Fugl-Meyer assessment

The FMA is an observational rating scale with ordinal data to assess the sensorimotor impairment in individuals who have had a stroke [7]. The scale is commonly used by physical therapists, but also by other health professionals with adequate training. The motor domain of the Fugl-Meyer Assessment for upper extremity (FMA-UE) includes four subscales: A. Upper Extremity (0–36), B. Wrist (0–10), C. Hand (0–14), D. Coordination/Speed (0–6) composing a total maximum score of 66 points. The FMA for lower extremity (FMA-LE) includes two subscales: E. Lower Extremity (0–28) and F. Coordination/Speed (0–6) composing a total score of 34 points. The rating is based on direct observation of the motor performance at each item using a 3-point ordinal scale (0 = cannot perform, 1 = performs partially, and 2 = performs fully) [7].

Forward translation and drafting the first version

Permission for translation was sought and approved from the University of Gothenburg. The translation process was initiated through a collaboration workshop at the University of Gothenburg that incorporated educational course on measurement methodology of the FMA scale.

The FMA-UE and FMA-LE test protocols and application manuals were translated to Colombian Spanish, based on the original version in English and Swedish. Two official experienced translators (native speakers of the target language and fluent in English or Swedish), certified by the Colombian Foreign Relations Ministry, performed the respective translations, emphasizing the conceptual translation. These translations were compared by the Colombian researchers in collaboration with Swedish researchers and the discrepancies were discussed until consensus was met for the first translated version (Version 1) in Colombian Spanish (Figure 1).

Backward translation, review and drafting of the second version

The first Spanish version (Version 1) was independently translated back to English by an official bilingual translator (native speaker of English and fluent in Colombian Spanish) to ensure that the content and the translation corresponded to the semantic characteristics of the original version.

Three bilingual experts (physical therapists) compared the original and back translation of the scale by rating whether the
Spanish translation was understandable and corresponded to the original English version in terms of conceptual and cultural equivalence. A 3-level Likert scale (3 = complete agreement, 2 = partial disagreement, 1 = complete disagreement) was used. When an item was scored as partial or complete disagreement, the experts had to record their observations and provide recommendations for improvement.

The Colombian researches performed the second review of the first Spanish version based on the bilingual experts’ ratings and recommendations until consensus was reached. The Swedish researchers familiar with the original FMA were consulted to ensure equivalence with the original version. A revised version (Version 2) of the FMA protocol and application manual was drafted for the pilot trial.

**Procedures for the pilot trial**

A pilot trial including 10 patients with stroke was carried out in order to pretest the translated Spanish version and support the sample size calculation for a larger reliability study. The inclusion criteria were: first ever stroke demonstrating sequelae of hemiplegia or hemiparesis admitted to the Central Military Hospital of Colombia, hospitalized at 9 and 10 days post stroke, age between 18 and 90 years. Exclusion criteria were: other disorders such as blindness, deafness, amputation of lower or upper limb, cerebellar stroke or multiple stroke episodes, declined cognitive function that would limit cooperation in FMA testing. Data collection for the pilot trial began in August 2014, and was ended in November 2014. Ethical approval was received from the ethical committee of the Central Military Hospital (Act No. 9, June 12, 2013) and a signed informed consent was obtained from all participants or their family members.

Background information, including demographic data, stroke diagnosis and disability (modified Ranking Scale) were recorded. The patient’s performance on the FMA was simultaneously but independently assessed by two trained physical therapists on two consecutive days (9 and 10 days post stroke). During the first assessment one of the assessors was acting as test leader (i.e., instructing the patient and scoring) and the other as observer (scoring by observing). These roles were switched on the second assessment day. The examiners did not communicate during the testing session or afterwards regarding the scoring. The scoring protocols of different colors were used for different days, and the filled protocols were stored in separate containers. This procedure was designed for a larger inter- and intra-rater reliability study and was pretested in this pilot study.

The statistical methodology of Elisabeth Svensson [21] was employed for the analysis of the inter-rater and intra-rater reliability of the pilot data. This method was selected because it uses a ranking approach and is particularly designed for analysis of disagreements in paired ordinal data. The degree of agreement was assessed at the item level using the percentage of disagreements in paired ordinal data. The degree of agreement uses a ranking approach and is particularly designed for analysis of relative position, concentration, and rank variation that would limit cooperation in FMA testing. The relative position and concentration values can vary from –1 to 1, where 0 means no difference between evaluators. Values within –0.1 and 0.1 were considered negligibly small with reference to clinical relevance, while values outside this range were considered as clinically relevant disagreement. Relative rank variation can vary from 0 to 1 and a value less than 0.1 means that the difference is negligible. Statistically significant disagreement of relative position, concentration, and rank variation values was indicated with a 95% confidence interval that did not include the value zero [21].

**Adjustments to final version**

The disagreements found in the statistical analysis were used as guidance for identification of items that might have been interpreted differently by two assessors. These items were screened to ensure whether the disagreements in those could have been caused by the translational discrepancies. All questions and uncertainties raised during the pilot trial were recorded and reviewed by the Colombian researches and confirmed with the researches at the University of Gothenburg to reach consensus and approve the final version of the protocol and application manual of the FMA.

**Results**

**Forward translation and drafting the first version**

The translations from the Swedish and English version to the Spanish showed smaller linguistic differences. In the review process, together with the researchers familiar with both the English and Swedish original FMA, it was found that the language and description of tasks were more clearly described in the English version. Some contextual modifications were made to this first Spanish version to improve comprehension. The changes primarily consisted of clarifications regarding anatomical terms and positions, like shoulder flexion and pronation/supination.

**Backward translation and drafting the second version**

The agreement ratings of the three bilingual experts between the original English version and the back translated version are reported in Table 1. Partial disagreement was rated at least by one bilingual expert for all subscales except for upper extremity Coordination/speed subscale. Complete disagreement was only rated by one expert for subscale of Movement in mixing synergies. All ratings showing partial or complete disagreement were reviewed in detail together with recommendations for improvements in language provided by the bilingual experts. For example, the experts recommended replacing sensomotora (sensorimotor in English) with sensoriomotora, posición de sentado (seated position) with posición sedente, the adjective en masa (mass) with masiva (massive), espina (spine) with columna (spinal column), sinergias de mezcla de movimiento voluntario (mixed synergies of voluntary movement) with movimiento voluntario mezclando sinergias (voluntary movement mixing synergies). The acronym PIP, referring to the proximal interphalangeal was replaced with IFP (for the Spanish term interfalángica proximal). It is important to note that, while some of the recommendations led to a greater comprehension of the FMA protocol and manual and were therefore accepted by the researchers, the others were rejected, if they changed the context of the assessment. For example, the experts recommendation on replacing the prepositions desde (from) and hasta (until or to) with their less precise synonyms de and al, respectively, were not followed in Section II: Mano desde rodilla contratralateral hasta oído ipsilateral (Hand from contralateral knee to ipsilateral ear).

**The pilot trial**

During the study period, 29 patients were screened for inclusion, of which 19 did not meet the inclusion or exclusion criteria: four patients were discharged before day 9 or 10 (including one...
The demographic and clinical characteristics of the included patients are shown in Table 2.,

The levels of agreement between and within assessors for the FMA-UE varied between 40 and 100% (agreement values for each item of the FMA can be found in the Supplementary Tables 1 and 2, published online). Most of the items demonstrated high level of agreement with values above 70%. Low percentage of agreement was only present for time item in the coordination subscale. The systematic disagreements were more frequent in comparisons between test occasions than between assessors. Individual variability (relative variance) was only present for time item in the coordination subscale in comparisons between test occasions.

For example, the abducción (abduction) was replaced with aducción (adduction) in the section of Voluntary movement within synergies (A.I); the shoulder position at 30 to 90° in the Repeated dorsal and volar flexion with elbow at 90 degrees in the Wrist subscale (B) was corrected to 0°; and the wording was changed to rango de movimiento activo completo, fluido (full range of motion, smoothly) for the full score of 2 points of the Repeated dorsal and volar flexion items and to circunducción completa y suave (complete and smooth circumduction) for the full score of 2 points of the Circumduction item. For the subscale of Coordination/Speed (D), the description was revised to indicate more clearly the movement start with hand on the knee, movement to the nose and end of the movement at the knee. The therapist counts mentally the number of attempts so that the speed of the movement would not be influenced by the counting. In the same subscale, the text for Time item for full score of 2 points was changed to menos de 2 segundos de diferencia (less than 2 s difference). For the subscale of Voluntary movement with little or no synergy (E.IV) in the item of the dorsal flexion it was clarified that the dorsiflexion should be performed smoothly for the full score of 2 points of the Repeated dorsal and volar flexion items and to circunducción completa y suave (complete and smooth circumduction) for the full score of 2 points of the Circumduction item. For the subscale of Coordination/Speed (D), the description was revised to indicate more clearly the movement start with hand on the knee, movement to the nose and end of the movement at the knee. The therapist counts mentally the number of attempts so that the speed of the movement would not be influenced by the counting. In the same subscale, the text for Time item for full score of 2 points was changed to menos de 2 segundos de diferencia (less than 2 s difference). For the subscale of Voluntary movement with little or no synergy (E.IV) in the item of the dorsal ankle flexion it was clarified that the dorsiflexion should be performed separately for each foot.

These final adjustments were sent to the researchers at the University of Gothenburg in November 2015, who reviewed and approved the proposed adjustments. This last revision of the FMA protocol and application manual resulted in the final Colombian Spanish version of the FMA.

Discussion

In this study, a rigorous standardized process was used to perform a translation and cultural validation of the original FMA from English and Swedish into Colombian Spanish. This work will have
a great impact on the physiotherapy practice of stroke rehabilitation in Colombia. An introduction and use of FMA in clinical practice in Colombia would have a major impact to the stroke rehabilitation, also in a wider perspective. Use of FMA would allow unified description of stroke severity and recovery patterns in individuals with stroke in Colombia and other Spanish speaking countries, which opens up new opportunities to make comparisons of rehabilitation outcomes with other countries and regions.

The current study provides a comprehensive description of a standardized methodological process that included not only translation, but also a stepwise drafting process and pilot testing of the final version. For this process, internationally established recommendations and guidelines for translation and cultural adaptation were followed [17–19,24]. This study can serve as an example on how this methodological process can be applied when clinical observational rating scales are translated to other languages. This information can be valuable for a broader community of Colombian and other Spanish speaking physiotherapists, researchers, and educators and used as guidance in future translations of clinical scales. This will in turn strengthen the translational collaboration with other regions and facilitate the research in Colombia and other Spanish speaking countries.

Currently, the original Fugl-Meyer scale [7] is officially translated and validated in English, Swedish, Norwegian, and now in Colombian Spanish. The protocol of the scale is freely available (www.neurophys.gu.se/rehabmed) in these languages. The FMA requires little equipment, is clear in scoring and both, the FMA-UE and FMA-LE, require approximately 10–15 min administration time, which makes the scale easily accessible and clinically feasible [3]. Recent studies have also shown that the FMA scores follow a continuum of motor ability, which means that people with less ability score low, while people with more ability get higher scores on FMA [25–27]. Although it has not been shown that the FMA score purely reflects manifestation of movement within or outside synergies, it has been confirmed in several studies that, in part of reflex items, the FMA scale is a unidimensional assessment with an item hierarchy consistent with contemporary motor control theory and evidence that suggests that an individual’s ability to perform a task is a dynamic interaction between neural factors and the task-specific difficulty of the task [26–29].

These studies strengthen the continuous use of FMA both in research and clinical practice. The FMA is also the only recommended assessment scale for upper and lower extremity impairment for stroke studies [4,6].

To our knowledge, the original FMA [30] and French-Canadian version of the FMA [31] has been translated to Brazilian Portuguese. Maki et al. [30] reported shortly that forward and backward translation was done by two bilingual translators and that two physiotherapists and one neurologist applied the scale on five patients to verify the understanding of the items, and establish the final version. In Michaelsen et al.’s study [31], the French-Canadian version was translated by one Brazilian physiotherapist and adapted after a pilot reliability testing of eight patients. Recently, the FMA-UE was translated to Danish from an English version by See et al. [32] following the forward and backward translation process, in which the back translated English version was approved by the corresponding author of the English version and the Danish version was pilot tested on 10 patients [14]. The translation process and cultural adaptation of FMA is, in general, only described briefly in previous studies. In the current study a rigorous process was applied, which included the training and approval provided by the researchers familiar with the original scale, official forward and backward translation, bilingual expert evaluation and stepwise revision by the Colombian and Swedish researchers and clinicians. This standardized process was undertaken to eliminate the researcher bias.

In the pilot test of this study, with 10 patients, the scale showed good reliability among evaluators. A statistical method, particularly designed for ordinal data showed that although the sample size was small, sufficient levels of agreement were observed. In the previous study by Maki et al. [30], the translated Brazilian version was only tested on five subjects to verify understanding of the items and in the study of Michaelsen et al. [31] the pilot reliability testing was performed on eight patients, and it is not clear whether the reliability or validity was also investigated later in a larger sample. A larger inter-rater and intra-rater reliability study to determine agreement at the item level using the Svensson’s methodology of the Colombian Spanish FMA will be performed at the Central Military Hospital of Colombia.

Another concern that arises from previous translation studies of FMA is the modifications made to the original scale. For example, in some translations, the items of the Hand subscale, the hook grasp and thumb adduction [31,33,34], and the spherical grasp [31] show differences from the original description. Another modification is that the patient is asked to make an active grasp of the tested object that is presented on the table [34]. In the original FMA, the objects are interposed close to the patient’s hand and the patient’s ability to open the finger and hand to the size of the object and ability to grasp around the object and hold the object is assessed. Coordination/speed items are in some translations done with eyes open [34], in contrary to original FMA in which the eyes are closed. In the same subscale, the scoring is different from original FMA for tremor and dysmetria in some translations [14,32,34], in which a patient with complete paralysis is able to obtain full score of 2 when tremor or dysmetria cannot be observed anywhere else on the body. Another modification that has been published in different translations of the FMA-LE applies on the reflex activity. In these translations, only the Patella and Achilles reflex activity is included [34]. In the original FMA, the Patella and Achilles tendon reflexes are both considered as reflex activity of the extensors, and the reflex activity in flexors is assessed in knee flexors. In all these translations, the modifications are not addressed as modification to the original FMA, even when the original reference to Fugl-Meyer et al. has often but not always been provided.

It is not clear how these modifications appear and what is the rationale behind it. In the current study, the Colombian researchers took contact with the Swedish researchers, who were the holders of the original approved FMA scale in English and Swedish. The translation process started with a joint theoretical and educational program for the Colombian researchers and incorporated a continuous process of discussion and reconciliation between both research groups. This process ensured that the conceptual, semantic, and operational equivalence in wording and item descriptions of the translated version were kept close to the original source version. The strength of this study is the rigorous methodological structure and process adapted for the translation process. The small sample size of 10 individuals in an early subacute phase after stroke might be a limitation of the current study. The reliability of the Colombian Spanish version of the FMA at the item level by using statistics appropriate for ordinal data will be investigated in a larger population of individuals with stroke at the Central Military Hospital of Colombia.

In summary, the translation and use of the FMA in Colombia and other Spanish speaking countries has great significance for the physiotherapists, researchers, and educators in the area of stroke rehabilitation. A large body of research has shown that FMA is easy to use, does not need any special equipment and has
excellent validity, reliability, and responsiveness, which makes this scale particularly suitable for different kinds of clinical settings world-wide. A wider use of FMA and other recommended outcome measures in Spanish speaking countries will open up possibility to compare stroke and rehabilitation outcomes with other countries and regions. This study also provides a comprehensive description of the methodological procedures, which can be used in other similar studies of translation and cultural adaptation of clinical observational rating scales.

Acknowledgements

The authors wish to express their appreciation to the Central Military Hospital of Colombia, the bilingual experts and the patients who participated in the pilot trial.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This project was funded by the Central Military Hospital of Colombia through Research Project No. 2013059, registered with the Research Unit.

ORCID

Katharina S. Sunnerhagen  http://orcid.org/0000-0002-5940-4400
Margit Alt Murphy  http://orcid.org/0000-0002-3192-7787

References

[22] Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull. 1968;70:213–220.


