0008-3194/98/141–149/$2.00/©JCCA 1998

Response set bias, internal consistency and construct validity of the Oswestry Low Back Pain Disability Questionnaire

Anthony C Tibbles, BSc, DC, FCCS(C)*
Judith K Waalen, PhD**
François Hains, DC, FCCS(C)†

Background: *The Oswestry Low Back Pain Disability Questionnaire (ODQ) is a widely used 10-item paper and pencil measure of disability resulting from low back pain. However, few studies have assessed the psychometric properties of the instrument. This study evaluated the response set bias, the internal consistency, and the construct validity of the ODQ.*

Objectives: *The original ODQ was compared to seven modified versions to examine whether a response set bias existed. The internal consistency of the ODQ was assessed using the Cronbach alpha. Finally, the relationship between scores on the ODQ and the Roland Morris Functional Disability Scale (RM) was examined.*

Methods: *Seven modified versions of the ODQ were developed from the original. One of the eight versions was randomly allocated to 102 adult patients presenting with low lack pain. There was no attempt to select patients on the basis of pain intensity or prior treatment so as to maximize the range and diversity of low back pain sufferers.*

Results: *Results suggest that the responses given on the eight versions of the ODQ are a function of content and not of the format in which the items are presented. The ODQ also has strong internal consistency (alpha = 0.85) and is strongly correlated to the RM ($r = .70$, $p = .0005$). The ODQ is a significant predictor of the RM scores ($T = 9.45$, $p = .0005$) and duration of symptoms ($T = -2.17$, $p = .0325$).*

Conclusion: *The ODQ appears to possess stable psychometric properties. The use of more than one*

Historique *: Le Questionnaire Oswestry sur l'invalidité causée par les lombalgies basses (ODQ), qui est largement utilisé, permet de mesurer par écrit, à l'aide de dix éléments différents, l'invalidité résultant de lombalgies basses. Peu d'études ont cependant évalué les propriétés psychométriques de cet outil de mesure. La présente étude a évalué la série de biais dans les réponses, la cohérence interne et la validité des notions du ODQ.*

Objectif *: Le ODQ original a été comparé à sept versions modifiées dans le but de vérifier s'il existait une série de biais dans les réponses. La cohérence interne a quant à elle été évaluée à l'aide de la méthode de l'alpha de Cronbach. Enfin, le lien entre les cotes du ODQ et celles de l'échelle d'invalidité fonctionnelle Roland Morris (RM) a été étudié.*

Méthodologie *: Sept versions modifiées du ODQ ont été élaborées à partir de la version originale. Une des huit versions a été soumise au hasard à 102 adultes souffrant de lombalgies basses. Les patients ont été choisis sans égard à l'intensité de leur douleur ou à un traitement antérieur afin qu'ils présentent la plus grande diversité possible.*

Résultats *: Les résultats semblent indiquer que les réponses données dans les huit versions du ODQ le sont en fonction du contenu et non du fromat dans lequel les éléments sont présentés. Le ODQ affiche également une cohérence interne élevée (alpha = 0,85) et une forte corrélation avec le RM ($r = 0,70$, $p = 0,0005$). Il constitue un élément prédictif important du RM*

* Associate Clinical Professor, Division of Clinical Education, Canadian Memorial Chiropractic College, Toronto, Ontario.
** Professor, Division of Postgraduate and Continuing Education, Canadian Memorial Chiropractic College, Toronto, Ontario, and Professor, Ryerson Polytechnic University, Toronto, Ontario.
† Research Associate, Canadian Memorial Chiropractic College, Toronto, Ontario, and Graduate Student, Department of Social and Preventive Medicine, University of Montreal, Montreal, Canada.
Reprint requests to: A. Tibbles, Canadian Memorial Chiropractic College, 1900 Bayview Ave., Toronto, Ontario M4G 3E6.

version provides practitioners with a means of repeatedly assessing the disability levels of patients suffering from low back pain over the course of treatment.
(JCCA 1998; 42(3):141–149)

KEY WORDS: low back pain, disability, psychometric properties, Oswestry Low Back Pain Disability Questionnaire.

(T = 9,45, p = 0,0005) et de la durée des symptômes (T = −2,17, p = 0,0325).

***Conclusion** : Le ODQ semble posséder des propriétés psychométriques stables. L'utilisation de plus d'une version permet aux praticiens d'évaluer à maintes reprises, tout au long du traitement, les niveaux d'invalidité des patients souffrant de lombalgies basses.*
(JACC 1988; 42(3):141–149)

MOTS CLÉS : lombalgies basses, invalidité, propriétés psychométriques, questionnaire Oswestry sur l'invalidité causée par les lombalgies basses.

Introduction

Patients and physicians often view disability and low back pain as synonymous when, in fact, they are very different.[1] Disability is a concept broader than pain or physical impairment. It refers to the loss of functional ability that results from the physical impairment.[2,3] Waddell and Main[2] define disability as a diminished capacity for everyday activities or the limitation of a patient's performance compared to a fit person's performance of the same age and gender. In contrast, pain is defined as an association of stimuli with responses which result in an unpleasant experience which hurts a person and from which they want to be freed[4] and physical impairment is an anatomical or pathological abnormality leading to a loss of normal body ability.[2] The distinction is underscored in the World Health Organization definition of disability which states that it is the loss of functional ability and activity consequent upon impairment.[5] These definitions acknowledge that disability includes a patient's perception and response to the physical change, and thus involves other components such as psychological and social factors.[6]

Studies have shown that physical signs alone are rather insensitive measures of disability.[7] Clinicians may often underestimate the impact of illness by missing its influence on a patient's daily activities.[8,9] In addition, there is growing treatment emphasis on minimizing the impact of illness on everyday activities. With this change in paradigm, efficacy and efficiency of care can no longer be judged by unidimensional methods such as morbidity rates.[10] Pain, although it is a subjective perception, is a measure of impairment rather than disability.[11] For these reasons, tools designed to measure disability must attempt to include patient reaction to the limitations imposed by the condition in order to accurately assess its impact. Tools such as the Oswestry Low Back Pain Disability Questionnaire (ODQ) are of potential value in this regard. There are, however, some caveats to their use. The use of a functional assessment tool that has not been subjected to accepted psychometric methods has been justly criticized.[12] As a result, researchers are increasingly striving to measure disability by applying psychometric methods to their evaluation procedures.[13–17]

An individual patient's perception of his or her disability is largely determined by its effect on the patient's activities of daily living. Functional status measures are instruments which assess limitations in performing usual human tasks of living.[18,19] The subjective data from these questionnaires can be as reproducible as more traditional measures,[20] which are generally subjective self-reports. The main practical advantages of self-reporting are the ease and cost effectiveness with which data can be obtained.[21] Not only do the activities of daily living scales focus on function, they also may identify discrete dimensions of disability.[18,22] These dimensions of disability are the different attributes, such as psychosocial, and physical, which may contribute to the patient's perception of his or her limitation. Sufficient detail may be garnered to distinguish activities which only increase pain from those which are actually limited by pain.[19]

The ODQ is an ordinal rating scale of home and work

functional disabilities resulting from back pain.[21] It attempts to measure the handicap and disability of pain rather than the nature of pain.[6,23] The instrument consists of 10 items, the first of which is a pain-rating scale. The other items concern a variety of normal daily activities agreed by a consensus panel to be relevant to low back disability.[4] The ODQ is a practical scale taking only 3–5 minutes to administer and one minute to score.[24] As a self-administered questionnaire, it avoids interviewer bias and allows for a uniform presentation.[24] The ODQ has been shown to be consistent over time (test-retest reliability).[24] In a recent study, it was shown to be sufficiently reliable and reponsive to clinical changes over time to be considered useful in a randomized clinical trial.[25] Internal consistency was also determined to be good in a study using chronic low back pain patients as subjects.[24] However, the internal consistency may have been high because, as Nelson et al.[26] point out, all item responses are written in order of increasing severity. The ODQ appears to be valid when used to observe expected improvements among patients with a first episode of low back pain[27] and with the Waddell Disability Index.[2] In spite of the ODQ's widespread use[27–30] few psychometric properties of the scale, other than the original test-retest reliability, have appeared in the literature.

Further, measuring dimensions of disability is important to fully understand an individual's problem and to define appropriate treatment. Therefore, instruments that measure activities of daily living should be assessed to determine which dimensions they capture. The Roland-Morris Scale (RM), for example, has been investigated in this way. The RM is a shorter version (24 item checklist) of the Sickness Impact Profile specifically designed to measure disability resulting from low back pain. Like the ODQ, it is an activity of daily living scale which has been found to be unidimensional, capturing the physical dimension of low back pain disability.[31] The RM is a more recently developed scale than the ODQ, but it has also found widespread use. It requires approximately 5 minutes to complete[7] and has been shown to be a sensitive and reliable measure of disability resulting from low back pain.[32] Its validity has been demonstrated by comparison to the Pain Rating Scale,[32] the ODQ,[27] and to clinical observation.[31]

Deyo,[18] McDowell and Newell,[6] Streiner and Norman[23] and Triano et al.[25] suggest that currently used useful questionnaires, such as the ODQ, require further investigation to determine their psychometric properties in order to have confidence in the results. These properties include the measurement of construct validity, internal consistency, factor structure, and response set bias.

The present study was designed to evaluate three elements of the ODQ: response set bias, internal consistency, and construct validity. The first question posed was whether a response set bias existed in the ODQ. This is a distinct possibility since the item responses are in increasing order from no disability to total disability.[24] A patient may, therefore, respond to later questions based on his or her earlier responses.[23] If present, such a bias could be reduced by altering the direction of the item responses.[34] In the second part of the study, the total ODQ score and its individual items were compared to each other and to the results on the RM in order to assess internal consistency using the Cronbach alpha technique.[34] Correlation among ODQ items also needs investigation. Highly correlated items, actually due to redundancies, may have falsely created the high internal consistency reported in the original research.[24] In the last part of the study, the construct validity of the ODQ was measured by comparing ODQ scores to the RM, patient age, and duration of low back pain using multiple regression.

Methods

Design

This was a descriptive study investigating the response set bias, internal consistency, and construct validity of the ODQ in a sample of subjects with low back pain. This type of design involves no intervention or treatment.

Materials

To investigate the psychometric properties of the design of the ODQ, different forms of the instrument were used. For this study, seven modified versions of the questionnaire were developed. Each of these versions differed from the original in one respect. Version 1 consisted of the original 10 ODQ items varying from no disability to total disability. The six responses were arranged in order of increasing disability. In Version 2, the order of the items was reversed, but the order of the six responses was retained in an increasing direction. In Version 3, the original item order was used, but the direction of the responses was

reversed to decreasing disability. In Version 4, both the order of the items and the direction of the responses was reversed from the original. In the next two versions, the original order of the items was retained, but the responses alternated the order of the intensity of the responses. That is, in Version 5, odd items were ordered in increasing disability, while the even items were listed in decreasing disability. In Version 6, the item order was arranged as in the original, but the odd items had their responses listed in order of decreasing disability while the even items had their responses ordered in increasing disability. In Versions 7 and 8, the item order was reversed from the original and the response order followed the pattern of Versions 5 and 6, respectively.

Subjects

Adult subjects 17 years of age or older with low back pain were used. A wide spectrum of the condition was desired. Non-consecutive subjects were selected from among the outpatients of chiropractic college clinics, as well from patients of chiropractors in private practice. Patients who presented with a complaint of low back pain were asked to participate in the study. The only inclusion criteria were low back pain and subjects of an age above 16 years. No differentiation was made as to intensity or prior treatment, to maximize the range and diversity of low back pain sufferers.

Protocol

After first obtaining written consent for participation in the investigation from each patient, demographic data (name, age, gender, occupation) were collected. The duration of low back pain for the present episode was also recorded. Using a random numbers table, one of the eight versions of the ODQ to be completed by the patient was chosen. To be able to measure ODQ construct validity, the RM was also completed by the patient. The order in which the subjects completed the two instruments was alternated to control for testing effects. The process of data collection spanned eight months.

Sample size estimate

The ODQ is composed of ten items. It is recommended that analyses of this type should be performed on no less than five to ten subjects per variable or item.[6,35] Although, a minimum of 50 subjects was required, 102 subjects were recruited. No subjects that had met the inclusion criteria dropped out of the study.

Statistical analyses

Although the ODQ is an ordinal scale, many researchers now use parametric statistics on these type of instruments. We have made the same decision. Using multivariate analysis of variance, the item scores and the total scores on the eight versions were analyzed to investigate response set bias, and/or interaction between the question order and the response order. Cronbach Alpha was calculated to determine internal consistency of the ODQ on this sample of patients. To determine the relationship between the two disability instruments, the ODQ score and its individual items were compared to the total RM score using correlation coefficients. Multiple regression was used to determine the variables most predictive of the RM. The significance level was set at .05. All analyses were performed using SPSS for Unix, Release 6.1 (AIX 3.2, IBM RS/6000).

Results

Sample characteristics

The study sample consisted of 102 patients experiencing low back pain. The subjects varied in age from 17 to 72 years (Mean = 35, SD = 11.4). Males predominated (59%). A broad selection of occupational categories was represented. Approximately 36% worked in occupations classified as professional or semi-professional, 19% were students, 15% were managerial, 14% were clerical, 11% were manual labourers, 3% were retired, and 2% were homemakers.

The duration of low back pain ranged from 1 day to over 3 years. Utilizing the Quebec task force categories for duration of pain,[36] the most represented group was the subchronic patients, accounting for 40%, who had pain lasting from 7 through 49 days. Chronic patients, with pain lasting more than 49 days, accounted for 31%. Acute patients represented 27% of the sample. Thus, a wide spectrum of low back pain duration was captured in the sample. The mean duration of the episode was 408 days (SD = 1389), but the median number of days was 14.5. About half (54%) of the patients attended college teaching clinics while 46% attended private practitioners for treatment of their condition.

Approximately equal numbers of the eight versions of the ODQ were used. Twelve patients each completed Versions 1 and 2; nine completed Version 3; ten completed Version 4; 17 completed Version 5; 14, 15, and 13 completed Versions 6, 7, and 8 respectively.

The distribution of scores on the individual ODQ items are summarized in Table 1. Overall, the individual item scores were skewed to the lower end of the disability scale. The total ODQ score is 24%. For example, the pain intensity ratings (Q1) of these low back patients varied from minimal to no pain (8.8%) to very severe pain (3.9%). The median pain intensity rating was 1.0 (very mild). Regarding limitations in personal care (Q2), most patients (61%) reported they could manage these functions without pain. Lifting (Q3) was rated by 38% as causing extra pain. One percent, however, reported that they could not lift. The majority of respondents (67%) reported that they could walk (Q4) without pain. Seven percent of patients could not sit (Q5) for less than an half an hour. Standing (Q6) most commonly caused at least some difficulty (72%). The most commonly answered response to Question 7 regarding sex life, was that it caused some pain (34%). Social activities (Q8) were most able to be performed with slight pain (47%). The majority of patients (56%) felt that they had no trouble sleeping (Q9). Travelling (Q10) was pain-free in 11% of cases. Most respondents said that travelling caused some back pain (35%). The scores of the 24 item Roland-Morris Disability checklist varied from 1 to 24, with the median score being 4.0 (Mean 6.1, SD = 5.6).

Overall, the sample had demographic heterogeneity and the duration of complaint was widely distributed. In terms of disability, however, the severely disabled were underrepresented.

Response set bias

A multivariate analysis of variance was performed to determine if the eight versions had order or sequencing effects. No significant differences between versions were found. Table 2 outlines the results. The order in which the questions were asked does not seem to alter how patients completed the form. Similarly, sequencing of the responses did not have an effect on how the ODQ was answered (F ratio = 0.59, p = .762). As a result of the lack of sequencing or order effects, all versions of the questionnaire were used in the subsequent analyses.

Item redundancy and construct validity

The correlations between the individual items on the ODQ (Table 3) indicate that the strength of the relationships

Table 1
Distribution of scores on the individual items of the ODQ (n = 102)

Item	mean	stddev	skewness	kurtosis
Q1 Pain Intensity	1.52	0.91	0.71	0.65
Q2 Personal Care	0.50	0.78	2.56	10.52
Q3 Lifting	1.78	1.00	0.38	–0.98
Q4 Walking	0.63	1.05	1.63	1.63
Q5 Sitting	1.54	1.17	0.56	–0.05
Q6 Standing	1.32	1.08	0.85	–0.06
Q7 Sex Life	0.75	0.84	1.03	1.02
Q8 Social Life	1.32	1.02	0.93	1.12
Q9 Sleeping	0.95	1.36	1.44	1.23
Q10 Traveling	1.64	0.97	0.26	–0.34
Total	11.94	6.87	0.70	–0.06
Total Disability Score = 11.92 × 2 = 23.88%				

Table 2
Univariate F tests on the item scores and total score by version (7,94 df)

Item	SS	MS	F	p value
Q1 Pain Intensity	8.22	1.17	1.47	.189
Q2 Personal Care	4.21	0.60	0.99	.446
Q3 Lifting	4.68	0.67	0.38	.914
Q4 Walking	5.73	0.82	0.73	.651
Q5 Sitting	7.05	1.01	0.73	.650
Q6 Standing	2.43	0.35	0.28	.960
Q7 Sex Life	3.05	0.44	0.60	.754
Q8 Social Life	6.48	0.93	0.89	.518
Q9 Sleeping	15.80	2.26	1.24	.288
Q10 Traveling	6.22	0.89	0.94	.483
Total Score	200.61	28.66	0.59	.762

Table 3
ODQ and RM correlation coefficients

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total
Q2	.13	—									
Q3	.36*	.28*	—								
Q4	.16	.43*	.40*	—							
Q5	.47*	.20*	.39*	.32*	—						
Q6	.33*	.12	.36*	.34*	.21*	—					
Q7	.28*	.41*	.51*	.52*	.31*	.31*	—				
Q8	.44*	.41*	.53*	.59*	.44*	.45*	.58*	—			
Q9	.24*	.19	.35*	.43*	.26*	.21*	.45*	.31*	—		
Q10	.36*	.32*	.36*	.42*	.55*	.31*	.56*	.51*	.29*	—	
Total	.57*	.50*	.71*	.70*	.64*	.56*	.74*	.79*	.60*	.70*	—
RM	.33*	.44*	.50*	.66*	.33*	.36*	.56*	.61*	.40*	.42*	.70*

* p value ≤ 0.05

between items varies from .13 (pain and personal care) to .59 (walking and social life). The concern that the individual items on the ODQ may be highly correlated, thereby artificially inflating internal consistency through redundancy, does not seem warranted.

A comparison of the ODQ total score with the RM score shows that they are strongly related ($r = .70$, $p = .0005$). The individual item correlations are lower than the total score, ranging from .33 (pain, sitting) to .66 (walking). Multiple regression (backward elimination) was used to identify which variables (age, gender, duration of episode and total ODQ score) were most predictive of the scores on the RM. The ODQ ($T = 9.45$, $p = .0005$) and duration of episode ($T = -2.17$, $p = .0325$) were the significant predictors.

Internal consistency

The assessment of internal consistency of the eight versions of the ODQ was measured by the Cronbach alpha. Alpha reliability was found to be high (0.852). Further, as illustrated in Table 4, the removal of any of the ten items does little to improve the alpha coefficient.

Table 4
Cronbach Alpha coefficient and alpha scores if individual items are deleted (*n* = 102)

Item	Alpha Scores if item deleted
Q1 Pain Intensity	0.836
Q2 Personal Care	0.841
Q3 Lifting	0.825
Q4 Walking	0.824
Q5 Sitting	0.832
Q6 Standing	0.840
Q7 Sex Life	0.822
Q8 Social Life	0.813
Q9 Sleeping	0.843
Q10 Traveling	0.824
Standardized item alpha = 0.852	

Discussion

The original ODQ consists of 10 items, each with six possible responses. The response choices are in ascending order of severity. When patients answer, they may possibly respond to later questions based on their earlier responses. That is, there may have been a response set bias in the instrument.[33] For example, patients may simply judge that since they choose the mildest detractor for the first question, which is pain intensity, then the mildest detractors for the rest of the questions also applies to their situation. This may give an inappropriate disability rating. Although their pain intensity is mild, perhaps impairment of another activity, such as social life, may be much higher.

This investigation into response set bias showed no significant effects, by either question order or response order. This means that all versions of the questionnaire were answered in a similar way, regardless of whether the questions or responses appeared in different orders. This finding may have clinical application. When assessing disability, it is often necessary to retest a patient's rating several times. This raises the possibility of familiarity with the instrument and response bias confounding a true assessment. Alternate versions of the ODQ could be used on subsequent patient assessments to address this potential problem.

Internal consistency in the ODQ was also addressed in this study. A questionnaire which displays high inter-item correlations is said to be homogeneous and is therefore likely to show consistent responses. This internal consistency is a measure of the extent to which items measure the same characteristic. Internal consistency is typically equated with Cronbach's coefficient alpha. The present study demonstrated internal consistency that is considered very good.[37] Fairbank et al.[24] showed good internal consistency for the original questionnaire in a study of 22 chronic low back pain patients. The present study adds to their previous work in two ways. First, the original paper on the ODQ used only one version and all item responses were written in order of increasing severity. In the present study, this potential problem was addressed by the inclusion of versions in which item order was reversed. Consistency was seen over all versions, thereby eliminating this concern. Also, the lack of any exceedingly strong inter-item correlations effectively dispenses with the concern that there may be redundancies creating the high internal consistency. A high level of internal consistency is appropriate for the ODQ because it measures a relatively narrow aspect of health, namely low back pain. Reliable internal structure allows confidence in the measurement of the current status of a patient's disability resulting from low back pain.[6] This finding adds to the Triano et al.[25] study which showed that the ODQ is reliable and responsive enough to be used in randomized clinical trials. Second, inclusion of a widely distributed sample of low back pain duration complements the original reliability investigation. The paper on the reliability of the ODQ included only chronic low back pain patients.[24] In this study, all three duration categories (acute, subacute, chronic) are almost equally represented. Thus, confidence in the reliability of the scale is now increased over a wider range of duration of low back pain.

This study also compared the ODQ and RM. The RM has been shown to represent physical disability[31] and this study has demonstrated that the ODQ is strongly correlated to it. Correlation between two scales that purport to measure disability provides a degree of construct validity to both measures. Construct validity was a quality that had not previously been established for the ODQ.[11]

The reported duration of the low back pain episode supports the work of Tait et al.[38] who found that pain duration correlated negatively with level of disability. They suggested that, perhaps, patients accommodate to pain when it persists over time. Their research was based on chronic pain patients exclusively. The present study included acute, sub-acute and chronic patients which allows for more confidence in this negative relationship.

Interpretation of these results, however, should be qualified by certain limitations of this study. The patients in this study were physically able to attend outpatient c linics or private practices for care. Thus, we have failed to capture those low back pain suffers who are too disabled to travel. Further, the average total disability score was 24% which suggests that the severely disabled have not been captured.

Conclusion

The functional status of patients is the most desirable outcome measure for both clinical use and research. Treatment for low back pain is directed at restoring function and actual performance in normal activities, not simply at reducing pain.[12,18,19] Thus, it is important to use reliable and valid measures of disability.

The ODQ is an instrument designed specifically for patients with back pain. It does not contain a response set bias. As such, alternate versions may be used on subsequent disability assessments. Psychometric evaluation of the ODQ has shown it to be an instrument with high internal consistency. The ODQ scores are strongly correlated to RM scores.

Acknowledgements

This study was funded by an internal grant from the Division of Postgraduate and Continuing Education, which we gratefully acknowledge. We also thank Ms. Carol Hagino for the construction of the versions of the questionnaire used in this study and Dr. Gordon Emslie for his statistical advice.

References

1 Waddell G, Newton M, Henderson I, Somerville D, Main CJ. A fear avoidance beliefs questionairre (FABQ) and the role of fear avoidance beliefs in chronic low back pain and disability. Pain 1993; 52:157–168.
2 Waddell G, Main CJ. Assessment of severity in low back disorders. Spine 1984; 9(2):204–208.
3 Healthy people 2000. National health promotion and disease prevention objectives. U.S. Department of health and human services. Public health service. 1990.
4 Vernon HT. Applying researched-based assessments of pain and loss of function to the issue of developing standards of care in chiropractic. Chiro Technique 1990; 2(3):121–126.
5 Bombardier C. Measurement of disability. Spondyloarthropathies. Calin A, ed. Orlando: Grune and Stratton, Inc., 1984:104.
6 McDowell I, Newell C. Measuring health: a guide to rating scales and questionnaires. Second Edition. Oxford University Press, New York. 1996:359–361.
7 Roland M, Morris R. 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 1983; 8(2):141–144.
8 Grady KE, Wallston BS. Research in health care settings. Applied social research methods series. Volume 14. Sage Publications, Newbury Park. 1988. 101–116.
9 Williams RC. Toward a set of reliable and valid measures for chronic pain assessment and outcome research. Pain 1988;35:239–251.
10 Bergner M, Bobbit RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. Med Care 1981; 19(8):787–805.
11 Ruta DA, Garratt AM, Wardlaw D, Russell IT. Developing a valid and reliable measure of health outcome for patients with low back pain. Spine 1994; 19(17):1887–1896.
12 Silverstein B, Kilgore KM, Fisher WP, Harley JP, Harvey RF. Applying psychometric criteria to functional assessment in medical rehabilitation: I Exploring unidimensionality. Arch Phys Med Rehabil 1991; 72:631–637.
13 Greenough CG, Fraser RD. Assessment of outcome in patients with low back pain. Spine 1992; 17(1):36–41.
14 Stratford PW, Binkley J, Solomon P, Gill C, Finch E. Assessing change over time in patients with low back pain. Phys Ther 1994; 74(6):528–533.
15 Beurskens AJ, de Vet HC, Koke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. Pain 1996; 65(1):71–6.
16 Kopec JA, Esdaile JM, Abrahamowicz M, Abenhain L, Wood-Dauphinee S. The Quebec Back Pain Disability Scale: measurement properties. Spine 1995; 20(3):341–352.

17 Strong J, Ashton R, Chant D. Pain intensity measurement in chronic low back pain. Clin J Pain 1991; 7(3):209–218.
18 Deyo RA. Measuring the functional status of patients with low back pain. Arch Phys Med Rehabil 1988; 69:1044–1053.
19 Waddell G. Occupational low-back pain, illness behavior, and disability. Spine 1991; 16(6):683–684.
20 Deyo RA, Andersson G, Bombardier C, Cherkin DC, Keller RB, Lee CK, Liang MH, Lipscomb B, Shekelle P, Spratt KF, Weinstein JN. Outcome measures for studying patients with low back pain. Spine 1994; 19(18S):2032S–2036S.
21 Andersson GBJ. Functional evaluation of the back. Biomechanics of human movement. Berme N, Cappozzo A, eds. Worthington: Bertec Corp., 1990:361–381.
22 Gronblad M, Jarvinen E, Hurri, Hupli M, Karaharju BO. Relationship of the Pain Disability Index (PDI) and the Oswestry Disability Questionnaire (ODQ) with three dynamic physical tests in a group of patients with chronic low-back and leg pain. Clin J Pain 1994; 10(3):197–203.
23 Hanson DT, Ayres JR. Chiropractic outcome measures. Chiro Technique 1991; 3(l):53
24 Fairbank JCT, Couper J, Davies JB, O'Brien JP. The Oswestry Low Back Pain Disability Questionnaire. Physiotherapy 1980; 66(8):271–273.
25 Triano JJ, McGregor M, Cramer GD, Emde DL. A comparison of outcome measures for use with back pain patients: results of a feasibility study. J Man Physiol Ther 1993; 16:67–73.
26 Nelson EC, Landraf JM, Hay RD, Kirk JW, Wasson JH, Keller A, Zubkoff M. The COOP charts: a system to measure patient function in physician's offices. In: Greenfield S ed. Functional status measurement in primary care, Wonca classification committee. New York: Springer-Verlag, 1990:97–124.
27 Leclaire R, Blier F, Fortin L, Proulx R. A cross-sectional study comparing the Oswestry and Roland-Morris functional disability scales in two populations of patients with low back pain of different levels of severity. Spine 1997; 12(1):68–71.
28 Haas M, Jacobs GE, Raphael R, Petzing, K. Low back pain outcome measurement assessment in chiropractic teaching clinics: Responsiveness and applicability of two functional disability questionnaires. J Manipulative Physiol Ther 1995; 18(2):79–87.
29 Mikail SF, DuBreuil S, D'Eton J. A comparative analysis of measurements used in the assessment of chronic pain patients. Psychol Assessment 1993; 5(1):117–120.
30 Little DG, MacDonald D. The use of the percentage change in Oswestry Disability Index score as an outcome measure in lumbar spinal surgery. Spine 1994; 19(19):2139–2143.
31 Deyo RA. Measuring the functional status of patients with low back pain. Arch Phys Med Rehabil 1988; 69:1044–1053.
32 Roland M, Morris R. A study of the natural history of low back pain; Part II: development of guidelines for trials of treatment in primary care. Spine 1983; 8(2):145–150.
33 Streiner DL, Norman GR. Health measurement scales; a practical guide to their development and use. Oxford University Press, Toronto. 1989.
34 Cronbach LJ. Essentials of psychological testing, 3rd ed. Harper and Row, New York. 1970:451.
35 Tabachnick BG, Fidell LS. Using multivariate statistics, 2nd ed. New York: Harper Collins, 1989:597–677.
36 Spitzer WO. Scientific approach to the assessment and management of activity-related spinal disorders. Spine 1987; 7(suppl):S17.
37 DeVillis R F. (1991) Scale development: theory and application. Newbury Park, CA, Sage Publications.
38 Tait RC, Pollard CA, Margolis RB, Duckro PN, Krause SJ. The pain disability index: psychometric and validity data. Arch Phys Med Rehabil 1987; 68:438–441