European Journal of Physical and Rehabilitation Medicine
EDIZIONI MINERVA MEDICA

Reliability, responsiveness and minimal clinically important difference of the two Fear Avoidance and Beliefs Questionnaire scales in Italian subjects with chronic low back pain undergoing multidisciplinary rehabilitation

Marco MONTICONE, Luca FRIGAU, Howard VERNON, Barbara ROCCA, Andrea GIORDANO, Salvatore SIMONE VULLO, Francesco MOLA, Franco FRANCHIGNONI

European Journal of Physical and Rehabilitation Medicine 2020 May 18
DOI: 10.23736/S1973-9087.20.06158-4

Article type: Original Article

Article first published online: May 18, 2020
Manuscript accepted: May 16, 2020
Manuscript revised: May 4, 2020
Manuscript received: December 30, 2019

Reliability, responsiveness and minimal clinically important difference of the two Fear Avoidance and Beliefs Questionnaire scales in Italian subjects with chronic low back pain undergoing multidisciplinary rehabilitation.

Marco Monticone [a], Luca Frigau [b], Howard Vernon [c], Barbara Rocca [d], Andrea Giordano [e], Salvatore Simone Vullo [a], Francesco Mola [b], Franco Franchignoni [d].

[a] Dept. Medical Sciences and Public Health, University of Cagliari, Cagliari; and Neurorehabilitation Unit, Dept. Neuroscience and Rehabilitation, G. Brotzu Hospital, Cagliari, Italy.

[b] Dept. Economics and Business Science, University of Cagliari, Italy.

[c] Canadian Memorial Chiropractic College, Toronto, Ontario, Canada.

[d] ICS Maugeri, IRCCS, Physical Medicine and Rehabilitation Unit, Institute of Lissone (MB), Italy.

[e] ICS Maugeri, IRCCS, Bioengineering Unit, Institute of Veruno (NO), Italy.

Corresponding author

Marco Monticone

Dept. of Medical Sciences and Public Health, University of Cagliari.

Cittadella Universitaria, Strada Statale, 554 - Monserrato, Cagliari, Italy.

Phone: +39 0706753109; Fax: +39 07051096487

Email: marco.monticone@unica.it

Declaration of Interest

The authors declare that they have no competing interests.

ABSTRACT

***Background*:** The Fear-Avoidance Beliefs Questionnaire (FABQ) is a widely used outcome measure. There is still a lack of information concerning responsiveness and minimal clinically important difference (MCID), limiting its use for clinical and research purposes.

Aim: To examine reliability, responsiveness and MCID of the two FABQ scales in subjects with chronic Low Back Pain.

***Design**:* Methodological research based on a prospective single-group observational study.

Setting: Outpatient, rehabilitation unit.

Population: Chronic non-specific Low Back Pain.

***Methods**:* At the beginning and the end of a multidisciplinary programme (8-week), 129 subjects completed the FABQ scales. Reliability was determined as internal consistency (Cronbach's alpha) and test-retest reliability (96-hour interval; n= 64). Responsiveness was calculated both by distribution-based and anchor-based methods, using as external criterion the global perceived effect scale (GPE: 7 levels), rated by each individual.

Results: Cronbach's alpha and ICC(2,1) were respectively: 0.75 and 0.90 for FABQ-Physical Activity scale (FABQ-PA), and 0.85 and 0.95 for FABQ-Work scale (FABQ-W). Minimum detectable change (MDC_{95}) values were 3.69 points for FABQ-PA, and 5.95 points for FABQ-W. In receiver-operating characteristics curves, splitting GPE data into null/minimal/moderate improvement vs. large improvement (GPE 0-2 vs. GPE 3): 1) for FABQ-PA, the area under the curve (AUC) was 0.97. The best cutoff score identifying meaningful change in fear-avoidance beliefs about physical activity was a change of 4 points; 2) for FABQ-W, the AUC was 0.97 and the best cutoff score for meaningful change in fear-avoidance beliefs about work activities was a change of 7 points.

***Conclusions**:* After triangulation of the above results, a change of 4 points for FABQ-PA and 7

points for FABQ-W were selected as MCID. These two values represent cutoffs that seem to accurately identify meaningful change in fear-avoidance beliefs, according to subject's judgement.

Clinical rehabilitation impact: The present study calculated - in a sample of people with chronic Low Back Pain – the minimal clinically important change of the two FABQ scales (FABQ-Physical Activity scale and FABQ-Work scale). These values increase confidence in interpreting score changes, thus enhancing their meaningful use in both research and clinical contexts.

KEYWORDS: Chronic low back pain; Rehabilitation; Responsiveness; Minimal clinically important difference; Fear-Avoidance Beliefs Questionnaire; Kinesiophobia.

Short title: Minimal clinically Important difference of the FABQ

Abstract word count: 340 words.

Manuscript word count: 3287 words.

Two tables, two figures.

INTRODUCTION

In people with chronic low back pain (LBP), psychological factors are important predictors for long-term disability, and pain-related fear is often more disabling than the pain itself [1]. Indeed, it is assumed that fear-avoidance beliefs prevent subjects from regaining normal function by promoting the development of guarded movements and contributing to disability onset and persistence [2–4]. Fear-avoidance is also an important causal mediator of the cognitive-behavioural treatment effect [5,6]. The fear-avoidance model suggests that subjects without fear-avoidance beliefs are more likely to confront pain problems and are more active in the coping process [2].

The Fear Avoidance and Beliefs Questionnaire (FABQ) is a commonly-used condition-specific health status measure for the assessment of fear of movement related to LBP. It was developed by Waddell G *et al* and published in 1993 as a means of focusing on subjects' beliefs about how physical activity and work affect pain [7]. The original 16-item FABQ showed acceptable psychometric properties such as internal consistency, reproducibility, and validity; however, factor analysis suggested dividing the FABQ into two scales, the first assessing fear-avoidance beliefs about work (FABQ-Work (W); 7 items, range: 0-42) and the second assessing fear-avoidance beliefs about physical activities (FABQ-Physical Activity (PA); 4 items, range: 0-24) [7].

The FABQ has been translated and cross-culturally adapted into many languages, including the Italian, finding psychometric properties similar to the original [8]. Some responsiveness indices of the two FABQ scales (i.e. indicators of the tool's ability to detect changes in the construct to be measured over time [9,10]) have been examined with a fairly large range of results [11-15]. However, the minimal clinically important difference (MCID, also known as minimally important change, i.e. the smallest change in score of the construct to be measured that subjects perceive to

be important [9,10]) of the FABQ-W and FABQ-PA has not been determined in literature, although it represents a key parameter for clinicians (in order to assess interventions effectiveness and guide decision making), as well as for researchers (in order to investigate treatment effectiveness in clinical trials, sample size estimates, and cost evaluations) [9,10,16-18].

Hence, we aimed to calculate, in a sample of Italian subjects with chronic LBP undergoing multidisciplinary rehabilitation, reliability and responsiveness (in terms of standard error of measurement [SEM], minimum detectable change [MDC]), and MCID) of the two FABQ scales (FABQ-W and FABQ-PA), to enhance confidence in their use in clinical practice and research.

MATERIALS AND METHODS

Subjects

Outpatients consecutively admitted to a Rehabilitation Unit were enrolled between January and June 2015. Inclusion criteria were: an established diagnosis of chronic non-specific LBP (i.e. documented history of pain lasting for more than twelve weeks, without a recognizable, specific pathoanatomical cause of the symptom [19,20]); good understanding of Italian language; and being working adults (i.e. employee, self-employed, or housewife). Exclusion criteria were: i) presence of one of the following conditions: systemic illness, cognitive impairment (Mini Mental State Examination of <24), recent (<12 weeks) myocardial infarction, cerebrovascular events, and having previously received any exercise therapy and/or cognitive-behavioural therapy: ii) refusal or inability to adhere to the treatment.

The subjects' sociodemographic and clinical characteristics were investigated using a specific form including information on: age, gender, body mass index, pain duration, education level, occupation, co-morbidities, and marital status.

To assess the test–retest reliability, 64 participants randomly selected (coin tossing) from the total sample were asked to complete the questionnaire twice, within a 96-hour interval, before start of the rehabilitation programme.

This research was part of a prospective single-group observational study approved by the Institutional Review Board of the local Hospital (date of approval: 22/12/2014; no. 26/14). All subjects gave their written consent to participate and the study was conducted in conformity with ethical and humane principles of research.

Outcome measures

THE 'FEAR-AVOIDANCE BELIEFS QUESTIONNAIRE' (FABQ) - It assesses subjects' beliefs in two domains examining how respectively physical activity and work affect LBP. The items are scored on a 7-point Likert scale (from 0 "completely disagree" to 6 "completely agree"). Higher values indicate higher degree of fear-avoidance beliefs. The score of each FABQ scale is used independently: in the FABQ-W (range 0-42) 7 out of the 11 original items are added to a sum score (# 6, 7, 9-12, and 15), and in the FABQ-PA (range 0-24) 4 out of the 5 original items (# 2-5). The 5 remaining questions are not used in scoring but just as delusive items, as proposed by Waddell et al. [7]. The satisfying psychometric properties of the Italian FABQ (FABQ-It) have been proved (Cronbach α of 0.82; test-retest reliability of 0.87; and moderate correlation with the Tampa Scale of Kinesiophobia of 0.440) [8].

THE GLOBAL PERCEIVED EFFECT – At the end of treatment subjects' global perceived effect of the intervention related to fear-avoidance beliefs (GPE) was evaluated using the following question: "*Overall, how much did the treatment you received help your fear of movement due to current pain?*". The perceived effect was scored using a 7-level Likert scale with following response

options: -3 “made things a lot worse”; -2 “made things worse”; -1 “made things partially worse”; 0 “did not help (unchanged)”; +1 “helped only a little”; + 2 ”helped”; + 3 “helped a lot” [21].

The physiotherapist (blind to FABQ results) who treated each participant was asked to answer a similar question (“*Overall, how much do you think that the treatment you delivered did help your subject’s fear of movement due to her/his current pain?”*), scored in the same way, in order to assess his/her perceived change in subjects’ fear-avoidance beliefs, over the treatment period (namely, GPE-PT).

Procedure

Participants underwent an individual 60-min motor training sessions twice a week for eight-week outpatient programme, that included exercises aimed at improving postural control, strengthening and stabilising the back muscles, and stretching. Subjects additionally received 60-min cognitive-behavioural therapy sessions once a week for eight-week aimed at modifying their fear of movement beliefs, and ensuring gradual reactions to illness behaviours, as well as information and education in ergonomic principles. This rehabilitation programme was the same for all of the subjects and was preliminary tested for efficacy by means of a randomized controlled pilot study [22]. During the study, mild analgesics (i.e. paracetamol) and non-steroidal anti-inflammatory drugs (NSAIDs) were permitted. No more than one tablet per day on demand for more than 5-7 days continuously was allowed; their intake as well as patients’ symptoms and needs were constantly monitored.

All subjects were required informed consent concerning research aim, questionnaires and procedures, before study began. Questionnaires were administered as part of the pre/post-rehabilitation assessment by our staff (one physiotherapist with a master’s degree, not involved in

any other part of the study), who checked them and returned any uncompleted part to the participants for completion; thus, there were no missing values.

Statistical analysis

In each of the two FABQ scales indexes of reliability and responsiveness have been calculated, being reliability a critical component of responsiveness.

Reliability has been determined as: i) internal consistency, calculating Cronbach's alpha; the closer to 1, the higher the internal consistency of the scale's items. Alpha values of >0.70 are recommended for group-level comparisons, whereas a minimum of 0.85 to 0.90 is desirable for individual judgments; ii) test-retest reliability of global scores, by means of the Intraclass Correlation Coefficient, with a 'two-way mixed effects, single measurement' model, ICC(2,1) [10]. Test-retest reliability was expected to be high and thus 64 subjects were judged as adequate to verify a test-retest correlation of 0.8 with a 95% CI width of 0.2 ($\alpha = 0.05$) [23].

Responsiveness was investigated both with distribution- and anchor-based methods. As for distribution-based methods (that analyze the ability to detect change in general), we calculated the standard error of measurement (SEM) based on the analysis of variance used to produce the test-retest ICC (equation 1) [9,10]

$$SEM = SD\sqrt{1 - ICC_{2,1}} \qquad \text{Equation 1}$$

Based on the SEM, we calculated the minimum detectable change (MDC) as in equation 2.

$$MDC = SEM * z\ value * \sqrt{2} \qquad \text{Equation 2}$$

The MDC represents the smallest change in score that probably reflects true change, and not simply measurement error. Its 90% confidence level (MDC_{90}) corresponds to a z value of 1.65, while the 95% confidence level (MDC_{95}) to a z value of 1.96 in Equation 2. The meaning of these values is that if, for instance, a subject has a change score larger than the MDC_{95} threshold, it is possible to state (with 95% confidence) that this change is real and not due to measurement error [10].

The anchor-based method used GPE as an external criterion (anchor) to determine whether changes in outcome scores were clinically meaningful [16-18]. We analyzed two parameters: 1) for the mean change approach, we calculated the mean change of participants graded on the GPE as not improved (GPE= 0 “did not help (unchanged)”), minimally improved (GPE= 1, “helped only a little”), moderately improved (GPE= 2, “helped”), and largely improved (GPE= 3, “helped a lot”); and 2) for the receiver operating characteristic (ROC) curve approach, we determined the best cutoff score and the area under the curve (AUC) after splitting the participants based on the cutoff point indicating large improvement (GPE= 0, 1 or 2 vs. GPE= 3).

A ROC curve plots sensitivity values (y-axis) against false-positive rates (1-specificity, x-axis) [10,24]. Sensitivity was calculated as the number of participants correctly identified as improved based on the cutoff value of the external criterion, divided by all participants identified as having undergone a large improvement (GPE= 3). Specificity refers to the number of participants who were correctly identified as not improved (based on the same cutoff value) divided by all participants who did not undergo a large change (GPE= 0-2). The optimal cutoff score was chosen as the point that jointly maximized sensitivity and specificity (being associated with the least amount of misclassification). The AUC can be interpreted as the probability of correctly identifying, in randomly selected pairs of subjects who have and have not shown an improvement, the one who has improved. This area theoretically ranges from 0.5 (no discriminating accuracy) to 1.0

(perfect accuracy): the greater the AUC, the greater a measure's ability to distinguish subjects who improve from those who do not improve; as a general rule, an AUC>0.80 is considered to have excellent discrimination [10, 24]. As an anchor-based method, ROC curves were selected, which are useful indicators of the relationship between a measure and an external indicator of change, such as the GPE. The optimal cutoff point was estimated from ROC curve, selecting the change score associated with the smallest residual sum of sensitivity and specificity.

In addition, in order to check the validity of our GPE as anchor, we examined its correlation with the pre-post treatment change score in both FABQ-W and FABQ-PA , as well as with the rating of perceived change in subjects' fear-avoidance as scored by his/her physiotherapist (GPE-PT), hypothesizing to find a nontrivial correlation (>0.30-0.50) with all of them [9].

RESULTS

Subjects

A flowchart of the enrolment of participants is in Figure 1. The study population consisted of 129 subjects (43 males, 33%) with an average age of 48 ± 16 years, an average pain duration of 21 ± 17 months, and an average body mass index of 23 ± 4 kg/m^2. At baseline, average pain intensity was of 5.8 ± 2 (on a 0-10 numerical rating scale) and an average lumbar disability was of 33.5 ± 7 (on the 0-100 Oswestry Disability Index). The main socio-demographic characteristics of the sample are shown in Table 1.

The median scores (interquartile range, IQR) were as follows: i) at admission, FABQ 41 points (IQR 17); FABQ-W 27 (IQR 10) and FABQ-PA 16 (IQR 7); ii) at discharge, FABQ 33 points (IQR 17); FABQ-W 21 (IQR 12) and FABQ-PA 12 (IQR 7); iii) as for the test-retest group (n=64), baseline: FABQ 44

points (IQR 18); FABQ-W 26.5 (IQR 17.5) and FABQ-PA 17 (IQR 5.5); retest: FABQ 44 points (IQR 17.5); FABQ-W 26.5 (IQR 18) and FABQ-PA 17 (IQR 6).

The values related to Cronbach's alpha and test-retest reliability – ICC(2,1) - in both scales are shown in Table 2, together with related SEMs and MDCs.

According to the GPE scores, 41.9% of subjects (n=54) had no or minimal improvement (GPE 0 or 1), 32.5% (n=42) had a moderate improvement (GPE= 2), and 24.8% (n=32) had a large improvement (GPE= 3). Just one person (0.8% of the sample) referred a minimal deterioration in fear-avoidance beliefs (GPE: -1 "made things partially worse") and was not inserted into calculations.

The score changes - median (IQR) - according to GPE levels were as follows: i) 0 points (IQR 2), for both FABQ-PA and FABQ-W, in those who rated a null or minimal improvement (GPE= 0 or 1); ii) 3 points (IQR 3) for FABQ-PA and 7 (IQR 6) for FABQ-W in those who rated moderate improvement (GPE= 2); and iii) 8 points (IQR 8) for FABQ-PA, and 10 points (IQR 8) for FABQ-W in those who rated a large improvement (GPE= 3).

Splitting GPE data into null/minimal/moderate improvement vs. large improvement (i.e. GPE= 0-2 vs. GPE= 3): 1) for FABQ-PA, the AUC was 0.97 (95% CI 0.95-0.99). The cutoff score that best identified meaningful change in fear-avoidance beliefs about physical activity was a change of 4 points (sensitivity 90.2; specificity 90.8); 2) for FABQ-W, the AUC was 0.97 (95% CI 0.95-0.99) (Figure 2). The cutoff score that best identified meaningful change in fear-avoidance beliefs about work activities was a change of 7 points (sensitivity 90.2; specificity 94.3).

Overall, a MCID value of 4 points for FABQ-PA and 7 points for FABQ-W was selected as the best triangulation of these results, with values higher that the respective MDC_{95} value for each scale.

The correlations between GPE and pre-post treatment change score were good both with FABQ-W (r_s= 0.71) and FABQ-PA (r_s= 0.68), while that between GPE and GPE-PT was excellent (r_s= 0.84).

DISCUSSION

Different approaches have been previously used to calculate responsiveness: in this study, we used both distribution-based and anchor-based methods.

It has been recommended that distribution-based methods should be used cautiously as they tend to measure the magnitude of change scores rather than their validity [9, 10].

In general, distribution-based approaches should act only as temporary surrogates, pending availability of empirically established anchor-based MCID values. Indeed, the results of anchor-based methods (and related values of MCID) should be considered more important than those of the distribution-based methods (including values of MDC) [9].

Our responsiveness indices (for FABQ-PA: MDC_{95}=3.69; and for FABQ-W: MDC_{95}=5.95) calculated with distribution-based methods showed values in line with the literature, that reported MDC_{95} values ranging from 2.1 [12] to 9 [14] for FABQ-PA, and from 2.75 [12] to 16 [14] for FABQ-W.

Concerning the point above, it is important to recognize that: (a) the MDC values reflect the specific ICC values (on equal terms, the lower the reliability coefficient, the greater the SEM and the higher the MDC will be) and standard deviation values in each particular study, and therefore different groups of individuals would be expected to generate different values; (b) changes below the MDC threshold could be interpreted at an individual level as random fluctuations in score rather than actual change [24].

As for anchor-based methods, the first issue related to the appropriateness of cutoff values is the selection of the anchor. We used a 7-point GPE (3 negatives, 3 positives, and an “unchanged”

category) and considered subjects with GPE values of +3 ('helped a lot') as largely improved and the others (GPE= 0, +1, +2) as not-significantly or moderately improved. In the literature, there is no agreement on the type of GPE to use, the threshold at which to dichotomize the GPE, or which groups to include in the analysis [25]. For instance, the choice to include in the "changed" group those who reported moderate improvement (GPE= 2, 'helped') would have inevitably led to a lower cutoff point in the ROC analysis. However, the value would have been lower than our MDCs; thus, this finding corroborates our cutoff choice.

Moreover, the validity of using GPE as an anchor (i.e. the ability of the respondents to reasonably estimate the change in their fear-avoidance beliefs) in this study has been corroborated by its high correlation with pre-post treatment change scores in both FABQ-W and FABQ-PA, as well as with the rating of change in subject's fear-avoidance perceived by his/her physiotherapist.

The cutoff score identified by receiver-operating characteristics curves (a change of 4 points in FABQ-PA and of 7 points FABQ-W) showed very good sensitivity, specificity, and classification accuracy (all of >90%).

Overall, in establishing the MCID, we have to select values that are above the MDC.

After triangulation of all our results, we argue that:

- as for FABQ-PA, a change of 4 points seems the most appropriate MCID value. This value is higher than MDC_{95} (3.69 points) and over the median value of those who rated themselves as moderately improved (GPE= 2), and represents the best cutoff in the ROC curve (accuracy of 93%).

- as for FABQ-W, a change of 7 points appears as the best MCID. This value represents the best cutoff in the ROC curve (and exactly the median of those who rated themselves as moderately improved), and is higher than MDC_{95} (5.95 points).

The use of these thresholds should be clinically driven. Adopting an higher threshold, fewer people would be identified by the instrument as having shown a minimal clinically important change in

their fear-avoidance beliefs. This could increase, for example, the risk of undertreating some psychological factors playing a major role in the disability levels exhibited by chronic LBP subjects, particularly in the absence of established FABQ cutoff values for low, moderate or high levels of fear-avoidance beliefs [26,27].

Overall, a change of 4 points represents a variation of 16.7% for FABQ-PA (maximum score: 24), and that of 7 points a variation of 16.7% for FABQ-W (maximum score: 42). To the Authors' knowledge this is the first time these estimates are made available.

However, caution in generalizing these results is warranted due to some study limitations:

α) measurement error (and parameters derived from it) are not often constant across different levels of function and related scores [28];

β) the MCID estimates are context specific, and thus they vary according to methods, anchor and sample characteristics (including disease severity and baseline values) [29,30];

χ) a subject's GPE is likely to include constructs additional to those measured by the questionnaire assessing the perceived effect of the intervention related to fear-avoidance beliefs [25,31]. Indeed, subjects often show a problematic ability to understand the context of the GPE question, and thus they take into account also global changes in pain, disability and/or quality of life when assessing their improvement (or deterioration). Likewise, it has been shown that also physiotherapists' ratings of perceived subject's fear-avoidance seem unable to discriminate fear-avoidance from other factors (such as catastrophizing or disability) [32]. Furthermore, the GPE (and the MCID values derived from it) may suffer from the problem of the subjective retrospective judgments of change (e.g., due to "recall bias") [33].

δ) the selection criteria of our convenience sample (recruited with a consecutive sampling method) may represent a threat to external validity. Our sample was a cross-section of

adults drawn from a single Italian rehabilitation facility and with different severities: further studies are recommended in other countries and contexts.

ε) the responsiveness of the two FABQ scales in subjects identified as worsening needs to be addressed yet.

In conclusion, this study findings show that both FABQ-W and FABQ-PA were sensitive in detecting clinical changes in subjects with chronic LBP undergoing multidisciplinary rehabilitation. The MCIDs provided by this study represent cutoffs that seem to accurately identify important change in fear-avoidance beliefs, according to subject's judgement. Thus, it is recommended taking these MCID estimates into account when assessing improvement or planning clinical studies on similar samples.

ACKNOWLEDGEMENTS

The authors would like to thank all participants and physiotherapists who took part in the study. Some research activities described in this paper have been conducted within the R&D project "Cagliari2020", partially supported by the Italian University and Research Ministry (grant# MIUR_PON04a2_00381).

REFERENCES

1. Pincus T, Vlaeyen JW, Kendall NA, Von Korff MR, Kalauokalani DA, Reis S. Cognitive-behavioral therapy and psychosocial factors in low back pain: directions for the future. Spine 2002; 27: E133–8.

2. Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000; 85: 317–32.

3. Grotle M, Vøllestad NK, Brox JI. Clinical course and impact of fear-avoidance beliefs in low back pain: prospective cohort study of acute and chronic low back pain: II. Spine 2006; 31: 1038–46.

4. Burke AL, Mathias JL, Denson LA. Psychological functioning of people living with chronic pain: A meta-analytic review. Br J Clin Psychol 2015; 54: 354-60.

5. Monticone M, Ambrosini E, Rocca B, Cazzaniga D, Liquori V, Pedrocchi A, et al. Group-based task-oriented exercises aimed at managing kinesiophobia improved disability in chronic low back pain. Eur J Pain 2016; 20: 541-51.

6. Fordham B, Ji C, Hansen Z, Lall R, Lamb SE. Explaining how cognitive behavioral approaches work for low back pain: Mediation analysis of the Back Skills Training Trial. Spine 2017; 42: E1031-9.

7. Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A Fear-Avoidance Beliefs Questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain 1993; 52:157–68.

8. Monticone M, Baiardi P, Bonetti F, Ferrari S, Foti C, Pillastrini P, et al. The Italian version of the Fear-Avoidance Beliefs Questionnaire (FABQ-I): cross-cultural adaptation, factor analysis, reliability, validity, and sensitivity to change. Spine 2012; 37: E374–80.

9. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008; 61:

102–9.

10. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. Health Qual Life Outcomes 2006; 4: 54.

11. Ibrahim AA, Akindele MO, Kaka B, Bello B. Translation, cross-cultural adaptation, and psychometric properties of the Hausa version of the Fear-Avoidance Beliefs Questionnaire in patients with low back pain. Scand J Pain 2019; 19: 83–92.

12. Bid DD, Soni NC, Rathod PV, Ramalingam AT, Sinha RK. Cross cultural adaptation, reliability and validity of Gujarati version of Fear Avoidance Belief Questionnaire in chronic low back pain. Natl J Integr Res Med 2016; 7: 1-8.

13.Terho H, Haapea M, Paananen M, Korniloff K, Häkkinen A, Karppinen J. Translation and validation of the Finnish version of the Fear-Avoidance Beliefs Questionnaire (FABQ). Scand J Pain 2016; 10: 113-8.

14. Grotle M, Brox JI, Vøllestad NK. Reliability, validity and responsiveness of the Fear-Avoidance Beliefs Questionnaire: methodological aspects of the Norwegian version. J Rehabil Med 2006; 38:346–53.

15. Chaory K, Fayad F, Rannou F, Lefèvre-Colau MM, Fermanian J, Revel M, et al. Validation of the French version of the fear avoidance belief questionnaire. Spine 2004; 29:908–13.

16. Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a “Core Outcome Set” – a practical guideline. Trials 2016; 17: 449.

17. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: a critical review and recommendations. J Clin Epidemiol 2000; 53:459–68.

18. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. The COSMIN

checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. BMC Med Res Methodol 2010; 10: 22.

19. Airaksinen O, Brox JI, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, et al. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur Spine J 2006; 15 Suppl 2; S192-300.

20. Maher C, Underwood M, Buchbinder R. Non-specific low back pain. Lancet 2017; 389: 736–47.

21. Kamper SJ, Ostelo RW, Knol DL, Maher CG, de Vet HC, Hancock MJ. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. J Clin Epidemiol 2010; 63: 760–6.

22. Monticone M, Ambrosini E, Rocca B, Magni S, Brivio F, Ferrante S. A multidisciplinary rehabilitation programme improves disability, kinesiophobia and walking ability in subjects with chronic low back pain: results of a randomised controlled pilot study. Eur Spine J 2014; 23: 2105–13.

23. Bonett DG. Sample size requirements for estimating intraclass correlations with desired precision. Stat Med. 2002; 21:1331–5.

24.Portney LG, Watkins MP. Foundations of clinical research: applications to practice, 3rd ed. Upper Saddle River, NJ: Prentice Hall Health; 2009.

25. Engel L, Beaton DE, Touma Z. Minimal Clinically Important Difference: A review of outcome measure score interpretation Rheum Dis Clin North Am 2018; 44: 177-88.

26. Wertli MM, Rasmussen-Barr E, Weiser S, Bachmann LM, Brunner F. The role of fear avoidance beliefs as a prognostic factor for outcome in patients with nonspecific low back pain: a systematic review. Spine J 2014; 14: 816-36.

27. Wertli MM, Rasmussen-Barr E, Held U, Weiser S, Bachmann LM, Brunner F. Fear-avoidance beliefs-a moderator of treatment efficacy in patients with low back pain: a systematic review.

Spine J. 2014; 14: 2658-78.

28. Romero S, Bishop MD, Velozo CA, Light K. Minimum detectable change of the Berg Balance Scale and Dynamic Gait Index in older persons at risk for falling. J Geriatr Phys Ther 2011; 34: 131-7.

29. Wang YC, Hart DL, Stratford PW, Mioduski JE. Baseline dependency of minimal clinically important improvement. Phys Ther 2011; 91: 675-88.

30. Ward MM, Guthrie LC, Alba M. Dependence of the minimal clinically important improvement on the baseline value is a consequence of floor and ceiling effects and not different expectations by patients J Clin Epidemiol 2014; 67: 689-96.

31. Kamper SJ, Ostelo RW, Knol DL, Maher CG, de Vet HC, Hancock MJ. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. J Clin Epidemiol. 2010; 63: 760-6.e1.

32. Calley DQ, Jackson S, Collins H, George SZ. Identifying patient fear-avoidance beliefs by physical therapists managing patients with low back pain. J Orthop Sports Phys Ther 2010; 40: 774-83.

33. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. J Man Manip Ther 2009; 17:163–70.

Table 1. Socio-demographic characteristics of the study sample (n=129).

Variable	No.	%
Marital status		
Unmarried	52	40
Married	70	54
Widowed	7	6
Employment		
Employee	70	54
Self-employed	45	35
Housewife	14	11
Education		
Elementary school	12	9
Middle school	31	24
High school	49	38
University	37	29
Smoking		
Yes	20	16
No	109	84
Comorbidities		
Cardiac diseases	30	24

Respiratory diseases	18	14
Gastroenteric diseases	7	5
Endocrine diseases	11	9
None	63	48

Table 2 – Indices of reliability (internal consistency; test-retest reliability) and of distribution-based responsiveness, related to FABQ-PA and FABQ-W.

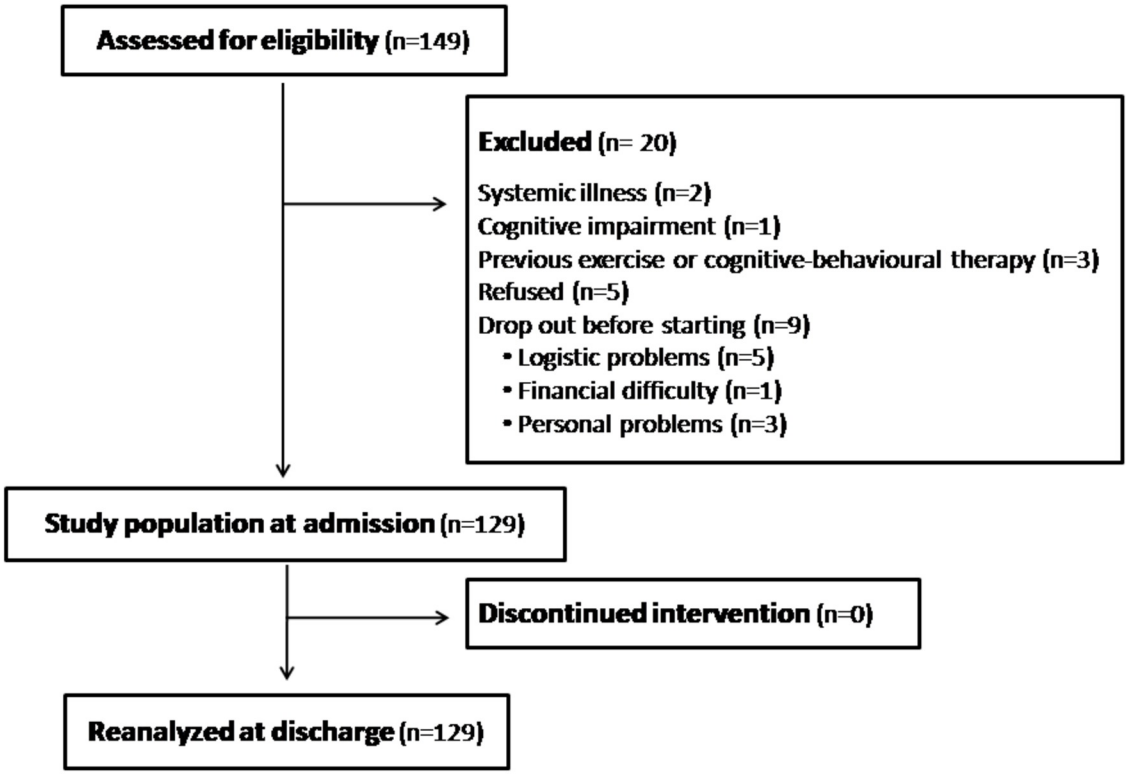
	FABQ -PA	**FABQ-W**
Cronbach's alpha	0.75	0.85
Test-retest reliability, ICC(2,1)	0.90	0.95
SEM (score points)	1.33	2.15
MDC_{90} (score points)	3.11	5.01
MDC_{95} (score points)	3.69	5.95

FABQ-PA: Fear-Avoidance Beliefs-Physical Activity; FABQ-W: Fear-Avoidance Beliefs-Work; ICC: Intraclass Coefficient Correlation; SEM: Standard Error of Measurement; MDC: Minimum Detectable Change.

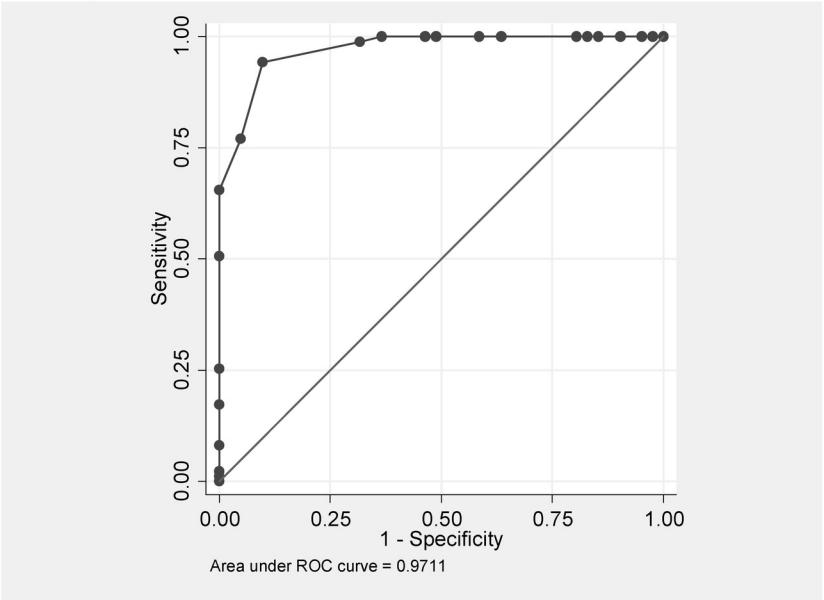
Figure legends

Figure 1 – Flowchart of enrolment procedure

Figure 2 - Receiver-operating-characteristic curves of the FABQ-PA and FABQ-W, showing their overall accuracy in identifying a large improvement (reduction) in fear-avoidance beliefs, according to the GPE at post-treatment (GPE 0, +1, +2 vs. GPE +3).



FABQ-PA



FABQ-W