Practical challenges and limitations using the Oswestry Disability Low Back Pain Questionnaire in a private practice setting in New Zealand. A clinical audit.

Nicola Nunn *DipPhty, Dip MT*

Principal Physiotherapist, The Back Clinic

ABSTRACT

The Oswestry Disability Low Back Pain Questionnaire (ODQ) is a validated outcome measure responsive to chronic low back pain. In New Zealand, a proportion of private practitioners use this disease specific outcome measure This clinical audit of a solo physiotherapy practice treating patients with chronic low back pain sought to outline the challenges and limitations of maintaining and interpreting the ODQ. Over a period of 9 years and 447 patients, 225(52.9%) completed the ODQ at discharge and 127(29.9%) at follow up of 2 months. Analysis of individual categories revealed smaller change scores in those with the highest baseline scores, i.e.; lifting, sitting and activities. Delay to treatment >3 months reduced change scores and increased treatment numbers (p=0.001).

Outcomes showed improvement of between 53% and 79% at discharge, which was maintained at follow-up. Minimal detectable change was calculated to be 11% but the absence of appropriate data to calculate a Minimal Important Clinical Difference (MICD) was a limitation. Collecting and maintaining outcome statistics impacts on the interpretation of the measure. Poor questionnaire return rate and administrative errors also reduced the validity of the data.

Nunn N (2012): Practical challenges and limitations using the Oswestry Disability Low Back Pain Questionnaire in a private practice setting in New Zealand. A clinical audit. New Zealand Journal of Physiotherapy 40(1) 24-28.

Key Words: Oswestry Disability Low Back Pain Questionnaire, Chronic Low Back Pain

INTRODUCTION

Outcome measures are recognised as a valuable tool when assessing response to intervention, yet little is known about the practical challenges involved in their collection and interpretation within private physiotherapy practice in New Zealand. Research usually requires a team effort in a university-funded setting (Dworkin et al 2005, Fairbank et al 2005, Stratford et al 1994, Wessels et al 2006) but over the last two decades there has been increased pressure by funders on individual physiotherapists to substantiate the effects of clinical treatment using outcome measures. The first experience in New Zealand of this was during the late 1990's with multiple funders for injury treatment and a perception that funding would be dependent on outcomes. As recently as 2009, the Accident Compensation Corporation (ACC) mandated the use of outcome measures, namely, the Visual Analogue Score and the Patient Specific Functional Scale (ACC News 2009), which must be submitted when requesting further treatment funding. However, there are significant limitations regarding what the outcome measure tells us because of confounding influences, for example, psycho-social factors. Selecting a treatment based on evidence from randomised controlled trials is the gold standard, but in clinical practice, there may not be such evidence to draw on. Thus using outcome measures, as a clinical audit to gauge response to treatment, becomes the only available alternative (Herbert et al 2005).

This is a study of a private solo practice specialising in the treatment of chronic non- specific low back pain, between the years of 2001 and 2009, during which the Oswestry Disability Low Back Pain Questionnaire (ODQ) was used as an outcome measure (Fairbanks et al 1980). The ODQ was chosen over the Roland Morris Disability Questionnaire, because it was considered to capture more thoroughly the parameters of interest to both therapist and patient (Beattie and Maher 1997). This is an internationally recognised outcome measure used by some New Zealand physiotherapists (Copeland et al 2008) and has been validated for reliability and responsiveness for chronic back pain (Bombardier 2000, Campbell et al 2006, Charted Society of Physiotherapy 2004, Deyo 1988, Frost et al 2008).

The ODQ has 10 questions, each broken into 6 levels of severity (0-5). Possible total scores range from 0 for no disability to 50 for the greatest disability. The total score is multiplied by 2 and divided by 100 to provide a percentage score. Question categories include personal care, lifting, sitting, standing, walking tolerance, social activity, travel, sex and pain severity. Fairbank et al (1980), the original developers of the ODQ, devised an arbitrary scale of disability: 0-20% minimal disability, 20-40% moderate, 40-60% severe, 60-80% crippled and 80-100% bed bound or exaggeration, thereby giving the therapist a benchmark for severity. To assess significant change in treatment, researchers have recommended a mean score improvement of 4-6 points or, alternatively, calculating the Minimal Detectable Change (MDC) (Chartered Society of Physiotherapy 2004, Davidson and Keating 2002). An improvement of 10.5 percentage points (range 4-17) is recommended for the Oswestry scale to be 95% confident that meaningful change has occurred (Davidson and Keating 2002). In the UK, a 10 percentage point improvement is accepted as significant change (Chartered Society of Physiotherapy 2004).

The treatment provided at this practice is focused on functional movement patterns. The patient is educated in recognising patterns of spine behaviour that cause pain and how to use efficient, balanced, pain free movement. Education is provided in conjunction with specific stabilisation exercise, manual treatment including myofascial release techniques, mobilisation/manipulation and balance/proprioceptive exercise when required.

The purpose of this paper is to elucidate the practical challenges of collecting data in the private practice context and to analyse how this can affect clinical decision making.

METHOD

In this study patients were included from 2001-2009 using the ODQ (version 2) (Chartered Society of Physiotherapy 2004). Each patient was given an ODQ form to complete 10 minutes prior to their initial treatment. In 2008, this process was adjusted to reduce the problem of late arrival of the patients, such that the ODQ was mailed to the patient, so it could be completed prior to their appointment. The second Questionnaire was completed at discharge, or soon after by letter, and the third questionnaire was completed by mail follow up at two months post discharge. The patient was sent the questionnaire with an explanatory letter and a stamped addressed return envelope.

In this study, 'chronic low back pain' is defined as pain persisting longer than 3 months as stated in the NZ Acute Low Back Pain Guidelines, (ACC 2001, Bogduk 2004) and 'interval' as the time period from injury to first treatment at this clinic. Clarification on the ethics of this study was sought from the Health and Disability Ethics Committee who confirmed a review was not required for this audit (NTY/11/EXP/069). However, informed patient consent was obtained.

DATA ANALYSIS

Data were collected using the Peak Software "Physio Office" Programme, which was uniquely configured to allow for collection and export of data to a relational database. Statistical analysis was undertaken using Statistica version 5.1. Chi squared analysis or students t test were used for discrete or continuous variables respectively. Significance was set at a p value of less than 0.05 (2-tailed). Anova was used for the multiple variables.

RESULTS

The number of patients treated with an initial diagnosis of low back pain who participated in the ODQ study was 447. This comprised 50% of the total number of patients (880) treated with low back pain over this period in the practice. The mean age of patients included in the ODQ study was 42.9 years and 61% were female. The reasons why patients were not included were: incorrect data entry resulting in failure to capture statistics (33%), declined to participate (11%), and incorrect diagnosis (6%)

The 'interval' ranged from 5 weeks to 10 years and 70% of patients were classified as chronic, with 34% having symptoms for more than 1 year. There was a significant difference in the number of treatments required by patients presenting less than 3 months post injury compared to those presenting greater than 3 months post injury (6.1 SD4.6 versus 7.2 SD4.7, respectively, p=0.001) (Table 1).

Table 1: Treatment numbers across total low back pain population

Average Number of Treatments	6.8 (SD 4.7)
Patients <3/12 from date of injury	6.1 (SD 4.6)**
Patients >3/12 from date of injury	7.2 (SD 4.7)
Mean Time from Date of Injury to First Treatment (Interval) (yrs)	2.2 (SD 4.8)

** p= 0.001 Patients <3/12 post injury versus >3/12 post injury

The ODQ was administered on three occasions; 225 patients (52.9%) completed a form at discharge and 127 (29.9 %) completed the third and final follow up questionnaire.

Neither age nor sex influenced completion of the ODQ data at discharge or at follow up (Figure 1). However, increasing delay in presentation to first treatment resulted in reduced return rates of the questionnaire (p<0.01) (Figure 2).Thus 127 patients comprised the group that underwent further detailed analysis.

Figure 1: Non Return of Questionaires by Age

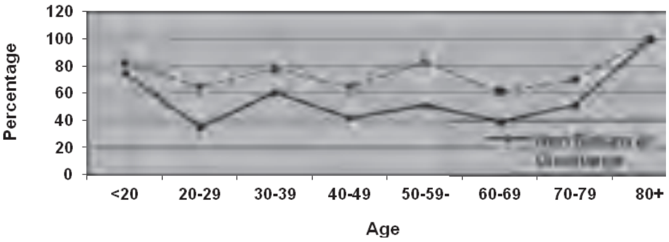
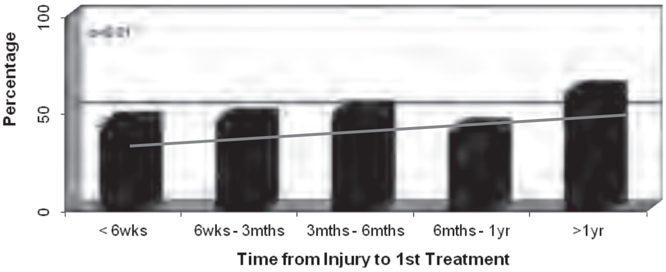


Figure 2: Non Return of Questionaires by Interval

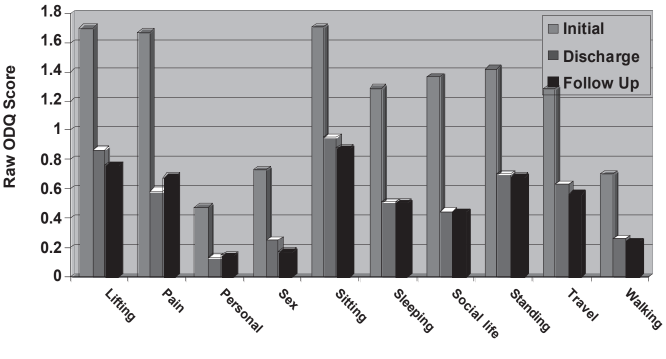


ODQ RESULTS

The total initial ODQ score was 24.5%, the total discharge score was 10%, and the total follow-up score was 10%. Improvement in total score from baseline to discharge was 42.5% and from baseline to follow-up was 42.8%. The average baseline raw score across all categories was 1.25 (SD1.0) (24.5% SD20%). This places patients in the moderately disabled category as classified by Fairbanks (1980).

Three records (2.3%) showed symptoms had worsened at discharge and remained the same at follow up. The improvement from baseline to discharge was significant across all categories (p<0.001) and was maintained at the 8 week follow up period (Figure 3).

Figure 3: ODQ Results by category



The mean improvement for those patients who completed all three questionnaires was 5.5. The MDC in this population was lower due to the moderately low average baseline score (24.5%)

and was calculated to be 5.5 (See Appendix 1) (Davidson and Keating 2002).

INDIVIDUAL CATEGORY ASSESSMENT

Figure 3 shows the response in the individual categories and the scores at initial assessment (baseline), discharge and follow-up but these include scores equal to zero. In order to assess the impact of treatment on those individual categories that had been the predominant complaint initially, data were re-analysed using categories with a baseline score greater than or equal to 1. At baseline, the categories most disabling to the patient were lifting, sitting, standing, pain and social activities. These categories improved by 57%, 53%, 56%, 58% and 69%, respectively. Greatest improvement from baseline to follow up occurred in sex (79%), personal (77%), and walking (71%) (Table 2).

outcome scores and was associated with reduced compliance in return of the questionnaires.

Collecting data in a private practice setting without the aid of research facilities or staff presents administrative and financial challenges and in this case required extra staff to be employed. The greatest difficulty was collecting data at discharge (52.9% questionnaire return rate) and follow up (29.9%). This potentially selects out those patients with favourable outcomes and limits interpretation of results. In a comparative study at the La Trobe University Physiotherapy School, 51% of the initial 207 participants returned the follow up questionnaire at 6 weeks following discharge (Davidson and Keating 2002). Varying periods of follow up between 1 week and 2 years are used in research (Fairbank et al 2005, Fritz et al 2001, Hagg et al 2002, Hagg et al 2003, Walsh et al 2003). Although the follow up questionnaire was scheduled to occur at 8 weeks, in actuality this varied up to 16 weeks, depending on the work-load of the administrative person.

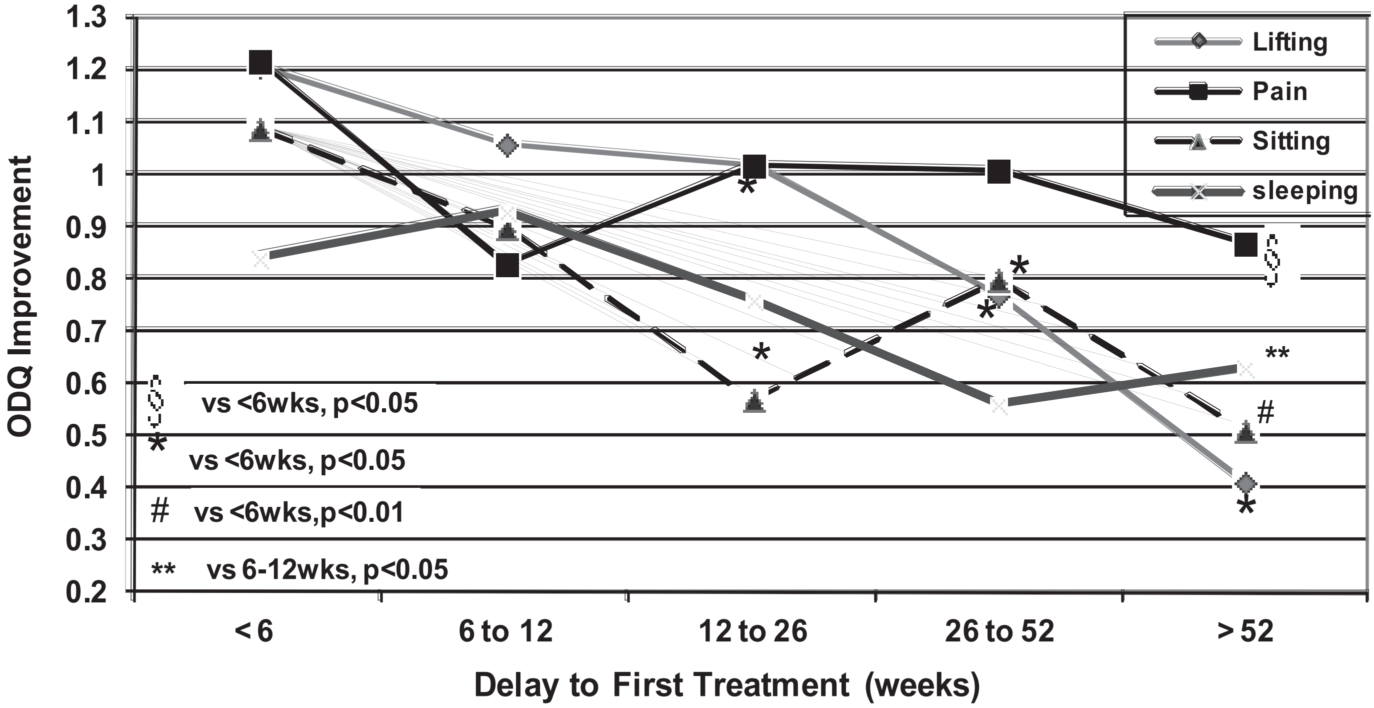
Table 2: Average Scores by individual category (n=127)

Category	Lift	Pain	Personal	Sex	Sitting	Sleep	Social	Standing	Travel	Walk
Baseline Score	2.0	1.7	1.2	1.6	2.1	1.4	1.9	1.7	1.6	1.6
Follow Up Score	0.9	0.7	0.3	0.3	1.0	0.6	0.5	0.8	0.6	0.4
Improvement (%) (SD)	57 (26)	58 (19)	77 (18)	79 (22)	53 (22)	58 (17)	69 (20)	56 (22)	60 (20)	71 (18)

INJURY CHRONICITY AND OUTCOMES

Delays to initial treatment impacted the overall ability to improve the ODQ score at discharge (p=0.013 for overall trend). This was most apparent in the areas of sitting, lifting, pain and sleeping (Figure 4).

Figure 4: Impact of Delay to Treatment on ODQ Outcomes



DISCUSSION

The most significant findings of this study were that delay between injury and treatment affected treatment numbers and

Questionnaire return rate also appeared to be influenced by injury 'interval' (Figure 2) and may have impacted on the accuracy of the outcome measure due to patient recall (Hagg et al 2002, Kamper et al 2009). The scores at discharge and follow up; however, were remarkably similar, suggesting that the improvement in those who did return the questionnaires was maintained despite varying collection intervals. Many studies take only two measurements, baseline and follow up; however, three data points are recommended, preferably one away from the direct influence of the therapist for improved accuracy and reproducibility (Kendal 1997).

INTERPRETING THE SCORE

Calculations are unique to each population studied (Beaton 2000). Davidson and Keating (2002), in their study of 106 low back pain ambulatory patients receiving physiotherapy, found the Minimal Detectable Change (MDC) value to be 10.5% at minimum and up to 15% when the baseline scores were higher. It is recognised that the less disabled the population the lower the MDC value (Beaton 2000, Stratford et al 1998). Our study has shown a value of 5.5 in this population (or 11%), consistent with the lower level of baseline disability. This score was met at discharge and maintained at follow up. However, a significant MDC value does not necessarily mean the change was clinically relevant and important to patients (Beaton 2000, Stratford et al 1998). This necessitated the calculation of the Minimal Important Clinical Difference (MICD), which has been defined as the 'smallest change that is important to patients' (Fritz 2001, Lauridson et al 2006, Stratford et al 1998). Calculating the MICD requires a gold standard question in the form of a global questionnaire to allow calculation of sensitivity and specificity (Lauridson et al 2006, Riddle et al 1998, Stratford et al 1998, Walsh et al 2003). A simultaneous global questionnaire was not administered in this study. Due to these limitations, no absolute conclusions can be drawn regarding the efficacy of the treatment provided in this study.

The ODQ has generally been scored across all questions, however this study looked at the results in each individual question in order to target specific patient needs. Hagg et al (2003, p. 13) state that "individual items of a multi item outcome measure cannot be expected to change to the same magnitude". The overall score may dilute a significant improvement in the presenting complaint if there was minimal change in other less relevant areas. Reviewing the individual category scores demonstrated that the greater the baseline score, the smaller the change score, which is consistent with other studies indicating that areas most affected by chronic pain were most resistant to change (Table 4) (Beaton 2000, Hagg et al 2003). Treatment plans were directly affected by the baseline scores and treatment was targeted to improve functional techniques, particularly in the areas of lifting, sitting, social activities and standing posture. One way of improving the applicability of the score to the individual patient might be to allow the patient to weight the individual categories according to their significance for that patient, as has been implemented in the Lower Limb Task Questionnaire (McNair et al 2007). Applying an MICD calculation to individual categories may further improve relevance. With further research in this area, a standardised mechanism of calculating change could be developed.

IMPACT OF DELAY TO TREATMENT

The 'interval' affected the overall scores revealing that the greater the time between injury and treatment the harder it was to make a difference. Correspondingly, there was an increase in the number of treatments required for those with pain of over 3 months duration. This is in line with other studies showing earlier intervention (within 3 months), improves outcomes with fewer treatments (CAHE 2009) and is endorsed in the Acute Low Back Pain Guidelines (ACC 2001). This was also evident in the individual functional categories of lifting, sitting and sleeping (Figure 4).

CONCLUSION

The practical challenge of collecting and maintaining outcome statistics impacts on the interpretation of the measure. Poor questionnaire return rate and administrative errors reduced the validity of the data. Improving the return rate of questionnaires remains a significant challenge in private practice and this requires more resources in the form of increased time allocation, staffing and finance. Analysing and interpreting outcomes in this clinical audit altered the procedures of administration at this practice and a global measure has now been included. Despite the limitations outlined, this study does support the benefit of earlier intervention for back complaints both in reducing the number of treatments required, and thus cost involved, and improving the magnitude of symptom relief gained.

Centralising and standardising the outcome measure assessment for the country would reduce the individual costs by clinicians and provide a more robust medium for auditing private practice treatment in New Zealand.

KEY POINTS

- Collecting and maintaining an outcome measure database in private practice presents administrative challenges which impact on the outcome measure's validity.
- Delay of treatment effects treatment numbers, outcome measure scores, and return of outcome measure forms.
- Calculating a significant improvement is unique to the patient population and requires the addition of a global questionnaire to enable calculation of an MICD.

ACKNOWLEDGEMENTS

I wish to acknowledge the help my husband Dr Chris Nunn gave me with the statistics.

FUNDING

This study was self funded.

ADDRESS FOR CORRESPONDENCE

Nicola Nunn, Principal Physiotherapist, The Back Clinic, 30 Maeroa Rd, Hamilton, New Zealand. Phone 64 7 839 1166; Fax 64 7 839 1686. Email; nicky@thebackclinic.co.nz

REFERENCES

ACC News (2009): www.acc.co.nz [accessed November 27, 2011]

ACC (2001): New Zealand Acute Low Back Pain Guidelines. http://www.acc.co.nz/PRD_EXT_CSMP/groups/external_ip/documents/internet/wcm002131.pdf [Accessed January 16 2012]

Beaton D (2000): Understanding the relevance of measured change through studies of responsiveness. Spine 25, 24: 3192-3199

Beattie P and Maher C (1997): The role of functional status questionnaires for low back pain *Australian Journal of Physiotherapy* 43, 1: 29-36

Bogduk N (2004): Management of chronic low back pain. Medical Journal of Australia 180:79-83

Bombardier C (2000): Outcome assessments in the evaluation of treatment of spinal disorders. summary and general recommendations. *Spine* 25, 24: 3100-31

Campbell H, Rivero-Arias O, Johnson K, Gray A, Fairbank J and Frost H (2006): Responsiveness of objective, disease-specific and generic outcome measures in patients with chronic low back pain: an assessment of improving stable and deteriorating patients. *Spine* 31, 7: 815-822

Copeland J, Taylor W and Dean S (2008): Factors influencing the use of outcome measures for patients with low back pain: a survey of physical therapists. *Physical Therapy* 88, 12: 1-14.

Davidson M and Keating J (2002): A comparison of five low back disability questionnaires: reliability and responsiveness. *Physical Therapy* 82: 1, 8-24

Deyo RA (1988): Measuring the functional status of patients with low back pain. *Archives of Physical Medicine and Rehabilitation* 69: 1044-53

Dworkin RH, Turk DC, Farrar J T, Haythornthwaite J, Jensen MP, Katz NP, et al (2005): Core outcome measures for chronic pain clinical trial: IMMPACT recommendations. *Pain* 113: 9-19

Evaluating Best Practice for Purchasing Strategies for Physiotherapy Services for Injury Related Conditions (2009): Centre for Allied Health Evidence (CAHE) Report for ACC, University of South Australia. http://www.acc.co.nz/search-results/index.htm?ssUserText=outcome+measures. [Accessed November 25, 2011]

Fairbank J, Frost H, Wilson-MacDonald J, Ly-Mee Yu, Barker K and Collins R (2005): Randomised controlled trial to compare surgical stabilization of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: The MRC Spine Stabilisation Trial. *British Medical Journal* 28: 330

Fairbanks J, Davies J, Couper J, and O Brien J (1980): The Oswestry low back pain disability questionnaire. *Physiotherapy* 66, 8: 271-273

Fritz J and Irrgang J (2001): A comparison of a modified Oswestry Low Back Pain Disability Questionnaire and the Quebec pain Disability Scale. *Physical Therapy* 81, 2: 776-787

Frost H, Lamb K and Brown S (2008): Responsiveness of a Patient Specific Outcome Measure compared with the Oswestry Disability Index v2.1 and Roland Morris Disability Questionnaire for patients with sub acute and chronic low back pain. *Spine* 33, 22: 2450-2457

Hagg O, Fritzell P, and Nordwall A (2003): The clinical importance of changes in outcome scores after treatment for chronic low back pain. *European Spine Journal* 12: 12-20.

Hagg O, Fritzell P, Odén A and Nordwall A (2002): Simplifying outcome measurement. *Spine* 27, 11: 1213-1222

Herbert R, Jamtvedt G, Mead J, Birger and Hagen K. (2005): Outcome measures measure outcomes not effects of intervention. *Australian Journal of Physiotherapy* 51: 3

Kamper S, Maher C and MacKay G (2009): Global rating of change scales: a review of the strengths and weaknesses and considerations for design. *Journal of Manual & Manipulative Therapy* 17: 3

Kendal N (1997): Developing outcome assessments: a step by step approach *New Zealand Journal of Physiotherapy* 25, 3: 11-17

Lauridson H, Hartvigsen J, Mannich C, Korsholm L and Grunnet-Nilsson N (2006): Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain sufferers. *BMC Musculoskeletal Disorders* 7: 82.

McNair PJ, Prapavessis H, Collier J, Bassett S F, Bryant A, and Larmer P J (2007): The lower-limb tasks questionnaire: an assessment of validity, reliability, responsiveness, and minimal important differences. *Archives of Physical Medicine and Rehabilitation* 88, 8: 993-1001

PPA Recommendations for low back pain related functional limitation outcome measures. (2004) The Chartered Society of Physiotherapy www.csp.org.uk July 2004. [Accessed May 10, 2010].

Riddle D, Stratford PW and Binkley J (1998): Sensitivity to change of the Roland Morris back pain questionnaire: Part 2. *Physical Therapy* 78: 11

Stratford P, Binkley J, Riddle D and Guyatt G (1998): Sensitivity to change of the Roland Morris Back Pain Questionnaire. Part 1. *Physical Therapy* 78: 11.

Stratford P, Binkley J, Solomon P, Gill C and Finch E. (1994): Assessing change over time in patients with low back pain. *Physical Therapy* 74, 6: 528-533

Walsh T, Hanscom B, Lurie J, and Weinstein J (2003): Is a condition-specific instrument for patients with low back pain/ leg symptoms really necessary? *Spine* 28, 6: 607-615

Wessels T, van Tulder M, Sigl T, Ewert T, Limm H, and Stucki G. (2006): What predicts outcome in non- operative treatment of chronic low back pain? A systematic review." *European Spine Journal* 15: 1633-1644

Appendix 1: Calculating the Minimal Detectable Change

The Minimal Detectable Change (90% CI of the error associated with the repeated measurements) for this population was calculated using the formulas described by Davidson and Keating (2002) and Beaton(2000):

$SEM = SD_{ave} \sqrt{1-R} = 1.0\sqrt{1-0.84} = 0.4$

Where SD was the average standard deviation of the scores initially and R is the test re-test reliability co-efficient. The reliability co-efficient was not calculated directly from the population in this study as there was no concurrent global questionnaire to use for this calculation. The reliability co-efficient value of 0.84 was therefore chosen from the Davidson et al (2002) study as the population and time frames were similar to the present study. The error associated with three measurements was calculated with the following formula:

$SEM\ repeat = \sqrt{3}x\ SEM = 3.4$

The 95% CI (the MDC) was calculated by multiplying the result by 1.64 (the tabled z value). MDC value = 5.5

(This z value is sometimes changed by researchers and the alternative z value of 1.96 used (Hagg et al 2003) giving an alternative MDC calculation of 6.6.)