



Original Research

The Spanish Version of the Functional Rating Index in Patients With Low Back Pain: Preliminary Results of the Validation Study



Ivan Medina-Porqueres, PT, RN, MSc ^a,
Manuel Torres-Fernandez, PR ^a,
Francisco J. Palenque-Lobato, PT ^b,
Nina Poljakovic-Kovacev, PR ^a,
Antonio Bellido-Da-Conceição, PT ^a,
Ronald J. Feise, DC, CEBC ^c,
Raquel Cantero-Tellez, PT, OT, PhD ^a

^a Department of Physical Therapy, Faculty of Health Sciences, University of Malaga, Malaga, Spain

^b Physical Therapy Unit, Malaga Regional Hospital, Malaga, Spain

^c Institute of Evidence-Based Chiropractic, Scottsdale, Arizona

KEYWORDS

Cross-cultural comparison;
Functional rating index;
Low back pain;
Psychometrics;
Rehabilitation;
Translations

Abstract Objective: To assess the reliability, validity, and the psychometric properties of the Spanish version of the Functional Rating Index (Sp-FRI) in a preliminary cohort of patients with low back pain (LBP).

Design: Prospective observational multicenter study.

Setting: Outpatient physical therapy clinics and units from public and private settings.

Participants: Patients with LBP (N=22; 52.5±12.5y) entered the study.

Interventions: The translation and cross-cultural adaptation were performed following international guidelines through a 5-step procedure.

Main Outcome Measures: The Sp-FRI was administered along with the Spanish version of Roland-Morris Disability Questionnaire (Sp-RMDQ) and numeric pain rating scale (NPRS) index. Preliminary testing included readability, comprehensibility, ceiling and floor effects, reliability, and validity. Statistical analysis was based on the Fernandez-Huerta index, and the calculation of Cronbach alpha, intraclass correlation coefficient (ICC), and Spearman's

List of abbreviations: 95% CI, 95% confidence interval; GPES, global perceived effect scale; ICC, intraclass coefficient; LBP, low back pain; NPRS, numeric pain rating scale; Sp-FRI, Spanish version of Functional Rating Index; Sp-RMDQ, Spanish version of Roland-Morris Disability Questionnaire.

Disclosures: none

Cite this article as: Arch Rehabil Res Clin Transl. 2019;1:100008.

<https://doi.org/10.1016/j.arrct.2019.100008>

2590-1095/© 2018 The Authors. Published by Elsevier Inc. on behalf of the American Congress of Rehabilitation Medicine. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

correlation coefficient, respectively. All patients completed the Sp-FRI again after 1-2 days to assess its test-retest reliability.

Results: None of the participants requested clarification of any of the items at the time of completion. The test-retest reliability of the FRI score was substantial (ICC 0.77). Cronbach alpha was 0.859. Spearman correlation coefficient between Sp-FRI and Sp-RMDQ was 0.66; $P < .0001$, and between Sp-FRI and NPRS was 0.66; $P < .0001$. No ceiling or floor effects were detected.

Conclusions: In light of these preliminary data, the Sp-FRI appears to be linguistically accurate and has been adapted to the Spanish-speaking population. It demonstrated reliability and validity and is suitable for clinical and research use among Spanish patients with LBP, with an acceptable degree of internal consistency and concurrent validity.

© 2018 The Authors. Published by Elsevier Inc. on behalf of the American Congress of Rehabilitation Medicine. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Low back pain (LBP) is a major musculoskeletal problem in the general population, affecting people of all ages worldwide,¹ and one of the most common reasons for medical consultation.² The high prevalence of this pathology and the associated costs result in a substantial socio-economic burden to society. A proper characterization and management of LBP is crucial to fight this condition. Making a reliable, specific pathoanatomical diagnosis for patients with LBP has become an ongoing challenge for clinicians. Self-reported questionnaires of clinical and functional status have become important outcome measures that can help achieve this objective.

The Functional Rating Index (FRI) is a patient self-report measure that uses both pain and function for a wider view of a patient's disability.³ This questionnaire derived from a combination of Neck Disability Index⁴ and Oswestry Disability Questionnaire,⁵ and it has been widely used to evaluate LBP patients in a number of populations with whom it has demonstrated satisfactory reliability, validity, and responsiveness.^{6,7} The FRI consists of 10 items that measure both pain and function of the spine. Of these 10 items, 8 refer to activities of daily living that can be adversely affected by a disease of the spinal system, and 2 refer to 2 different attributes of pain. Using a 5-point scale for each item (where 4 = worst possible pain and/or unable to perform this function and 0 = no pain or full ability to perform a function and/or the severity of pain at the present time).³ The range of scores is 0% (no disability) to 100% (severe disability). When all 10 items of the FRI are completed, the score is calculated as follows: $(\text{total score} / 40) \times 100\%$. When only 9 items are completed, the score is calculated as follows: $(\text{total score} / 36) \times 100\%$.⁸

The FRI was originally developed in English language for population suffering from LBP.³ The reliability and validity of this instrument have been tested in a limited number of studies, and validated translations have been chronologically published in Turkish (2006),⁹ Brazilian-Portuguese (2007),¹⁰ Persian (2011),¹¹ Chinese (2012),¹² and Thai (2015).¹³ Prior to this study, no previous version of this tool existed for the Spanish population with LBP. We conducted this study to adapt and validate the Spanish version of the Functional Rating Index (Sp-FRI) in patients with LBP, and

this article reports the preliminary psychometric properties of this instrument administered to these people prior to their intervention treatments.

Methods

Translation process

The translation was performed using a 5-step process based on the guidelines for the cross-cultural adaptation process written by the American Association of Orthopaedic Surgeons¹⁴: (1) initial translation; (2) synthesis; (3) back translation; (4) expert committee; and (5) pilot study. Two independent translators whose mother tongue was Spanish translated the instrument from English to Spanish for the initial translation. These 2 translators were required for the synthesis step to meet, discuss, and compose a synthesized version of the translated FRI. Two other independent translators, whose native language was English and who were blinded to the original instrument, then translated

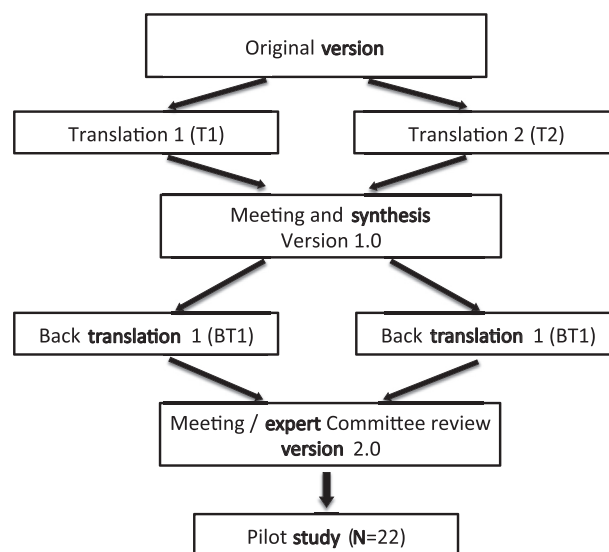


Fig 1 Translation-back translation method.

A

Índice de Tasa Funcional

Uso exclusivo en **problemas de cuello y espalda**.

Con el fin de valorar convenientemente su lesión, debemos comprender cuánto afectan sus **problemas de cuello y espalda** a su capacidad para afrontar las actividades de la vida diaria. En cada apartado de los siguientes, **rodée por favor con un círculo el número que de forma más aproximada describe actualmente su lesión** :

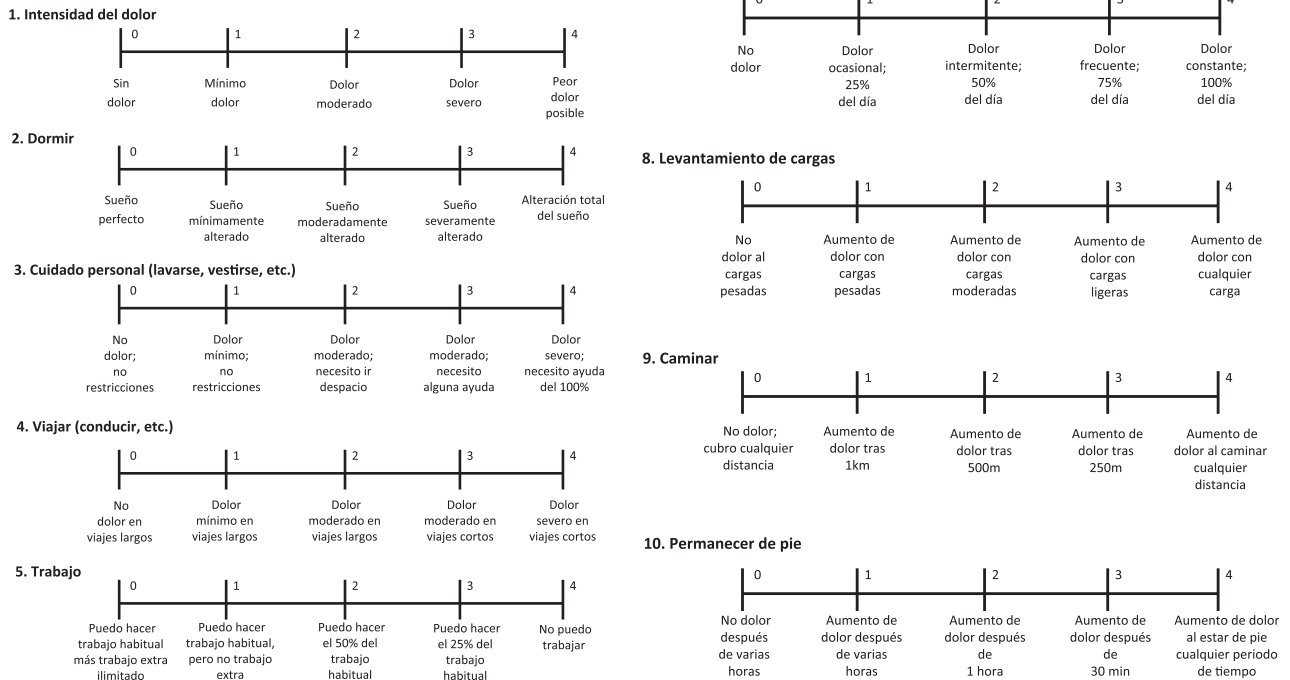


Fig 2 (A) Spanish Functional Rating Index. (B) Original Functional Rating Index (Institute of Evidence-Based Chiropractic).

the synthesized version back to English. An expert committee composed of a linguist, 2 clinicians, and the 4 translators revised the whole process and consolidated the prefinal version. Figure 1 summarizes the translation process. We pilot tested the prefinal version of the Sp-FRI on 20 healthy participants to find any difficult, upsetting, or confusing items. No difficulties encountered by the respondents were noted in the pilot study. A consolidated version of the Sp-FRI was then created (fig 2). This last version was sent to the developer to evaluate the conceptual equivalence of backward version with the original one.

Study participants

Patients were recruited nonselectively and consecutively in the period from October 2015 to January 2016 from 5 different physical therapy units (public and private setting) in Malaga, Spain. Participants were recruited according to the following eligibility criteria: (1) patients were required to be fluent in Spanish (minimum grade 6 reading level) and to be able unable to complete the questionnaires; (2) aged between 18 and 80; (3) with acute or chronic, nonspecific LBP with no radiating pain to lower limbs; and (4) no neurologic signs. Exclusion criteria included not being fluent in Spanish or being unable to complete the questionnaires, previous spinal surgery, presence of tumor, systemic rheumatic disease, ankylosing spondylitis, or neurologic disorders.

Setting

Five different clinical centers in Malaga, Spain volunteered to participate in a multicenter validation study. The selected centers were outpatient physical therapy clinics and units (public and private), all chosen to represent different social and cultural contexts within different areas of this region.

Ethical considerations

Ethical approval for the study protocol was granted by the Regional Research Ethics Committee of Malaga in Malaga, Spain. The original FRI authors³ were contacted, and they provided authorization to conduct this work. The study was conducted in accordance with medical professional codes and Ethical Principles for Medical Research Involving Human Subjects (Declaration of Helsinki 2008).

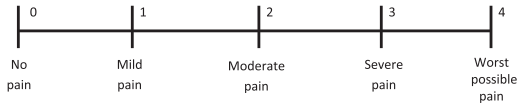
Participants completed the questionnaires voluntarily, and the aims and objectives of the study were explained to each patient before participation. Written informed consent and verbal assent were given by all participants prior to the interview. Prior to study participation, patients received written and oral information about the content and extent of the study. No financial incentives were provided to any study participant. The protection of

B Functional Rating Index

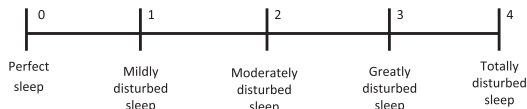
Four use with **Neck and/or Back Problems** only.

In order to properly assess your condition, we must understand how much **your neck and/or back problems** have affected your ability to manage everyday activities. For each item below, **please circle the number which most closely describes your condition right now.**

1. Pain intensity



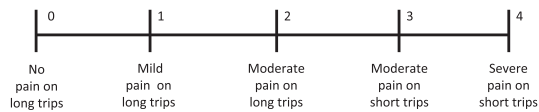
2. Sleeping



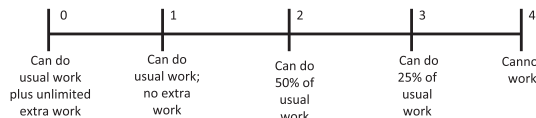
3. Personal care(washing, dressing, etc.)



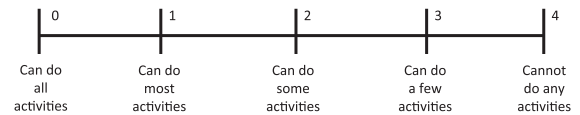
4. Travel (driving, etc.)



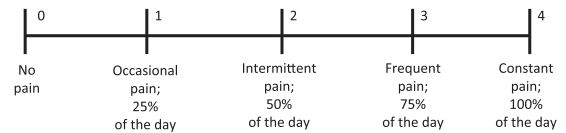
5. Work



6. Recreation



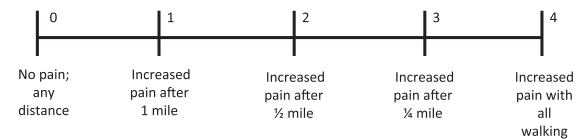
7. Frequency of pain



8. Lifting



9. Walking



10. Standing



Fig. 2 Continued.

participants' personal data was performed according to the Spanish Organic Law of Protection of Personal Data 15/99.

Instruments

The instruments selected for this study include the Spanish version of Roland-Morris Disability Questionnaire (Sp-RMDQ)¹⁵ and numeric pain rating scale (NPRS).¹⁶ The choice of these instruments was guided by the availability, established psychometric properties, and nonsuperiority of other instruments.

To accommodate scoring irregularities, the following rules were established: (1) when an individual scores 2 responses on the same item, responses are averaged; (2) when an individual scores response between 2 numbers, the answer is the average of the 2 numbers; and (3) when a participant does not respond to an item, the missing value was imputed with the middle score value of the scale.

Demographic data

Age, gender, weight, height, educational level, self-report health status (excellent, very good, good, fair, poor), occupational workload (no work, light, moderate, heavy), past year impairment (days), and the presence of actual

LBP were recorded for each participant prior to the other measures. The duration of LBP was also noted (<3mo, 3mo to 1y, >1y). Pain was assessed by use of a numeric rating scale (0-10, NPRS).

We used the level of education as proxy indicator of socioeconomic status, because information on income was not available. The level of education was described according to the Spanish education system into 4 categories: less than elementary school degree, elementary school degree, high school degree, and college/university degree equivalent to less than 9 years of school, between 11 and 13, and more than 17, respectively.

The occupations of participants were grouped into 4 categories, according to their physical demands: (1) light work or tasks; (2) moderate work or tasks; (3) heavy work or tasks; (4) no work or tasks. Health status during last 12 months was categorized into 5 levels: (1) very bad; (2) bad; (3) normal; (4) good; (5) very good.

Roland-Morris Disability Questionnaire

The Roland-Morris Disability Questionnaire (RMDQ) is a 24-item, self-administered questionnaire with *yes/no* answers. It is simple and fast and can be filled out by the patient. The RMDQ is scored by adding up the number of items checked by

the patient. The score can, therefore, vary from 0 to 24. If patients indicate in any way that an item is not applicable to them, the item is scored *No*, ie, the denominator remains 24. The range of scores is 0% (no disability) to 100% (severe disability). The higher the number, the greater the perceived pain and dysfunction; the lower the number, the lower the perceived pain and dysfunction.

Numeric pain rating scale

The 11-point NPRS allows patients to rate their pain ranging from 0 (no pain) to 10 (worst imaginable pain).^{17,18} This instrument has been shown to have concurrent and predictive validity as a measure of pain intensity. It has been widely used as an outcome measure in clinical trials involving patients with LBP.¹⁹

Global perceived effect scale

Global perceived effect scale (GPES) is an 11-point Likert-type scale evaluating *global perceived effect* of the physical therapy treatment. It serves as an external criterion of clinically important change. The response options are +5 (completely recovered), 0 (no change), and -5 (vastly worse).

At the first visit, the participants with back pain were asked to complete the Sp-FRI, Sp-RMDQ, and NPRS. At the second visit, 1-2 days later, all participants were asked to complete the Sp-FRI and the GPES. The readability, comprehensibility, ceiling and floor effects, reliability, and validity of the Sp-FRI were analyzed.

Statistical methods

Descriptive statistics were used to describe the demographic and disease-related data for the study participants.

Feasibility assesses the ease with which patients complete and researchers administer a questionnaire. Grammar and language difficulty of the Sp-FRI were assessed with the Fernandez-Huerta Index²⁰ using available word processing software (Microsoft Word 2011). As a measure of data quality, we examined the number of questions where respondents needed some clarification.

Cronbach alpha and alpha if item deleted were calculated as a measure of internal consistency.²¹ The Sp-FRI was considered internally consistent when the items correlated moderately both with each other, and with the total score coefficients between 0.70 and 0.95 will be considered adequate.²² In addition, the Cronbach alpha was calculated for the questionnaire in case an item would be removed, to see if a given item negatively influenced the Cronbach alpha.²³

Test-retest reliability was evaluated by calculating intraclass correlation coefficient (ICC) for the total score. ICC values were interpreted following Landis and Koch's scale²⁴ of strength for reliability coefficients: poor (0.00-0.20), fair (0.21-0.40), moderate (0.41-0.60), substantial (0.61-0.80), and almost Perfect (0.81-1.00).

Construct validity was measured by analyzing the correlations between the Sp-FRI and the previously described

reference instruments (Sp-RMDQ, GPES, NPRS) using Spearman correlation coefficients. Cohen criteria for correlation strength in psychometric validation was employed: <0.30 weak, 0.30-0.49 moderate, >0.50 strong.²⁵

All statistical analyses were performed using SPSS for Windows statistical package (v22.0).^a Significance was accepted at an alpha level of 0.05.

Results

Descriptive statistics

The original sample included 24 participants. Two participants who had 2 or more missing values on their Sp-FRI or Sp-RMDQ responses were excluded. Another participant with only 1 missing value was not excluded, and the missing value was arbitrarily filled with the middle score value of the scale. Therefore, the sample was established in 22 participants (14 females, 63.7%; 8 males, 36.3%) with a mean age of 52.5 years ranging from 23 to 76. More than one-third of participants (36.3%) had an elementary school educational level. The sickness absence rate for the selected sample was 5.6 ± 19.7 days. Almost half of participants (45%) experienced pain for more than a year, and more than 36% of patients considered their health status as *bad*. Demographic and clinical data are presented in table 1.

Comprehensibility

None of the participants requested clarification of any of the questionnaire items at the time of completion. The

Table 1 Descriptive characteristics

Characteristics	n (%) or Mean \pm SD
Sex	
Men	8 (36.3)
Women	14 (63.7)
Age (y)	52.5 \pm 12.5
Weight (kg)	77.5 \pm 14.1
Height (cm)	167.9 \pm 9
Academic degree	< Elementary school 0 (0) Elementary school 8 (36.4) High school or professional education 8 (36.4) University level 6 (27.2)
Occupational workload	Light work 7 (31.8) Moderate work 6 (27.2) Heavy work 4 (18.1) No work 5 (22.7)
Past year impairment (d)	5.6 \pm 19.7
Current pain duration	<3 mo 8 (36.4) >3 mo <1 y 4 (18.2) >1 y 10 (45.4)
Satisfaction level	9.3 \pm 1.2
Health status	Very bad 0 (0) Bad 8 (36.4) Normal 7 (31.8) Good 7 (31.8) Very good 0 (0)

Table 2 Mean and range of raw scores for evaluated instruments

Measurement Instrument	Mean Scores \pm SD	Range Scores
Roland-Morris Questionnaire	13 \pm 5	3-22
Functional Rating Index Test	20.2 \pm 6.9	5-35
Functional Rating Index ReTest	19.1 \pm 7.3	5-35
Numeric Pain Rating Scale	5.9 \pm 1.9	1-8

committee also gave their consent to the semantic comprehensibility of the questionnaire without remarkable comments.

Readability or feasibility

Readability statistics for the Sp-FRI were excellent with a Fernandez-Huertas Index of 77. None of the 24 original participants required additional clarification. Data were coded and interpreted by researchers with no difficulty.

Score distribution

Ceiling and floor analyses were performed on the total score. The Sp-FRI scores were well distributed, and no respondent achieved neither the lowest (floor effect) nor the highest (ceiling effect) score for the FRI. Thus, no floor or ceiling effects were noted in the studied population. Table 2 reflects scores for every evaluated instrument.

Reliability

Test-retest reliability was substantial (ICC 0.772; 95% confidence interval [95% CI] 0.53-0.89), and the mean score was 20.22 \pm 6.94. The minimum score recorded was 12.5% (for only 1 participant), and the maximum score was 87.5% (for 1 participant). None of the participants had the minimum possible score (0 points) or the maximum possible score (100%).

Cronbach alpha was 0.85, indicating good reliability of the instrument. Cronbach alpha (if 1 item deleted) ranged between 0.82 and 0.86 for the test. Cronbach alpha was 0.90 and Cronbach alpha (if 1 item deleted) ranged between 0.89 and 0.90 for the retest. Similarly, internal consistency for Sp-RMDQ was found to be 0.85, showing the same good reliability than Sp-FRI itself.

Construct validity

Values of Sp-FRI correlated equally with Sp-RMDQ ($r=0.66$; $P<.0001$) and NPRS ($r=0.66$; $P<.0001$), and all correlations were significant. According to Cohen's criteria, a strong correlation exists in both cases.²⁵

Discussion

The purpose of this study was to obtain preliminary data concerning the psychometric properties of the Sp-FRI. Various translations of the FRI have been validated, and the instrument has been demonstrated to be psychometrically strong and appropriate for use in Persian, Chinese, Brazilian-Portuguese, Thai, and Turkish people with LBP.¹⁰⁻¹³ The instrument has been tested in various

populations, including athletes⁷ and elderly populations.²⁶ In summary, the Sp-FRI scale was internally consistent and had acceptable concurrent validity among Spanish patients with LBP. However, the results are only preliminary, and further studies are needed to test the properties of the scale in larger populations.

The readability statistics indicate that the questionnaire language and grammar are appropriate for the population studied. The Fernandez-Huertas score as well as the fact that no participants needed assistance or additional clarification suggests that the instrument was not difficult to understand or complete.

The Cronbach alpha coefficient for the Sp-FRI (0.85) demonstrated sufficient internal consistency and was similar to the original and other language versions: 0.92 for the original,³ 0.86 for the Thai,¹³ 0.92 for the Brazilian-Portuguese,¹⁰ 0.96 for the Turkish,⁹ 0.90 for the Chinese,¹² and 0.89 for the Persian¹¹ versions. The recommended Cronbach alpha is approximately 0.80. If the internal consistency is too high, it may indicate that the items are too homogeneous.²⁷ Although these results could not be compared directly, because the study nature differed and the methods varied, the similar values suggest that the Sp-FRI is internally consistent and no items are redundant. In addition, when most of the items were deleted the alpha value decreased, hence providing evidence that all the items are important to the establishment of the index.

The test-retest reliability was substantial (0.77) for the total score, which is consistent with the previous versions, such as Persian (ICC 0.81, $P<.0001$),¹¹ Chinese (ICC 0.95; 95% CI 0.92-0.97),¹² Thai (ICC 0.82; 95% CI 0.70-0.91),¹³ and Brazilian-Portuguese (ICC 0.95; 95% CI 0.93-0.97).¹⁰ In the original study, a 0.99 ICC_{3,k} test-retest reliability was calculated.

To assess concurrent validity, we analyzed the correlations between the Sp-FRI and reference instruments (Sp-RMDQ, NPRS). Overall, the Spearman correlation coefficients indicated that the total score was moderately correlated with these existing scales. In particular, the Sp-FRI score was moderately correlated with the Sp-RMDQ ($r=0.66$). Similar results were observed in Thai ($r=0.55$),¹³ Chinese ($r=0.74$),¹² and Persian ($r=0.61$)¹¹ versions, only the Brazilian-Portuguese one showing a strong correlation ($r=0.80$).¹⁰ Although a direct comparison cannot easily be made, these similar results reinforce the concurrent validity of the Sp-FRI.

Study limitations

Our study indicated preliminary psychometric evidence that the translated and culturally adapted Sp-FRI was equivalent to the original instrument to be used among Spanish population presenting LBP. However, aspects such as responsiveness, factor structure, or minimally clinically important change in the Sp-FRI score still remain to be analyzed. In addition, a number of further limitations should be stressed out regarding our work. First, only patients with LBP participated in this study. Because FRI has been used for assessing participants suffering from cervical conditions in the past, the researchers suggest evaluation and a new validation with adaptations for patients with

cervical disorders. Second, the small population was selected from urban, coast areas. It may be, therefore, suitable to undertake further studies including people from internal and/or rural areas to increase the generalizability of the results. Third, our participant selection was not random, which may lead to a sample bias. Finally, an obvious small sample size for evaluating distinct psychometric properties in the Sp-FRI final version is the fourth and clearest limitation. A further, large-sized study is being developed where a minimum of 100 participants will be recruited.

Conclusions

Preliminary assessment of the Spanish-language version of the FRI tool suggests similar measurement properties to the previously validated English-language version of the FRI. The Sp-FRI has initially demonstrated to be linguistically accurate and has been adapted to the Spanish-speaking population. These preliminary results indicate that this tool appears to be reliable and valid, which would accordingly make it suitable for clinical and research use with Spanish people with LBP. However, inherent limitations attributable to preliminary studies with small sample sizes should be kept in mind. Undoubtedly, there is a need to undertake new studies with larger, more heterogeneous samples of patients with LBP to confirm these findings.

Supplier

a. SPSS for Windows; IBM Corporation.

Corresponding author

Ivan Medina-Porqueres, PT, RN, MSc, Department of Physical Therapy, Faculty of Health Sciences, University of Malaga, Arquitecto Francisco Peñalosa St, 29010 Malaga, Spain. *E-mail address:* imp@uma.es.

Acknowledgments

We thank the staff of the participating centers and units, especially all those clinical practitioners who participated in patient recruitment. We also thank F.J. Baron-Lopez, PhD, for his valuable methodologic critiques of the study.

References

- van Tulder MW, Furlan AD, Gagnier JJ. Complementary and alternative therapies for low back pain. *Best Pract Res Clin Rheumatol* 2005;19:639-54.
- Borenstein D. Epidemiology, etiology, diagnostic evaluation, and treatment of low back pain. *Curr Opin Rheumatol* 1996;8:124-9.
- Feise RJ, Michael Menke J. Functional Rating Index: a new valid and reliable instrument to measure the magnitude of clinical change in spinal conditions. *Spine* 2001;26:78-86.
- Vernon H, Mior S. The Neck Disability Index: a study of reliability and validity. *J Manipulative Physiol Ther* 1991;14:409-15.
- Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66:271-3.
- Childs JD, Piva SR. Psychometric properties of the functional rating index in patients with low back pain. *Eur Spine J* 2005;14:1008-12.
- Naghdi S, Nakhostin Ansari N, Yazdanpanah M, Feise RJ, Fakhari Z. The validity and reliability of the functional rating index for evaluating low back pain in athletes. *Scand J Med Sci Sports* 2015;25:840-5.
- Feise RJ, Menke JM. Functional Rating Index: literature review. *Med Sci Monit* 2010;16:RA25-36.
- Ceran F, Ozcan A. The relationship of the Functional Rating Index with disability, pain, and quality of life in patients with low back pain. *Med Sci Monit* 2006;12:CR435-C439.
- Costa LO, Maher CG, Latimer J, Ferreira PH, Pozzi GC, Ribeiro RN. Psychometric characteristics of the Brazilian-Portuguese versions of the Functional Rating Index and the Roland Morris Disability Questionnaire. *Spine* 2007;32:1902-7.
- Ansari NN, Feise RJ, Naghdi S, Ebadi S, Yoosefinejad AK. The functional rating index: reliability and validity of the Persian language version in patients with low back pain. *Spine* 2011;36:E1573-7.
- Wei X, Chen Z, Bai Y, et al. Validation of the simplified Chinese version of the Functional Rating Index for patients with low back pain. *Spine* 2012;37:1602-8.
- Chansirinukor W. Thai version of the Functional Rating Index for patients with back and neck pain: part 1 cross-cultural adaptation, reliability and validity. *J Med Assoc Thai Chot-maihet Thangphaet* 2015;98(Suppl 5):S97-105.
- Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25:3186-91.
- Kovacs FM, Llobera J, Gil Del Real MT, Abaira V, Gestoso M, Fernández C, et al. Validation of the spanish version of the Roland-Morris questionnaire. *Spine* 2002;27:538-42.
- McCaffery M, Pasero C. Teaching patients to use a numerical pain-rating scale. *Am J Nurs* 1999;99:22.
- Katz J, Melzack R. Measurement of pain. *Surg Clin North Am* 1999;79:231-52.
- Price DD, Bush FM, Long S, Harkins SW. A comparison of pain measurement characteristics of mechanical visual analogue and simple numerical rating scales. *Pain* 1994;56:217-26.
- Childs JD, Piva SR, Fritz JM. Responsiveness of the numeric pain rating scale in patients with low back pain. *Spine* 2005;30:1331-4.
- Fernandez-Huerta J. Medidas sencillas de lecturabilidad. *Consigna* 1959;(214):29-32.
- Streiner D, Norman G. Health measurement scales: a practical guide to their development and use. Oxford: Oxford University Press; 2003.
- Bravo G, Potvin L. Estimating the reliability of continuous measures with Cronbach's alpha or the intraclass correlation coefficient: toward the integration of two traditions. *J Clin Epidemiol* 1991;44:381-90.
- Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34-42.
- Landis RJ, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.
- Cohen JA. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960;20:37-46.
- Bayar B, Bayar K, Yakut E, Yakut Y. Reliability and validity of the Functional Rating Index in older people with low back pain: preliminary report. *Aging Clin Exp Res* 2004;16:49-52.
- Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ* 2011;2:53-5.