The Oswestry Disability Index

Jeremy C. T. Fairbank, MD, FRCS,* and Paul B. Pynsent, PhD†

Study Design. The Oswestry Disability Index (ODI) has become one of the principal condition-specific outcome measures used in the management of spinal disorders. This review is based on publications using the ODI identified from the authors’ personal databases, the Science Citation Index, and hand searches of Spine and current textbooks of spinal disorders.

Objectives. To review the versions of this instrument, document methods by which it has been validated, collate data from scores found in normal and back pain populations, provide curves for power calculations in studies using the ODI, and maintain the ODI as a gold standard outcome measure.

Summary of Background Data. It has now been 20 years since its original publication. More than 200 citations exist in the Science Citation Index. The authors have a large correspondence file relating to the ODI, that is cited in most of the large textbooks related to spinal disorders.

Methods. All the published versions of the questionnaire were identified. A systematic review of this literature was made. The various reports of validation were collated and related to a version.

Results. Four versions of the ODI are available in English and nine in other languages. Some published versions contain misprints, and many omit the scoring system. At least 114 studies contain usable data. These data provide both validation and standards for other users and indicate the power of the instrument for detecting change in sample populations.

Conclusions. The ODI remains a valid and vigorous measure and has been a worthwhile outcome measure. The process of using the ODI is reviewed and should be the subject of further research. The receiver operating characteristics should be explored in a population with higher self-report reliability. The behavior of the instrument is incompletely understood, particularly in sensitivity to real change. [Key words: back pain, Oswestry Disability Index, outcome measures, validity] Spine 2000;25:2940–2953

The Oswestry Disability Index (ODI) and the Roland–Morris disability questionnaire (R-M) have emerged as the most commonly recommended condition-specific outcome measures for spinal disorders.122,28,34,153

The development of the Oswestry Disability Index was initiated by John O’Brien in 1976. Patients with back pain were interviewed by an orthopedic surgeon (Stephen Eisenstein), and an occupational therapist (Judith Couper). Various drafts of the questionnaire were tried. The questionnaire had been published in 198038 and widely disseminated from the 1981 meeting of the International Society for the Study of the Lumbar Spine (ISSLS) in Paris.

The objects of this article are:
- To present the various versions of the ODI instrument for comparison
- To review the various efforts that have been made to validate this questionnaire
- To compare the scores obtained in studies of different patient population both before and after treatment
- To review the methodology of outcome measurement
- To consider what is actually measured by this and similar instruments

Search Methodology

Citations were identified from the authors’ personal databases, the Science Citation Index (searching for the original reference38), and hand searches of Spine and current textbooks of spinal disorders. There are well over 200 citations of the ODI in the Science Citation Index alone.

Versions of the ODI

Table 1 shows four versions of the ODI. Version 1.0 is the original,38 reproduced by Hupli et al66 (with a scoring system) and Boden6 without one. It has also been published omitting a single item from both section 8 (sex life) and section 9 (social life).8 The American Academy of Orthopedic Surgeons (AAOS) and other spine societies have adapted version 1.0 into their spine outcome instruments. This version reflects American rather than British usage. It omits sections 1, 8, and 9. It scores the remaining sections from 1 to 6 (rather than 0–5), which leads to confusion when comparing scores obtained with other versions.37,28

Version 2.0 was a modification of the ODI made by a Medical Research Council group in the United Kingdom.1,104,105,116 It has been widely distributed by correspondence and is available as part of a computer interview in the United Kingdom (slightly modified)1,117 or in the United States (through MODEMS; available at PO Box 2354, Des Plaines, IL 60017-2354).

A revised Oswestry Disability Questionnaire was published by a chiropractic study group in the United Kingdom in 1989.62 Its objective was to increase the sensitivity of the scale for less disabled patients, but it confuses impairment with disability. The sex question is omitted.62,91,168 In the authors’ view, this version is not
Table 1.

<table>
<thead>
<tr>
<th>ODI Version 1.0</th>
<th>ODI Version 2.0</th>
<th>AAOS/MODEMS</th>
<th>Chiropractic “Revised Oswestry pain questionnaire”</th>
</tr>
</thead>
<tbody>
<tr>
<td>This questionnaire has been designed to give the doctor information as to how your back pain has affected your ability to manage in every day life. Please answer every section, and mark in each section only the one box which applies to you. We realize you may consider that two of the statements in any one section relate to you, but please just mark the box which most closely describes your problem.</td>
<td>Could you please complete this questionnaire? It is designed to give us information as to how your back (or leg) trouble has affected your ability to manage in everyday life. Please answer every section. Mark one box only in each section that most closely describes you today.</td>
<td>In the past week, please tell us how pain has affected your ability to perform the following activities. (Circle the one statement that best describes your average ability.)</td>
<td>Please read: This Questionnaire It is designed to enable us to understand how much your low back pain has affected your ability to manage your everyday activities.</td>
</tr>
<tr>
<td>Section 1—Pain intensity</td>
<td>Section 1—Pain intensity</td>
<td>Section 1—Pain intensity</td>
<td>Section 1—Pain intensity</td>
</tr>
<tr>
<td>I can tolerate the pain I have without having to use painkillers.</td>
<td>I have no pain at the moment.</td>
<td>The pain is very mild at the moment.</td>
<td>The pain comes and goes and is very mild.</td>
</tr>
<tr>
<td>The pain is bad but I manage without taking painkillers.</td>
<td>The pain is moderate at the moment.</td>
<td>The pain is fairly severe at the moment.</td>
<td>The pain comes and goes and is moderate.</td>
</tr>
<tr>
<td>Painkillers give complete relief from pain.</td>
<td>The pain is very severe at the moment.</td>
<td>The pain is the worst imaginable at the moment.</td>
<td>The pain is moderate and does not vary much.</td>
</tr>
<tr>
<td>Painkillers give moderate relief from pain.</td>
<td>Painkillers give very little relief from pain.</td>
<td>Painkillers have no effect on the pain and I do not use them.</td>
<td>Painkillers have no effect on the pain and I do not use them.</td>
</tr>
<tr>
<td>Section 2—Personal care (washing, dressing, etc.)</td>
<td>Section 2—Personal care (washing, dressing, etc.)</td>
<td>Section 2—Personal care</td>
<td>Section 2—Personal care</td>
</tr>
<tr>
<td>I can look after myself normally without causing extra pain.</td>
<td>I can look after myself normally without causing extra pain.</td>
<td>I can look after myself normally.</td>
<td>I would not have to change my way of washing or dressing in order to avoid pain.</td>
</tr>
<tr>
<td>I can look after myself normally but it causes extra pain.</td>
<td>I can look after myself normally but it is very painful.</td>
<td>I can look after myself but it is painful.</td>
<td>I do not normally change my way of doing it.</td>
</tr>
<tr>
<td>It is painful to look after myself and I am slow and careful.</td>
<td>It is painful to look after myself and I am slow and careful.</td>
<td>I find it necessary to change my way of doing it.</td>
<td>I find it necessary to change my way of doing it.</td>
</tr>
<tr>
<td>I need some help but manage most of my personal care.</td>
<td>I need some help but manage most of my personal care.</td>
<td>I need some help but manage most of my personal care.</td>
<td>I need some help but manage most of my personal care.</td>
</tr>
<tr>
<td>I need help every day in most aspects of self-care.</td>
<td>I need help every day in most aspects of self-care.</td>
<td>I need help every day in most aspects of self-care.</td>
<td>I need help every day in most aspects of self-care.</td>
</tr>
<tr>
<td>I do not get dressed, wash with difficulty and stay in bed.</td>
<td>I do not get dressed, wash with difficulty and stay in bed.</td>
<td>I do not get dressed, wash with difficulty and stay in bed.</td>
<td>I do not get dressed, wash with difficulty and stay in bed.</td>
</tr>
<tr>
<td>Section 3—Lifting</td>
<td>Section 3—Lifting</td>
<td>Section 3—Lifting</td>
<td>Section 3—Lifting</td>
</tr>
<tr>
<td>I can lift heavy weights without extra pain.</td>
<td>I can lift heavy weights without extra pain.</td>
<td>I can lift heavy weights without extra pain.</td>
<td>I can lift heavy weights without extra pain.</td>
</tr>
<tr>
<td>I can lift heavy weights but it gives extra pain.</td>
<td>I can lift heavy objects but it is painful.</td>
<td>I can lift heavy objects but it is painful.</td>
<td>I can lift heavy weights but it gives extra pain.</td>
</tr>
<tr>
<td>Pain prevents me from lifting heavy weights off the floor, but I can manage if they are conveniently positioned, e.g. on a table.</td>
<td>Pain prevents me from lifting heavy objects off the floor, but I can lift heavy objects if they are on a table.</td>
<td>Pain prevents me from lifting heavy objects off the floor, but I can lift heavy objects if they are on a table.</td>
<td>Pain prevents me from lifting heavy weights off the floor.</td>
</tr>
<tr>
<td>Pain prevents me from lifting heavy weights but I can manage light to medium weights if they are conveniently positioned.</td>
<td>Pain prevents me from lifting heavy objects but I can lift medium weights if they are on a table.</td>
<td>Pain prevents me from lifting heavy objects but I can manage light to medium weights if they are on a table.</td>
<td>Pain prevents me from lifting heavy objects off the floor but I can manage if they are conveniently positioned, e.g. on a table.</td>
</tr>
<tr>
<td>I can lift only very light weights.</td>
<td>I can lift only light objects.</td>
<td>I can only lift light objects.</td>
<td>I can only lift light objects.</td>
</tr>
<tr>
<td>I cannot lift or carry anything at all.</td>
<td>I cannot lift anything.</td>
<td>I cannot lift anything.</td>
<td>I cannot lift anything.</td>
</tr>
</tbody>
</table>

The Oswestry Disability Index: Fairbank and Pynsent 2941
Table 1. Continued

<table>
<thead>
<tr>
<th>ODI Version 1.0</th>
<th>ODI Version 2.0</th>
<th>AAOS/MODEMS</th>
<th>Chiropractic “Revised Oswestry pain questionnaire”</th>
</tr>
</thead>
</table>

Section 4—Walking
Pain does not prevent my walking any distance.
Pain prevents me walking more than 1 mile.
Pain prevents me walking more than 1/2 mile.
Pain prevents me walking more than 1/4 mile.
I only walk using a stick or crutches.
I am in bed most of the time and have to crawl to the toilet.

Section 5—Sitting
I can sit in any chair as long as I like.
I can sit in my favourite chair as long as I like.
Pain prevents me sitting for more than 1 hour.
Pain prevents me sitting more than 1/2 an hour.
Pain prevents me sitting for more than 10 minutes.
Pain prevents me from sitting at all.

Section 6—Standing
I can stand as long as I want without extra pain.
I can stand as long as I want but it gives me extra pain.
Pain prevents me from standing for more than 30 minutes.
Pain prevents me from standing for more than 10 minutes.
Pain prevents me from standing at all.

Section 7—Sleeping
Pain does not prevent me from sleeping well.
I can sleep well only by using tablets.
Even when I take tablets I have less than 6 hours sleep.
Even when I take tablets I have less than 4 hours sleep.
I am in bed most of the time and have to crawl to the toilet.

Section 4—Walking
I can run or walk without pain.
I can walk comfortably, but running is painful.
Pain prevents me from walking more than 1 hour.
Pain prevents me from walking more than 30 minutes.
Pain prevents me from walking more than 10 minutes.
I am unable to walk or can walk only a few steps at a time.

Section 4—Walking and Running (in the past week).
I can run or walk without pain.
I can walk comfortably, but running is painful.
Pain prevents me from walking more than 1 hour.
Pain prevents me from walking more than 30 minutes.
Pain prevents me from walking more than 10 minutes.
I am unable to walk or can walk only a few steps at a time.

Section 4—Walking
I have no pain on walking.
I have some pain with walking but it does not increase with distance.
I cannot walk more than One Mile without increasing pain.
I cannot walk more than 1/2 Mile without increasing pain.
I cannot walk at all without increasing pain.

Section 5—Sitting
I can sit in any chair as long as I like.
I can sit only in my favourite chair as long as I like.
Pain prevents me from sitting for more than 1 hour.
Pain prevents me from sitting for more than 30 minutes.
Pain prevents me from sitting for more than 10 minutes.
Pain prevents me from sitting at all.

Section 7—Sleeping
I sleep well.
Pain occasionally interrupts my sleep.
Pain interrupts my sleep half of the time.
I never sleep well.

Section 4—Walking
I have some pain with walking but it does not increase with distance.
I cannot walk more than One Mile without increasing pain.
I cannot walk more than 1/2 Mile without increasing pain.
I cannot walk at all without increasing pain.

Section 5—Sitting
I can sit in any chair as long as I like.
I can sit only in my favourite chair as long as I like.
Pain prevents me from sitting for more than 1 hour.
Pain prevents me from sitting for more than 1/2 an hour.
Pain prevents me from sitting for more than 10 minutes.
Pain prevents me from sitting at all.

Section 7—Sleeping
I get no pain in bed.
I get pain in bed but it does not prevent me from sleeping well.
Because of pain my normal nights sleep is reduced by less than 1/2.
Because of pain my normal nights sleep is reduced by less than 3/4.
Pain prevents (me) from sleeping at all.
<table>
<thead>
<tr>
<th>ODI Version 1.0</th>
<th>ODI Version 2.0</th>
<th>AAOS/MODEMS</th>
<th>Chiropractic “Revised Oswestry pain questionnaire”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Section 8—Sex life</strong></td>
<td><strong>Section 8—Sex life (if applicable)</strong></td>
<td><strong>Section 8—Social life</strong></td>
<td><strong>Section 8—Social life</strong></td>
</tr>
<tr>
<td>My sex life is normal and causes no extra pain. My sex life is normal but causes some extra pain. My sex life is nearly normal but is very painful. My sex life is severely restricted by pain. My sex life is nearly absent because of pain. Pain prevents any sex life at all.</td>
<td>My sex life is normal and causes no extra pain. My sex life is normal but causes some extra pain. My sex life is nearly normal but is very painful. My sex life is severely restricted by pain. My sex life is nearly absent because of pain. Pain prevents any sex life at all.</td>
<td>My social life is normal and gives me no extra pain. My social life is normal but increases the degree of pain. Pain has no significant effect on my social life apart from limiting my more energetic interests, e.g. dancing, etc. Pain has restricted my social life and I do not go out as often. Pain has restricted social life to my home. I have no social life because of pain.</td>
<td><strong>Section 8—Social life</strong> My social life is normal and gives me no pain. My social life is normal but increases the degree of my pain. Pain has no significant effect on my social life apart from limiting my more energetic interests, e.g. dancing, etc. Pain has restricted my social life and I do not go out very often. Pain has restricted social life to my home. I have hardly any social life because of the pain.</td>
</tr>
<tr>
<td><strong>Section 9—Social life</strong></td>
<td><strong>Section 9—Social life</strong></td>
<td><strong>Section 9—Travelling</strong></td>
<td><strong>Section 9—Travelling</strong></td>
</tr>
<tr>
<td>My social life is normal and gives me no extra pain. My social life is normal but increases the degree of pain. Pain has no significant effect on my social life apart from limiting my more energetic interests, e.g. dancing, etc. Pain has restricted my social life and I do not go out as often. Pain has restricted social life to my home. I have no social life because of pain.</td>
<td>My social life is normal and causes me no extra pain. My social life is normal but increases the degree of pain. Pain has no significant effect on my social life apart from limiting my more energetic interests, e.g. sport, etc. Pain has restricted my social life and I do not go out as often. Pain has restricted social life to my home. I have no social life because of pain.</td>
<td>I can travel anywhere without extra pain. I can travel anywhere but it gives me pain. Pain is bad but I manage journeys over two hours. Pain restricts me to journeys of less than one hour. Pain prevents me to short necessary journeys under 30 minutes. Pain prevents travel except to the doctor or hospital.</td>
<td>I get no pain whilst travelling. I get some pain whilst travelling but none of my usual sorts of travel make it any worse. I get extra pain whilst travelling but it does not compel me to seek alternative forms of travel. I get extra pain whilst travelling which compels me to seek alternative forms of travel. Pain restricts all forms of travel. Pain prevents all forms of travel except that done lying down.</td>
</tr>
<tr>
<td><strong>Section 10—Travelling</strong></td>
<td><strong>Section 10—Travelling</strong></td>
<td><strong>Section 10—Changing degree of pain</strong></td>
<td><strong>Section 10—Changing degree of pain</strong></td>
</tr>
<tr>
<td>I can travel anywhere without extra pain. I can travel anywhere but it gives me pain. Pain is bad but I manage journeys over two hours. Pain restricts me to journeys of less than one hour. Pain restricts me to short necessary journeys under 30 minutes. Pain prevents me from travelling except to receive treatment.</td>
<td>I can travel anywhere. I can travel anywhere but it gives me pain. Pain is bad but I can manage to travel over two hours. Pain restricts me to journeys of less than one hour. Pain restricts me to trips less than 30 minutes. Pain prevents me from travelling.</td>
<td>My pain is rapidly getting better. My pain fluctuates but overall is definitely getting better. My pain seems to be getting better but improvement is slow at present. My pain is neither getting better or worse. My pain is gradually worsening. My pain is rapidly worsening.</td>
<td>52. Travelling (in the past week): I can travel anywhere. I can travel anywhere but it gives me pain. Pain is bad but I can manage to travel over two hours. Pain restricts me to journeys of less than one hour. Pain restricts me to trips less than 30 minutes. Pain prevents me from travelling.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Continued
have to be made for this. MODEMS now includes both versions 1.0 and 2.0, as well as the AAOS instruments. Some researchers have used versions of the instrument scoring each section from 1 to 6 (notably the Ste. Justine Group44 and the AAOS and North American Spine Society [NASS]). This can make hand scoring more difficult and unreliable (Table 3).37,103 If the first answer of each section is scored zero, then it can be ignored when summing the score. If it is scored 1 then it must be counted and subtracted from the total to calculate the final score. This can and does lead to errors: Orr et al.112 reported to NASS in 1998 a series of 25 spinal fusion patients with a preoperative ODI score of 18.7 ± 5.5 (SD). This is inconsistent with all other publications in Table 4 and Figure 1.

**Definitions of Disability**

The World Health Organization (WHO) definitions of impairment, disability, and handicap are now widely accepted.167 Discussion groups in 1991146 and 199261 reviewed the available outcome measures for patients with back pain. Both groups concluded that the ODI was reasonably confined to disability according to the WHO definition. Many of the alternatives also attempted to measure impairment (pain) and some ranged into the areas of handicap as well. Handicap has been extremely difficult to measure by questionnaire.118 The ODI was deliberately focused on physical activities and not the psychological consequences of acute or chronic pain.

**Validity and Reliability**

In the 20 years since the ODI was published, there have been considerable advances in the understanding of instrument validation.140

**Face and Content Validity.** This means that the scale appears to be assessing the desired qualities. In the author’s original study, a group of 25 patients was reported in the first episode of low back pain who might be reasonably expected to improve with passing time.38 The ODI tracked this process. Beuerskins et al8 performed a more sophisticated analysis of 81 patients during a 5-week period confirming an expected improvement in ODI scores. Their study design allows calculation of an effect size of 0.8.18 However, Kopec et al75 reported an effect size of only 0.07. Fisher and Johnson40 conducted one of the most detailed validations of the questionnaire (version 2.0). They related patient behavior while they were completing this and other questionnaires to their responses within the questionnaires. Two sections of the questionnaires (sitting and walking) correlated with patient response, but correlation was less satisfactory for a third (lifting).

**Test–Retest.** In the original study patients with chronic low back pain were tested twice at a 24-hour interval (n = 22, r = 0.99).38 This may include a memory effect. If the test–retest interval is extended to 4 days, the correlation of scores decreases to n = 22, r = 0.9117 and, if retested after a week, n = 22, r = 0.83.51 The disadvan-
tage of increasing the time interval is that natural symptom fluctuation may also be an influence. Grevitt (personal communication May, 1997) found a poorer test–retest correlation in a study in which he mailed versions of the questionnaire to patients to complete and then asked them to complete the questionnaire again in a different format when they attended as outpatients. The authors compared a hard copy version 2.0 with a computer version of the questionnaire and found a high correlation ($n = 183, r = 0.89$).1

![Figure 1. Weighted means (95% confidence intervals) for the Oswestry Disability Index calculated from pooled data for various categories of patients. The number of patients for each category is also marked. See Table 4 for more information.](image-url)
Time Frame. The original questionnaire did not specify the time frame in which the subject should answer the questionnaire, although it is implicit that it means “now.” Version 2.0 specifically asks about “now.” The NASS modification emphasizes a review of symptoms in the past week. Recently, the authors’ work on outcomes in neurogenic claudication has included the use of version 2.0 of the ODI and the instrument described by Stucki et al.42 using a time frame of a week. Patients prefer the format in which they are asked for symptoms “now.”24

Internal Consistency. Cronbach’s α is a measure of all the various components of a questionnaire moving together. Strong et al.141 (using version 1.0) found Cronbach’s α to be 0.71, Fisher and Johnson40 (using version 2.0) 0.76, and Kopec et al.75 0.87. All these investigations show an acceptable degree of internal consistency. In the original study38 and the 1989 study1 the current authors found that all the sections tended to move with Section 1 (pain scale) although Fisher and Johnson40 did not confirm this in patients who had chronic pain.

Validation by Comparison With Other Tests
The ODI shows moderate correlation with pain measures such as a visual analogue scale (n = 94, r = 0.62)51 and the McGill Pain Questionnaire.53,107 The ODI has been used to validate the Pain Disability Index,51,51,124,141 the Low Back Outcome Score,49 Manniche,99,100 the Aberdeen score,123 a new German language scale,3 the Curtin Scale,57 and a functional capacity evaluation.74 The ODI correlates with the Short Form (SF)36.50 ODI is a better predictor of return to work than two different mechanical methods of lumbar spine assessment94,109 It predicts isokinetic performance,71 isometric endurance,55 and pain with sitting and standing (but not lifting) in a secret observation study.40 In the Mackenzie system of evaluation, “centralizers” show improving ODI scores.143 Physical tests correlate with the ODI54 but range of movement does not.53

The ODI has been used in questions by the designers of other instruments.2,20,21,39,46,49,87,100 The ODI has been used in at least one study of neck pain patients.169

Identification of Patient Populations
The ODI has been used to identify populations of patients for research projects.17,23,43,55,59–61,65,102 but is unlikely that this approach has much clinical application.

Categoric Versus Dimensional Scales
The data gathered in the ODI is in a categoric format, but each category is ordinal. This is converted to discrete quantitative data by summing (a dimensional scale). This assumes that disability can be viewed as a continuum from “not disabled” to “severely disabled.” Many view the ODI as having a linear correlation with disability, and thus a person with a score of 40 is twice as disabled as one with a score of 20. This is unlikely to be true, because the structure of some sections are not linear (e.g., Section 5: 1 hour, 30 minutes, and 10 minutes). Similar arguments and assumptions apply to change in score. Most users readily apply statistical tests to before and after treatment value with no regard to the starting point of the first value. Little and MacDonald92 have expressed this change as a percentage of the original score, arguing that it is better to shift a patient from 20% to 10% than to go from 60% to 50%. No other investigator has used this scheme.

An alternative is to aggregate the index into several categories. In the original paper five levels of the score were suggested (0–20%, 21–40%, 41–60%, 61–80%, and 81–100%). Some investigators have used this system to categorize their patients.9,95,96 Others have divided their patient population into two groups above and below a criterion, such as 40%.125 A further possibility is to disaggregate the ODI.7 The issues of disaggregation are discussed by Scott et al.127

Clinically Significant Change
Meade107 chose 4 points as the minimum difference in mean scores between groups that carried clinical significance. The U. S. Food and Drug Administration (FDA) has chosen a minimum 15-point change in patients who undergo spinal fusion before surgery and at follow-up (Lipscombe, personal communication, May, 1999). Figure 2 and Table 5 show change in weighted means calculated from publications reporting ODI before and after treatment in various subgroups of patients. Large changes in score are seen in patients with primary back pain and the least in those with spinal metastases. Table 4 and Figure 1 demonstrate the large standard deviations seen in clinically homogenous populations with various back pain syndromes. More work is needed in this area.

Analysis of Changing Scores
The change in the total score and change in components of the ODI have been investigated.110,120,121,136,139 Sources of error include inconsistencies in the answering of a questionnaire, the natural fluctuations of symptoms as well as clinical improvements (Figure 2 and Table 5).

The Oswestry Disability Index and Roland–Morris Scores
The ODI has been directly compared with R-M in several studies.1,88,89 The two scales correlate (n = 500, ρ = 0.77).7 The scatter chart from Baker et al1 (Figure 3A) was obtained when both questionnaires were simultaneously presented in a computer questionnaire. The results reflect the imprecision of these scales. The ODI tends to score higher than the R-M score (Figure 3). Thus it is likely that the ODI is better at detecting change in the more seriously disabled patients, whereas the R-M score may well have an advantage in patients with minor disability.

Receiver Operating Characteristic
This is a concept used to explore the diagnostic test performance of an instrument or the ability of the instrument to detect change,25 where its sensitivity is plotted against 1 minus specificity. This allows the ability of the instrument to detect change to be investigated. The ROC
index (D) for the ODI was found to be 0.76, a score that is acceptable but not as good as the R-M scale. This is perhaps not surprising in a population of patients who are not severely affected (mean ODI = 26.2 ± 13.5 [SD]). \(^4,94\) The ROC index has not been calculated for the ODI in a group of more severely affected patients. Because the ROC curve depends on sensitivity and specificity, there is an inherent assumption that a “true disability” is known. This may be difficult to justify. \(^25\)

**Male–Female Differences**

Some researchers have reported consistently higher ODI scores in women than men, but others have not confirmed this. \(^75,119\)

**Normal Subjects**

There are few published reports of ODI scores in the “normal population.” Two small biomechanical studies used back pain–free control subjects. \(^72,73\) The Ste. Justine study of idiopathic scoliosis includes a telephone survey of normal subjects published separately from the parent studies in subsequent correspondence in *Spine*. \(^103\) The current authors have used the ODI in control subjects age-matched to a patient population with neurogenic claudication. \(^24\) Together the results in these studies yield a mean score of 10.19 (range, 2.2–12; Table 4).

**Citations and MBAOT**

It is always frustrating for a reader to find that articles are inaccurately cited. Sometimes the inaccuracies are such that the reference is impossible to find. In the current investigation, approximately half the papers contained minor spelling errors in their reference to the original work, such as Deyo and Centor in 1986. \(^25\) A number of authors substituted Judith Couper’s Occupational Therapy qualification (MBAOT) MBAOT or even Mboat for her surname. The order of the authors’ names has been altered frequently. One reason for this is the poor typographical layout in the original journal. Another is the copying of unread references from one paper to another. An incomplete list of these papers is cited. \(^6,20,21,24–33,70,93,94,113,152–161\)

**Discussion**

The ODI has proved to be a versatile questionnaire, although unfortunately a single version no longer exists. It is also unfortunate that the time frame for symptoms has been varied by others outside versions 1.0 and 2.0.

---

**Table 5.**

<table>
<thead>
<tr>
<th>Group</th>
<th>Weighted Mean Difference</th>
<th>Presample</th>
<th>Postsample</th>
<th>Sample S.D.</th>
<th>Range</th>
<th>F</th>
<th>Number of Groups</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spinal metastases</td>
<td>9</td>
<td>34</td>
<td>34</td>
<td>17.4–18.1</td>
<td>−</td>
<td>1</td>
<td>115</td>
<td></td>
</tr>
<tr>
<td>Sciatica with p.i.d.</td>
<td>15.1</td>
<td>330</td>
<td>330</td>
<td>10.5–17.2</td>
<td>0.58</td>
<td>4</td>
<td>48, 47, 150</td>
<td></td>
</tr>
<tr>
<td>Chronic back pain</td>
<td>15.4</td>
<td>793</td>
<td>702</td>
<td>10.0–26.0</td>
<td>0.02</td>
<td>12</td>
<td>131, 31, 126, 32, 50, 80, 138, 14, 97, 149, 147, 148, 12, 11, 10</td>
<td></td>
</tr>
<tr>
<td>Spondylolisthesis</td>
<td>14.4</td>
<td>51</td>
<td>51</td>
<td>14.0–20.1</td>
<td>−</td>
<td>1</td>
<td>133</td>
<td></td>
</tr>
<tr>
<td>Primary back pain</td>
<td>25.5</td>
<td>168</td>
<td>168</td>
<td>9.6–14.8</td>
<td>0.004</td>
<td>2</td>
<td>89</td>
<td></td>
</tr>
</tbody>
</table>

---

**Figure 2.** Weighted means with 95% confidence intervals for measured change of the ODI calculated from pooled data for various categories of patients. The number of patients for each category is also marked. See Table 5 for more information.
mately, there can be no absolute measure of disability and the score of this or similar instruments takes on a life of its own. The results published by the diversity of investigators gives an indication of the likely responses that can be expected in a particular patient subgroup. If the ODI is to be used as an outcome measure, this can assist in power calculations for a planned trial (Figure 4).

The choice of which condition-specific disability questionnaire to use must be an individual one. The ODI has found favor in studies of patients with more severe symptoms, although it also appears to provide a robust indication of those with minor symptoms. The R-M has been used more frequently in the primary care environment and in the elderly. There is no questionnaire that can be used to measure handicap.

The time scale of the ODI and the R-M questionnaire is “now.” The authors believe that this is more robust than asking subjects to average their symptoms over the previous week, as is the choice of the AAOS.

The measurement of disability is an important component of the management of patients with back pain.

Figure 3. A comparison of the Oswestry Disability Index (ODI) and Roland-Morris (R-M) questionnaire. Note the R-M questionnaire has been adjusted to a percentage so that it can be directly compared; normally it scores between 0 and 24. A, A scatterplot adapted from Baker et al. for 183 patients. The mean ODI score is 34.9 whereas the mean for the R-M questionnaire is 52.45. Thus, on average the R-M scores higher. The mean difference of 17.5 is significant (P < 0.001) using a paired t test. The dotted line shows the line of identity where the ODI and R-M scores would take the same values. B, The use of a Bland and Altman plot to exemplify differences. The solid line marks the mean difference, and the two dashed lines mark two standard deviations on either side of the mean (the upper and lower limits of agreement). The trend, in which the greater the mean values the more negative the differences, is clear.
Self-report questionnaires have been better than so-called “objective” measures, such as range of movement and various measures of functional capacity, in achieving this. This has meant that some of self-report disability scores have become, in their own right, a dimension of disability, in the same way that the Glasgow Coma Scale has become a measure of head injury status in its own right. It is not possible to define the mathematical behavior of these scales, although many investigators have used them as if they behave in a linear fashion. The collation of scores from different diagnostic groups recorded in different cultures speaks for the robustness of this concept.

In spite of the inadequacies of physical measures, the authors do not believe that such measures should be abandoned as outcome measures, because they may well be measuring a dimension different from that measured by the questionnaires. The current authors have used the Shuttle Walking Test as an outcome measure in a physiotherapy study, in patients with neurogenic claudication, and, currently, in the Spine Stabilization Trial. Further work is needed to optimize physical measures and the correlation of responses with the questionnaires.

Because so many researchers have recommended the use of the ODI, it is important that the structure and the scoring systems be adhered to. This opens the possibilities of aggregating the results of studies and comparing outcomes. Doubtless, scales will evolve in the future, and new ones will be presented. However, the authors believe that until a method is developed that is clearly superior, the ODI, the R-M, or both should be used as condition-specific outcome measures in studies of patients with back pain.

Conclusions

The ODI remains a valid and vigorous measure of condition-specific disability. The authors recommend the use of version 2.0. The data presented are a guide to the power of the instrument to detect meaningful changes in disability status. More work is needed in this area. The ROCs of the curve should be explored in populations with higher self-report disabilities. More studies are needed to explore the response to change in an individual. The work started by the Ste. Justine group in extracting more information from disability instruments should be developed in more disabled populations than those with idiopathic scoliosis. The statisticians of the Ste. Justine Group have argued that additional information can be obtained by disaggregating the score and using sophisticated statistical techniques. The ODI has been published in at least four formats in English and in nine other languages. The four versions in English are presented in full.

Key Points

- The ODI has been published in at least four formats in English and in nine other languages. The four versions in English are presented in full.
- The authors recommend the use of version 2.0.
- The ODI has stood the test of time and many reviews. It is usable in a wide variety of applications as a condition-specific outcome measure of spine-related disability.
- Results of a meta-analysis show variations in estimated population means of ODI scores for different spinal diseases and changes after treatment consistent with clinical experience.


The 1980 publication of the Oswestry Disability Index (ODI) in the journal *Physiotherapy* foreshadowed an explosion in the field of pain and disability scales for patients with back pain by nearly a decade. The authors’ opening comments point out an unfortunate situation that can arise from early development of a useful tool, namely, multiple versions and idiosyncratic scoring methods. The recommended version from the authors is their ODI 2.0 and the scoring method is outlined. This is not a trivial concern when one considers that both a score of 0 and a score of 100 have been used to define normal.

A stated objective of this review was to “maintain a gold-standard” in the field. Dr. Fairbank’s seminal piece from 1980 concluded that the ODI was a reliable instrument based on the high test-retest correlation in 22 patients over a 24-hour period. The validation portion consisted of following 25 patients with their first episode of LBP over a 3-week period. As the patients reported improvement in their backache (as was expected), their ODI scores improved also.

Dr. Fairbank proposes in this piece, “The wide use of the ODI is part of the validation process.” The thought that wide use and reasonable performance as expected on a small sample are synonymous with validation and a rigorous review is one that falls short of current capabilities in the field. It should no longer be enough to simply report findings that turned out as expected, or that a gold-standard measure is crowned as a result of widespread use. Good validation studies should state a clear hypothesis and test it using a rigorous design and statistical analysis. This review article nicely compiles a wide range of work utilizing the ODI over the past 20 years. While the breadth of this compilation is notable, and the validation steps taken at various times have raised interesting questions, it has not, in my opinion, established a gold-standard measure. In conclusion, a fitting statement from McDowell and Newell: “It is possible to use statistically correct procedures to refine an instrument whose content is based on clinical wisdom and common sense.”

References