Title

Cross-cultural adaptation and validation of the Manchester Foot Pain and Disability Index into Spanish.

Authors

Gijon-Nogueron G[1]; Ndosi M[2]; Luque-Suarez A[3]; Alcacer-Pitarch B[2]; Munuera PV[4]; Garrow A[5]; Redmond AC[2]

Affiliations

1. Department of Nursing and Podiatry, University of Málaga, Málaga, Spain
2. Section of Clinical Biomechanics and Physical Medicine, Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, United Kingdom.
3. Department of Physiotherapy, University of Malaga, Malaga, Spain
4. Department of Podiatry, University of Seville, Seville, Spain
5. The School of Nursing, Midwifery &Social Work. University of Manchester, Manchester, United Kingdom

Correspondingauthor

Gabriel Gijon-Nogueron

Facultad de Ciencias de la Salud.Paseo de Martiricos 3.Malaga 29009. Spain.

gagijon@uma.es/+0034626449254

Keywords

Cross-cultural validation; Patient outcome measures; Foot Pain; Rasch analysis

Abstract

Purpose: The Manchester Foot Pain and Disability Index (MFPDI) is a self-assessment 19-item questionnaire developed in the UK to measure foot pain and disability. This study aimed at conducting cross-cultural adaptation and validation of the MFPDI for use in Spain.

Methods: Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes measures were followed in the MFPDI adaptation into Spanish. The cross-cultural validation, involved Rasch analysis of pooled datasets from Spain and the UK.

Results: Spanish dataset comprised 338 patients, five used in the adaptation phase and 333 in the cross-cultural validation phase, mean age(SD) = 55.2(16.7) and 248(74.5%) were female. A UK dataset (n=682) added in the cross-cultural validation phase; mean age(SD) = 51.6(15.2%) and 416 (61.0%) were female.

A preliminary analysis of the 17-item MFPDI revealed significant local dependency of items causing significant deviation from the Rasch model. Grouping all items into testlets and re-analysing the MFPDI as a 3-testlet scale resulted in an adequate fit to the Rasch model, χ^2(df) = 15.945(12) p=0.194, excellent reliability and unidimensionality. Lack of cross-cultural invariance was evident on the Functional and Personal appearance testlets. Splitting the affected testlets discounted the cross-cultural bias and satisfied requirements of the Rasch model. Subsequently, the MFPDI was calibrated into interval-level scales, fully adjusted to allow parametric analyses and cross-cultural data comparisons when required.

Conclusions: Rasch analysis has confirmed that the MFPDI is a robust 3-subscale measure of foot pain, function and appearance in both its English and Spanish versions.

Introduction

Foot pain and disability associated with foot pain may affect both younger and older populations [1-4],with the prevalence of foot pain increasing with age [5], being affected by gender [6] and also being associated with obesity [7]. Up to 80% of older people report foot problems [2]and the importance of foot pain in the population relates to its frequent association with physical impairments, such as poor balance [8], risk of falling [9; 10] and locomotor disability [11].

Foot problems can have a profoundly negative impact on function and health-related quality of life (QoL) [12],and assessing the presence and consequences of foot pain and disability is essential for the quantification of impairment and evaluation of treatments. Several measurement tools have been developed to assess the impact of foot pain: these include the American Orthopaedic Foot and Ankle Society scales [13], the Foot Function Index [14], the Foot Health Status Questionnaire [15],the Self-Administered Foot Evaluation Questionnaire, SAFE-Q version 2[16] and the Manchester Foot Pain and Disability Index (MFPDI) [17]. These tools are widely used in clinical and research practice [6; 18; 19], particularly the MFPDI, which has become a popular tool for the measurement of foot pain and disability because of a growing body of literature supporting its psychometric/clinimetric properties and independent evidence of its validity and versatility [20; 21].

The MFPDI is a self-assessment questionnaire developed in the UK [17] to measure foot pain and disability in the general population. It is based on 19 statements, two of which are related to the difficulty in performing work or leisure activities ("I am unable to carry out my previous work" and "I no longer do all my previous activities") and are excluded from the questionnaire if the respondent is of a retirement age. The remaining 17 items constitute three constructs (sub-scales): functional limitation (10 items), pain intensity (5 items) and concern with personal appearance (2 items). Each of the 17 statements has a 3-category response structure: 'none of the time' = 0, 'on some days' = 1, and 'on most/every day(s)' = 2. The MFPDI has been validated in middle-aged and older populations [17; 18], and has been used as an outcome measure both in large population-based surveys and in clinical studies, including randomized controlled trials [11; 19; 22]. The MFPDI has also been used in studies of various diseases with foot manifestations including systemic sclerosis [23], Ehlers-Danlos Syndrome [24] and rheumatoid arthritis [25].

For a questionnaire to be used in a different country, its items must not only be translated well linguistically, but also must be adapted culturally to maintain the content validity at a conceptual level across different cultures [26]. The term "cross-cultural adaptation" is used to describe this process, which ensures a

consistency in the content and face validity between the original and the new versions of the questionnaire. In addition to cross-cultural adaptation, if data from different cultures are to be compared subsequently, the new questionnaire must be demonstrated to have retained the measurement properties (construct validity and reliability) and the cross-cultural invariance [27]. The term "cross-cultural validation" is used here to describe the process of testing for construct validity, reliability and cross-cultural invariance of the adapted questionnaire.

Given the increasing need to undertake multinational studies, potentially useful tools such as the MFPDI need to demonstrate cross-cultural invariance, in order to ensure equivalence when data pooling or comparisons are made across countries [28-30]. The MFPDI has been translated and cross-culturally adapted and validated into Greek, [31] Portuguese [32] and Italian [33] but there was no Spanish version until now. Cross-cultural adaptation and validation of the MFPDI into Spanish will contribute towards establishing the MFPDI as a cross-cultural measure and improving the standardization of data capture in research and treatment evaluation in clinical settings across countries. The objective of the current study was to undertake cross-cultural adaptation and validation of a Spanish language version of the MFPDI for use in Spain.

Methods

Patients and Design

Ethical approval was obtained from the Experimentation Ethics Committee of the University of Seville, Spain. All participants in the study gave a written informed consent. A license for use of the MFPDI was obtained from the copyright holder (Isis Outcomes, Oxford, UK). Participants were recruited from seven podiatry clinics participating in the Private Podiatry Clinics Network located in the south of Spain. Participating patients met the following inclusion criteria: i) aged 18 or over, ii) required podiatry treatment, iii) were in current employment iv) were mobile and capable of walking household distances unaided, v) were native

Spanish speakers. Participants were excluded from the study if they did not have the capacity to comprehend the questionnaire.

The study consisted of two phases: the cross-cultural adaptation of the MFPDI and the cross-cultural validation of the resulting adapted MFPDI. The cross-cultural adaptation process was undertaken using the guidelines methodology specified by the licence holder Isis Outcomes and derived from the methodology recommended by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) for the translation and validation of patients reported outcome measures [33]. Cross-cultural validation of the adapted

MFPDI was then undertaken using Rasch analysis [29; 35]. Figure 1 summarises the cross-cultural adaptation process.

Cross-cultural adaptation

Cross-cultural adaptation involved eight stages: i) forward translation, ii) reconciliation; iii) back translation, iv) back translation review, v) harmonisation, vi) pilot testing/ cognitive debriefing, vii) pilot testing review/reviev of cognitive debriefing results, and viii) proofreading.

Forward Translation: Two forward translations in Spanish were undertaken from the original English language version of the MFPDI. The translations were undertaken by two independent health professionals who were native residents of Spain and fluent in both Spanish and English. Both translators based their translations on the guidelines provided by Isis Outcomes for the Translation and Linguistic Validation Process.

Forward Translation Reconciliation: The two forward translations were reconciled into one version (draft 1) by two original translators, a third independent translator and with input from the project lead.

Back Translation: The reconciled Spanish language version (draft 1) was back translated into English independently by two professional English native translators resident in Spain. The translators had no prior knowledge of the MFPDI and were not given sight of the original wording of the English version of the MFPDI.

Back translation review: The senior investigator and a native Spanish speaker resident in England and fluent in both languages reviewed the back translation for any discrepancies in meaning or terminology used. Any problematic item was discussed until the discrepancies were resolved. This process resulted in a refined second draft of the Spanish translation (draft 2).

Harmonisation: To produce the final Spanish language translation, a harmonization meeting was undertaken involving three Spanish translators, the senior investigator and the developer of the original UK version of the MFPDI. During this meeting any discrepancies or issues that were highlighted from the back-translation were discussed, the translated version of the MFPDI was evaluated and a final version agreed.

Pilot Testing/Cognitive Debriefing: Once the translation process was completed, the translation was formatted to match precisely the licensed English language version. The translated MFPDI version was initially assessed for

comprehensibility in five patient participants, who were Spanish residents and native speakers. At this stage each participant was asked by the in-country investigator to carry out the following tasks:

- To complete a copy of the translated MFPDI (the time taken to completed the MFPDI was also documented).
- To comment on the response options within the back-translated MFPDI.
- To comment on any wording that was difficult to understand.
- To suggest alternative wording/phrasing for any wording that was difficult to understand.
- To describe in their own words what the wording meant to them. These responses were recorded verbatim and translated into English.

The five patients' responses were summarised by the in-country investigator. This summary also contained any changes, recommendations or suggestions indicated by the participants and in-country investigator.

Pilot testing review/reviev of cognitive debriefing results: To improve the performance of the translated questionnaire the pilot testing results were reviewed by the project manager and the in-country investigator. At this stage any item that caused comprehensibility difficulties for ≥ 40 % of the participants was review, and any modifications suggested by the respondent's comments were incorporate to the final translated version.

Proof-reading: The in-country investigator and a translator, who was not involved in the translation process previously, independently proofread the final formatted translation and any suggested changes were discussed with the senior investigator.

Following this process a final draft of the MFPDI translated and culturally adapted into "Spanish for Spain" was locked down and entered into the cross-cultural validation phase (final draft).

Cross-cultural validation

The validation phase tested the MFPDI as a 17-item questionnaire (excluding the two items which are not applicable to all individuals) [17]. To assess cross-cultural validity of the MFPDI, two datasets were used: a new Spanish dataset and an existing UK dataset. The Spanish dataset comprised the MFPDI data from 333 Spanish participants who completed the tool during routine clinic appointments, while the UK dataset comprised the original data used in the development of the MFPDI [17] (used with the permission of the University of Manchester). Rasch analysis was used to assess the construct validity, reliability and the cross-cultural invariance of the translated (Spanish) MFPDI [29; 35].

The data were analysed descriptively using IBM SPSS software version 19 [36] then evaluated for fit to the Rasch model using RUMM2020 software [37]. Rasch analysis is a mathematical modelling technique used to assess properties of outcome measures against a measurement model developed by the Danish mathematician Georg Rasch[38]. In Rasch analysis, the observed data from questionnaires are measured against the model to assess their goodness of fit, with 'good fit' indicating a criterion-related construct validity, reliability and statistical sufficiency[39-41]. The overall fit statistics are given in terms of a χ^2 interaction and its associated probability, which is expected to be non-significant (i.e. not deviating from the Rasch model). For the model fit, the observed value for the residuals of each item is expected to lie within lie within ±2.5, and to have a mean of zero and standard deviation of one. A more detailed description of the Rasch analysis approach, its use in rheumatology and the interpretation of fit statistics is given elsewhere [35].

For this study, Rasch analysis took the following sequence: (i) examination of threshold ordering, (ii) assessment of overall fit, (iii) assessment of individual item fit and local dependence, (iv) strict assessment of unidimensionality and reliability (vi) assessment of cross-cultural invariance and invariance to age and gender. The Master's Partial Credit Model [42] parameterisation was employed, as the response format of the MFPDI is polytomous.

Threshold maps and category probability curves were examined for any disordering, then the residual correlation matrix was examined and items having a correlation of greater than ±0.3 were identified as displaying a local dependency [43]. Locally dependent items were then combined into subscales and each subscale treated as a "testlet". A testlet is defined as a subset of items that is treated as a measurement unit in test construction, administration and/or scoring [44]. The testlets were eventually tested for fit, unidimensionality and invariance. Strict unidimensionality was assessed by using the independent t-test method suggested by Smith[45]. The reliability of the MFPDI was reported using Person Separation Index (PSI), which provides the estimate of the internal consistency of the scale equivalent to Cronbach's alpha, only using the logit value as opposed to the raw score in the same formulae. A minimum value of 0.7 is required for group use and 0.85 for individual use [35]. Invariance was tested using the differential item functioning (DIF) facility built-in to RUMM2020, where Item characteristic curves for each testlet are checked for any significant DIF with respect to individual properties. Cross-cultural invariance was performed by pooling the Spanish and the UK datasets and testing for DIF by country. If uniform DIF was observed in any testlet, it was "split" for DIF in order to adjust for the cross-cultural bias [46]. In addition, splitting for DIF allows calibration of a DIF-adjusted conversion table with which to transform the ordinal scores into interval-level scores for parametric analyses and cross-cultural comparison of

data. The p-values for fit statistics and DIF analyses were Bonferroni-adjusted to the alpha level (i.e. p = 0.05/number of tests carried out) in order to avoid type I errors resulting from to multiple testing [47]. The expected values for perfect model fit are presented at the bottom of table 2.

Results

Patient characteristics

A total of 338 Spanish participants were recruited into this study, five for the pilot phase and 333 for the cross-cultural validation phase. Their mean age (SD) was 55.2 (16.7) and 248 (74.5%) were female. An existing UK dataset (n = 682) was used for cross-cultural validation phase and their mean age (SD) was 51.6 (15.2%) and 416 (61.0%) were female. Table 1 shows the characteristics of the validation study participants.

Cross-cultural adaptation phase

In the translation of the MFPDI, 'standard' Spanish was used in the translation (avoiding colloquial language) in order to improve comprehension. In this phase, there were some minor difficulties associated with two items: item 1, "I avoid walking outside at all" presented idiomatic difficulty in translating '*outside'*; while item 17 "I get shooting pains in my feet", presented conceptual difficulty in defining *shooting pains*. These were solved in the harmonisation meeting carried out by three Spanish translators, the senior investigator and the original developer of the MFPDI. The pilot study results showed no discrepancies in meaning or terminology used in the translated (Spanish) version of the MFPDI, therefore no modification was required. No respondents had requested aid in the interpretation of any item of the questionnaire. The mean (SD) time in minutes required to complete the MFPDI was1 min 47 secs (107 secs).

Cross-cultural validation phase

In both data sets (Spain and UK), the thresholds were adequately ordered, indicating that the 3-category response structure ('none of the time', 'on some days' and 'on most/every day(s)) was working as expected. The preliminary analysis of the 17 items as a single scale resulted in a significant deviation from the Rasch model χ^2(df) = 179.197 (68), p <0.001; and 292.526 (153), p<0.001for both the Spanish and the UK datasets respectively. Examination of the correlation matrices revealed that most of the items belonging to the same subscale were significantly locally dependent. These locally dependent items were combined into three corresponding subscales noted previously (functional, personal appearance, and pain) and the MFPDI was then treated as a three-testlet scale in all subsequent analyses.

Rasch analysis of the MFPDI as a three-testlet scale resulted in an adequate fit to the Rasch model χ^2(df) = 15.945 (12) p = 0.194 and 31.024 (21), p = 0.073, for the Spanish and the UK datasets respectively. The test of strict unidimensionality revealed proportions of significant t-tests (95%CI) were 0.016 (-0.006, 0.040) and 0.010 (-0.011, 0.032) for the Spanish and the UK datasets respectively, indicating that the three-testlet scale forms a unidimensional scale. The scale reliability (person separation indices - PSI) were excellent, i.e. .848 and .823 for the Spanish and the UK versions of the MFPDI respectively. Table 2 presents the results of Rasch analysis for the Spanish, the UK and the pooled datasets.

Pooling the Spanish and the UK data resulted again in a lack of the overall fit to the model. Differential item functioning (DIF) analysis revealed a lack of cross-cultural invariance on the first and the second testlets (functional and personal appearance). Splitting these two items resulted in creation of cultural-specific (emic) testlets for each country and cultural-general (etic) testlets in order to account for the cross-cultural bias. Following 'split for DIF', fit to the Rasch model was satisfied in each testlet and the overall scale χ^2(df) = 57.944 (45), p = 0.093, and the reliability remained excellent (PSI = .844) (Table 2). Table 3 presents fit statistics for each testlet after adjusting for cross-cultural invariance.

Following fit to the model, the MFPDI ordinal scores were Rasch-transformed and interval-level scales were calibrated for the Spanish and UK versions of the MFPDI, taking account of the cross-cultural DIF. These Rasch-transformed scores are suitable for use subsequently in parametric analyses if required and for data pooling or comparison between the UK and Spanish datasets. Table 4 is a DIF-adjusted conversion table of MFPDI raw scores into Rasch-transformed scores.

Discussion

This study used standardized methods for cross-cultural adaptation and validation of outcome measures to develop a Spanish version of the Manchester Foot Pain and Disability Index. The MFPDI translated well into Spanish without needing significant changes, only minor corrections in items 1 and 17 that were solved by the harmonisation meeting. The translated version was well-accepted by patients. Kaoulla et al., [31] in adapting their Greek version of the MFPDI had found difficulties in translating items 16 and 17. The pilot study did not yield any further discrepancies in meaning or terminology used in the second draft of the Spanish translation of the MFPDI. No participants requested aid in interpretation of any item of the questionnaire and so no further

modification was required. This is in accordance with Marinozziet al., [22]who found few problems in their study translating the MFPDI into Italian.

The preliminary assessment of the 17 MFPDI items revealed significant residual correlations (local dependency) among domain-specific items, which significantly affected the overall fit to the model. Local dependency artificially inflates the overall reliability of the scale, affecting its unidimensionality and resulting in biased parameter estimation [43]. This was the main issue causing deviation from the Rasch model in the UK and the Spanish datasets. Grouping the items into testlets (corresponding to subscales), corrected the local dependency and resulted in a unidimensional scale. The original scale [17] was not developed using Rasch analysis. A previous study by Muller and Roddy [21] used Rasch analysis to validate the MFPDI but they analysed only two subscales 'function' and 'pain' as the third subscale 'appearance' did not have enough responses. This present study has confirmed that all the three subscales represent independent constructs, which when added together form a unidimensional scale.

DIF analysis revealed a significant DIF-by-country i.e. cross-cultural non-invariance in the data. Cross-cultural invariance is a requirement for questionnaires intended for multinational use [28; 48]. Splitting the DIF-affected items resulted in creation of cultural-specific (emic) items for each country and cultural-general (etic) items. This permits the scale to be culturally relevant while permitting comparisons across cultures and languages on the basis of the common etic items. Adjustment to cross-cultural DIF, adequately accounts for the cross-cultural bias, meaning that the data between the two countries can be compared. When cross-cultural comparisons between the two countries are performed, the DIF-adjusted conversion table (Table 4) will need to be used. While we have provided a DIF-adjusted table for completeness, some caution should be exercised in applying the conversions in practice. If future datasets are small in size or are not normally distributed then erroneous conclusions could be drawn if simple conversions are applied.

The new (Spanish) version of the MFPDI was shown to be a cross-culturally valid measure of foot pain and disability. Its reliability was excellent (PSI = .85) achieving the thresholds required for both group and individual use [35]. Absence of DIF by age, gender and education also indicated that the translated MFPDI worked in the same way across different groups of patients.

Conversion of the raw (ordinal) scores into interval level (Rasch-transformed) values enables the use of MFPDI scores in parametric analyses [49] alongside other measures, given adequate sample sizes and normal

distribution. When data pooling or comparisons are undertaken, the cross-cultural, DIF-adjusted convertion chart helps to ensure accurate estimates and adjust for cross-cultural bias. When data comparison with another country is required, then a different conversion chart will have to be used.

This study has three main limitations. Firstly, convenience sampling was used and patients were recruited from private clinics only, as the Spanish public health system does not offer podiatry services. While the sample was not pre-specified to be explicitly representative, it comprises patients from seven different clinics and includes representation of patients with different age groups, educational backgrounds and employment status. Indeed this study has demonstrated that the MFPDI has retained its construct validity following its adaptation into Spanish. The sampling approach is unlikely therefore, to have had a significant influence on the conclusions of this study. Secondly, since the MFPDI has been validated at the testlet-level, it may not be possible to make cross-cultural comparisons of the foot disability at the individual item-level. Lastly, since this study has validated this Spanish version of the MFPDI for use in Spain, further cross-cultural validation would be required if this questionnaire is to be used in other Spanish-speaking countries such as those in South America.

Since the validity and unidimensionality of the MFPDI have been confirmed, future research can make cross-cultural comparisons in foot pain and disability between patients in the UK and Spain. In addition, further research is needed to determine the minimal clinically important difference (MCID) in each MFPDI subscale and to assess how this correlates with other measures of disability.

Conclusion

A gold standard translation process has been used to develop a Spanish version of a widely used foot-specific patient-reported outcome measure. Rasch analysis has confirmed that the MFPDI is a robust, 3-subscale measure of foot pain, function and appearance in both its English and Spanish versions. Future work can make cross-cultural comparisons using the calibrated scales.

Authors contribution

GGN and ALS were the in-country investigators. PVM and BAP were responsible for the cross-cultural adaptation into Spanish. MN undertook the cross-cultural validation using Rasch analysis and interpreted the results together with ACR. AG supplied the UK MFPDI dataset and contributed in the cross-cultural validation. MN had full access to all the data in the study and took full responsibility for the integrity of the accuracy of the data analysis. All authors contributed in the writing and review of the manuscripts and approved the final version.

Conflict of Interest Statements

All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available from the corresponding author) and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work

Acknowledgements

Firstly, the authors acknowledge the assistance of Isis Outcomes (Isis Innovation Ltd, Oxford UK) for granting permission to use the MFPDI in this study. Permissions to use the MFPDI in all translations can be sought directly from Isis Outcomes (http://isis-innovation.com/outcomes/index.html). Secondly, we would also like to acknowledge the assistance of the University of Manchester for access to the original MFPDI UK dataset. Thirdly, we would like to thank our colleagues who helped with data collection: Javier Ramos, Bárbara Pineda, María Reina, Cristina Algaba, Raquel Cintado, Encarnacion Cortes Jeronimo, Irene Garcia Paya and Guillermo Lafuente. Lastly, we would like to thank all patients that took part in this study.

References

1. Spahn, G., Schiele, R., Hell, A., Klinger, H., Jung, R., & Langlotz, A. (2004). The prevalence of pain and deformities in the feet of adolescents. Results of a cross-sectional study. Zeitschrift fur Orthopadie und ihre Grenzgebiete, 142(4), 389-396.

2. Benvenuti, F., Ferrucci, L., Guralnik, J., Gangemi, S., & Baroni, A. (1995). Foot pain and disability in older persons: an epidemiological study. Am Geriatrics Soc, 43, 479-484.

3. Badlissi, F., Dunn, J., Link, C., Keysor, J., McKinlay, J., & Felson, D. (2005). Foot musculoskeletal disorders, pain, and foot-related functional limitation in older persons. Journal of the American Geriatrics Society, 53(6), 1029.

4. Menz, H., Barr, E., & Brown, W. (2011). Predictors and persistence of foot problems in women aged 70 years and over: a prospective study. Maturitas, 68(1), 83.

5. Parsons, S., Breen, A., Foster, N., Letley, L., Pincus, T., Vogel, S., & Underwood, M. (2007). Prevalence and comparative troublesomeness by age of musculoskeletal pain in different body locations. Family Practice, 24(4), 308.

6. Menz, H. B., Gill, T. K., Taylor, A. W., & Hill, C. L. (2011). Age and gender differences in disabling foot pain using different definitions of the manchester foot pain and disability index. BMC musculoskeletal disorders, 12, 243.

7. Tanamas, S., Wluka, A., Berry, P., Menz, H., Strauss, B., Davies-Tuck, M., Proietto, J., Dixon, J., Jones, G., & Cicuttini, F. (2012). Relationship between obesity and foot pain and its association with fat mass, fat distribution, and muscle mass. Arthritis Care & Research, 64(2), 262.

8. Menz, H., & Lord, S. (2001). Foot pain impairs balance and functional ability in community-dwelling older people. Journal of the American Podiatric Medical Association, 91(5), 222.

9. Menz, H., Morris, M., & Lord, S. (2006). Foot and ankle risk factors for falls in older people: a prospective study. The journals of gerontology. Series A, Biological sciences and medical sciences, 61(8), 866.

10. Mickle, K., Munro, B., Lord, S., Menz, H., & Steele, J. (2009). ISB Clinical Biomechanics Award 2009: toe weakness and deformity increase the risk of falls in older people. Clinical biomechanics (Bristol, Avon), 24(10), 787.

11. Mickle, K., Munro, B., Lord, S., Menz, H., & Steele, J. (2011). Cross-sectional analysis of foot function, functional ability, and health-related quality of life in older people with disabling foot pain. Arthritis Care & Research, 63(11), 1592.

12. Kaoulla, P., Frescos, N., & Menz, H. B. (2011). A survey of foot problems in community-dwelling older Greek Australians. Journal of Foot and Ankle Research, 4, 23.

13. Kitaoka, H., Alexander, I., Adelaar, R., Nunley, J., Myerson, M., & Sanders, M. (1994). Clinical rating systems for the ankle-hindfoot, midfoot, hallux, and lesser toes. Foot & ankle international/American Orthopaedic Foot and Ankle Society [and] Swiss Foot and Ankle Society, 15(7), 349.

14. Budiman-Mak, E., Conrad, K., & Roach, K. (1991). The Foot Function Index: a measure of foot pain and disability. Journal of clinical epidemiology, 44(6), 561.

15. Bennett, P., Patterson, C., Wearing, S., & Baglioni, T. (1998). Development and validation of a questionnaire designed to measure foot-health status. Journal of the American Podiatric Medical Association, 88(9), 419.

16. Niki, H., Tatsunami, S., Haraguchi, N., Aoki, T., Okuda, R., Suda, Y., Takao, M., & Tanaka, Y. (2013). Validity and reliability of a self-administered foot evaluation questionnaire (SAFE-Q). Journal of Orthopaedic Science, 18(2), 298-320.

17. Garrow, A. P., Papageorgiou, A. C., Silman, A. J., Thomas, E., Jayson IV, M., & Macfarlane, G. J. (2000). Development and validation of a questionnaire to assess disabling foot pain. Pain, 85(1), 107-113.

18. Menz, H., Tiedemann, A., Kwan, M., Plumb, K., & Lord, S. R. (2006). Foot pain in community-dwelling older people: an evaluation of the Manchester Foot Pain and Disability Index. Rheumatology, 45(7), 863-867.

19. Roddy, E., Muller, S., & Thomas, E. (2011). Onset and persistence of disabling foot pain in community-dwelling older adults over a 3-year period: a prospective cohort study. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 66(4), 474.

20. Cook, C. E., Cleland, J., Pietrobon, R., Garrow, A., & Macfarlane, G. (2007). Calibration of an item pool for assessing the disability associated with foot pain: an application of item response theory to the Manchester Foot Pain and Disability Index. Physiotherapy, 93(2), 89-95.

21. Muller, S., & Roddy, E. (2009). A rasch analysis of the Manchester foot pain and disability index. Journal of Foot and Ankle Research, 2, 29.

22. Spink, M. J., Menz, H. B., & Lord, S. R. (2008). Efficacy of a multifaceted podiatry intervention to improve balance and prevent falls in older people: study protocol for a randomised trial. BMC Geriatrics, 8(1), 30.

23. Alcacer-Pitarch, B., Buch, M. H., Gray, J., Denton, C. P., Herrick, A., Navarro-Coy, N., Collier, H., Loughrey, L., Pavitt, S., & Siddle, H. J. (2012). Pressure and pain In Systemic sclerosis/Scleroderma-an evaluation of a simple intervention (PISCES): randomised controlled trial protocol. BMC Musculoskeletal Disorders, 13(1), 11.

24. Berglund, B., Nordström, G., Hagberg, C., & Mattiasson, A. C. (2005). Foot pain and disability in individuals with Ehlers-Danlos syndrome (EDS): impact on daily life activities. Disability & Rehabilitation, 27(4), 164-169.

25. Bowen, C. J., Edwards, C. J., Hooper, L., Dewbury, K., Sampson, M., Sawyer, S., Burridge, J., & Arden, N. K. (2010). Improvement in symptoms and signs in the forefoot of patients with rheumatoid arthritis treated with anti-TNF therapy. Journal of Foot and Ankle Research, 3, 10.

26. Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. Spine, 25(24), 3186.

27. Ware, J. J., Gandek, B., Keller, S., & IQOLA Group. (1996). Quality of life and pharmacoeconomics in clinical trials. In B. Spilker (Ed.), Quality of Life and Pharmacoeconomics in Clinical Trials (2nd ed., pp. 681–692.). Philadelphia: Lippincott-Raven.

28. Petersen, M., Groenvold, M., Bjorner, J., Aaronson, N., Conroy, T., Cull, A., Fayers, P., Hjermstad, M., Sprangers, M., & Sullivan, M. (2003). Use of differential item functioning analysis to assess the equivalence of translations of a questionnaire. Quality of Life Research, 12(4), 373-385.

29. Tennant, A., Penta, M., Tesio, L., Grimby, G., Thonnard, J.-L., Slade, A., Lawton, G., Simone, A., Carter, J., Lundgren-Nilsson, Ã. s., Tripolski, M., Ring, H., Biering-SÃ¸rensen, F., Marincek, C., Burger, H., & Phillips, S. (2004). Assessing and Adjusting for Cross-Cultural Validity of Impairment and Activity Limitation Scales Through Differential Item Functioning Within the Framework of the Rasch Model: The PRO-ESOR Project. Medical Care, 42(1), I-37.

30. Scott, N., Fayers, P., Aaronson, N., Bottomley, A., de Graeff, A., Groenvold, M., Koller, M., Petersen, M., & Sprangers, M. (2007). The use of differential item functioning analyses to identify cultural differences in responses to the EORTC QLQ-C30. Quality of Life Research, 16(1), 115-129.

31. Kaoulla, P., Frescos, N., & Menz, H. B. (2008). Development and validation of a Greek language version of the Manchester Foot Pain and Disability Index. Health Qual Life Outcomes, 6, 39.

32. Ferrari, S. C., Santos, F. C., Guarnieri, A. P., Salvador, N., Correa, A., Hala, A. Z. A., C√∫stodio, O., & Trevisani, V. F. M. (2008). Indice Manchester de incapacidade associada ao pé doloroso no idoso: traducao, adaptacao cultural e validasao para a lingua portuguesa. Brazilian Journal of Rheumatology, 48(6), 335-341.

33. Marinozzi, A., Martinelli, N., Panasc√¨, M., Cancilleri, F., Franceschetti, E., Vincenzi, B., Di Martino, A., & Denaro, V. (2009). Italian translation of the Manchester-Oxford Foot Questionnaire, with re-assessment of reliability and validity. Quality of Life Research, 18(7), 923-927.

34. Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. (2005). Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: report of the ISPOR Task Force for Translation and Cultural Adaptation. Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research, 8(2), 94.

35. Tennant, A., & Conaghan, P. (2007). The Rasch Measurement Model in Rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper. Arthritis and Rheumatism, 57(8), 1358-1362.

36. IBM. (2010). IBM SPSS Statistics 19.0.1: SPSS Inc.

37. RUMM Laboratory. (1998). Rasch Unidimensional Measurement Models. Perth: RUMM Laboratory Pty Ltd.

38. Rasch, G. (1967). An informal report on a theory of objectivity in comparisons. Leyden: University of Leyden.

39. Anderen, E. (1977). Sufficient statistics and latent trait models. Psychometrika, 42, 69-81.

40. Bond, T. G., & Fox, C. M. (2001). Applying the Rasch model: Fundamental measurement in the human sciences. London: Lawrence Arlbaum Associates

41. Rosenbaum, P. R. (1989). Criterion-related construct validity. Psychometrika, 54(4), 625-633.

42. Masters, G. A. (1982). Rasch model for partial credit scoring. Psychometrika, 47, 149-174.

43. Wright, B. (1996). Local Dependency, Correlations and Principal Components. Rasch Measurement Transactions, 10(3), 509-511.

44. Guemin, L., Robert, L. B., & David, A. F. (2000). Incorporating the Testlet Concept in Test Score Analyses. Educational Measurement: Issues and Practice, 19(4), 9-15.

45. Smith Jr, E. (2002). Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. Journal of Applied Measurement, 3(2), 205.

46. Brodersen, J., Meads, D., Kreiner, S., Thorsen, H., Doward, L., & McKenna, S. (2007). Methodological aspects of differential item functioning in the Rasch model. Journal of Medical Economics, 10(3), 309-324.

47. Bland, J. M., & Altman, D. G. (1995). Multiple significance tests: the Bonferroni method. BMJ, 310(6973), 170.

48. Johnson, T. (1998). Approaches to equivalence in cross-cultural and cross-national survey research. ZUMA Nachrichten Spezial, 3, 1-40.

49. Wright, B., & Linacre, J. (1989). Observations are always ordinal; measurements, however, must be interval. Archives of physical medicine and rehabilitation, 70(12), 857.

Table 1: Sample characteristics by country

Country		***Spain***	***UK***
Sample		333	682
Age	Mean (SD)	55.2 (16.7)	51.6 (15.2)
Gender	Female (%)	248 (74.5)	416 (61.0)
Education background	No formal education (%)	15 (4.5)	
	Primary school (%)	98 (29.4)	
	High school (%)	113 (33.9)	
	Higher education (%)	107 (32.1)	
Work status	Unemployed (%)	80(24.0)	
	Employed (%)	137 (41.1)	
	Retired (%)	116 (34.8)	

SD, standard deviation

Table 2: Results of Rasch analysis for the Spanish, the UK and the pooled datasets

	Item Fit Residual		Person Fit Residual		Chi Square Interaction		PSI		Unidimensionality
Analysis	Mean	SD	Mean	SD	Value (DF)	p		N	Independent t-tests (95%CI)
Spain (17-item scale)	-0.019	1.823	-0.158	1.214	179.197 (68)	<0.001	0.908	319	
Spain (3-testlet level)	-0.164	3.067	-0.364	0.931	15.945 (12)	0.194	0.848	319	0.016 (-0.006, 0.040)
UK (17-item scale)	-0.709	2.487	-0.559	1.983	292.526 (153)	<0.001	0.899	554	
UK (3-testlet level)	-0.366	2.799	-0.429	0.986	31.024 (21)	0.073	0.823	452	0.010 (-0.011, 0.032)
Pooled (17-item scale)	-0.654	2.989	-0.453	2.072	451.031 (153)	<0.001	0.903	873	
Pooled (3-testlet level)	-0.766	4.397	-0.441	1.059	49.169 (27)	0.006	0.835	771	0.010 (-0.011, 0.032)
Pooled (adjusted for DIF)	-0.420	2.979	-0.415	0.983	57.944 (45)	0.093	0.844	771	
Expected values	0	1	0	1		>0.05	>0.7		Lower-bound CI <0.05

SD, standard deviation; DF,degrees of freedom; P,χ^2probability, where non-significant P impliesfit to the model; PSI,person separation index; CI,confidence interval

Table 3: Item fit for each testlet after adjusting for cross-cultural DIF

Testlet	Location	SE	Fit Residuals	χ^2	DF	*P-value**
Functional - Spain	-0.292	0.020	-2.684	5.498	203.510	0.789
Functional - UK	-0.141	0.023	-4.219	18.475	245.610	0.030
Appearance - Spain	0.400	0.056	3.096	9.633	203.510	0.381
Appearance - UK	0.074	0.054	1.267	7.693	283.250	0.565
Pain	-0.041	0.025	0.439	16.645	456.130	0.055

SE,standard error; DF, degrees of freedom; *P,Bonferroni-adjusted χ^2probability where non-significant P impliesfit to the model.

Table 4: Conversion of MFPDI raw (subscale) scores into Rasch-transformed scores for Spain and the UK.

Raw (observed) scores	Rasch transformed, DIF-adjusted scores*				
	Functional (Spain)	Functional (UK)	Appearance (Spain)	Appearance (UK)	Pain (Common)
0	0.0	0.0	0.0	0.0	0.0
1	3.2	2.6	1.1	1.2	2.0
2	5.6	4.5	1.9	2.1	3.4
3	7.4	5.8	2.9	2.9	4.4
4	8.8	6.7	4.0	4.0	5.1
5	10.0	7.4			5.7
6	10.8	7.9			6.3
7	11.5	8.4			6.9
8	12.0	8.8			7.6
9	12.5	9.1			8.6
10	12.8	9.4			10.0
11	13.2	9.7			
12	13.5	10.1			
13	13.9	10.5			
14	14.3	11.0			
15	14.6	11.6			
16	15.1	12.5			
17	15.7	13.6			
18	16.5	15.2			
19	17.8	17.2			
20	20.0	20.0			

*The Rasch-transformed scores are at interval-level