

# Oswestry Disability Questionnaire (ODQ)

## Description

The Oswestry Disability Questionnaire (ODQ) was developed to measure disability in people with low back pain (Fairbank et al 1980). The ODQ has ten sections: pain, personal care, lifting, walking, sitting, standing, sleeping, sex life, social life, and travelling. Modified versions have deleted references to 'pain killers' and 'tablets' from the Pain and Sleeping sections. In one version the 'Sex Life' section has been replaced by 'Changing Degree of Pain' (Hudson-Cook et al 1989) and in another by 'Employment/Homemaking' (Fritz and Irrgang 2001). The version used by Davidson and Keating (2002) replaces miles with kilometres in the Walking section. The developers recommend Version 2.0 of the Oswestry (Fairbank and Pynsent 2000), which instructs patients to answer the questions in relation to how their back problem is affecting them 'today', rather than the original instructions which do not specify a time-frame.

**Instructions to the client and scoring** It takes 5 minutes for the patient to complete the ODQ and less than one minute to score. Respondents choose one of six statements that most applies to them in each section. The first statement is scored 0, the second is scored 1, and so on to 5 for the sixth statement. The sum of the section scores is transformed to a percentage score, adjusted for missed sections. The total

possible score ranges 0–100 and a higher score indicates worse function. Scores from 0–20% are claimed to indicate 'minimal disability', 20–40% 'moderate disability', 40–60% 'severe disability', 60–80% 'crippled', and 80–100% 'bedbound or exaggerating' (Fairbank et al 1980). The term 'housebound' is suggested as a more appropriate contemporary descriptor for the 60–80% score range.

**Reliability, Validity and Sensitivity to Change** Reliability coefficients ranging 0.83 to 0.99 have been reported in 11 studies of test-retest reliability. The standard error of measurement ranges from 4.5 to 6 points. There is a large body of evidence to support the construct validity of the ODQ and the instrument is able to detect statistically significant and clinically important change over time in group data.

**Comparison with alternative questionnaires** A Canadian study reported similar reliability and responsiveness of the Oswestry, the Quebec, and the Roland-Morris scales, with all these superior to the SF-36 Physical Functioning Scale (Kopec et al 1995). An Australian study reported similar responsiveness of the Oswestry, Quebec, Roland-Morris, Waddell, and SF-36 Physical Functioning scale, but superior reliability of the Oswestry, Quebec, and SF-36 Physical Functioning scales (Davidson and Keating 2002).

## Commentary

Although it was developed 25 years ago the ODQ remains one of the most widely used low back disability questionnaires. It is easy for both patients and clinicians to use and its clinimetric properties are well established. Australian physiotherapy data suggest that a change over a 6-week period of 10 points allows one to be 90% confident that the observed change is beyond measurement error, and a change of between 5 and 9 points is considered by patients to be clinically important (Davidson and Keating 2002). Change of less than 10 points should not be entirely discounted; however, the smaller the change in score the greater the likelihood that variation in scores is due to measurement error. For individual patients the change in section scores as well as in total scores should be inspected to determine which activities are improving. When the initial score is less than 10 points there is insufficient range remaining to detect improvement in scores beyond measurement error (at 90% confidence) and another instrument should be selected for patient assessment. A very high score (>80%) in ambulatory patients is rare and clinicians should explore personal and environmental factors that may magnify patient self-reporting of disability.

There is some evidence to suggest that when the initial score

on health status measures is very high, a relatively greater amount of score change may be required before people typically perceive the change to be clinically important, and when the initial score is very low, very small changes may be perceived as important. Oswestry scores are ordinal data and one therefore cannot say, for example, that a person with a score of 60 is twice as disabled as someone with a score of 30, or that a change of 10 points at different points of the scale are necessarily equivalent.

There is no evidence that any particular modified version of the ODQ is superior to another. However, the version with 'Changing degree of pain' is not recommended because this is a question about a change in health status while the other questions are about current health status. A limitation of the Oswestry is that, if redirecting patient attention from pain is an important treatment goal, the pain-focused language may be counterproductive.

**Megan Davidson**

*La Trobe University*

**Jennifer Keating**

*Monash University*

## References

- Davidson M and Keating JL (2002): *Phys Ther* 82: 8–24.  
 Fairbank JC et al (1980): *Physiotherapy* 66: 271–273.  
 Fairbank JC and Pynsent PB (2000): *Spine* 25: 2940–2952; Discussion 2952.  
 Fritz JM and Irrgang JJ (2001): *Phys Ther* 81: 776–788.  
 Hudson-Cook N et al (1989): In Roland M and Jenner JR (Eds.) *Back Pain: New Approaches to Rehabilitation and Education*. Manchester: Manchester University Press, pp. 187–204.  
 Kopec JA et al (1995): The Quebec Back Pain Disability Scale. *Spine* 20: 341–352.

# The Neck Disability Index (NDI)

## Description

The Neck Disability Index (NDI) is a ten-item questionnaire based on the Oswestry Low Back Pain Index that assesses disability associated with neck pain and whiplash (Vernon and Mior 1991). There are four items that relate to subjective symptomatology (pain intensity, headache, concentration, sleeping) and six items that relate to activities of daily living (lifting, work, driving, recreation, personal care, reading).

The questionnaire requires only 5–10 minutes to complete and score, and requires no special training to administer.

**Instructions to the client and scoring** Clients select from one of six potential responses for each item ranging from no disability (0) to total disability (5). The ten items are summed to gain the total score thus ranging from 0 (no disability) to 50 (maximum disability). Some authors convert this score to a percentage.

Vernon and Mior (1991) propose that a score of less than 4

indicates no disability, 5–14 mild disability, 15–24 moderate disability, 25–34 severe disability, and scores greater than 35 complete disability.

**Reliability and validity** Test-retest reliability has been found to be high over a two day period in whiplash injured participants (Pearson's  $r = 0.89$ ,  $p < 0.05$ ) (Vernon and Mior 1991). Ackelman et al (2002) have demonstrated similarly high test-retest reliability between time periods of three weeks and three months (Pearson's  $r = 0.95$  and  $0.94$  respectively) in a cohort of chronic neck pain participants.

The NDI has been shown to correlate with visual analogue scores of pain (correlation coefficient 0.6) and with the McGill Pain Questionnaire (correlation coefficient 0.7) (Pietroben et al 2002).

**Michele Sterling**

*The University of Queensland*

## Commentary

With the increasing requirement for physiotherapists to provide measures of clinical outcome that reflect function, the NDI provides a useful, reliable and valid way to do this for patients with neck pain. The NDI is sensitive to change and correlates well with measures such as the Visual Analogue Scale, the McGill Pain Questionnaire (Vernon and Mior 1991), and the General Health Questionnaire (Lankester et al 2004). Clinicians can consider a 10% change in the NDI (or a least 1 point improvement in 5 out of 10 items) to be clinically relevant (Stratford et al 1999).

Comparisons of the treatment effect between studies can be made when the NDI is used consistently as an outcome measure. Ylinen et al (2003) found that specific neck strengthening exercises reduced NDI scores of chronic neck pain patients by 9 points over 12 months when compared to a non-significant 4 point change over 12 months in the control group comprising instruction on aerobic and stretching exercises. In comparison, Evans et al (2002) demonstrated that specific neck strengthening reduced NDI scores by 10.7 points at 2 years compared with spinal manipulative therapy alone (7.4 point change). Both authors concluded that neck strengthening exercises were more likely to provide functional outcomes for neck pain patients, using the NDI as the measure.

The NDI may also be used to predict poor outcome in Whiplash Associated Disorders (WAD). High initial scores on the NDI together with other factors have been shown to

predict those who would not recover by 6 months after a whiplash injury (Sterling et al 2005). Similarly, Miettinen et al (2004) found a baseline NDI score of  $> 20$  to be significantly related to poor outcome for WAD at 3 years ( $p < 0.01$ ). Hence the administration of the NDI at baseline may assist clinicians and researchers in identifying whiplash sufferers with a poorer prognosis.

Clinicians should be aware that there are some limitations of the NDI. These include the tendency for a ceiling effect, where patients who are very sick may reach a maximal score, hence reducing the ability to detect further decline in function (Riddle and Stratford 1998). Furthermore, Hoving et al (2003) concluded that the NDI has limitations in assessing disability in WAD, partly because it does not measure emotional and social functioning, considered important in this group. Therefore, clinical situations in which the NDI may not reflect the full impact of the disease may include the 'very sick' and in WAD.

In summary the NDI is a reliable, comprehensively validated (Pietroben et al 2002), and clinically useful tool to measure disability due to neck pain. Clinicians may use the NDI to evaluate functional outcomes in patients with neck pain and whiplash, expecting that a change of at least 10% in the score would be clinically relevant.

**Trudy Rebbeck**

*The University of Sydney*

## References

- Ackelman E et al (2002): *J Rehabil Med* 34: 284–287.  
 Evans R et al (2002): *Spine* 27: 2383–2389.  
 Hains F et al (1998): *J Manipulative Physiol Ther* 21: 75–80.  
 Hoving J et al (2003): *Pain* 102: 273–281.  
 Lankester B et al (2004): *Eur Spine J* 13: 605–609.  
 Miettinen T et al (2004): *Spine* 29: E47–51.  
 Pietroben R et al (2002): *Spine* 27: 515–522.  
 Riddle DL and Stratford PW (1998): *Phys Ther* 9: 951–963.  
 Sterling M et al (2005): *Pain* 114: 141–148.  
 Stratford P et al (1999): *Physiother Can* 51: 107–112.  
 Vernon H and Mior S (1991): *J Manipulative Physiol Ther* 14: 409–415.  
 Ylinen J et al (2003): *JAMA* 289: 2509–2516.