J Rehabil Med 2006; 38: 346–353

Taylor & Francis
Taylor & Francis Group

RELIABILITY, VALIDITY AND RESPONSIVENESS OF THE FEAR-AVOIDANCE BELIEFS QUESTIONNAIRE: METHODOLOGICAL ASPECTS OF THE NORWEGIAN VERSION

Margreth Grotle, PT, PhD[1,2], Jens I Brox, MD, PhD[3] and Nina K Vøllestad, PhD[2]

From the [1]National Resource Center for Rehabilitation in Rheumatology, Diakonhjemmet Hospital, [2]Section for Health Science, University of Oslo and [3]Department of Orthopaedics, Section for Physical Medicine and Rehabilitation, National Hospital, Oslo, Norway

Objective: **To evaluate reliability, validity and responsiveness of the Fear-Avoidance Beliefs Questionnaire (FABQ) for use in Norwegian patients with low back pain.**
Design: **A prospective cohort study with 2 groups.**
Patients: **The questionnaire was tested in 123 patients with acute low back pain and 50 patients with chronic low back pain.**
Methods: **A translation and cross-cultural adaptation was performed. Test-retest reliability was assessed in 28 patients with chronic low back pain. Responsiveness was assessed in acute low back pain.**
Results: **Two factors for the FABQ were confirmed; fear-avoidance beliefs about work (FABQ–Work) and physical activity (FABQ–PA), accounting for 60% and 54% of the total variance in acute and chronic low back pain, respectively. For FABQ–Work and FABQ–PA internal consistency was 0.90 and 0.79, intra-class correlation coefficients 0.82 and 0.66, minimal detectable changes 12 and 9 points, and coefficients of variation were 16% and 23%. The FABQ correlated weakly to moderately with pain, disability, distress, and clinical variables. Standardized response means were low for FABQ–Work (0.32) and moderate (0.56) for FABQ–PA. Both FABQ subscales showed initially floor and/or ceiling effects.**
Conclusion: **The Norwegian FABQ version had acceptable factor structure, internal consistency, test-retest reliability and construct validity. The responsiveness of the FABQ–Work was low, and for the FABQ–PA moderate, in the acute sample.**

Key words: low back pain, Fear-Avoidance Beliefs Questionnaire, translation, reliability, minimal detectable change, construct validity, factor analysis, responsiveness.

J Rehabil Med 2006; 38: 346–353

Correspondence address: Margreth Grotle, PT, PhD, National Resource Center for Rehabilitation in Rheumatology, Diakonhjemmet Hospital, Oslo, Norway. E-mail: margreth.grotle@medisin.uio.no

Submitted November 1, 2005; accepted March 22, 2006

INTRODUCTION

The role of pain-related fear and avoidance behaviour in the development of chronic low back pain (LBP) and disability has received much attention. Most studies, which are cross-sectional and involve patients with chronic LBP, show moderate associations between fear-avoidance beliefs and self-reported disability (1–4). Some studies show similar concurrent associations in samples with acute LBP (5–7) and in the general population (8). In addition, prospective studies indicate that fear-avoidance beliefs may contribute at an early stage to the development of chronic LBP and associated disability (9–13).

Fear-avoidance beliefs are usually assessed using questionnaires, in particular the Fear Avoidance Beliefs Questionnaire (FABQ) (14). So far, only a few reports on the translation procedure and methodological properties of the FABQ have been published (2, 4, 15). As far as we know, the FABQ has been translated to Arabic (2), German (4) and French (15). There exists no such report for a Norwegian version. Most of the previous methodological studies show that the FABQ has good test-retest reliability when used for patients with chronic LBP (7). One study has shown that the FABQ are also reliable in a population with acute LBP (7). Since there is an increasing use of the FABQ as an outcome measure in intervention studies (17–20), it is important to explore the minimal detectable change (MDC) and the responsiveness of the FABQ. To our knowledge, values for the MDC and responsiveness for the FABQ are reported only a few times. One previous study among patients with chronic LBP found that the responsiveness of the FABQ was low (4). Hence, there is a need for further studies on the methodological properties of the FABQ, in particular when used among patients with acute LBP. Since the responsiveness of a measure is highly dependent on floor and ceiling effects, there is a need for investigating these properties for the FABQ.

The aims of this study were therefore to explore the methodological properties of the Norwegian FABQ version in terms of factor structure, internal consistency, test-retest reliability, construct validity, and responsiveness when used in two different subgroups of the LBP population. In addition, floor and ceiling effects were investigated.

MATERIAL AND METHODS

Fear-Avoidance Beliefs Questionnaire

The FABQ consists of 16 items, which is divided into 2 subscales: fear-avoidance beliefs for work (FABQ–Work) with 11 items and fear-avoidance beliefs for physical activity (FABQ–PA) with 5 items. The

 ISSN 1650-1977
DOI: 10.1080/16501970600722403

items are scored on a 7-point Likert scale (strongly disagree to strongly agree). The score of each subscale is used independently: In the FABQ–Work (range 0–42) 7 of the 11 items are added to a sum score (#6, 7, 9–12, and 15), and in the FABQ–PA (range 0–24) 4 of the 5 items (#2–5). The 5 remaining questions are used as delusive items as proposed by Waddell et al. (14).

Translation and cross-cultural adaptation

The FABQ was translated with a forward and backward translation procedure according to recommended guidelines (21). The English version of the FABQ was translated into Norwegian by 2 different and independent bilingual translators, whose first language was Norwegian. The backward translation was carried out by 2 other bilingual translators, whose first language was English. The translations and back-translations were discussed by a review committee (the translators and the research group). Discrepancies between the various versions were resolved by consensus to achieve conceptual equivalence between the pre-final Norwegian version and the original English version of the FABQ. The pre-final FABQ-version was pre-tested in 10 patients with acute LBP and 10 patients with chronic LBP. The patients were asked to report any problems when completing the questionnaire. None of the patients had comments that made changes necessary. The final version of the Norwegian FABQ was further tested in 2 Norwegian samples. Modified instructions were given when filling in the work subscale of the FABQ, for example, housewives were given instructions to regard their household duties as their work.

Patients

Patients aged between 18 and 60 years with acute or chronic LBP were included. A total of 123 patients who, for the first time due to acute LBP of less than 3 weeks' duration, sought help from a medical doctor or chiropractor in primary health were recruited to a prospective cohort study. The patients with chronic LBP ($n = 50$) had back pain of at least 3 months' duration and had been referred for specialist examination at the Back Clinic at Østfold Hospital. In both samples, pregnant patients and patients with "red flags" (e.g. cauda equina syndrome, progressive paresis, suspected tumour or local infection, ankylosing spondylitis, rheumatoid arthritis or other inflammatory diseases, fracture, or other symptoms that required urgent attention and further referral to specialists) were excluded. All patients gave informed consent after receiving both written and oral information about the project. The Ethics Committee for Medical Research in Health Region I of Norway approved the study.

Measurements

The FABQ was administered to all patients as part of a comprehensive questionnaire used in the cohorts. The comprehensive questionnaire collected sociodemographic data and medical history and included different self-report measures of pain, psychological distress and functional status. The sociodemographic variables were age, gender, education (<12 or ≥ 12 years), smoking status (yes/no) and work status (ordinary working situation, sick leave, or receiving rehabilitation or disability pension).

Pain variables consisted of pain localization (localized to the lower back with or without radiation to the one of the lower extremities), duration of present episode (days), use of pain medication during the last month (yes/no), and pain intensity rated on a numerical pain rating scale where 0 = no pain and 10 = pain as bad as it could be (22). Disability in daily activities was measured using the Oswestry Disability Index, version 2.0 (23). Disability days concerned how many days the patients had been restricted from participating in their daily work (employed work, homework, or school) due to LBP during the last 4 weeks (24).

Psychological distress was measured using the Hopkin's Symptom Check List (25) and psychosomatic awareness by the Modified Somatic Perception Questionnaire (26).

A clinical examination was carried out by a physical therapist, including a neurological examination (straight leg raising test, ankle and patellar reflexes, sensory loss, weakness in muscles of foot or thigh), lumbar spinal mobility (finger–floor distance in forward and side bending), and non-organic signs according to Waddell et al. (27). Three physical performance tests were included: the Shuttle Walk Test (28), sock test (29) and pick-up test (30).

Treatment and follow-up procedure

After the clinical examination, the examining physical therapist provided the patients with acute LBP with information according to clinical guidelines for acute LBP (31). The patients with chronic LBP were treated according to the model of Indahl et al. (32), which includes a thorough clinical examination by a medical doctor and physical therapist, a brief education program, and recommendations of staying physically active. Outcomes at 3-month follow-up were assessed using a postal questionnaire.

Test-retest reliability

For test-retest analyses, the FABQ was administered to a sub-sample of patients with chronic LBP. It was assumed that the back condition of patients with chronic LBP was more stable than that of patients with acute LBP. Thirty consecutive patients with chronic LBP were asked to complete the questionnaire after 2 days and return it by post.

Statistical analysis

The Statistical Package for Social Science (SPSS), version 10.0 (SPSS Inc., Chicago, IL, USA), was used to analyse the data. Kolmogorov-Smirnov tests for normality and distribution plots suggested that the FABQ-scores were normally distributed. Parametric statistics were therefore used in the analyses. Mean and standard deviation (SD) or frequencies were calculated for numerical and categorical variables, respectively. Differences between mean scores were compared by paired t-tests (numerical data) and χ^2 tests (categorical data).

The factor structure in the FABQ was evaluated using principal components analysis and was performed separately in the 2 samples. The retained components (or factors) in each scale had eigenvalue >1, and independent factors were obtained by using the Varimax rotation method.

Test-retest reliability was assessed by calculating the intra-class correlation coefficient from a one-way random effects model (1, 1) (33). An intra-class correlation coefficient of at least 0.75 was considered high, between 0.75 and 0.40 was considered moderate, and less than 0.40 was considered low (35). As recommended by Bland & Altman (34), test-retest reliability was also presented in terms of repeatability and by plots of the difference between the first and second scoring against a mean of the sum scores. Repeatability, which is similar to the MDC, was based on the standard error of measurement (SEM) between test and retest scores, and was calculated by multiplying the SEM by 2.77 to correspond with the 95% confidence intervals (CI). The difference between 2 measurements for the same subject was expected to be less than $2.77 \times \text{SEM}$ for 95% of pairs of observations, and defined the smallest difference that could be detected between the 2 measurements. To provide a unitless measurement error, the average coefficient of variance for paired measurements was calculated as the ratio of the SD divided by the mean and multiplied by 100.

The internal consistency for the questionnaires was assessed using Cronbach's α. A Cronbach's α of at least 0.80 was considered good, between 0.80 and 0.70 was considered moderate, and less than 0.70 was considered low (35).

Construct validity was assessed using the Pearson correlation coefficient in numerical scales and Spearman rank correlation coefficient in categorical scales. The analyses were carried out separately for acute and chronic LBP. The relationships were interpreted as being highly correlated when r was at least 0.60, moderately correlated when r was between 0.30 and 0.60, and weakly correlated when r was 0.30 or less (34).

Responsiveness (sensitivity to change) was assessed by standardized response mean (SRM) (Cohen's effect size) and were calculated by dividing the mean change by the SD of the mean change scores (36). The change scores were calculated by subtracting the follow-up scores from the baseline scores. Statistical significance of differences in the change scores was tested by independent samples t-test. An SRM of at least 0.80 was considered a good effect size, 0.40–0.80 was considered a moderate effect size, and less than 0.40 was considered a small effect size (36). Due to the low sample size in the chronic sample and the expectancies of

small changes in outcome measures during a 3-month follow-up, responsiveness was only calculated for the acute sample.

Potential floor and ceiling effects was explored by calculating the proportion of patients with insufficient initial score to reliably detect improvement or deterioration, respectively. Floor and ceiling effects were defined with the same criterion as in a previous study (37), which suggested a floor and ceiling effect when more than 15% of the respondents scored within the MDC limits (on both end of the scales) at baseline. The MDC values for the numeric pain rating scale and the Oswestry Disability Index were taken from previous studies using the same material as in the present study (38, 39).

RESULTS

The patients in the 2 groups represented different subgroups of the LBP population as indicated by the significant differences in many of the baseline variables in Table I. Overall, patients were able to complete the FABQ questionnaire without help. There were few missing values. Only 8 patients (5%) had some missing values in the physical activity scale and 11 (7%) on the work scale. For 15 patients (9%), the work scale was not relevant. When examining the distribution of each item, only item 8 (“I have a claim for compensation for my pain”) demonstrated a clearly skewed distribution with approximately 95% of the patients scoring 0 (“completely disagree”). Furthermore, item 8 showed no correlation with any of the other items and was excluded from further analyses.

The scored items were subjected to factorial analysis. In the acute sample the principal components analysis revealed that 2 components clearly exceeded eigenvalues of 1 (4.8 and 1.8), whereas a third component had an eigenvalue of 1.006. The 3 components accounted for 69% of the variance, whereas the 2 first components explained 60%. The 2 first components were retained in the Varimax rotation and the results are presented in Table II. The rotated solution gave a simple structure, with both components showing a number of strong loadings and all variables except one (#15) loaded substantially on only one component. The first component represented work-related beliefs and the second represented beliefs regarding physical activities. The 2-factor solution explained a total of 60% of the variance, with the work-related component contributing 36% and the physical activity-component contributing 24%.

In the chronic sample also 2 components with eigenvalues exceeding 1 (4.4 and 1.5) were extracted, which represented 54% of the total variance. The rotated solution showed that the strongest component in the chronic sample was component 1, which represented beliefs mainly related to physical activity, but also beliefs relating to work (Table II). As Table II shows, 4 of the variables (#9, 11, 12, 15) loaded on both components, but with highest values on component 2. Component 2 represented only beliefs related to the work and contributed with 19% of the variance. Component 1 explained 35% of the variance.

Twenty-eight of the 30 patients returned the retest questionnaire. No significant differences in demographic data were observed between these patients and the total chronic sample. There were no statistically significant differences between the first and second completions of the questionnaires. Test-retest reliability was high for the FABQ–Work and moderate for the FABQ–PA (Table III).

The FABQ scales and baseline variables were in general weakly to moderately correlated (Table IV). The strongest correlations were found in the chronic sample, in particular between the 2 FABQ scales and between the FABQ–Work and smoking, psychological distress and disability days, respectively. In general, both FABQ scales correlated weakly with the clinical variables and the physical performance tests.

At baseline, the mean FABQ–Work was significantly higher in the chronic sample compared with the acute sample, whereas there was no statistically significant difference in FABQ–PA (Table I). The mean change scores of both FABQ scales were significantly larger in the acute compared with the chronic sample, however (Table V). In the acute sample a moderate level of responsiveness of 0.56 (SRM) was found for the FABQ–PA, whereas the responsiveness was only 0.32 for the FABQ–Work. Furthermore, high responsiveness was observed for both pain and disability with SRM values of 1.49 and 0.99, respectively.

By comparing the initial score distributions of the FABQ and the MDC levels, the proportion of patients with insufficient initial scores to reliably detect improvement (floor effect)/deterioration (ceiling effect) were estimated. The results presented in Table VI show that in both back pain groups and on both FABQ-subscales more than 15% of the patients had initial scores that hampered the ability to detect improvement and/or deterioration on these scales, in particular for the FABQ–Work. When excluding the patients with initially no fear-avoidance beliefs (baseline scores below the MDC at the lower end of the scale), the mean change scores increased to 5.8 (SD 7.1) for the FABQ–PA and to 6.5 (SD 9.6) for the FABQ–Work, hence, increasing the responsiveness to 0.82 and 0.67, respectively. A similar trend was found for the FABQ–PA in the chronic sample as the mean change score increased to 2.4 (SD 4.8). For the FABQ–Work, however, the mean change score remained unchanged with a mean of −1.2 (SD 7.8). The numeric pain rating scale showed no floor or ceiling effects in the acute sample, but exceeded the 15% criterion for detection of deterioration in the chronic sample. The Oswestry Disability Index seemed to be able to detect improvement as well as deterioration in both samples.

DISCUSSION

This is the first report of a cross-cultural translation and adaptation of the Norwegian FABQ version. The Norwegian FABQ was clearly understood by the patients and few data were missing. Item 8 was excluded due to its skewed distribution in which 95% of patients scored 0. A skewed distribution of item 8 has also been reported in other studies (2, 14). The principal components analysis confirmed that the FABQ should be scored in 2 separate scales. In the acute sample all items in the 2 components had salient loadings on their respective factor with no considerable loading to any of the other factors in the first solution. This factorial structure is similar to the findings

Table I. *Baseline characteristics of patients with acute and chronic low back pain (LBP)*

	Acute LBP ($n = 123$)	Chronic LBP ($n = 50$)	p-value (difference)
Sociodemographic variables			
Age (years)	37.8 (10.4)	40.4 (9.5)	0.04
Gender (men)	55 (45%)	19 (38%)	0.34
Education (< 12 years)	56 (46%)	38 (76%)	0.001
Smoking (yes)	47 (38%)	25 (50%)	0.09
Work status			
Ordinary working situation	108 (88%)	13 (26%)	<0.001
Sick leave	8 (7%)*	34 (68%)	
Rehabilitation or disability pension	7 (6%)*	3 (6%)	
Pain and disability variables			
Pain localization (with radiating pain)	52 (42%)	43 (86%)	<0.001
Duration of back pain episode (days)	8.1 (6.6)	579.9 (785.1)	<0.001
Pain medication last month (yes)	85 (70%)	40 (80%)	0.16
Pain intensity (0–10)	6.7 (1.8)	6.1 (2.4)	0.06
Oswestry Disability Index (0–100)	28.0 (5.1)	31.8 (11.2)	0.11
Disability days last month	2.5 (5.2)	19.2 (12.4)	<0.001
Psychological variables			
FABQ–Work	13.2 (11.0)	24.6 (11.7)	<0.001
FABQ–Physical Activity	12.3 (6.1)	12.9 (6.4)	0.57
Psychosomatic awareness (MSPQ)	5.1 (4.8)	11.0 (6.9)	<0.001
Distress (HSCL25)	1.3 (0.4)	1.7 (0.4)	<0.001
Clinical tests			
Straight leg raising (degrees)	83.1 (12.6)	73.9 (11.5)	<0.001
Forward lumbar bending (fingertip–floor distance, cm)	12.0 (20.3)	12.2 (16.3)	0.93
Neurological signs ($\geq$ 2)	30 (24%)	29 (58%)	<0.001
Non-organic signs ($\geq$ 3)	2 (2%)	7 (14%)	0.001
Pick-up test (with problems)	46 (49%)‡	30 (61%)	0.16
Sock-test (with problems)	49 (52%)‡	27 (55%)	0.74
Walking test (seconds)	11.8 (2.1)	13.4 (2.8)	<0.001

Continuous variables are presented as means with standard deviation within parentheses and categorical variables as frequencies with percentages in parentheses.
*Patients who received economical compensation due to other reasons than back problems.
‡$n = 94$.
FABQ = Fear Avoidance Beliefs Questionnaire; MSPQ = Modified Somatic Perception Questionnaire; HSCL25 = Hopkin's Symptom Check List, 25-items.

reported in the original study of Waddell et al. (14), as well as in the German (2) and French (4) versions. For example, Waddell et al. report that FABQ–Work explains 42% and FABQ–PA explains 18%, a total of 60% (14). In the present chronic sample, however, the component relating mainly to physical activities, explained more of the variance than the second component representing work-related beliefs. This finding may be explained by the fact that most of the patients with chronic LBP were out of work (sick leave, rehabilitation or disability pension), and probably the items related to work were not relevant for them.

Table II. *Results of the principal component analysis in acute and chronic low back pain (LBP)*

Acute LBP ($n = 123$)			Chronic LBP ($n = 50$)		
Item	Component 1	Component 2	Item	Component 1	Component 2
7	0.861		3	0.841	
10	0.835		5	0.827	
9	0.819		6	0.752	
11	0.771		4	0.721	
6	0.754		2	0.507	
12	0.658		12	0.552	0.656
3		0.808	9	0.521	0.622
5		0.720	11	0.503	0.602
2		0.700	15	0.440	0.581
4		0.673	10		−0.580
15	0.406	0.497	7		0.452
% of variance explained Component 1 36.0% and Component 2 23.7% (= 59.7%)			% of variance explained Component 1 34.6% and Component 2 19.4% (= 54.0%)		

Table III. *Test-retest reliability and internal consistency of the Fear Avoidance Beliefs Questionnaire (FABQ)*

	FABQ–Work ($n=28$)	FABQ–Physical Activity ($n=28$)
Test, mean (SD)	26.4 (10.4)	13.8 (5.2)
Retest, mean (SD)	27.6 (10.5)	14.8 (5.8)
Mean difference (SD)	−1.2 (6.2)	−1 (4.6)
Test, range	6–41	5–24
Retest, range	6–42	4–24
ICC 1,1 (95% CI)	0.82 (0.64;0.92)	0.66 (0.38;0.83)
Minimal detectable change	12.1	8.95
Coefficient of variance (%)	16.2	22.6
Chronbach's alpha test/retest	0.90	0.79

Time interval between test and retest was 2–4 days. Test-retest reliability is expressed by single measure intraclass correlation coefficient (ICC 1,1) with 95% confidence intervals (CIs), minimal detectable change, and coefficient of variance.
SD = standard deviation.

Second, the sample size of the chronic group was low for a factor analysis, and therefore should be replicated in another study.

A high internal consistency was found for FABQ–Work, which is consistent with other reports (2, 14, 40). The moderate level of internal consistency for FABQ–PA might partly be explained by the low number of items. The present internal consistency of 0.79 is higher than in most previous publications. The range in those is between 0.57 and 0.77 (2, 5, 14, 40). The moderate correlation between FABQ–Work and FABQ–PA also suggests that the 2 scales assess different constructs. Construct validity is usually assessed by comparing the questionnaire with other measures assessing the same construct (convergent validity) and in addition by comparing with measures expected to assess other, although related, constructs (divergent validity). In the present study, it was not possible to compare the FABQ with another fear-avoidance beliefs questionnaire, because no such measurement validated in Norwegian was available. However, the FABQ was compared with several pain, distress, and disability variables, in addition to physical performance tests. Overall, these results suggest that the construct assessed by the FABQ differs from these variables, and are in line with most other reports regarding the construct validity of the FABQ (1, 2, 4, 14, 40).

Test-retest reliability consistently showed better results for FABQ–Work than for FABQ–PA. These results are in line with other studies (2, 4, 14). MDC, which provides an error estimate given in the scale's unit, was approximately 12 of a possible 42 points in FABQ–Work and 9 of a possible 24 points in FABQ–PA. Another Norwegian study reported 16 points for FABQ–Work and 8 points for FABQ–PA (16). For the French version, somewhat lower values have been reported, 10 points for FABQ–Work and 7 points for FABQ–PA (4). Since a change less than the MDC is indistinguishable from the point estimate of the measurement error, this limit provides an important estimate when using the FABQ as an outcome measure in repeated measurements. Similarly, the coefficient of variation, which reflects the relative measurement error as a percentage score, are important when comparing the error limits of several measures. For example, this study shows that 16% and 23% of changes in the FABQ–Work and FABQ–PA, respectively, should be attributed to measurement error. Information regarding the measurement error of an outcome measure is necessary when interpreting the amount of change achieved and the responsiveness of the measure.

The present results showed low responsiveness for FABQ–Work and moderate for FABQ–PA in the acute sample. The poor results on the responsiveness of the FABQ–Work are similar to the findings of Chaory et al. (4), who report SRMs at

Table IV. *Correlation matrix of fear-avoidance beliefs and baseline variables in patients with acute and chronic low back pain*

	FABQ–Work Acute	FABQ–Work Chronic	FABQ–Physical Activity Acute	FABQ–Physical Activity Chronic
FABQ–Physical Activity	0.36**	0.56**	–	–
Age	−0.16	−0.36*	−0.31**	−0.12
Gender	−0.18	−0.15	−0.08	−0.16
Education	−0.18	−0.15	0.05	−0.02
Smoking	0.06	0.52**	0.03	0.45**
Pain localization	<0.001	0.25	−0.08	**0.15**
Pain medication	−0.08	0.16	−0.06	**0.14**
Pain intensity	0.20*	0.28	0.16	**0.23**
MSPQ	0.31**	0.59**	0.11	**0.32***
HSCL25	0.30**	0.43**	0.14	**0.36***
Oswestry Disability Index	0.08	0.34*	0.34**	**0.39****
Disability days	0.37**	0.52**	0.08	**0.33***
Straight leg raising	0.02	−0.31*	−0.22*	**−0.36***
Forward lumbar bending	−0.03	0.14	0.08	**0.28**
Neurological signs	0.11	−0.01	−0.001	**0.12**
Non-organic signs	0.03	0.32*	−0.12	**0.27**
Pick-up test	0.001	−0.11	0.13	**0.11**
Sock-test	−0.16	0.03	0.03	**0.16**
Walking test	−0.04	0.22	0.05	**0.31***

Pearson correlation coefficients for continuous variables, Spearman for categorical variables. Correlations >0.50 in bold print. FABQ = Fear Avoidance Beliefs Questionnaire; MSPQ = Modified Somatic Perception Questionnaire; HSCL25 = Hopkin's Symptom Check List, 25-items.
*p <0.05; **p <0.01.

Table V. *Absolute scores at baseline and 3-months follow-up and mean change scores of the Fear Avoidance Beliefs Questionnaire (FABQ), pain and disability*

	Baseline, mean (SD)	3 months, mean (SD)	Change scores, mean (SD)	*p*
FABQ–Work (0–42)				
Acute LBP	13.2 (11.0)	9.8 (11.5)	2.8 (8.7)	0.008
Chronic LBP	24.6 (11.7)	25.3 (12.3)	−1.4 (8.0)	
FABQ–Physical Activity (0–24)				
Acute LBP	12.3 (6.1)	8.2 (6.4)	4.0 (7.2)	0.002
Chronic LBP	12.9 (6.4)	12.7 (6.1)	0.3 (5.8)	
Pain intensity (0–10)				
Acute LBP	6.7 (1.8)	2.6 (2.6)	4.1 (2.7)	<0.001
Chronic LBP	6.1 (2.4)	5.6 (2.4)	0.7 (2.2)	
Oswestry Disability Index (0–100)				
Acute LBP	28.0 (5.1)	10.9 (12.5)	17.3 (17.5)	<0.001
Chronic LBP	31.8 (11.2)	27.8 (13.5)	3.8 (8.9)	

Positive change values indicate improvement, negative values indicate deterioration. The *p*-values refer to statistical differences in the change scores between the 2 back groups (independent *t*-test).
SD: standard deviation, LBP: low back pain.

0.30 for FABQ–Work and 0.34 for FABQ–PA. One possible explanation for the low SRMs on the FABQ–Work in the acute sample may probably be the rather low level of fear-avoidance beliefs observed at baseline, which was lower than in other studies including patients with acute LBP (6, 13). The large proportion of patients falling within the MDC at the lower end of the scale at baseline (floor effect), explains the low ability to detect changes in the FABQ-scales in this sample. The responsiveness of both FABQ scales was considerable improved when analysing the sample without the patients with no fear-avoidance beliefs at baseline. Hence, in patient samples with initially high scores on the FABQ scales, the responsiveness will probably be higher than reported in the present study. The responsiveness of the FABQ should therefore be further studied in other patient samples.

As expected, in the chronic sample there were small changes in pain and disability, as well as fear-avoidance beliefs, during the relatively short follow-up period of 3 months. Although the patients reported high levels of pain and disability initially, the level of fear-avoidance beliefs was only moderate, in particular for the FABQ–PA. The results showing an obvious floor effect of the FABQ–PA scale are important when using this scale as an outcome measure. However, the results regarding floor and ceiling effects in the chronic sample should be interpreted carefully. Firstly, the MDCs used in the present study were based on estimates for the whole scale width. They may vary in the upper and lower end of the scores, and should therefore be further explored in other studies. Secondly, our criterion for floor/ceiling effects were more strict than in most earlier studies as we used 95% limits for the calculation of the MDCs in contrast to 90% limits used by both Davidson & Keating (37) and Stratford et al. (41) (90% limits means lower MDC estimates). Hence, distribution scores and responsiveness of the FABQ in chronic LBP should be further explored.

An obvious limitation of this study is that the main outcome measures were based on self-report. Some clinical tests were therefore included. However, they were only weakly correlated with fear-avoidance beliefs in the present study. More studies

Table VI. *Patients with insufficient initial score of Fear Avoidance Beliefs Questionnaire (FABQ), pain and disability to reliably detect change (%)*

	Acute LBP ($n=123$)		Chronic LBP ($n=50$)	
	Patients with insufficient initial score to reliably detect improvement (scores < MDC)	Patients with insufficient initial score to reliably detect deterioration (scores > [max score − MDC])	Patients with insufficient initial score to reliably detect improvement (scores < MDC)	Patients with insufficient initial score to reliably detect deterioration (scores > [max score − MDC])
FABQ–Work (MDC = 12)	58 (48)	10 (8)	9 (18)	21 (42)
FABQ–Physical Activity (MDC = 9)	31 (26)	39 (33)	13 (26)	18 (36)
Pain intensity (MDC = 2)	0	18 (15)	4 (8)	9 (18)
Oswestry Disability Index (MDC = 11)	13 (11)	0	2 (4)	0

MDC = minimal detectable change. The MDC values for the numeric pain scale and Oswestry Disability Index come from a previous study on methodological properties in the same samples as used here.

are clearly needed on the relationship between self-reported fear-avoidance beliefs, as measured by a questionnaire, and physical tests for relevant activities. One of the main strengths of the present study is the comparison of fear-avoidance beliefs in the 2 different subgroups of the back population.

This study showed that the Norwegian FABQ version had acceptable factor structure, internal consistency, test-retest reliability, and construct validity. The responsiveness of the FABQ–Work was low and for the FABQ–PA moderate, in the acute sample. These results suggest that the FABQ can be recommended for assessing self-reported fear-avoidance beliefs in patients with acute and chronic LBP. The use of the FABQ as an outcome to measure change should be interpreted carefully in relation to expected error variance.

ACKNOWLEDGEMENTS

Support for this study was provided by the Norwegian Fund for Postgraduate Training in Physiotherapy and the Norwegian Back Pain Network.

REFERENCES

1. Vlaeyen JWS, Kole-Snijders AMJ, Boeren RGB, van Eek H. Fear of movement/(re)injury in chronic low back pain and its relation to behavioral performance. Pain 1995; 62: 363–372.
2. Pfingsten M, Kroner-Herwig B, Leibing E, Kronshage U, Hildebrandt J. Validation of the German version of the Fear-Avoidance Beliefs Questionnaire (FABQ). Eur J Pain 2000; 4: 259–266.
3. Severeijns R, Vlaeyen JW, van den Hout MA, Weber WE. Pain catastrophizing predicts pain intensity, disability, and psychological distress independent of the level of physical impairment. Clin J Pain 2001; 17: 165–172.
4. Chaory K, Fayad F, Rannou F, Lefevre-Colau MM, Fermanian J, Revel M, et al. Validation of the French Version of the Fear Avoidance Belief Questionnaire. Spine 2004; 29: 908–913.
5. Symonds TL, Burton AK, Tillotson KM, Main CJ. Do attitudes and beliefs influence work loss due to low back trouble? Occup Med 1996; 46: 25–32.
6. George SZ, Fritz JM, Erhard RE. A comparison of fear-avoidance beliefs in patients with lumbar spine pain and cervical spine pain. Spine 2001; 26: 2139–2145.
7. Swinkels-Meewisse IE, Roelofs J, Verbeek AL, Oostendorp RA, Vlaeyen JW. Fear of movement/(re)injury, disability and participation in acute low back pain. Pain 2003; 105: 371–379.
8. Buer N, Linton SJ. Fear-avoidance beliefs and catastrophizing: occurrence and risk factor in back pain and ADL in the general population. Pain 2002; 99: 485–491.
9. Burton AK, Tillotson KM, Main CJ, Hollis S. Psychosocial predictors of outcome in acute and subchronic low back trouble. Spine 1995; 20: 722–728.
10. Linton SJ, Buer N, Vlaeyen JW, Hellsing AL. Are fear-avoidance beliefs related to the inception of an episode of back pain? A prospective study. Psychol Health 2000; 14: 1051–1059.
11. Picavet HSJ, Vlaeyen JWS, Schouten JSAG. Pain catastrophizing and kinesiophobia: Predictors of chronic low back pain. Am J Epidemiol 2002; 156: 1028–1034.
12. Klenerman L, Slade PD, Stanley IM, Pennie B, Reilly JP, Atchison LE, et al. The prediction of chronicity in patients with an acute attack of low back pain in a general practice setting. Spine 1995; 20: 478–484.
13. Fritz JM, George SZ, Delitto A. The role of fear-avoidance beliefs in acute low back pain: relationships with current and future disability and work status. Pain 2001; 94: 7–15.
14. Waddel 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: 157–168.
15. Al Obaidi SM, Nelson RM, Al Awadhi S, Shuwaie AI. The role of anticipation and fear of pain in the persistence of avoidance behavior in patients with chronic low back pain. Spine 2000; 25: 1126–1131.
16. Holm I, Friis A, Storheim K, Brox JI. Measuring self-reported functional status and pain in patients with chronic low back pain by postal questionnaires. A reliability study. Spine 2003; 28: 828–833.
17. George SZ, Fritz JM, Bialosky JE, Donald DA. The effect of a fear-avoidance-based physical therapy intervention for patients with acute low back pain: results of a randomized clinical trial. Spine 2003; 28: 2551–2560.
18. Storheim K, Brox JI, Holm I, Koller AK, Bø K. Intensive group training versus cognitive intervention in sub-acute low back pain: short-term results of a single-blind randomized controlled trial. J Rehabil Med 2003; 35: 132–140.
19. Brox JI, Sorensen R, Friis A, Nygaard O, Indahl A, Ingebrigtsen T, et al. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. Spine 2003; 28: 1913–1921.
20. Walsh DA, Radcliffe JC. Pain beliefs and perceived physical disability of patients with chronic low back pain. Pain 2002; 97: 23–31.
21. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality-of-life measures: literature review and proposed guidelines. J Clin Epidemiol 1993; 46: 1417–1432.
22. Jensen MP, Karoly P. Self-report scales and procedures for assessing pain in adults. In: Turk DC, Melzack R, eds. Handbook of Pain Assessment, 1st edn. New York: The Guilford Press; 1992, p. 135–151.
23. Baker DJ, Pynsent PB, Fairbank J. The Oswestry disability index revisited: its reliability, repeatability and validity, and a comparison with the St Thomas's disability index. In: Roland MO, Jenner JR, eds. Back pain. New approaches to rehabilitation and education. Manchester: Manchester University Press; 1990, p. 175–181.
24. Deyo RA, Battie M, Beurskens AJ, Bombardier C, Croft B, Koes A, et al. Outcome measures for low back pain research. A proposal for standardized use. Spine 1998; 23: 2003–2013.
25. Derogatis LR, Lipman RS, Rickels K, Uhlenhut EH, Cori L. The Hopkins Symptom Checklist (HSCL): a self-report inventory. Behav Science 1974; 1: 1–15.
26. Main CJ. The Modified Somatic Perception Questionnaire (MSPQ). J Psychosom Res 1983; 27: 503–514.
27. Waddel G, McCullogh JA, Kummel E, Venner RM. Inappropriate symptoms and Non-organic signs tests. Spine 1980; 5: 117–125.
28. Ljungquist T, Fransson B, Harmsringdahl K, Björnham Å, Nygren Å. A physiotherapy test package for assessing back and neck dysfunction – discriminative ability for patients versus healthy control subjects. Physiother Res Int 1999; 4: 123–140.
29. Strand LI, Wie SL. The sock test for evaluating activity limitation in patients with musculoskeletal pain. Phys Ther 1999; 79: 136–145.
30. Strand LI. The pick up test for evaluating activity limitation in patients with musculoskeletal back pain. 3rd edn. 2001, p. 17–27.
31. Nasjonalt ryggnettverk – Formidlingsenheten. Akutte korsryggsmerter. Tverrfaglige, kliniske retningslinjer. In: Nasjonalt ryggnettverk, ed. Oslo: 2002, p. 1–48.
32. Indahl A, Velund L, Reikeraas O. Good Prognosis for low back pain when left untampered. A randomized clinical trial. Spine 1995; 20: 473–477.
33. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 2002; 86: 420–428.
34. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999; 8: 135–160.
35. Andresen EM. Criteria for assessing the tools of disability outcomes research. Arch Phys Med Rehabil 2000; 81: s15–s20.
36. Cohen J. Statistical power analysis for the behavioral sciences. Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1988.
37. Davidson M, Keating JL. A comparison of five low back disability questionnaires: reliability and responsiveness. Phys Ther 2002; 82: 8–24.

38. Grotle M, Brox JI, Vøllestad NK. Concurrent comparison of responsiveness in pain and functional status measurements used for patients with low back pain. Spine 2004; 29: E492–E501.
39. Grotle M, Brox JI, Vollestad NK. Cross-cultural adaptation of the Norwegian versions of the Roland-Morris Disability Questionnaire and the Oswestry Disability Index. J Rehabil Med 2003; 35: 241–247.
40. Crombez G, Vlaeyen JWS, Heuts PHTG, Lysens R. Pain-related fear is more disabling than pain itself: evidence on the role of pain-related fear in chronic back pain disability. Pain 1999; 80: 329–339.
41. Stratford PW, Binkley J, Solomon P, Finch E, Gill C, Moreland J. Defining the minimum level of detectable change for the Roland-Morris Questionnaire. Phys Ther 1996; 76: 359–368.