Eur Spine J (2010) 19:633–640
DOI 10.1007/s00586-009-1180-9

ORIGINAL ARTICLE

Self-reported assessment of disability and performance-based assessment of disability are influenced by different patient characteristics in acute low back pain

Benedict Martin Wand · Lara A. Chiffelle · Neil Edward O'Connell · James Henry McAuley · Lorraine Hilary DeSouza

Received: 9 June 2009 / Revised: 27 July 2009 / Accepted: 15 September 2009 / Published online: 23 October 2009

Abstract For an individual, the functional consequences of an episode of low back pain is a key measure of their clinical status. Self-reported disability measures are commonly used to capture this component of the back pain experience. In non-acute low back pain there is some uncertainty of the validity of this approach. It appears that self-reported assessment of disability and direct measurements of functional status are only moderately related. In this cross-sectional study, we investigated this relationship in a sample of 94 acute low back pain patients. Both self-reported disability and a performance-based assessment of disability were assessed, along with extensive profiling of patient characteristics. Scale consistency of the performance-based assessment was investigated using Cronbach's alpha, the relationship between self-reported and performance-based assessment of disability was investigated using Pearson's correlation. The relationship between clinical profile and each of the disability measures were examined using Pearson's correlations and multivariate linear regression. Our results demonstrate that the battery of tests used are internally reliable (Cronbach's alpha = 0.86). We found only moderate correlations between the two disability measures ($r = 0.471$, $p < 0.001$). Self-reported disability was significantly correlated with symptom distribution, medication use, physical well-being, pain intensity, depression, somatic distress and anxiety. The only significant correlations with the performance-based measure were symptom distribution, physical well-being and pain intensity. In the multivariate analyses no psychological measure made a significant unique contribution to the prediction of the performance-based measure, whereas depression made a unique contribution to the prediction of the self-reported measure. Our results suggest that self-reported and performance-based assessments of disability are influenced by different patient characteristics. In particular, it appears self-reported measures of disability are more influenced by the patient's psychological status than performance-based measures of disability.

Keywords Low back pain · Disability · Functional performance · Validity · Outcome

B. M. Wand (✉)
School of Health Sciences, The University of Notre Dame Australia, 19 Mouat St, Fremantle, WA 6959, Australia
e-mail: bwand@nd.edu.au

L. A. Chiffelle
Progress Physiotherapy Services, 22 Harvey St, Collie, WA 6225, Australia

N. E. O'Connell · L. H. DeSouza
Department of Health Sciences and Social Care, Brunel University, Kingston Lane, Uxbridge, Middlesex UB8-3PH, UK

J. H. McAuley
Musculoskeletal Division, The George Institute for International Health, PO Box M201, Missenden Rd, Sydney, NSW 2050, Australia

Introduction

Investigations into the nature of the low back pain (LBP) experience and evaluation of treatment approaches for LBP require sound and meaningful ways of measuring patient status [1–4]. It is generally agreed that quantifying the functional consequences of LBP for the individual is of prime importance in describing clinical status [2]. Measurement at this level is dominated by various self-reported disability questionnaires [3, 5]. These tools are cheap, quick and easy to administer and demonstrate good levels of reliability and responsiveness [3, 5]. However, recent research has begun to raise some questions about the validity of self-reported disability measures for quantifying actual function. Studies on LBP patients have indicated that there is often discrepancy between self-reported disability and an individual's actual daily activity level [6, 7] and the level of agreement between performance-based measures of disability and self-reported measures of disability is not strong [6, 8–15], suggesting that these measures might describe different aspects of functional capacity.

Establishing how best to capture the functional consequences of LBP for an individual is not a trivial issue. Models that have been developed to explain the LBP experience (for e.g. [16–18]) have evolved principally through investigating which factors contribute to self-reported measures of disability, and the endorsement of any therapy is largely dependent on its ability to meaningfully change scores on self-reported disability questionnaires. If self-reported disability does not accurately reflect the functional consequences of LBP for an individual then our understanding of the LBP experience and how it should be managed might be flawed.

One approach to exploring the association between self-reported and actual functional status is via analysing the relationship between self-reported disability and direct measurements of functional performance. Whilst data in this area are accumulating in chronic LBP [8, 9, 11], it is unclear whether the discrepancy between self-reported disability and performance-based measures suggested in chronic LBP is found in the acute LBP (ALBP) population. We were interested in exploring this issue. In the current study we developed a battery of tests that included activities that we felt could be tolerated by patients with ALBP, that were likely to be reflective of functional problems in this group and that have demonstrated high levels of reliability in other clinical populations [8, 10, 19–25]. Subjects were timed on the performance of each of these tasks and the scores summed to produce a timed functional tasks total score (TFTTS).

The purpose of this study was to establish the feasibility of using these tests in the ALBP population, to determine whether these tests measure a single construct, to investigate the relationship between self-reported disability and performance-based evaluation of disability and determine which patient characteristics are associated with each disability measure.

Methods

Participants

This study was run alongside a randomised, controlled trial of physiotherapy care for ALBP [26]. Subjects were 94 acute non-specific LBP patients referred to the Physiotherapy Department of a suburban district hospital in the United Kingdom by either their general practitioner or the Hospital Accident and Emergency Department. To be eligible for inclusion, patients had to report an episode of non-specific LBP of less than 6 weeks duration, be aged between 20 and 55 years and provide written, informed consent. Those with recurrent pain needed to have been pain free for at least 3 months prior to the onset of the current episode.

Participants were excluded if they presented with nerve root pain, evidence of specific spinal pathology such as malignancy, fracture, infection, inflammatory joint or bone disease, were pregnant or less than 3 months postpartum, were involved in litigation related to their back problem, had a coexisting major medical disease, were already involved in active physical therapy for their problem, or had undergone previous spinal surgery. The study was approved by the Health Authority's Research Ethics Committee and no patient received financial remuneration for participating in the study.

Procedure

On initial presentation, all eligible subjects were seen by a research assistant, given information about the project and signed a consent form. Subjects had their height and weight measured and completed a series of questionnaires related to their demographic and clinical status. Subjects then undertook a series of functional tasks, the performance of which were timed by the research assistant. All assessments were undertaken on the same day.

Measures

Demographic questionnaire

Participants completed a questionnaire that solicited information regarding age, gender, and work status. The research assistant then established the duration of the problem, symptom distribution [27], medication use related to their LBP episode; whether this episode was the first attack of LBP or a recurrence and whether onset was sudden or gradual. For the correlation and regression analyses the categorical variables were dummy coded (see Table 1).

Pain

Pain intensity was measured using a 0–10 numerical rating scale. The anchors used were 0 = 'no pain' and 10 = 'pain as bad as it could be'. Subjects completed the same scale for, 'usual pain during the last week' and 'present pain'.

Table 1 Baseline demographic and clinical characteristics of the study sample ($n = 94$)

Variable	*N* (%) or mean (SD)
Gender (female), (female = 0, male = 1) *n* (%)	47 (50%)
Age (years), mean (SD)	34.77 (8.51)
BMI, mean (SD)	25.60 (4.43)
Work status	
Off work due to LBP, *n* (%)	38 (40.4)
Working, *n* (%)	46 (48.9)
Not applicable, *n* (%)	10 (10.6)
Past history (first episode = 0, recurrence = 1)	
First episode, *n* (%)	40 (42.6)
Recurrent episode, *n* (%)	54 (57.4)
Onset type (sudden = 0, gradual = 1)	
Sudden onset, *n* (%)	67 (71.3)
Gradual onset, *n* (%)	27 (28.7)
Duration of LBP (weeks), mean (SD)	2.95 (1.48)
Use of analgesics for LBP, (yes = 0, no = 1)	
Yes, *n* (%)	52 (55.3)
No, *n* (%)	42 (44.7)
Symptom distribution	
No symptoms, (1) *n* (%)	1 (1.1)
LBP without radiation, (2) *n* (%)	52 (55.3)
Proximal radiation, (3) *n* (%)	19 (20.2)
Distal radiation, (4) *n* (%)	22 (23.4)
Usual pain (0–10), mean (SD)	5.45 (2.27)
Present pain (0–10), mean (SD)	3.79 (2.44)
RMDQ (0–24), mean (SD)	11.32 (6.21)
MCS (0–50), mean (SD)	23.58 (4.26
PCS (0–50) mean (SD)	18.40 (3.54)
ZUNG (0–69) mean (SD)	21.91 (10.77)
MSPQ (0–39) mean (SD)	7.35 (5.18)
STAIS (6–24) mean (SD)	12.73 (3.95)

BMI Body mass index, *Usual pain* numerical rating scale for usual pain intensity, *Present pain* numerical rating scale for present pain intensity, *RMDQ* Roland and Morris disability questionnaire, *MCS* SF-36 mental component score, *PCS* SF-36 physical component score, *ZUNG* modified Zung self-reported depression inventory, *MSPQ* modified somatic perceptions questionnaire, *STAIS* Spielberger state-trait anxiety inventory score

Psychological characteristics

State anxiety was estimated using six items from the Spielberger State-trait Anxiety Inventory (STAIS) [28]. The presence of depressive symptoms was determined using the Modified Zung Self Rated Depression Inventory (ZUNG) [29] and somatic distress was estimated using the Modified Somatic Perception Questionnaire (MSPQ) [30].

Health-related quality of life

Health-related quality of life was measured using the Short Form-36 health survey (SF-36) [31]. The subscales were collapsed to calculate a physical well-being component score (PCS) and a mental well-being component score (MCS).

Self-reported disability

Self-reported LBP-related disability was measured using the Roland and Morris Disability Questionnaire (RMDQ) [32]. This is a LBP-specific scale in which subjects are presented with a list of 24 functional activities and asked to indicate if their LBP interferes with performance of each task at present.

Performance-based assessment of disability

Subjects were then taken to a room that had been specifically set up and marked for the performance of the timed functional tasks. The same research assistant was responsible for all assessments and used a stop watch with an accuracy of 1/100th of a second to record the time taken to complete each task. Subjects were asked to perform each activity at a comfortable speed, and were advised they could stop the test if they experienced too much pain or discomfort. Standardised instructions for the tasks were given to each person. Before beginning the timed test, subjects were shown a demonstration of the task to eliminate problems with the timing due to unfamiliarity with the test. Once all of the tasks were completed, each subject's times were added together to give a total score. The tasks were performed in the following order.

Timed sit to stand test (TSTS)

The time taken, in seconds, to stand up and sit down five times from a seated position was recorded. A standard 45-cm office chair with armrests was used for this test. The patient was instructed to perform the task at a comfortable speed with or without using the armrests and the examiner counted the repetitions aloud. Time was recorded from when the instruction to "go" was given and ceased when the subject sat down on the fifth occasion.

Timed up-and-go test (TUGT)

The time taken to stand up from a 45-cm chair with armrests, walk 3 m, turn, walk back to the chair and sit back down again was measured for each patient. A strip of red tape on the floor indicated the 3-m distance. The patient was instructed to perform the task at a comfortable speed and time was recorded from when the instruction to "go" was given and ceased when the subject sat down.

Timed 5-m walk test (5MWT)

The 5MWT measured the time taken in seconds by a patient to walk a distance of 5 m at a comfortable pace. Two strips of tape were placed on the floor 5 m apart. Patients started at a mark 2 m behind this point and were instructed to walk to a point 2 m beyond the 5-m line, thus allowing a 2-m warm-up and cool-down to minimise variations due to acceleration and deceleration. Time was recorded when the subject's foot crossed the first line on the floor and ceased when they crossed the second line.

Timed lying to stand test (TLTS)

Patients were positioned in supine lying with one pillow on an adjustable height examination bed positioned in the middle of the room. The bed was set at a height of 45 cm off the floor. Patients were asked to get up from lying on the bed into a standing position at a comfortable speed, using whatever means they liked and from whichever side they preferred. Time was recorded from the time when the examiner said "go" and stopped when the patient was standing upright.

Data analysis

The analyses were undertaken using SPSS for windows version 17 (SPSS, Chicago IL, USA). Baseline characteristics were summarised for descriptive purposes with mean and standard deviations used for continuous measures and percentages for categorical measures. To investigate if the timed tests chosen represent a single construct, Pearson's product-moment correlations were calculated to examine the relationship between each item and between each item and the total score, internal consistency was examined by use of Cronbach's alpha. Portney and Watkins [33] suggests a result between 0.70 and 0.90 is ideal, indicating a scale that is probably measuring the same trait, but without any redundant items.

The relationship between self-reported disability and the performance-based assessment of disability was explored using Pearson's correlation. Relationships were explored between the RMDQ and the TFTTS as well as between the RMDQ and each of the test items. The relationship between patient profile and RMDQ and the relationship between patient profile and the TFTTS were similarly investigated with Pearson's correlations. The significance level was set at $p < 0.01$ to account for multiple comparisons and the strengths of the univariate relationships were defined according to Cohen's rule of thumb [34]. To assess the combined and unique strength of the associations between patient profile and the two disability measures, multivariate linear regressions were performed with RMDQ and the TFTTS as the dependent variable and significantly correlated ($p < 0.1$) aspects of the patient profile as predictor variables.

Results

Sample characteristics

Table 1 provides a summary of the demographic characteristics and baseline clinical profile of the subjects participating in the study. Of the 94 ALBP patients that were tested, one subject was unable to complete the testing and for one subject only the total score was recorded.

Scale construction

The test scores obtained for each item and the total score are shown in Table 2. Bivariate correlations between individual items and between each item and the total score were all statistically significant (all $p > 0.001$). These data are summarised in Table 3. Cronbach's alpha for the four individual tasks was 0.86, this is above 0.7, indicating that the scale can be considered reliable with our sample, and below 0.9 suggesting that none of the items are redundant. When TUGT, TSTS, TLTS and 5MWT tasks were individually deleted, Cronbach's alpha measured 0.78, 0.90, 0.78, and 0.81, respectively.

Relationship between self-reported and performance-based measurement of disability

The RMDQ score and the TFTTS were significantly correlated ($r = 0.471$, $p < 0.001$), indicating a moderate relationship between these two measures of function. There were moderate correlations between the RMDQ and the

Table 2 Individual and total scores for timed functional tasks

	Timed up-and-go ($n = 92$)	Timed sit to stand ($n = 92$)	Timed supine to stand ($n = 92$)	Timed 5-m walk ($n = 92$)	Timed functional tasks total score ($n = 93$)
Mean, s (range)	10.95 (2.4–57.3)	19.79 (8.3–68.4)	6.20 (1.4–60.0)	5.57 (2.4–60.0)	43.43 (21.2–212.7)

Table 3 Results of Pearson's correlation tests between individual items and timed functional tasks total score ($n = 92$)

	Timed up-and-go	Timed sit to stand	Timed supine to stand	Timed 5-m walk	Timed functional tasks total score
Timed up-and-go		$r = 0.564$*	$r = 0.719$*	$r = 0.856$*	$r = 0.896$*
Timed sit to stand	$r = 0.564$*		$r = 0.575$*	$r = 0.384$*	$r = 0.784$*
Timed supine to stand	$r = 0.719$*	$r = 0.575$*		$r = 0.744$*	$r = 0.890$*
Timed 5-m walk	$r = 0.856$*	$r = 0.384$*	$r = 0.744$*		$r = 0.838$*
Timed functional tasks total score	$r = 0.896$*	$r = 0.784$*	$r = 0.890$*	$r = 0.838$*	

* Correlation is significant at $p < 0.001$

TSTS ($r = 0.482$, $p < 0.001$), the TLTS ($r = 0.408$, $p < 0.001$) and the TUGT ($r = 0.372$, $p < 0.001$). The RMDQ was only weakly correlated with the 5MWT ($r = 0.285$, $p = 0.006$).

Patient profiles

The correlations between patient characteristics and the RMDQ score are presented in Table 4. Of the 16 relationships analysed, 8 were statistically significant ($p < 0.01$). Strong relationships were noted for usual pain intensity ($r = 0.501$) and present pain intensity ($r = 0.592$).

The correlations between patient characteristics and the TFTTS are presented in Table 5. Of the 16 relationships analysed, three were statistically significant ($p < 0.01$). None of the associations explored met the criteria for a strong relationship.

Table 4 Results of Pearson's correlation tests between patient characteristics and self-reported disability (RMDQ) ($n = 93$)

Patient characteristics	Roland Morris disability questionnaire
Age	$r = 0.127$
Gender	$r = 0.186$*
BMI	$r = -0.161$
Symptom distribution	$r = 0.392$**
First/recurrent episode	$r = -0.105$
Sudden/gradual onset	$r = -0.234$*
Duration of LBP	$r = -0.210$*
Use of analgesics for LBP	$r = 0.303$**
Off work due to LBP	$r = 0.230$*
SF-36 PCS	$r = -0.489$**
SF-36 MCS	$r = -0.070$
Usual pain intensity	$r = 0.501$**
Present pain intensity	$r = 0.592$**
ZUNG	$r = 0.475$**
MSPQ	$r = 0.282$**
STAIS	$r = 0.383$**

* Correlation is significant at $p < 0.1$; ** Correlation is significant at $p < 0.001$

Table 5 Results of Pearson's correlation tests between patient characteristics and performance-based assessment of disability (TFTTS) ($n = 93$)

Patient characteristics	Timed functional tasks total score
Age	$r = 0.257$*
Gender	$r = 0.082$
BMI	$r = -0.005$
Symptom distribution	$r = 0.287$**
First/recurrent episode	$r = -0.041$
Sudden/gradual onset	$r = -0.111$
Duration of LBP	$r = 0.010$
Use of analgesics for LBP	$r = 0.212$*
Off work due to LBP	$r = 0.196$
SF-36 PCS	$r = -0.376$**
SF-36 MCS	$r = 0.115$
Usual pain intensity	$r = 0.265$*
Present pain intensity	$r = 0.457$**
ZUNG	$r = 0.144$
MSPQ	$r = 0.087$
STAIS	$r = 0.239$*

* Correlation is significant at $p < 0.1$; ** Correlation is significant at $p < 0.001$

Relationship between patient profile and the disability measures

All variables that were associated significantly with the RMDQ ($p < 0.10$, see Table 4) were entered as a group into a multivariate regression analysis with RMDQ score as the dependent variable. The model generated with this analysis explained 62% of the variance in RMDQ scores. Inspection of the standardised beta coefficients indicate that present pain intensity ($\beta = 0.285$, $p = 0.004$), the SF-36 PCS ($\beta = -0.264$, $p = 0.001$) and depression ($\beta = 0.254$, $p = 0.008$) make a significant unique contribution to the prediction of RMDQ score.

All variables that were associated significantly with TFTTS ($p < 0.10$, see Table 5) were entered as a group into a multivariate regression analysis with TFTTS as the dependent variable. The model generated with this analysis explained 41% of the variance in TFTTS. Inspection of the standardised beta coefficients indicate that present pain intensity ($\beta = 0.456$, $p < 0.001$), age ($\beta = 0.350$, $p < 0.001$) and the SF-36 PCS ($\beta = -0.221$, $p = 0.016$) make a significant unique contribution to the prediction of TFTTS.

Discussion

We were interested in exploring the feasibility of using a performance-based assessment of disability in the ALBP population, and in particular examining the relationships between performance-based and self-reported disability. Based on our preliminary investigations the protocol used would seem to be an acceptable way of measuring functional performance in this group. We combined a number of functional tasks using protocols that have demonstrated excellent reliability within clinical populations [8, 10, 19–25]. The tests are simple to administer, require little in the way of equipment and can be completed in a very short period of time. The testing protocol was well tolerated by this patient group with only one of 94 patients tested unable to complete the set tasks. Finally, the four tests that make up the TFTTS demonstrated associations with each other, and each test contributed to high internal consistency. This suggests that all items are measuring the same underlying construct.

As has been observed in non-acute LBP populations [6, 8–15], we found only a moderate relationship between self-reported disability and our performance-based assessment of disability. To further explore this finding, we looked at the univariate and multivariate relationships between patient profile and the two disability measures. The univariate analyses indicate a number of parameters that are significantly related to both disability measures. The distribution of symptoms, present pain intensity and self-reported physical well-being were significantly correlated with both disability measures.

While these three features were the only characteristics related to the performance-based measure of disability, we found a number of additional features that were significantly related to self-reported disability. Most notably, the self-reported measure was positively correlated with the subjects' level of distress, depression and anxiety. Multivariate analyses demonstrated a similar pattern. Using the same components of the patient profile we were able to explain more of the variance in self-reported disability than the performance-based measure. Furthermore, no psychosocial measure was uniquely predictive of the TFTTS, yet the level of depression made a significant, unique contribution to the prediction of the RMDQ score.

The moderate relationship between the two disability measures suggest that self-reported and performance-based assessments may be measuring somewhat different aspects of disability. The construct of self-reported disability appears to be influenced by more features of the clinical profile than the performance-based assessment, which suggests that clinicians and researchers should be cautious about assuming that self-report provides an unbiased assessment of the actual functional consequences of ALBP. Specifically, it seems that self-reported disability assessment is influenced by patients' psychological status to a greater degree than a performance-based assessment. Self-efficacy, or an individual's perception of their functional ability, may explain this finding. Self-efficacy is highly correlated with self-reported disability in LBP patients [35, 36], but is only weakly related to performance in functional capacity tests [37]. Self-reported measures may capture not what the patient can do, but what they think they can do and these two things might not always be the same, particularly in those patients with poorer psychological functioning. In addition, Rasch analyses of the most commonly used disability questionnaires indicate that they might not be wholly uni-dimensional tools [38, 39] so may capture information outside of the functional disability domain. It is possible that the timed functional tasks may offer a more precise model for evaluating the actual functional consequences of an ALBP episode. These findings are only preliminary and until more data can clarify this point, clinicians and researchers may find it useful to use both self-reported and performance-based assessments of disability to profile patients.

This investigation has several strengths. We sampled patients in the clinical setting, ensuring a clinically relevant cohort. Clinical status was extensively profiled using a number of reliable and validated outcome measures and the stringent inclusion and exclusion criteria ensured a relatively homogeneous group. Finally, we included tasks that move and load the spine in a variety of ways and chose activities of daily living rather than work related activities, ensuring relevance of the testing to a greater number of patients.

Consideration must also be given to some of the limitations of this research. We did not evaluate the reliability of this protocol in the ALBP population although the tasks chosen have demonstrated excellent reliability in other clinical conditions, including chronic LBP. It may be that four tasks is insufficient to adequately profile functional ability in this patient group, and different relationships may have been apparent if a more extensive protocol had been used. Both the univariate and multivariate analyses noted a significant relationship between the TFTTS and age, and this may need to be considered in interpretation of these tests. Finally, measurement of self-efficacy and effort [40] would have been desirable and potentially useful in further exploring the results seen.

We investigated a protocol of performance-based assessment of disability in an ALBP population. The protocol was quick and easy to perform and required little equipment. The testing was well tolerated by this population and the protocol demonstrated good internal consistency. In this sample of ALBP patients it appears that self-reported measures are more influenced by the patients' psychological status than performance-based assessments and may give an imprecise view of the actual functional consequences of an ALBP episode. To fully understand the functional consequences of ALBP, it may be advisable to use both self-report and performance-based assessments of disability.

Acknowledgments We would like to thank Christien Bird, Maureen MacDowell, Anne Golden and Mary Sexton for their contribution to the original investigation and the patients who participated in the study. The original trial was supported by the NHS Executive, South West Regional Office, Physical and Complex Disabilities National Programme. No additional funding was provided for this analysis. Ethical approval was obtained from the Brent and Harrow Health Authorities Research Ethics Committee

Conflict of interest statement No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subjects of this manuscript.

References

1. Flores L, Gatchel RJ, Polatin PB (1997) Objectification of functional improvement after nonoperative care. Spine 22:1622–1633
2. Deyo RA, Battie M, Beurskens AJH, Bombardier C, Croft P, Koes B, Malmivaara A, Roland M, Von Korff M, Waddel G (1998) Outcome measures for low back pain research: a proposal for standardized use. Spine 23:2003–2013
3. Bombardier C (2000) Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. Spine 25:3100–3103
4. Gatchel RJ, Turk DC (2008) Criticisms of the biopsychosocial model in spine care: creating and then attacking a straw person. Spine 33:2831–2836
5. Grotle M, Brox JI, Vollestad NK (2004) Concurrent comparison of responsiveness in pain and functional status measurements used for patients with low back pain. Spine 29:E492–E501
6. Verbunt JA, Westerterp KR, van der Heijden GJ, Seelen HA, Vlaeyen JW, Knottnerus JA (2001) Physical activity in daily life in patients with chronic low back pain. Arch Phys Med Rehabil 82:726–730
7. Verbunt JA, Sieben JM, Seelen HA, Vlaeyen JW, Bousema EJ, van der Heijden GJ, Knottnerus JA (2005) Decline in physical activity, disability and pain-related fear in sub-acute low back pain. Eur J Pain 9:417–425
8. Simmonds MJ, Olson SL, Jones S, Hussein T, Lee CE, Novy D, Radwan H (1998) Psychometric characteristics and clinical usefulness of physical performance tests in patients with low back pain. Spine 23:2412–2421
9. Lee CE, Simmonds MJ, Novy DM, Jones S (2001) Self-reports and clinician-measured physical function among patients with low back pain: a comparison. Arch Phys Med Rehabil 82:227–231
10. Cunha IT, Simmonds MJ, Protas EJ, Jones S (2002) Back pain, physical function, and estimates of aerobic capacity: what are the relationships among methods and measures? Am J Phys Med Rehabil 81:913–920
11. Novy DM, Simmonds MJ, Lee CE (2002) Physical performance tasks: what are the underlying constructs? Arch Phys Med Rehabil 83:44–47
12. Reneman MF, Jorritsma W, Schellekens JMH, Goeken LNH (2002) Concurrent validity of questionnaire and performance-based disability measurements in patients with chronic nonspecific low back pain. J Occup Rehabil 12:119–129
13. Gross DP, Battie MC (2003) Construct validity of a kinesiophysical functional capacity evaluation administered within a worker's compensation environment. J Occup Rehabil 13:287–295
14. Brouwer S, Dijkstra PU, Stewart RE, Goeken LNH, Groothoff JW, Geertzen JHB (2005) Comparing self-report, clinical examination and functional testing in the assessment of work-related limitations in patients with chronic low back pain. Disabil Rehabil 27:999–1005
15. Alschuler KN, Theisen-Goodvich ME, Haig AJ, Geisser ME (2008) A comparison of the relationship between depression, perceived disability, and physical performance in persons with chronic pain. Eur J Pain 12:757–764
16. Waddell G (1987) 1987 Volvo award in clinical sciences. A new clinical model for the treatment of low-back pain. Spine 12:632–644
17. Vlaeyen JWS, Kole-Snijders AMJ, Rotteveel AM, Ruesink R, Heuts PHT (1995) The role of fear of movement/(re)injury in pain disability. J Occup Rehabil 5:235–252
18. Pincus T, Vogel S, Burton AK, Santos R, Field AP (2006) Fear avoidance and prognosis in back pain: a systematic review and synthesis of current evidence. Arthritis Rheum 54:3999–4010
19. Carr JH, Shepherd RB, Nordholm L, Lynne D (1985) Investigation of a new motor assessment scale for stroke patients. Phys Ther 65:175–180
20. Harding VR, Williams AC, Richardson PH, Nicholas MK, Jackson JL, Richardson IH, Pither CE (1994) The development of a battery of measures for assessing physical functioning of chronic pain patients. Pain 58:367–375
21. Taylor S, Frost H, Taylor A, Barker K (2001) Reliability and responsiveness of the shuttle walking test in patients with chronic low back pain. Physiother Res Int 6:170
22. Steffen TM, Hacker TA, Mollinger L (2002) Age- and gender-related test performance in community-dwelling elderly people: six-minute walk test, Berg balance scale, timed up and go test, and gait speeds. Phys Ther 82:128–137
23. Schaubert K, Bohannon RW (2005) Reliability of the sit-to-stand test over dispersed test sessions. Isokinet Exerc Sci 13:119–122
24. Sherrington C, Lord SR (2005) Reliability of simple portable tests of physical performance in older people after hip fracture. Clin Rehabil 19:496–504
25. van Hedel HJ, Wirz M, Dietz V (2005) Assessing walking ability in subjects with spinal cord injury: validity and reliability of 3 walking tests. Arch Phys Med Rehabil 86:190–196
26. Wand BM, Bird C, McAuley JH, Dore CJ, MacDowell M, De Souza LH (2004) Early intervention for the management of acute low back pain: a single-blind randomized controlled trial of biopsychosocial education, manual therapy, and exercise. Spine 29:2350–2356
27. Spitzer W, LeBlanc F, Dupuis M (1987) Scientific approach to the assessment and management of activity-related spinal

disorders. A monograph for clinicians. Report of the Quebec Task Force on Spinal Disorders. Spine 12:S1–S59
28. Spielberger C, Gorsuch R, Lushene R (1970) The state-trait anxiety inventory manual. Consulting Psychologists Press, Palo Alto, CA
29. Zung WW (1965) A self-rating depression scale. Arch Gen Psychiatry 12:63–70
30. Main CJ (1983) The modified somatic perception questionnaire (MSPQ). J Psychosom Res 27:503–514
31. Ware JE, Sherbourne CD (1992) The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 30:473–483
32. Roland M, Morris R (1983) A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine 8:141–144
33. Portney LG, Watkins MP (2000) Foundations of clinical research: applications to practice. Prentice Hall Health, Upper Saddle River, NJ
34. Cohen JW (1988) Statistical power analysis for the behavioural sciences. Lawrence Erlbaum Associates, Hillsdale, NJ
35. Ayre M, Tyson GA (2001) The role of self-efficacy and fear-avoidance beliefs in the prediction of disability. Aust Psychol 36:250–253
36. Denison E, Asenlof P, Lindberg P (2004) Self-efficacy, fear avoidance, and pain intensity as predictors of disability in subacute and chronic musculoskeletal pain patients in primary health care. Pain 111:245–252
37. Reneman MF, Geertzen JHB, Groothoff JW, Brouwer S (2008) General and specific self-efficacy reports of patients with chronic low back pain: are they related to performances in a functional capacity evaluation? J Occup Rehabil 18:183–189
38. Page SJ, Shawaryn MA, Cernich AN, Linacre JM (2002) Scaling of the revised Oswestry low back pain questionnaire. Arch Phys Med Rehabil 83:1579–1584
39. Garratt AM (2003) Rasch analysis of the Roland disability questionnaire. Spine 28:79–84
40. Kaplan GM, Wurtele SK, Gillis D (1996) Maximal effort during functional capacity evaluations: an examination of psychological factors. Arch Phys Med Rehabil 77:161–164