LIISA PEKKANEN

Health-Related Quality of Life and Functioning in Patients with Spinal Fusion

■

UNIVERSITY
OF TAMPERE

Supervised by
Professor Arja Häkkinen
University of Jyväskylä
Finland
Docent Marko Neva
University of Tampere
Finland

Reviewed by
Docent Jari Arokoski
University of Eastern Finland
Finland
Docent Ilkka Helenius
University of Helsinki
Finland

Cover design by
Mikko Reinikka

441 729
Painotuote

To Heta

Abstract

The objective of this study was to assess disability, Health-Related Quality of Life (HRQoL) and health utility in patients undergoing spinal fusion surgery.

Since 2008 all patients undergoing non-urgent spinal fusion in Tampere University Hospital and Jyväskylä Central Hospital have been invited to the spinal database. Patients´ data were collected prospectively, prior to surgery and at three months, one year and two years postoperatively and there are different follow-up periods in sub-studies. A general population sample was also drawn and matched for age, gender and residential area for purposes of comparison with the fusion patients.

Three months post surgery the 173 patients showed a significant improvement in disability when assessed with the translated, validated and psychometrically tested Finnish version 2.0 of the Oswestry Disability Index (ODI). The mean preoperative total ODI score was 45 (SD 17) and the mean decrease at three months was -19 (95% CI: -22 to -17). At the one-year disability analysis only minor addition in improvement was observed as compared to the ODI at three months. In the general population the ODI was 15 (SD 17) in females and 9 (SD 13) in males. Despite the improvement in disability among the patients, both sexes still had higher mean ODI values at one year than the general population (p<0.001).

The changes in HRQoL were assessed with the 36-item Short Form Health Survey (SF-36). The SF-36 was also used divided in two component scores: the Physical Component Summary Score (PCS) and the Mental Component Summary Score (MCS). At three months the positive changes were significant in both the PCS and the MCS. In addition the relationship between disability and HRQoL was significant. Although at one year both the female and male patients attained the general population level in the MCS, in the PCS the patients fell behind the general population.

In the health utility analysis, the data of 242 fusion patients were analysed stratified into five surgical indication groups (degenerative olisthesis, isthmic olisthesis, spinal stenosis, degenerative disc disease or disc herniation and postoperative conditions). At two years the improvements in the SF-6D scores were significant in all groups (p<0.001). Furthermore, the changes in the SF-6D scores did not differ significantly between the groups (p=0.40).

In conclusion the evaluation of the present prospectively collected unselected patient material confirms that although the patients do not reach the level of the general population the patients get significant benefit from the fusion procedure in the current practise. This benefit is apparent already in the early phase of recovery and the positive changes stay stable during the two-year follow-up.

Tiivistelmä

Tutkimuksen tavoitteena oli arvioida selän jäykistysleikkauspotilaiden toimintakykyä, elämänlaatua ja terveyshyötyä.

Vuodesta 2008 on Tampereen Yliopistosairaalassa ja Keski-Suomen Keskussairaalassa kerätty kaikkien elektiivisten selän jäykistysleikkauspotilaiden tieto yhteiseen selkätietokantaan. Osatöissä seuranta-ajat vaihtelivat kolmesta kuukaudesta kahteen vuoteen. Potilaiden tiedot kerättiin prospektiivisesti sekä ennen leikkausta, että kolme kuukautta, vuosi ja kaksi vuotta leikkauksen jälkeen. Potilaiden tietoja verrattiin myös iän, sukupuolen ja asuinpaikan mukaan kaltaistettuun väestöotokseen.

Toimintakyvyn arvioimiseksi Oswestryn toimintakykyindeksin (ODI) versio 2.0 käännettiin suomeksi, validoitiin ja mittarin psykometriset ominaisuudet testattiin. Kolme kuukautta leikkauksen jälkeen oli käytössä 173:n potilaan aineisto ja ODI-indeksin perusteella toimintakyvyn paraneminen oli merkitsevää. Keskimääräinen ODI ennen leikkausta oli 45 (SD 17), viitaten alkuperäisen luokituksen mukaan vaikeaan haittaan, ja keskimääräinen muutos parempaan eli ODI:n pieneneminen oli 19 (95% CI: 17-22). Vuoden kohdalla 252 potilaan aineistossa haitta väheni edelleen vain vähän verrattuna kolmen kuukauden tuloksiin. Potilaiden mukaan kaltaistetussa naisväestössä ODI oli keskimäärin 15 (SD 17) ja miesväestössä 9 (SD 13). Huolimatta merkitsevästä toimintakyvyn paranemisesta, sekä nais- että miespotilailla oli väestöön verrattuna merkittävästi enemmän haittaa vuoden kohdalla leikkauksesta (p<0.001).

Terveyteen liittyvää elämänlaatua tutkittiin 36-item Short Form Health Survey-kyselyllä (SF-36) ja kolmen kuukauden kohdalla parannus elämänlaadun sekä fyysisellä (Physical Component Summary Score, PCS), että psyykkisellä (Mental Component Summary Score, MCS) osa-alueella oli merkitsevää. Kolmen kuukauden kohdalla havaittiin myös merkitsevä yhteys haitan vähenemisen ja elämänlaadun parantumisen välillä. Vuoden kohdalla potilaat, sekä naiset, että miehet, saavuttivat kaltaistetun väestön elämälaadun psyykkisessä ulottuvuudessa, mutta jäivät jälkeen elämänlaadun fyysisessä ulottuvuudessa.

Kahden vuoden kohdalla potilaiden (N=242) saavuttamaa terveyshyötyä eli utiliteettia tutkittaessa potilaat jaettiin viiteen kirurgisen leikkausindikaation mukaiseen ryhmään (degeneratiivinen olisteesi, istminen olisteesi, spinaalistenoosi, degeneratiivinen välilevytauti tai välilevytyrä, leikkauksenjälkeiset tilat). Kahden vuoden kohdalla kaikki ryhmät saavuttivat terveyshyötyä SF-6D mittarilla

arvioituna (p<0.001). Indikaatioryhmien välillä ei ollut eroa terveyshyödyssä saavutettujen muutosten välillä (p=0.40).

Yhteenvetona tässä prospektiivisessa tutkimuksessa, valikoimattomalla potilasaineistolla, havaittiin, että vaikka potilaat eivät saavuta kaltaistaan väestöä, niin nykyisillä leikkausindikaatioilla potilaat hyötyvät merkittävästi selän jäykistysleikkauksesta. Tämä hyöty on nähtävissä jo paranemisen varhaisessa vaiheessa, kolme kuukautta leikkauksen jälkeen ja tulos säilyy koko kahden vuoden seurannan ajan.

Table of contents

List of original publications

I   Pekkanen L, Kautiainen H, Ylinen J, Salo P, Häkkinen A. Reliability and validity study of the Finnish version 2.0 of the Oswestry Disability Index. Spine 2011; 36(4): 332-8.

II Pekkanen L, Neva M, Kautiainen H, Vihtonen K, Kyrölä K, Marttinen I, Wahlman M, Häkkinen A. Decreased disability is associated with improved perceived quality of life following spinal fusion. Disabil Rehabil. 2013 Aug; 35(16): 1364-70.

III Pekkanen L, Neva MH, Kautiainen H, Dekker J, Piitulainen K, Wahlman M, Häkkinen A. Disability and Health-Related Quality of Life in patients undergoing spinal fusion. A comparison with a general population sample. BMC Musculoskeletal Disorders 2013; 14: 211, 2474-14-211.

IV Pekkanen L, Neva MH, Kautiainen H, Kyrölä K, Marttinen I, Häkkinen A. Health utility in spinal fusion stratified by specific indications two years after surgery. Manuscript submitted.

The original publications are reproduced with the permissions of the copyright holders.

Abbreviations

BMI	Body Mass Index
BP	Bodily Pain
CI	Confidence interval
CLBP	Chronic low back pain
COPD	Chronic obstructive pulmonary disease
EQ-5D	EuroQol-5D, Measure of health status from the EuroQol Group
HRQoL	Health-Related Quality of Life
HUI	Health Utilities Index
ICF	International Classification of Functioning, Disability and Health
LBP	Low back pain
LTPA	Leasure time physical activity
MCID	Minimum Clinically Important Difference
MCS	Mental Component Summary Score
NHP	Nottingham Health Profile
NSAID	Non steroid anti-inflammatory drug
ODI	Oswestry Disability Index
PCS	Physical Component Summary Score
PF	Physical Function
PLF	Posterolateral fusion
PLIF	Posterior Lumbar Interbody Fusion
PRO	Patient-reported outcome
RAND-36	RAND 36-Item Health Survey
RCT	Randomized controlled trial
SF-12	The 12-item Short Form Health Survey
SF-36	The 36-item Short Form Health Survey
SF-6D	Six-dimensional health state classification from the SF-36
TDR	Total disc replacement
TLIF	Transforaminal Lumbar Interbody Fusion
VAS_{back}	Visual Analogue Scale; back pain
VAS_{leg}	Visual Analogue Scale; leg pain
WHO	World Health Organization

1. Introduction

Most people suffer from low back pain to some extent during their lives; a life-long prevalence of back pain has been reported to be as high as 84% (Walker 2000). Back-associated problems impose a huge burden on society; for example they are the leading cause of disability and early retirement. The vast majority of back pain patients recover during a short period of time, and only small proportion have to seek medical help owing to recurrent episodes of pain or more chronic pain (Coste et al. 1994, Lucas 2012, Waddell 1987).

Much effort has been expended on constructing a useful classification for low back pain. Despite these efforts, 85-95% of low back pain is non-specific (Waddell 2005). Surgical treatment, however, focuses mainly on specific back pain, i.e. nerve root pain caused by disc herniation or stenosis, and further, by degenerative olithesis, isthmic olisthesis or degenerative scoliosis.

Spinal fusion surgery in its present form, with pedicle screw instrumentation, has existed for at least for five decades (Roy-Camille et al. 1986). Fusion surgery has been shown to be effective in some high quality studies, for example in treating isthmic olisthesis and degenerative olisthesis (Ekman et al. 2005, Weinstein et al. 2009, Weinstein et al. 2007). However, controversy remains over the value of fusion surgery. In the treatment of non-specific back pain, the superiority of fusion over conservative treatment has not been satisfactorily demonstrated and the utility of this costly and risky surgical intervention remains doubtful.

The focus in assessing the outcome of spinal fusion surgery has moved from judging the possible fusion radiologically towards a patient-based evaluation of disability and quality of life. To be able to report and compare outcomes globally, validated outcome measurement tools are needed. In the field of spine surgery a disease-specific Oswestry Disability Index (ODI) (Fairbank et al. 1980) and generic 36-item Short Form Health Survey (SF-36) (Hays et al. 1993, Ware and Sherbourne 1992) questionnaires are widely used both in clinical and scientific work.

The primary aim of the present study was to offer novel information on the quality of life and functioning of spinal fusion patients during recovery in a real clinical setting using a longitudinal prospective database. This doctoral thesis thereafter aimed to translate and culturally adapt as well as psychometrically test the ODI for the national use in Finland.

2. Review of the literature

2.1 Low back pain

Low back pain (LBP) is a common and increasing problem in the western countries. Although prevalence studies show methodological heterogenity (Dionne et al. 2008), the prevalence of LBP has been reported to be 12-33% (Airaksinen et al. 2006). The prevalence of back pain peaks between ages 35 and 55 (Anderson 1997). Most patients, 80-90%, will recover regardless of treatment during a short period of time and only a small proportion will experience chronic pain or recurrent episodes (Coste et al. 1994, Lucas 2012, Waddell 1987). Back problems impose a huge burden on the health care system and society, as they are the leading course of disability (Bener et al. 2013) and early retirement (Andersson 1999). Some prognostic factors associated with a longer return-to-work time after a episode of a back pain have been identified: older age, greater disability, female gender, heavy work, diagnosis of specific back pain, existence of social dysfunction and isolation, and receiving higher compensation (Steenstra et al. 2005).

Back pain is a multi-dimensional problem comprising pathoanatomical, as well as physical and psychosocial factors (Waddell 2005). A specific diagnosis can be distinguished only in a minority of cases, but there is a strong consensus on the importance of achieving a diagnosis already in the acute phase of low back pain (Koes et al. 2001).

Many attempts have been made to classify back pain. The duration of back pain is classified into acute, lasting less than six weeks, subacute, lasting six to 12 weeks, and chronic, lasting more than 12 weeks (Koes et al. 2010). Another classification is to divide back pain into three categories: pain caused by a specific cause, nerve root or radicular pain, and non-specific pain (Waddell 1996). Specific and serious causes for back pain, so called red flags, are fracture, tumor, infection, or cauda equina syndrome, and these represent 1 to 2% of the entity. In approximately 5% back pain stems from nerve root pain caused by disc herniation or spinal stenosis (Waddell 2005). Degenerative olisthesis and degenerative scoliosis may also be a cause of nerve root pain; the prevalence of degenerative scoliosis has been reported to be from 3% in middle-aged individuals to as high as 50% in individuals aged 90

years or older (Kebaish et al. 2011). Isthmic olisthesis may be a reason for radiculating pain and this pathoanatomical feature has a prevalence of 6 to 11% in the adult population (Kalichman et al. 2009, Virta et al. 1992). In 85-95% of back pain cases a specific diagnosis is not possible, leaving the condition nonspecific (Waddell 2005). See Figure 1.

The biopsychosocial model was introduced by Peter O´Sullivan and represents a treatment-based grouping model where patients´ reactions to functional tests are assessed (O'Sullivan 2005). In this classification, the non-specific back pain group is divided into three equal-sized non-mechanical and mechanical subgroups. The first, non-mechanical, group consists of individuals whose pain is associated with psychosocial factors like fear avoidance. The second group is characterisized by mechanical etiology and includes individuals with movement impairments. The third group comprises individuals with directional or multi-directional movement control impairment (O'Sullivan 2005). Non-specific back pain can be termed mechanical back pain of musculosceletal origin, and the symptoms can vary along with the physical activity in question (Waddell 1996). This type of pain should be benign and self-limiting (Waddell 1996).

The sources of non-specific back pain have been studied widely. Degerative disc disease and facet joint osteoarthrosis are thought to be possible causes of back pain (Cavanaugh et al. 2006, Hurri and Karppinen 2004). An association has been reported between degenerative Modic changes in discs and chronic pain, between disc herniation and pain, and between disc herniation and the severity of pain (Takatalo et al. 2012). However, disc herniations have been reported to be most probable where the pain is recent or persistent, while disc herniations and radiological disc degeneration have also been reported in asymptomatic patients (Takatalo et al. 2012, Takatalo et al. 2011). There is also postulated that low-grade infection might be connected to LBP (Agarwal et al. 2010, Albert et al. 2013, Albert et al. 2008, Corsia et al. 2003).

The possible relation between non-specific back pain and motion has also been studied and it has been shown that motion patterns between the normal population and back pain patients differ from each other (Lund et al. 2002). The role of muscle atrophy has been assessed, and the findings have shown no associations between low back problems and either incresed fat content or a smaller cross-sectional area of the lumbar paraspinal muscles (Paalanne et al. 2011).

```
BACK PAIN

        Specific                                    Non-specific

Red flags        Nerve root pain and
                 radiculating pain

tumor        disc herniation or disc pathology    mechanical

infection    spinal stenosis                      non-mechanical

fracture     degenerative olisthesis

cauda equina degenerative scoliosis

             isthmic olisthesis
```
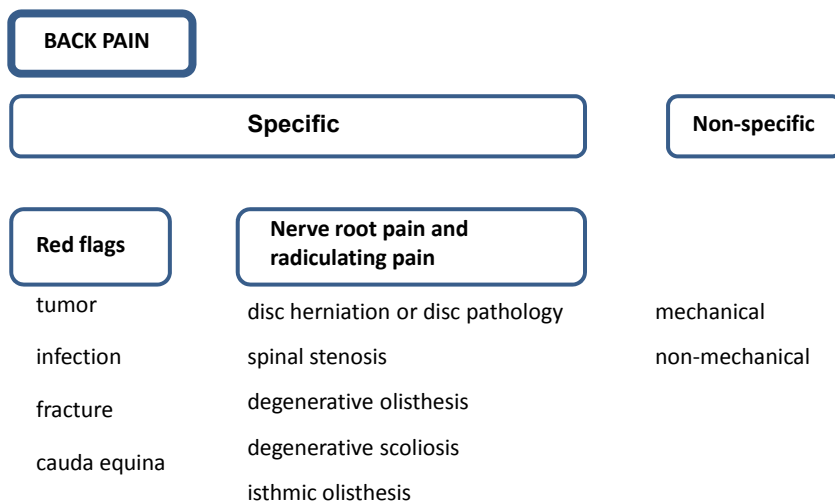
Figure 1. Classification of back pain. Modified from Waddell (Waddell 2005, Waddell 1996).

2.2   Treatment of low back pain

2.2.1   Non-operative treatment

According to current quideline the treatment of acute LBP lasting less than six weeks is based on patient information, medical treatment and some specific treatments to reduce pain like heat therapy and manipulation (Adult Low Back Disease: Current Care Summary 2008). The aims at patient information are among others to return to work as soon as possible and to continue the activities of daily living.

The majority of LBP can be managed with non-operative treatment. Many treatments are available, and previous reviews have shown that at least exercises and behavioural and multimodal treatment programs are able to induce improvement in LBP (Gutzman et al. 2004, Hayden et al. 2005, van Tulder et al. 2004, van Tulder et al. 2002). With exercises, the focus is improvement in physical capacity (Wessels et al. 2006), which is based on the reported association between deficient back muscle function and pain (Cassisi et al. 1993, Latimer et al. 1999, Mayer et al. 1994). Behavioral treatment in turn, tries to remodel behavioral practices assuming that, in addition to somatic factors, psychological and social

factors also affect pain and disability (McCracken and Turk 2002, van Tulder et al. 2004). Several cognitive concepts, such as fear-avoidance and illness beliefs, are thought to play a role in behaviour, and hence in pain and self-efficacy (de Rooij et al. 2011). Multimodal treatment is based on the theory that physical, psychosocial and social factors might have a role in diminishing pain and disability and promote the patient´s return to work (Gutzman et al. 2004). According to O´Sullivan, a physiotherapy intervention based on the biopsychosocial model and targeted at the underlying pain-driving mechanisms (pathological process, psychological and social factors, movement impairments, control impairments) may be able to adjust these disorders, and so remove both the physical and cognitive causes of pain (O'Sullivan 2005).

In chronic low back pain (CLBP) the symptoms are lasting more than 12 weeks. According to the current quideline the basis of treatment is rehabilitation where the patient is activated to participate in the treatment and to improve his or her functional ability (Adult Low Back Disease: Current Care Summary 2008). The rehabilitation is carried out by physical exercise therapy, multi-disciplinary rehabilitation and cognitive-behavioural therapy.

Medication is a basic non-operative method of treatment both in acute and chronic back pain. In the national recommendation, paracetamol is stated to be the first and safest choice, if the pain is not severe (Chou et al. 2007). Non-steroid anti-inflammatory drugs (NSAID) and a combination of NSAID and mild opiate are also beneficial (Chou et al. 2007, Roelofs et al. 2008). Gabapentin and topiramate have a role in neuropathic pain relief and tricyclic antidepressants have shown efficacy in chronic pain (Chou et al. 2007). Muscle relaxants have shown short-term effectiveness in case of back pain (van Tulder et al. 2003).

2.2.2 Operative treatment

In the field of spinal surgery, the fusion is the most demanding and costly procedure compared to lesser procedures like extirpation of disc herniation or microdecompression.

Spinal fusion surgery was first described in connection with the stabilisation of tuberculotic spine (Pott´s disease) and first performed formally by Russell A Hibbs 1911 (Hibbs 2007). The spinal fusion as a treatment of CLBP was described in the late 1920s (Hibbs and Swift 1929). Historically, since the early 1900s, posterior spinal fusion has been performed without instrumentation. The foundation of the technique is the tight packing of the cancellous bone chips removed from the

posterior iliac crest, which were recommended to be two millimetres thick and two centimetres wide (Smith-Petersen et al. 1945). Postoperative care included a paper corset, that was worn from four to six months, and followed by an ordinary corset for the same length of time (Eie 1966).

In Europe, the use of pedicle screw fixation started in the 1960s (Roy-Camille et al. 1986). The use of pedicle-screw fixation allows early mobilization without external supports like corsets. (Figure 2). The pedicle screws are placed with bony landmarks (Kim et al. 2004), tensifier imaging or computer-based navigation (Laine et al. 2000). With the posterior approach 360° fusion can also be performed using posterior interbody cages via the posterior lumbar interbody fusion technique (PLIF) or transforaminal lumbar interbody fusion technique (TLIF) (Cloward 1953, Lin et al. 1983, Lowe et al. 2002). The PLIF technique was first introduced in the 1950s and later modified by Lin (Lin et al. 1983) while a further developed version, TLIF, was described in the early 2000s (Lowe et al. 2002). The anterior technique was first described in the 1940s (Lane and Moore 1948) and later by several authors (Connor et al. 1967, Gumbs et al. 2007, Harmon 1960). Anterior fusion is done by using cages some of which are even of the stand alone type, needing no posterior fixation. Much debate has taken place during the past decades over whether fusion should be done by using the posterior or anterior approaches or combination of the two approaches. In a study of Helenius et al. with in situ fusions for high grade isthmic olisthesis in children and adolescents the circumferential fusion, i.e. with combined approach, provided significantly better long-term outcome both clinically and radiologically (Helenius et al. 2006). The large part of the material used in pedicle screws and rods in adults is titanium, although stainless steel and cobalt chromium alloy are also used.

In addition to fusion surgery, motion-preserving techniques have been introduced in the field of spinal surgical interventions. These techniques are attempts to preserve motion in the segment and to avoid the side effects of fusion, such as adjacent segment problems (Ghiselli et al. 2004, Park et al. 2004). Total disc replacement (TDR) is based on maintaining motion in the painful segment of the spine and removing the pain-generating disc (Freeman and Davenport 2006). The use of TDR was initiated in Europe 1988, and the first RCT study reported equal results between the TDR and anterior lumbar interbody fusion at two years (Geisler et al. 2004). A randomized study by Berg et al. reported better outcome for the TDR than fusion at one year, but the difference disappeared at two years (Berg et al. 2009). However, the long-term outcome and the possible superiority of TDR compared to fusion remain to be proven (Freeman and Davenport 2006). Another motion-preserving technique is dynamic semirigid stabilization of the lumbar spine, although the reported four-year results are less good than those obtained in the

fusion (Haddad et al. 2013). An interspinosus device, the x-stop, is used to achieve indirect decompression and yielded results similar to those obtained by conventional decompression at two years. However, more re-operations were needed after the x-stop (Strömqvist et al. 2013).

Indications for fusion surgery

Spinal fusion surgery has been used for various diagnostic indications among those with low back pain, both specific and non-specific. In specific LBP there is evidence favouring fusion in symptomatic isthmic olisthesis (Wood et al. 2011). A similar conclusion was reported by Möller and Hedlund in a prospective randomized study of 111 isthmic olisthesis patients: the functional outcome and the pain reduction were better in the operative group at one- and two-year follow-up (Möller and Hedlund 2000). The long-term analysis of the same study material showed some loss of the short-term improvement, but still the global outcome remained clearly better in the surgically treated group (Ekman et al. 2005). In spinal stenosis with degenerative olisthesis, fusion has been found to be more advantagous than conservative treatment or decompression alone (Malmivaara et al. 2007, Weinstein et al. 2009, Weinstein et al. 2008, Weinstein et al. 2007).

The treatment of back pain has been evaluated in three randomized controlled studies (RCT). In a Volvo Award Winner randomized controlled trial based on the Swedish Spine Study Group, surgery was more succesful than conservative treatment in patients with chronic back pain. This trial was conducted in 19 spine centres and with a total of 294 patients (Fritzell et al. 2001). In another randomized clinical trial, conducted in Norway with 64 patients, no difference was observed between fusion surgery and a combination of cognitive interventions and exercises (Brox et al. 2003). In a multicenter randomized study conducted by Fairbank et al. the benefit of fusion over intensive rehabilitation at two years remained unclear. The majority of the patients (total 349 participants, 176 surgical, 173 conservative) were operated on for chronic back pain (80%) and the rest for spondylolisthesis (11%) or after prior laminectomy (8%) (Fairbank et al. 2005). There are some methodological concerns in this study, e.g. the use of a flexible stabilisation method was included in the fusion surgery. In a very recent review of 26 articles, the authors point out that historically reviews of fusion as treatment have focused solely on limited randomized controlled trials (RCT) (Phillips et al. 2013). The full body of the literature also includes nonrandomized studies contributing real world findings alongside RCTs. The review concluded that fusion is a viable option with low back pain patients when the diagnosis is pain related to degeneration of the

motion segment, i.e. disc degeneration (Phillips et al. 2013). In several other reviews on the benefit of spinal fusion for low back pain, the conclusions have been either, that no conclusions can be drawn or that fusion is no better than intensive rehabilitation (Chou et al. 2009, Gibson and Waddell. 2005, Mirza and Deyo 2007).
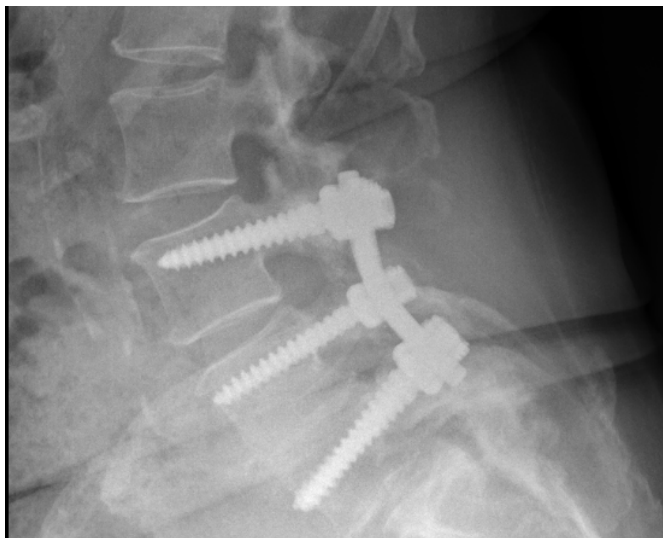
Figure 2. Lateral view x-ray of posterolateral fusion. X-ray from Jyväskylä Central Hospital.

2.3   Spinal fusion and outcome

In the literature published during the past years, the focus in assessing the outcome of spinal fusion surgery has shifted towards patient-reported outcomes (PRO) and quality of life scales (Becker et al. 2010, Carragee and Cheng 2010, Copay et al. 2008, Dimar et al. 2009, Djurasovic at al. 2011, Fritzell et al. 2001, Glassman et al. 2009, Glassman et al. 2009, Glassman et al. 2006) instead of evaluating whether radiologically solid fusion has been achieved or not. PROs are often the outcomes of the greatest importance to the patient, because they correspond the direct benefit of treatment rather than survival, the disease or physiological markers. The most commonly used instruments in outcome studies in the field of spine surgery are the Oswestry Disability Index (ODI) (Fairbank et al. 1980, Fairbank and Pynsent 2000), the 36-item Short Form Health Survey (SF-36) (Hays et al. 1993, Ware and Sherbourne 1992) and the Visual Analogue Scale (VAS) (Price et al. 1983).

2.3.1 Disability

Disability is a term enclosing impairments, activity limitations, and restrictions on social participation. Disability may lead to physical, cognitive, mental, sensory, emotional or developmental impairments or to some combination of these (World Health Organization, 2012). Back-specific measurements assess the aspects of a patient´s health that are affected by a specific disease and the connection of disability and symptoms to that certain disease (Kopec 2000). The ODI is one of the most widely used back-specific questionnaires, both clinically and in research in the field of spine surgery (Deyo et al. 1998, Osthus et al. 2006). It is also one of the back-specific instruments recommended by the World Health Organization (WHO) (Stucki and Sigl 2003). The aim of the ODI is to indicate the extent to which a person´s disability restricts his or her functional level. Other measurements used to assess disability are the Roland-Morris Disability Questionnaire (Roland and Morris 1983), the Million Visual Analog Scale (Million-VAS) (Million et al. 1982) and the Waddell Disability Index (Waddell and Main 1984).

John O'Brien began developing the ODI in 1976, and ODI version 1.0 was published in 1980 under the original name "The Oswestry Low Back Pain Disability Questionnaire" (Fairbank et al. 1980). Nowadays there are several versions of the ODI, and also revised versions, such as that published by a chiropractic study group in the United Kingdom (Hudson-Cook et al. 1989). The ODI contains ten items each comprising six statements graded from zero (lowest disability) to five (greatest disability). The section on the sex life of the patient may be omitted in some revised versions for cultural reasons (Fairbank and Pynsent 2000). Translations of versions 2.0 or 2.1, have been validated in several countries. The authors of the original ODI consider that all versions in different languages should be independently validated and, ideally, that there should be just one version in use for each language (Fairbank and Pynsent 2000).

Several studies have assessed the "psychometric" properties of the ODI. The main issues regarding the measurement quality are reliability, validity and responsiveness. Reliability is quality of method of measurement which consistently gives the same result (Pereira-Maxwell 1998). Reliability is tested by repeatability and reproducibility. Validity is defined as the ability of the instrument to measure what it was intended to measure (Finch et al. 2002, Lohr 2002). Validity includes both construct and content validity, and one part of construct validity is convergent validity.

Using a time interval of one week, the reliability of the Finnish ODI version 1.0 has been reported to be 0.83 (Grönblad et al. 1993). With a four-day interval only

the Pearson's correlation coefficient has been reported to be 0.91 (Kopec et al. 1996). Internal consistency using Cronbach´s alpha as a measure has been reported to be 0.76 with ODI version 2.0 (Fisher and Johnson 1997). For convergent validity correlations between the ODI and other instruments have been reported between the ODI and the Roland Morris Questionnaire (0.77) (Co et al. 1993), and between the ODI and Waddell Disability Index (0.70) (Waddell and Main 1984). Assessments of content validity have indicated that the ODI is more sensitive in detecting severe disabilities than the Roland Morris Questionnaire, reaching maximum scores earlier and tending to score higher also in the lower ranges of disability (Baker et al. 1989, Co et al. 1993). In responsiveness, a recent study by Johnsen et al. (Johnsen et al. 2013) found the Norwegian version 2.0 of the ODI to be very sensitive in detecting change among a group of 172 patients with diagnosed chronic low back pain. Similar results have been reported by the developers of the ODI, Fairbank et al. (Fairbank et al. 1980). The study group consisted of 25 patients suffering from their first attack of low back pain, and the expected improvement in their health status was also seen in their ODI score.

During the past five years many studies have reported the outcome of spinal fusion by using the ODI. In the previous literature there is one study with childhood and adolescent high-grade isthmic olisthesis. At the end of the 17-year follow-up the ODI varied from 3 to 10 between the groups of different surgical technique and the difference in the ODI between the posterolateral group and the circumferential group was significant (p=0.035). As conclusion circumferential fusion showed slightly better long-term results than posterolateral or anterior in situ fusion (Lamberg et al. 2007). To review the recent literature on disability in spinal fusion surgery, a systematic literature search was performed from January 2009 to May 2013 with keywords "spinal fusion" and "Oswestry Disability Index". A total of 381 articles were found and after reading the titles and abstracts, 136 of them were selected. Further selection criteria were a follow-up approximately 2 years, no reviews or case series, and no studies concerning trauma. Studies with 100 or more patients were selected and after examining the full text, 25 articles were included in the analysis. These studies are presented in Table 1.

All studies revealed the same trend, i.e., that the mean ODI scores were better postoperatively than preoperatively. In one study, patients with an absolute increase in their ODI score were compared with patients whose ODI score had decreased postoperatively. The proportion of patients with an absolute increase was 2.6 % (Gum et al. 2013). In five studies, the ODI was studied when one method of surgery was compared with another. Two of these were compared the conventional procedure with a minimal invasive procedure, one study finding the minimal invasive method more favourable, while the other did not (Kasis et al.

2009, Lee et al. 2012). Two studies reported that posterolateral fusion did not show greater improvement in disability compared with TLIF or PLIF (Høy et al. 2013, Wu et al. 2011). One study reported that whether or not the iliac crest autograph was used in fusion made no difference (Radcliff et al. 2012). There were also studies comparing the patients with different demographic data, such as smoking, co-morbidities and age (Cho et al. 2012, Glassman et al. 2010, Sanden et al. 2011, Wu et al. 2012). A large study showed that patients who smoked showed significantly less improvement postoperatively (Sanden et al. 2011). Primary operations were compared to revision operations in four studies. In three of these four the improvement in disability was poorer in the revision cases of surgery (Carreon et al. 2012, Glassman et al. 2009, Radcliff et al. 2013) while in the remaining study the revision patients achieved an equal improvement in disability (Carreon et al. 2013).

Table 1. Studies reporting disability by Oswestry Disability Index (ODI) in spinal fusion surgery

Authors	Patients	Follow-up	Study characteristics	Preoperative ODI values, mean (SD)	Postoperative ODI values, mean (SD)	Significance tests
Radcliff et al. 2013	413 patients; 54 with reoperation, 359 without reoperation, mean age 63 and 64 y	2 y	patients with reoperation after spinal stenosis surgery within 4 y	reoper 46 (18) no reoper 46 (18)	mean change reoper ~ -14 no reoper ~ -23	p<0.001 btw groups
Carreon et al. 2013	1055 patients, 722 primary vs. 333 revision, mean age 61 vs. 51 y	1 y	primary surgery compared to revision	primary 51 revision 54	primary 34 revision 40	p<0.001 pre vs. postop both groups
Min et al. 2013	172 patients, mean age 57 y	1 y	MIS TLIF, comparing 1, 2 and more than 3 levels	1 level 24 (2) 2 level 26 (3) ≥ 3 26 (4)	1 level 7 (1) 2 level 9 (2) ≥ 3 10 (5)	ns. btw groups
Lambat et al. 2013	1144 patients, 78 in each 3 groups: no complications, minor compications, major complications, mean age 48, 48, 49 y	2 y	evaluation of the influence of complications to the outcome of fusion surgery	no 53 minor 53 major 55	mean change no -13 minor -15 major -10	ns. btw groups
Høy et al. 2013	100 patients, mean age 50 y	2 y	comparison of TLIF and PLF	TLIF ~ 43 PLF ~ 41	TLIF ~ 30 PLF ~ 27	ns. btw groups

Table 1 continued

Gum et al. 2013	1054 patients, mean age 43 and 57 y	2 y	comparison of patients with worsening or improvement in the ODI	-	mean change worsening 8 mean change improvement -20	p<0.001 btw groups
Carreon et al. 2012	1104 patients, 8 diagnostic groups, mean age 57 y	2 y	assessing disability in diagnostic groups after decompression and fusion	mean 53 (14) max 56 (13) in adjacent-level degeneration min 50 (16) in scoliosis	mean change -13 (18) max -17 (19) in spondylolisthesis, min 8 (19) in nonunion	p<0.001 btw diagnostic groups
Cho et al. 2012	5119 patients, 23 with NIDDM and 23 controls, mean 61 vs. 59 y	2 y	influence of NIDDM to the result of fusion	NIDDM 43 (12) controls 44 (1)	NIDDM 30 (17) controls 30 (18)	ns. btw groups
Lee at al. 2012	144 patients, 72 in open surgery group and 72 in MIS group, mean age 57 and 52 y	2 y	comparing open and MIS TLIF surgery	open 44 (18) MIS 48 (19)	open 21 MIS 21	ns. btw groups
Radcliff et al. 2012	354, 108 with iliac crest autograft and 246 without	2 y	comparing the effect of iliac crest autograft usage in fusion surgery	-	without -24 (1) with -27 (2)	ns. btw groups
Wu et al. 2012	151 patients under or ovver 65 y, mean age 58 and 72 y	6 mo	comparing MIS TLIF in 2 age groups	< 65 y 46 (4) >65 y 47 (4)	< 65 y 15 (3) > 65 y 16 (4)	ns. btw groups

Table 1 continued

Djurasovic et al. 2012	1104 patients with decompression and fusion, mean age 57 y	2 y	assessing disability in lumbar decompression and instrumented fusion	53	40	p<0.001 pre vs. postoperative
Sandén et al. 2011	4555 patients, 758 smokers and 3797 nonsmokers with spinal stenosis surgery, from which 23 with fusion in both groups, mean 65 y and 70 y	2 y	assessing the effect of smoking	smokers 46 (15) non-smokers 45 (16)	smokers 33 (95% CI: 31-35) non-smokers 29 (95% CI: 28-29)	p<0.001 btw groubs
Inage et al. 2011	122 patients with spondylolisthesis in 1, 2 or 3 levels, mean age 64-68 y	2 y	reporting disability in spondylolisthesis	1 level 47 (17) 2 levels 33 (10) 3 levels 41 (12)	1 level 20 (6) 2 levels 25 (6) 3 levels 29 (6)	ns. btw groups
Tobler et al. 2011	156 patients undergoing MIS axial presacral interbody fusion, mean age 44 y	2 y	reporting disability in MIS axial interbody fusion	37 (15)	19 (19)	p<0.001 pre vs. postoperative
Wu et al. 2011	170 patients in PLIF and PLF groups, mean age 45 and 45 y	2 y	comparing PLIF and PLF	PLIF 36 (range 18-80) PLF 35 (range 16-78)	PLIF 16 (range 2-30) PLF 14 (range 0-26)	ns. btw groups

Table 1 continued

Cobo Soriano et al. 2010	203 patients with other lumbar disorder or herniation, mean age 52 y	1 y	comparing disability in decompression and instrumented PLF in 2 diagnostic groups	other lumb disord 44 (18) herniation 46 (18)	mean change other lumb disord -15 (21) herniation -29 (22)	p<0.001 both groups pre vs. postoperative
Carreon et al. 2010	783 patients with worker´s compensation and controls undergoing PLF, mean age 48 y in both groups	2 y	comparing outcome on patients with worker´s compensation and matched controls	non-compensation 58 (13) compensation 60 (12)	mean change non-compensation -13 (17) compensation -5 (14)	p=0.009 btw groups
Becker et al. 2010	195 patients with instrumented lumbar fusion, mean age 70-89 y	2 y	reporting disability in fusion surgery	53 (18)	2 y ~ 39	-
Glassman et al. 2009	283 adult deformity patients, 17-78 y, mean 50 y	2 y	assessing disability in adult deformity	37	23	-
Glassman et al. 2009	428 patients in 8 diagnostic groups, mean age 58 y	2 y	assessing disability in different diagnostic groups	-	mean change from nonunion -6 (14) to scoliosis -21 (20)	p=0.010 btw groups
Kasis et al. 2009	114 patients in standard PLIF and less invasive PLIF, mean age 49 and 46 y in 2 groups	2 y	comparing standard and less invasive PLIF operation	standard 46 (1) less invasive 47 (1)	standard 24 (2) less invasive 19 (1)	p<0.001 btw groups

Table 1 continued

Carreon et al. 2009	489 lumbar fusion patients, 18-87 y, mean age 56 y	2 y	assessing disability in lumbar fusion surgery	53 (14)	39 (21)	-
Dimar et al. 2009	224 patients with single level instrumented PLF, mean age 52 y	2 y	assessing disability in PLF	52 (range 30-94)	26 (range 0-82)	p<0.001 pre vs. postoperative
Glassman et al. 2009	224 patients with single level PLF, 174 younger than 65 y, 50 older than 65 y, mean ages 71 and 47 y	2 y	comparing younger and older patients with PLF	similar values younger and older, exact values not reported	mean change younger -25 (21) older -29 (20)	ns. btw groups

TLIF transforaminal lumbar interbody fusion; PLF posterolateral fusion; NIDDM non-insulin-dependent diabetes; BMI body mass index; MIS minimal invasive surgery; PLIF posterior lumbar interbody fusion; btw between; ns. not significant

2.3.2   Health-Related Quality of Life

The WHO has defined quality of life as a perception of individuals of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns" (The World Health Organization quality of life assessment 1995). Health-Related Quality of Life (HRQoL) is the degree to which individual´s usual or expected well-being is affected by a certain illness or its treatment (Cartwright 1999, Cella and Bonomi 1995) and thus is linked to factors of the environment, family and work (The World Health Organization quality of life assessment 1995). HRQoL can be assessed by using generic instruments. The SF-36 questionnaire represents a profile-type HRQoL measure describing the health state along various physical and emotional dimensions (Ware et al. 1994, Räsänen et al. 2006). The SF-36 is among most widely used HRQoL measures (Hays et al. 1993, Ware and Gandek 1998). Other HRQoL instruments are for example the 15D (Sintonen 2001) and Nottingham Health Profile (NHP) (Hunt et al. 1981). According to a recent recommendation (DeVine et al. 2011), some HRQoL tool should be used when evaluating the outcome of spinal fusion surgery in the clinical-research setting, and almost invariably recent studies report the SF-36 or either the Physical Component Summary Score (PCS) or the Mental Component Summary Score (MCS).

In the previous literature there is a long-term study of the HRQoL in surgically treated adolescent idiopathic scoliosis and isthmic olisthesis patients. In that study the scoliosis group had significantly higher values in the PCS than the isthmic olisthesis group (median 53.8 vs. 53.4 points, range 32-61 vs. 27-59, p=0.01) and the isthmic olisthesis group scored slightly better for the general health dimension than the olisthesis group (Helenius et al. 2008). In a Swedish study of long-term results with isthmic olisthesis patients there was no difference between operatively and conservatively treated patients when the dimensions of SF-36 were analyzed (Ekman et al. 2005). It was shown in that study, however, that the SF-36 scores in both groups were clearly below the scores of Swedish population.

For the purpose of evaluating the most recent literature, a systematic search was performed. Original articles reporting HRQoL as an outcome measure in adult spinal fusion procedures were included. A total of 112 articles were found in a systematic seach using "spinal fusion" and "health related quality of life" as search

words. After delimiting the search to publications from 2009 to April 2013 and after screening titles and abstracts, 24 articles remained and, after examining the full text of these articles, 14 articles were finally included. The search was supplemented with reference lists and the present author´s own files so that the final number of studies is 20 studies altogether (Table 2). The overall trend in the HRQoL studies was a significant improvement postoperatively from preoperative values. Two studies compared patients on unemployment benefit with patients on a disability compensation and matched controls. A parallel finding was that patients on worker´s compensation demonstrated less improvement than controls (Carreon et al. 2010, Gum et al. 2013). One study, which compared smokers and non-smokers, found clearly that the non-smokers showed significantly greater improvement (Sanden et al. 2011). In one study patients were stratified by age and both the younger and older patients improved with no significant difference between the groups (Glassman et al. 2009).

Table 2. Studies reporting Health Related Quality of Life (HRQoL) in spinal fusion surgery

Authors	Subjects	Follow-up	Study characteristics	Preoperative values, mean (SD)	Postoperative values, mean (SD)	Significance tests
Gum et al. 2013	1054 patients, 28 with absolute ODI increase, 1026 with ODI improvement, mean age 43 and 57 y	2 y	comparison of patients with worsening or improvement in the ODI	-	PCS worsening mean change -2 improvement mean change 7	p<0.001 btw groups
Lambat et al. 2013	1144 patients, 78 in each 3 groups: no complications, minor compications, major complications, mean age 48, 48, 49 y	2 y	evaluation of the influence of complications to the outcome of fusion surgery	PCS no 28 minor 27 major 26 MCS no 37 minor 39 major 38	mean change PCS no 5 minor 7 major 5 MCS no 4 minor 5 major 2	PCS and. MCS ns. btw groups
Radcliff et al. 2013	413 patients; 54 with reoperation, 359 without reoperation, mean age 63 and 64 y	4 y	patients with reoperation after spinal stenosis surgery within 4 years	reoperation SF-36 BP 30 (15) SF-36 PF 31 (21) no reoperation SF-36 BP 29 (16) SF-36 PF 32 (22)	mean change reoperation SF-36 BP 18 SF-36 PF 14 no reoperation SF-36 BP 28 SF-36 PF 25	p<0.001 btw groups in SF-36 BP and SF-36 PF

Table 2 continued

Gum et al. 2013	97 patients; 51 with disability compensation (with 51 matched controls), mean 42 y and 37 with worker´s compensation (with 37 matched controls) , mean age 42 y	2 y	comparing PLF in patients with disability compensation or worker´s compensation	PCS disability compensation 27 (control 27) worker´s compensation 27 (control 29) MCS disability compensation 28 (control 29) worker´s compensation 31 (control 30)	PCS change disability compensation 4 (control 6) worker´s compensation 2 (control 4) MCS change disability compensation 6 (control 5) worker´s compensation 5 (control 7)	ns. btw disability compensation vs. controls ns. btw worker´s compensation vs. controls
Carreon et al. 2013	1055 patients, 722 primary vs. 333 revision, mean age 61 vs. 51 y	1 y	primary surgery compared to revision surgery	PCS primary 28 revision 27 MCS primary 39 revision 38	PCS primary 36 revision 32 MCS primary 44 revision 41	PCS and MCS p<0.001 both groups pre vs. postoperative
Lee at al. 2012	144 patients in open and MIS TLIF, mean age 57 and 52 y	2 y	comparing open and MIS TLIF	PCS open 43 (27) MIS 43 (27)	PCS open 65 (25) MIS 68 (27)	-
Djurasovic et al. 2012	1104 patients with lumbar decompression and instrumented fusion, mean age 57 y	2 y	assessing HRQoL in lumbar fusion	PCS 28 MCS 38	mean change PCS 5 MCS 4	p<0.001 pre vs. postoperative
Parker et al. 2012	150 revision surgery patients, mean age 57 y	2 y	assessing HRQoL in revision surgery	SF-12 PCS 25 (7) MCS 41 (12)	mean change PCS 7 (11) MCS 7 (12)	p<0.01 both groups pre vs. postoperative

Table 2 continued

Rampersaud et al. 2011	214 patients with instrumented fusion, mean age 62 y	2 y	assessing HRQoL in fusion	PCS 30 (7) MCS 48 (11)	PCS change 11 (95% CI: 10-13) MCS change 4 (95% CI: 2-5)	significant improvement in PCS and MCS pre vas. postoperative
Sanden et al. 2011	4555 patients undergoing spinal stenosis surgery (23% with fusion), 758 smokers and 3797 nonsmokers, mean age 65 and 70 y	2 y	assessing the effect of smoking in spinal surgery	SF-36 PF smokers 30 (18) non-smokers 32 (21) SF-36 BP smokers 25 (15) non-smokers 27 (16)	SF-36 PF smokers 44 (95% CI: 42-47 non-smokers 52 (95% CI: 51-53) SF-36 BP smokers 46 (95% CI: 43-47) non-smokers 51 (95% CI: 50-52)	p<0.001 all differences btw smokers vs. non-smokers
Djurasovic et al. 2011	171 patients with revision after decompression, due to adjacent level degeneration or non union, mean age 56 y	2 y	assessing HRQoL after revision due to different diagnosis	PCS post decompression 26 adjacent level degeneration 27 non union 26	PCS post decompression 31 adjacent level deg 30 non union 29	-
Carreon et al. 2010	783 patients, from which 60 with worker´s compensation and 60 matched controls, mean ages 48 in both groups	2 y	comparing the effect of worker´s compensation in PLF	non-compensation PCS 27 (6) MCS 30 (11) compensation PCS 27 (5) MCS 31 (12)	mean change non-compensation PCS 4 (9) MCS 6 (11) compensation PCS -1 (10) MCS 7 (13)	PCS p=0.007 btw groups MCS ns. btw groups

Table 2 continued

Becker et al. 2010	195 patients with instrumented lumbar fusion, age 70-89	2 y	assessing HRQoL in lumbar fusion	PCS 28 (7) MCS 38 (12)	PCS ~ 32 MCS ~ 41	-
Mokhtar et al. 2010	105 patients with laminectomy and single level PLIF, mean age 67 y	2 y	assessing HRQoL in PLIF	SF-12 PCS 28 (95% CI: 27-30) SF-12 MCS 47 (95% CI: 46-50)	PCS 39 (95% CI: 37-42) MCS 52 (95% CI: 50-55)	PCS-12 and MCS-12 p<0.001 pre vs. postoperative
Dimar et al. 2009	224 patients with single level instrumented PLF, mean age 52 y	2 y	assessing HRQoL in PLF	PCS 27 (range 9-45)	PCS 40 (range 16-60)	p<0.001 pre vs. postoperative
Glassman et al. 2009	224 patients with single level PLF, 174 younger than 65 y, 50 older than 65 y, mean ages 71 and 47 y	2 y	assessing HRQoL in patients in 2 age groups	PCS younger 28 older 25	PCS younger 12 older 14	ns. btw groups
Glassman et al. 2009	283 adult deformity patients, mean age 50 y	2 y	assessing HRQoL in adult deformity	SF-12 PCS 34 SF-12 MCS 51	PCS 41 MCS 54	-
Li et al. 2009	104 patients with non-operative or operative treatment, mean ages r 76 y and 75 y	30-33 mo	comparing operative and non-oper treatment in adult scoliosis	SF-12 total operative group 31 (11)	SF-12 total operative group 42 (13)	p=0.027 operative group pre vs. postoperative

Table 2 continued

Carreon et al. 2009	489 patients with lumbar fusion, mean age 56 y	2 y	reporting HRQoL in lumbar fusion	PCS 28 (6) MCS 37 (13)	PCS 32 (12) MCS 40 (16)	-
Glassman et al. 2009	428 patients in 8 diagnostic groups, mean age 58 y	2 y	comparing fusion outcome in different diagnostic groups	-	PCS varying from adjacent level degeneration-group 3 (7) to disc pathology-group 8 (10)	ns. btw groups pre vs. postoperative

PLF posterolateral spinal fusion; PF Physical Function; BP Bodily Pain; MCS mental component summary score; PCS physical component summary score; MIS minimal invasive surgery; btw between; ns. not significant

2.3.3   Health utility

Along with limitations in resources in health care, and also in the field of spine surgery, demands concerning the effectiveness of different healthcare interventions have increased (Brazier et al. 1998). Along with disease-specific and generic outcome tools assessing disability and HRQoL, the health utility measures have also been introduced in the field of spine surgery. Assessing health utility also enables spine surgery outcomes to be compared to those of other disease states (Chapman et al. 2011). In the field of orthopaedics, total hip replacement, has already become a gold standard in these respects.

Single-index measures of health status preference can be used to judge the utility of a specific health care intervention (Carreon et al. 2012). Single index scores are anchored at zero for death and one for perfect health, and are also weighted for the relative desirability of the health state (Carreon et al. 2012, Gold et al. 1996). Preference-based measures for utility reported in the literature include the EuroQol-5D (EQ-5D) (EuroQol 1990), SF-6D (Brazier et al. 2002) and Health Utilities Index (HUI) (Feeny et al. 2002). Of these, the EQ-5D was the most frequently used (47.5%), while the proportion of using the SF-6D was 5% (Räsänen et al. 2006). The SF-6D is derived from the SF-36 using 6 of its health dimensions and 11 of its items (Brazier et al. 2002, Brazier et al. 1998). The SF-6D ranges from 0.29 to 1.00, with 1.00 indicating full health (Brazier et al. 2002, Brazier et al. 1998). The SF-6D has been psychometrically tested and its properties of construct validity, reliability and practicality have been proven in separate studies (Hollingworth et al. 2002, Solberg et al. 2005).

Population norms for the SF-6D are reported in the literature. In a study of 22,166 respondents from the United Kingdom, the mean SF-6D utility scores for males and females were 0.81 and 0.79, respectively (van den Berg 2012). Fryback et al. reported US norms for six different utility scores based on a telephone interview with 3,844 respondents. The SF-6D score varied from 0.80 in the younger age groups to 0.76 in the older age groups (Fryback et al. 2007). According to a review, the minimally important difference in the SF-6D has been investigated in seven studies and nine patient groups, varying from 60 to almost 5,000 individuals and patients, and it has ranged from 0.010 to 0.048, with a weighted mean 0.033 (95% CI: 0.029 to 0.037) (Walters and Brazier 2003). The lowest values were found among COPD patients and the highest among irritable bowel syndrome patients.

The differences concerning the minimally important difference and descriptive systems between the SF-6D and other preference-based outcome measures have also been studied. It has been concluded that these measurements are not interchangeable, that the values for minimally important differences are not equal, and, that the measurements give wholly different utility values for different groups of individuals (Barton et al. 2008, Bharmal and Thomas 2006, McDonough et al. 2005, Søgaard et al. 2009, Walters and Brazier 2005). It has been suggested that one of the advantages of the SF-6D over EQ-5D is that its descriptive system is broader and hence could provide a greater degree of sensitivity (Brazier et al. 2002).

Thus far only a few studies have assessed the utility of spinal fusion in a clinical setting. In a study of 1,104 patients with lumbar spinal fusion and two-year follow-up, eight diagnostic groups were classified: disc pathology, spondylolisthesis (both isthmic and degenerative olisthesis), instability, stenosis, scoliosis, nonunion, adjacent-level degeneration or postdiscectomy revision (Carreon et al. 2012). A statistically significant difference was observed in baseline SF-6D scores: the lowest (0.492) was in the non-union revision group and the highest (0.530) in the scoliosis group. During the 2-year follow-up, the biggest mean change, 0.088 was in the stenosis group and the lowest, 0.050 in the nonunion revision group. The mean changes in the other groups were 0.085 in the spondylolisthesis, 0.073 in the instability, 0.076 in the scoliosis, 0.076 in the disc pathology, 0.070 in the postdiscectomy revision, and 0.066 in the adjacent-level degeneration groups, respectively. The authors point out that these back patients reported the same level of utility values after spinal fusion than patients treated with unstable angina pectoris (Kim et al. 2005). In another study of single-level lumbar spinal fusion with 96 patients, health utility values were reported during a 5 year of follow-up (Glassman et al. 2012). A gradual increase throughout the follow-up period in utility was found: the baseline mean SF-6D score was 0.500, at two years the mean score was 0.638 and at 5 years it was 0.653. In a study of 45 consecutive TLIF-patients, the EQ-5D score was calculated as a health utility score. At the 2-year follow-up the mean improvement in the EQ-5D was 0.43 (SD 0.44) (Parker et al. 2011).

3. Aims of the research

The general aim of the study was first to validate a back-specific measurement tool for disability translated in to the Finnish language and thereafter assess disability, HRQoL and health utility in patients undergoing spinal fusion surgery.

The specific aims of the study were as follows (Roman numerals refer to original publications):

1. To assess psychometric properties and feasibility of the ODI version 2.0 translated into the Finnish language (I).
2. To assess changes at three months in disability and HRQoL after spinal fusion surgery and to examine the ODI in the framework of International Classification of Functioning, Disability and Health (ICF) (II).
3. To compare the level of disability and HRQoL among spinal fusion patients with the values of the general population (III).
4. To assess the changes in disability, HRQoL and thereby health utility of the treatment in spinal fusion patients according to different diagnostic indications for surgery during a two-year follow-up (IV).

4.   Methods

This doctoral dissertation consists of four separate studies conducted in Tampere University Hospital (2012 adult population 415 000) and Jyväskylä Central Hospital (2012 adult population 200 000) between the years 2006 and 2012. All studies were approved by the ethical committees of both hospitals and a written consent was obtained from all study subjects.

4.1   Subjects

The mean insidence of non-urgent spinal fusion operations between 2008 and 2012 in adult patients has been 25/100 000 (95% CI: 23 to 27) in the district of Tampere University Hospital and 30/100 000 (95% CI: 27 to 34) in the district of Jyväskylä Central Hospital.
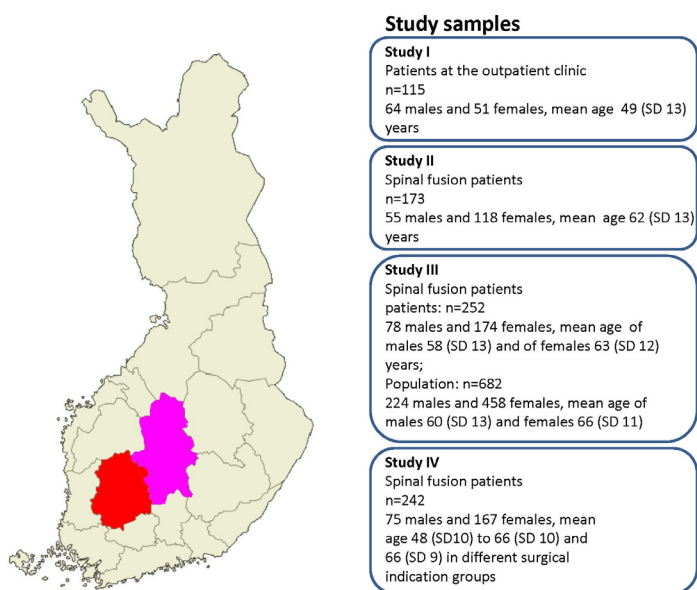

Study samples

Study I
Patients at the outpatient clinic
n=115
64 males and 51 females, mean age 49 (SD 13) years

Study II
Spinal fusion patients
n=173
55 males and 118 females, mean age 62 (SD 13) years

Study III
Spinal fusion patients
patients: n=252
78 males and 174 females, mean age of males 58 (SD 13) and of females 63 (SD 12) years;
Population: n=682
224 males and 458 females, mean age of males 60 (SD 13) and females 66 (SD 11)

Study IV
Spinal fusion patients
n=242
75 males and 167 females, mean age 48 (SD10) to 66 (SD 10) and 66 (SD 9) in different surgical indication groups

Figure 3. Study samples and patient characteristics.

4.1.1   Patients

The psychometric and feasibility study (I) comprised 115 patients with a clinically diagnosed low back problem who had been referred to the outpatient clinic of the Department of Physical Medicine and Rehabilitation in Jyväskylä Central Hospital, Finland. The inclusion criteria were age 18 years, low back pain with or without radiculating pain in the lower extremities and the ability to communicate in writing in the Finnish language. No specific exclusion criteria were used.

In the studies with different follow-up times (II-IV) the subject groups consisted of patients undergoing a non-urgent spinal fusion operation in either Tampere University Hospital or Jyväskylä Central Hospital. The data of all the voluntary adult patients were included in this prospective longitudinal spinal database. During the whole recruitment only 10 patients refused to participate altogether in these two participating hospitals. Patients were recruited through the whole period of these studies. The cohorts in the studies reflect the numbers of patients, who had met the follow-up requirements at the time of the analyses.

In the study of the early recovery phase (II) data of 220 had been gathered on of whom 173 patients (79%) had complete preoperative and postoperative data for all the outcome measures and were thus included in the final analysis. The non-participants did not differ by age (included 62 (SD 13) years and not included 65 (SD 12) years, p=0.16), gender (females 68%/62%, p=0.89) or level of pain (VAS_{back} 66 (SD 27)/60 (SD 24), p=0.23) from the included patients.

The one-year follow-up study, including the comparison with the general population sample (III), consisted of 285 patients with the 6 most common diagnoses for elective spinal fusion. These diagnoses were degenerative olisthesis, isthmic olisthesis, spinal stenosis, disc herniation or degeneration, postoperative conditions and degenerative scoliosis. Disability and HRQoL measures were available preoperatively for all patients, and 3 and 12 months postoperatively for 252 of these patients (88%), all of whom were included in the study.

The health utility study (IV) comprised 259 patients with two-year follow-up data. For 242 of these patients (93%), complete preoperative and postoperative data were available as well as data on diagnostic indication for surgery. For this patient cohort the data were stratified by the diagnostic indication for surgery: degenerative olisthesis, isthmic olisthesis, spinal stenosis, degenerative disc disease or disc herniation and postoperative conditions.

4.1.2   General population sample

The outcome of the spinal fusion patients was compared to that of a general population sample (III). The general population sample was age-, sex- and residential area-specific. Four controls for each of the fusion patients were drawn from the Finnish Population Register and the sampling was performed by Statistics Finland. A questionnaire was mailed to 1,140 controls in September 2010. After one reminder letter 2 months later, the percentage of returned answers was 61% (n=691) and the number of acceptably filled-in questionnaires was 682.

4.2   Study design

In the psychometric and feasibility study (I), the patients answered the mailed questionnaire package 2 weeks before arrival at the outpatient clinic. The participants answered the questionnaires a second time on arrival at the outpatient clinic. The questionnaire package included the Finnish translation of the ODI version 2.0, the Million Visual Analog Scale (Million-VAS), the Visual Analogue Scale for back and leg pain (VAS $_{back}$, VAS $_{leg}$), the depression scale (DEPS), questions eliciting standard sociodemographic data and the location and the duration of the pain, and a transition question to determine whether their health status had been stable, improved or worsened during the last two weeks. The first item from the Short Form-36 questionnaire ("In general, would you say your health is excellent, very good, good, fair or poor") was also asked as an independent question.

The data in the other studies (II, III, IV) were drawn from uniform spinal database in two Finnish hospitals: Tampere University Hospital and Jyväskylä Central Hospital. This database was started at the beginning of 2008 in both hospitals. The use of a spinal database and data collection has been a part of normal clinical practice. For patients, participation in the database has been voluntary. The patients were informed about the database when they visited the orthopaedic surgeon in the outpatient clinic before the operation. Informed written consents were received from the patients at the latest on arrival for the operation but most often in the outpatient clinic during the preoperative visit. Patients filled in a questionnaire booklet one to two weeks prior to the operation. The booklet consisted of questions concerning, for example, sociodemographic data and clinical information, the ODI, the SF-36, the VAS$_{back}$ and the VAS$_{leg}$. The postoperative timepoints for the data collection were 3, 12 and 24 months after the spinal fusion

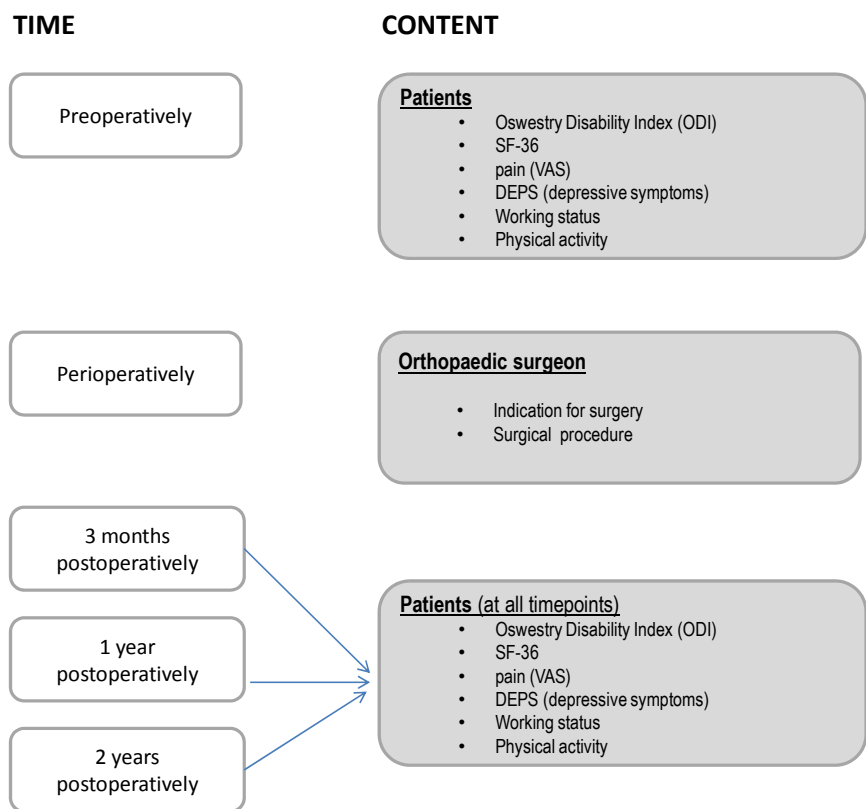operation. The orthopaedic surgeon filled in the diagnostic indication for the operation.

Figure 4. Data collection time points and the content of the data presented in this thesis and in the original articles.

4.3   Measurements

4.3.1   Oswestry Disability Index (ODI) (I-IV)

The ODI contains ten items each comprising six statements graded from zero (lowest disability) to five (greatest disability) and the time frame for the answers to the questions is "today" (Fairbank and Pynsent 2000). The total score is calculated as a sum of each completed item and expressed as a percentage of the maximum

number of possible points, i.e. related to the number of items the patient has given answers to. Scores are defined by the scale in the original publication: 0-20 minimal, 20-40 moderate and 40-60 severe disability. A score of 60-80 indicates a crippled patient and 80-100 indicates that patient is either bed-bound or exaggerating their symptoms.

4.3.2   Translation and cross-cultural adaptation of the ODI version 2.0 (I)

The cross-cultural adaptation process involves both translating the language and taking cultural adaptation issues into account when preparing the questionnaire for use in another setting (Beaton et al. 2000). When aiming at equivalence using the measure in another cultural context or in another language some form of cross-cultural adaptation is required (Beaton et al. 2000). The process of cross-cultural adaptation produces equivalence between the original source measure and the target (Beaton et al. 2000). Lack of equivalence limits the comparability of responses across different languages and cultures (Beaton et al. 2000). The present translation and cross-cultural adaptation followed the recommended guidelines (Beaton et al. 2000, Guillemin et al. 1993). The first step was an independent translation of the ODI version 2.0 in to the Finnish language by two professionals who had Finnish as their first language and who were professionals in rehabilitation medicine with a clinical orientation and understanding of the practical use of the questionnaire. Next, a consensus version of these two translated versions and the original English version was made. In the third step an English-speaking person from a non-medical field, back-translated the version into English without any prior knowledge of the ODI and this English version was compared to the original English version 2.0 (Fairbank and Pynsent 2000). When a consensus was reached between the translated Finnish version and the original English version, the Finnish version was evaluated by a Finnish language expert who was a member of the Finnish Medical Society Duodecim and the Finnish ODI version 2.0 was formally ready. This version was pilot tested with 30 patients before the patient recruitment for this trial was undertaken. The preliminary pilot testing demonstrated no concerns or reasons for any change to be made.

4.3.3 The ODI linked to the International Classification of Functioning, Disability and Health (ICF) (II)

In the study of the early recovery phase (II), the ODI was also studied in connection with the International Classification of Functioning, Disability and Health (ICF). The ICF offers a unified and standard language and framework for the description of functioning, disability, health and health-related states (Andresen 2000, Andresen et al. 2000, Dijkers et al. 2000, Gray and Hendershot 2000, WHO 2001). The linkage between the ODI and the ICF was done according the published guidelines (Cieza et al. 2005), first by separate health-care professionals independently and then in joint discussions to arrive a consensus. In this linking Part 1 of the ICF, titled Functioning and Disability, was used. Part 2 titled Contextual Factors, which encompasses environmental and personal factors, was not used (WHO 2001). The first component of Part 1, i.e., Body Functions and Structures was used as such: the domains were Mental functions and Sensory functions and pain. The latter component, Activities and Participation, was divided into two, so that the domains Mobility, Self-care and Interpersonal interactions and relationships were assigned to the Activities component and the domains Mobility and Community, social and civic life were assigned to the Participation component. The option of a partial overlap between the Activities and Participation components was chosen so, that the Mobility domain was left as a domain common to both. The ODI items were linked according to the main functional issue, since pain concerns every item, and the total score was expressed on a relative scale from zero to one hundred. See figure 5.
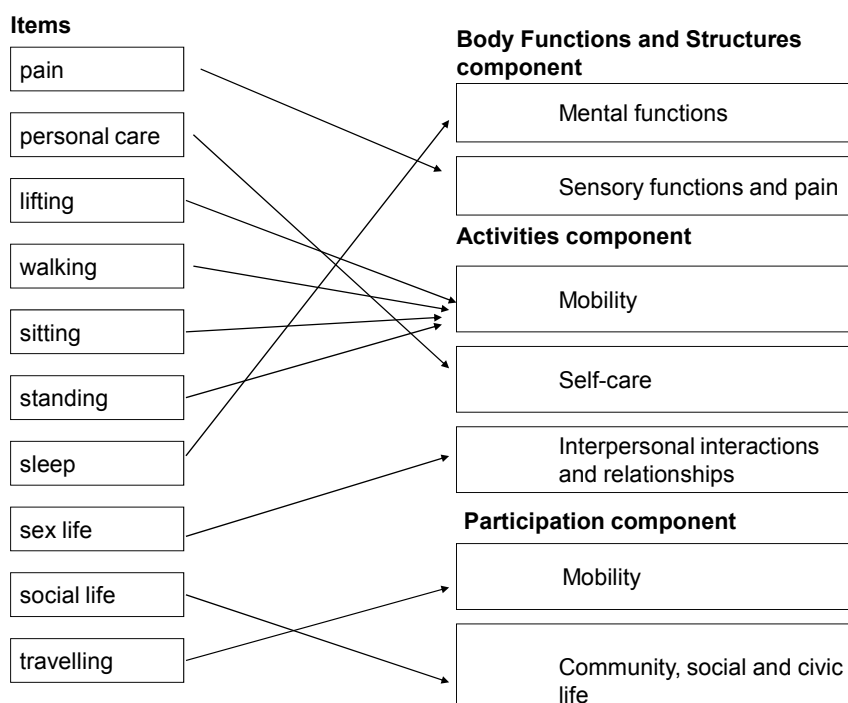
Items

pain	
personal care	
lifting	
walking	
sitting	
standing	
sleep	
sex life	
social life	
travelling	

Body Functions and Structures component

Mental functions

Sensory functions and pain

Activities component

Mobility

Self-care

Interpersonal interactions and relationships

Participation component

Mobility

Community, social and civic life

Figure 5. The linking between the Oswestry Disability Index (ODI) items and the International Classification of Functioning, Disability and Health (ICF).

4.3.4   The 36-item Short Form Health Survey (SF-36) (II-IV)

The SF-36 is the most widely evaluated generic, profile type of patient-assessed health outcome measure (Garratt et al. 2002). It reflects the patient´s health status through eight health concepts: Physical Functioning (PF), Role Physical (RP), Bodily Pain (BP), General Health (GH), Vitality (VT), Social Functioning (SF), Role Emotional (RE), and Mental Health (MH). The raw scores are transformed onto a 0 to 100 scale; a higher score is associated with better health. The eight domains can be aggregated into two distinct summary measures by using US reference population (1990) for standardization of the domains and for factor score coefficients: the physical component summary score, PCS, and the mental component summary score, MCS. The PCS and MCS are finally standardized using a mean of fifty and a standard deviation of ten (Ware and Kosinski 2001).

4.3.5   The SF-6D (IV)

The SF-6D is a classification for describing health derived from 11 items of the SF-36 (Brazier et al. 2002, Brazier et al. 1998). All together 18,000 health states can be defined with this method. The SF-6D values range from 0.29 to 1.00, with 1.00 indicating "full health" (Walters and Brazier 2005). The SF-6D as a preference-based health state score does not have a natural unit of measurement. The SF-6D scores are generated by using a set of parametric preference weights that are acquired using the distinguished valuation technique, standard gample, in a general population sample (Brazier and Roberts 2004).

4.3.6   Clinical findings (I-IV)

In the psychometric and feasibility study (I), disability of the patients was also assessed by the 15-item Million Visual Analog Scale (Million-VAS) (Million et al. 1982). In the Visual Analogue Scale (VAS) pain is assessed along a horizontal line 100 mm in length. The patient marks on the line the point that they feel represents their level of back or leg pain during the last week. The VAS score is expressed in millimetres (Price et al. 1983). Possible depressive symptoms were assessed using the depression scale (DEPS) (Salokangas et al. 1995). Patients were asked about their weight and height in order to calculate their body mass index (BMI) and demographic questions such as their employment status, their level of formal education, and their level of weekly physical activity (leasure time physical activity, LTPA). Patients were asked to rate the overall condition of back two years after the spinal fusion as "better", "unchanged" or "worse", than it was before the operation. Patients were also asked to state, whether they would go through the operation again.

4.3.7   Statistical methods

The results are expressed as mean or median, standard deviation (SD), interquartile range (IQR), as counts with percentages, or frequency distributions. Areas under the curves (AUC) were calculated with the trapezoidal method; AUC was divided by the total time of study and values are depicted in time-weighted mean scores. The 95% confidence intervals (95% CI) were obtained by generally or using bias-corrected bootstrapping (5000 replications), when appropriate. The reliability of the scales was evaluated by calculating the intraclass correlation coefficient (ICC)

and coefficient of repeatability, with the bias corrected and accelerated bootstrapping (5000 replications) confidence intervals. Internal consistency was estimated by calculating Cronbach's alpha. Construct validity was studied by maximum likelihood factor analysis with promax rotation for the ODI items matrix of polychoric correlations. Item analysis of the ODI scales was performed by analyzing item discriminating power (corrected item correlation) and item difficulty (item mean) depicted by the explanatory data analysis. Corrected item correlation was estimated using polyserial correlations.

The statistical significance between groups was evaluated by using generalized linear models with appropriate distribution and link function, with appropriate contrast. In the case of violation of the assumptions (e.g. non-normality), a bootstrap-type test was used. This method is significantly helpful when the theoretical distribution of the test statistic is unknown.

Repeated measures were analyzed using linear mixed models; fixed effects were group, time, and group-time interaction. Multivariate approach was made by using permutation type Hotelling's T-squared test. Regression lines with 95% CI were used to illustrate the relationships between variables. Correlation coefficients were calculated by the Pearson or Spearman methods with bootstrap type confidence intervals. The normality of the variables was tested by using the Shapiro-Wilk W test. All statistical tests were two sided, with a α-level of 0.05. No adjustment was made for multiple testing, but this information can be obtained by multiplying the actual p value by the number of comparison made.

Effect size was calculated by the method of Cohen's d (mean baseline scores minus mean follow-up scores, divided by the pooled standard deviation (Cohen J. 1988). Effect size of ≥ 0.20 was considered small, ≥ 0.50 medium, and ≥ 0.80 large. 95% CI for the effect sizes were obtained by bias-corrected bootstrapping (5000 replications).

5. Results

5.1 Disability

5.1.1 Psychometric properties and feasibility study of the Finnish version 2.0 of the ODI (I)

Altogether 115 patients participated in the study of disability assessment by using the ODI and Million-VAS scores. The mean age of the patients was 49 years (SD 13). The diagnoses of the outpatient-clinic patients were spondylarthrosis 29%, disc degeneration or herniation 23%, isthmic olisthesis 9%, spinal stenosis 8% and other disorders 31%. Other disorders were mainly back pain with non-defined origin or back pain with non-specific radiculating pain. The response rate in the ODI items varied from 85% for the sex life item to 99% for the items of personal care and social life. Floor effects (i.e. minimum score) were most often found in items concerning personal care, walking, sex life and social life. Ceiling effects in items (i.e. maximum score) were rare (Figure 6). Four patients scored 10 or lower on the first occasion and five on the second occasion. No patients scored the maximum on either the first or second occasion; the highest total score was 70 at the initial occasion.

On the first occasion the questionnaire was filled in, the mean (SD) ODI score was 33 (14) and the mean Million-VAS score was 53 (18) for patients who turned out to be stable in their subjective health between the first and second survey. For these patients, reproducibility (ICC) was 0.90 (95% CI: 0.85 to 0.94) for the ODI and 0.85 (95% CI: 0.76 to 0.91) for the Million-VAS. In patients who reported improvement or worsening in their symptoms, both instruments showed poorer reproducibility. When the patients were grouped according to change in their symptoms, the change in the ODI (P<0.001) and the change in the Million-VAS (P=0.003) decreased linearly across the groups (i.e. worsened-stable-improved); see Table 3.

The internal consistency of the ODI was 0.86 (95% CI: 0.81 to 0.90) and the internal consistency of the Million-VAS was 0.90 (95% CI: 0.86 to 0.93).

Item analysis of the ODI (Figure 7) showed that all the items had a good overall corrected item correlation. Item 9 (social life) had a high corrected item correlation and item 3 (lifting) had a low correlation. Items 2 (personal care) and 4 (walking) showed the lowest item mean values and item 3 (lifting) the highest.

Factor analysis performed for construct validity showed that the ODI scale was loaded on two factors. Factor 1 characterizes activities of daily living (sleeping, personal care, social life, sex life, lifting, sitting). Factor 2 characterizes activity in the upright position (walking, travelling, standing) and pain. These factors explained 51% of the total phenomenon.
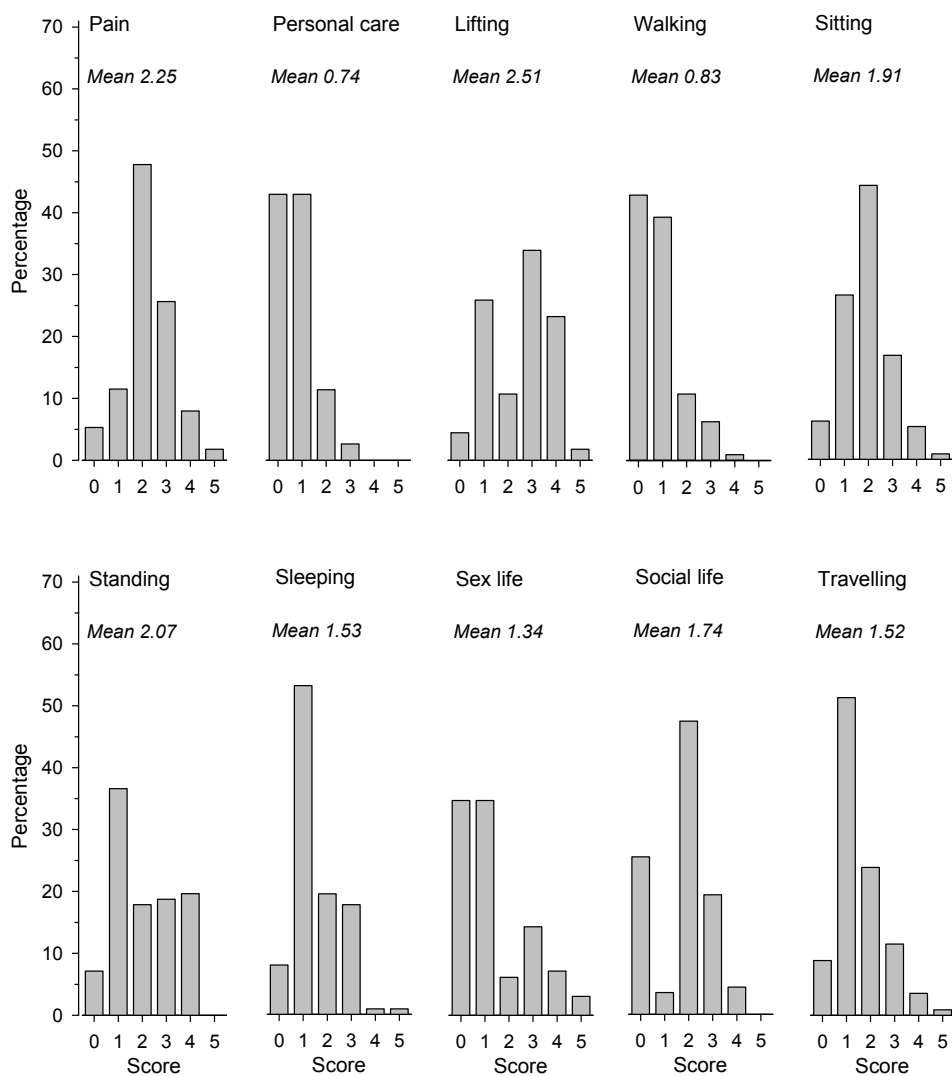
Figure 6. Distribution of the ODI items.

Table 3. Reproducibility of the Oswestry Disability Index (ODI) and Million Visual Analogue Scale (Million-VAS) Scores

Symptoms	n	First measurement, mean % (SD)	Change from first to second measurement, mean (95% Cl)	Reproducibility	
				ICC (95% Cl)	CR (95% Cl)
ODI					
improved	12	29.4 (15.3)	-9.3 (-15.2 to -4.4)	0.60 (0.12 to 0.82)	19.4 (9.4 to 29.6)
stable	86	33.1 (15.0)	-1.7 (-3.1 to -0.4)	0.90 (0.85 to 0.94)	12.5 (10.7 to 14.4)
worsened	16	35.5 (11.4)	+5.1 (1.1 to 9.5)	0.69 (0.41 to 0.83)	17.6 (12.1 to 23.2)
Million-VAS					
improved	12	50.9 (13.8)	-16.5 (-25.4 to -8.6)	0.17 (-0.34 to 0.54)	29.2 (20.1 to 38.3)
stable	82	52.6 (18.8)	-4.8 (-6.9 to -2.8)	0.85 (0.76 to 0.91)	18.5 (14.6 to 22.5)
worsened	13	56.9 (20.3)	+1.0 (-6.2 to +7.7)	0.69 (0.29 to 0.84)	26.4 (16.2 to 36.5)

ICC: Intraclass correlation coefficient. CR: Coefficient of repeatibility. Expresses the expected maximum size of 95% of the absolute differences between paired observations. 95% Cl obtained by bias corrected and accelerated bootstrapping.
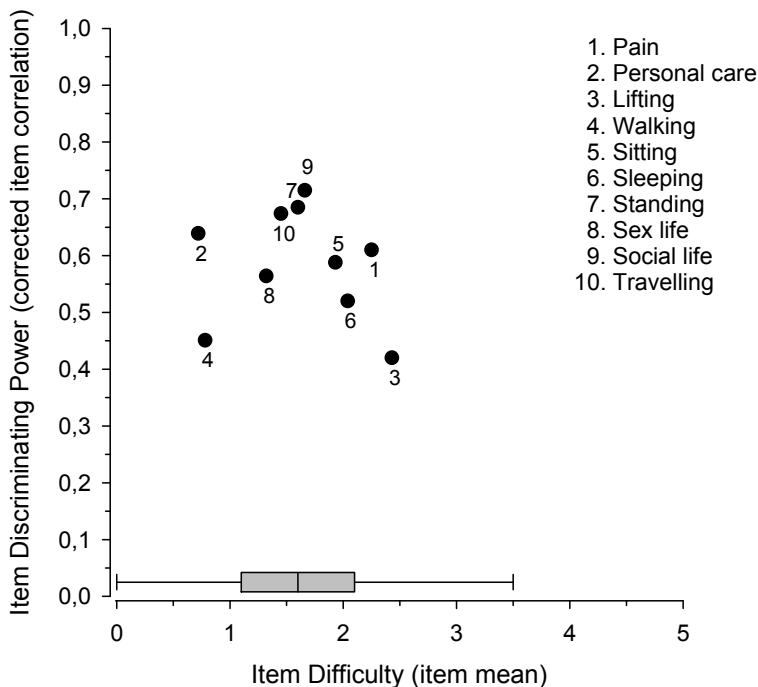
Figure 7. Item analysis for the Oswestry Disability Index (ODI). The bar denotes median with interquartile range. Numbers indicate the corresponding items in the ODI scale.

The correlation coefficients between the ODI and the Million-VAS were 0.75 (95 % CI: 0.64 to 0.84) and between the ODI and VAS_{back} and VAS_{leg} 0.48 (95% CI: 0.32 to 0.62) and 0.41 (95% CI: 0.23 to 0.57), respectively.

The ability of the ODI to distinguish between three different clinical states in a patient´s self-perceived health (excellent, good, poor) was studied at baseline. After adjusting for age, the relationships remained linear between the ODI and self-perceived health (p<0.001) (Figure 8). Between the ODI and duration of symptoms there was also linearity (p=0.011) but not between the ODI and the BMI (p=0.16).

The correlations between the ODI and age and the DEPS-score were 0.24 (95% CI: 0.07 to 0.40) and 0.52 (95% CI: 0.36 to 0.64), respectively.

The ODI score indicated no statistically significant difference in disability between the diagnostic groups. The mean ODI score ranged from 31 to 35 between the 5 diagnostic groups (i.e. disc degeneration or herniation, spinal stenosis, isthmic spondylolisthesis, spondylarthrosis and other disorders like non-specific back pain).
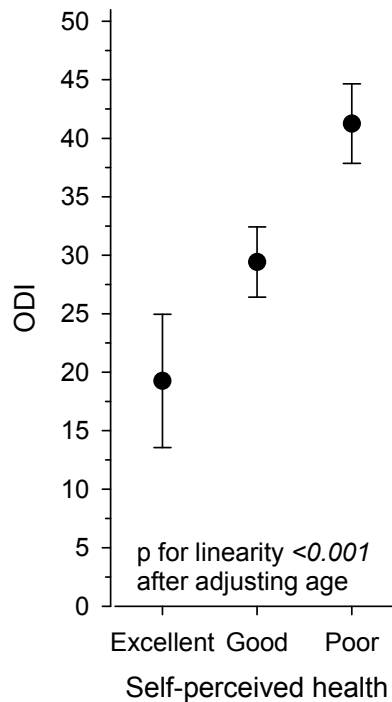
Figure 8. Relationship between the Oswestry Disability Index (ODI) and self-perceived health.

5.1.2   The ODI in the early recovery phase after spinal fusion surgery (II)

In the analysis of the early recovery phase (II), the preoperative mean ODI score was 45 (SD 17) and at the three-month follow-up the improvement was significant; the mean change was -19 (95% CI: -22 to -17). Preoperatively, 65% of patients had a total ODI score of 40 or more and three months postoperatively every fourth patient (25%) had a ODI score at this level.

When analyzing pain, the mean preoperative VAS_{back}, and VAS_{leg} values were 66 (SD 27) and 67 (SD 25), respectively. Mean postoperative change at three months was -42 (95% CI: -47 to -37) for VAS_{back} and -45 (95% CI: -51 to -40) for VAS_{leg}; these changes were statistically significant (p<0.001). The correlation between the changes in the ODI and VAS_{back} values was 0.56 (95% CI: 0.55 to 0.57) and

between the changes of the ODI and VAS_{leg} values 0.44 (95% CI: 0.44 to 0.45). For patients with an ODI score below 40 at three months after the fusion operation, the mean VAS_{back} value was 19 (SD 20) and patients having the ODI over 40, the mean VAS_{back} value was 40 (SD 26) (p<0.001). The corresponding values for VAS_{leg} were 15 (SD 20) and 45 (SD 30) (p<0.001).

5.1.3   The ODI in the ICF framework (II)

The ODI items were linked to the ICF framework. The linkage between the ODI items and the ICF is shown in Figure 5. The worst preoperative disability was in the Sensory functions and pain domains of the Body Functions and Structures component; the lowest disability was reported for the Self-care domain of the Activities component (Table 4).

Three months postoperatively, the improvements in all seven ICF domains were statistically significant. The effect sizes of the changes were large in all domains except in those of Self-care and Interpersonal interactions and relationships, which showed moderate effect-sizes. See Figure 9.

Table 4. Linkage between the Oswestry Disability Index (ODI) and the International Classification of Functioning, Disability and Health (ICF)

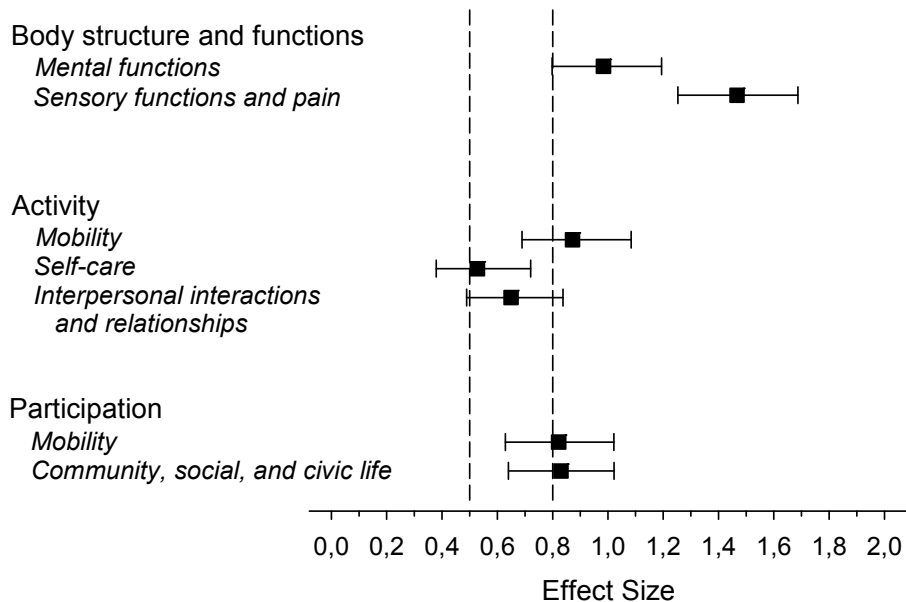| ICF Components Domains	ODI item	Preoperative, mean (SD)	Change at 3 months, mean (95% CI)	P-value
Body Functions and Structures				
Mental functions	7	36 (23)	-20 (-16 to -23)	<0.001
Sensory functions and pain	1	56 (24)	-31 (-27 to -35)	<0.001
Activities				
Mobility	3,4,5,6	51 (19)	-17 (-14 to -20)	<0.001
Self-care	2	25 (23)	-11 (-8 to -15)	<0.001
Interpersonal interactions and relationships	8	36 (36)	-21 (-16 to -26)	<0.001
Participation				
Mobility	10	46 (27)	-21 (-17 to -26)	<0.001
Community, social, and civic life	9	52 (25)	-22 (-17 to -27)	<0.001

Figure 9. Effect-sizes of changes in the Oswestry Disability Index (ODI) linked to the components and domains of the International Classification of Functioning, Disability and Health (ICF).

5.1.4 The ODI of the patients in comparison to the general population (III)

The one-year follow-up study (III) was performed for 252 patients (174 females and 78 males). The preoperative ODI score in the female patients was 47 (SD 16) and in the male patients 40 (SD 15). The general population sample of 682 individuals was age-, sex- and residential area-specific with the patients. Of the population sample, 67% were females. The mean ODI in the female population was 15 (SD 17) and in the male population 9 (SD 13) ($p<0.001$). One year post fusion the mean change in the ODI compared to the preoperative values was -25 (95% CI: -28 to -22) in females and -17 (95% CI -21 to -13) in males. Although a significant positive postoperative change in the ODI was observed in both sexes, the ODI of the patients was statistically significantly worse than the ODI of the general population at one year. The postoperative change in the ODI between

three months and one year was minor and not significant in males while in females the improvement was still significant (Figure 10).
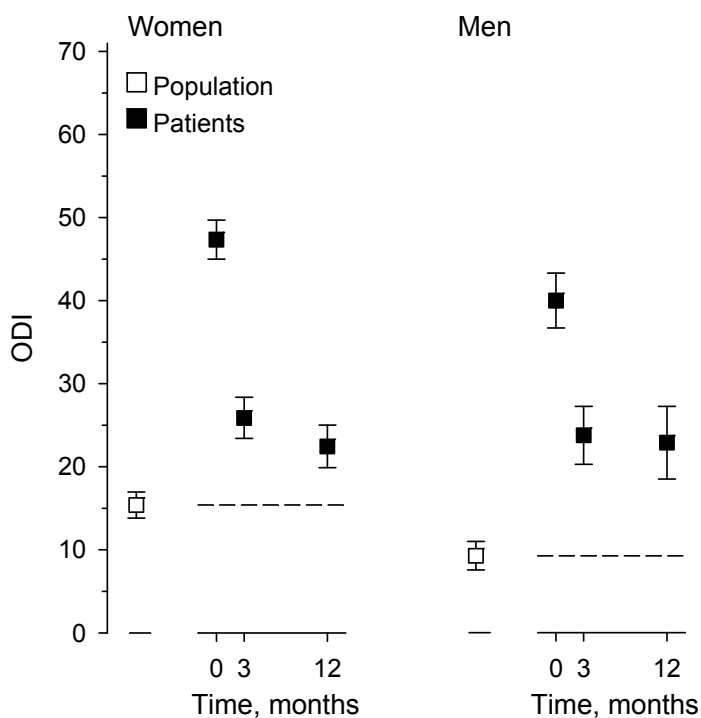

Figure 10. The mean (95% CI) Oswestry Disability Index (ODI) in the patients and in the general population (dashed line).

5.1.5 The ODI stratified by the surgical indication (IV)

In the two-year follow-up results (IV), shown in the Figure 11, the 242 patients were stratified by their surgical diagnostic indication into five groups. The groups were: degenerative olisthesis, isthmic olisthesis, spinal stenosis, degenerative disc disease or disc herniation and postoperative conditions. The ODI scores varied significantly between the groups over time (p<0.001). The ODI improved statistically significantly in all groups (p<0.001). The main change of the ODI was apparent at three months. During the 24-month follow-up, the mean positive change in the ODI ranged from 21 (95% CI: 18 to 24) in the degenerative olisthesis group to 27 (95% CI: 18 to 36) in the degenerative disc disease or disc

herniation group. The interaction between groups and time was not statistically significant (p=0.64)

At the 24-month follow-up the relationship between the change in the ODI compared to the preoperative values was also studied (Figure 12). Interestingly, only 16% of the change was explained by the baseline ODI values.
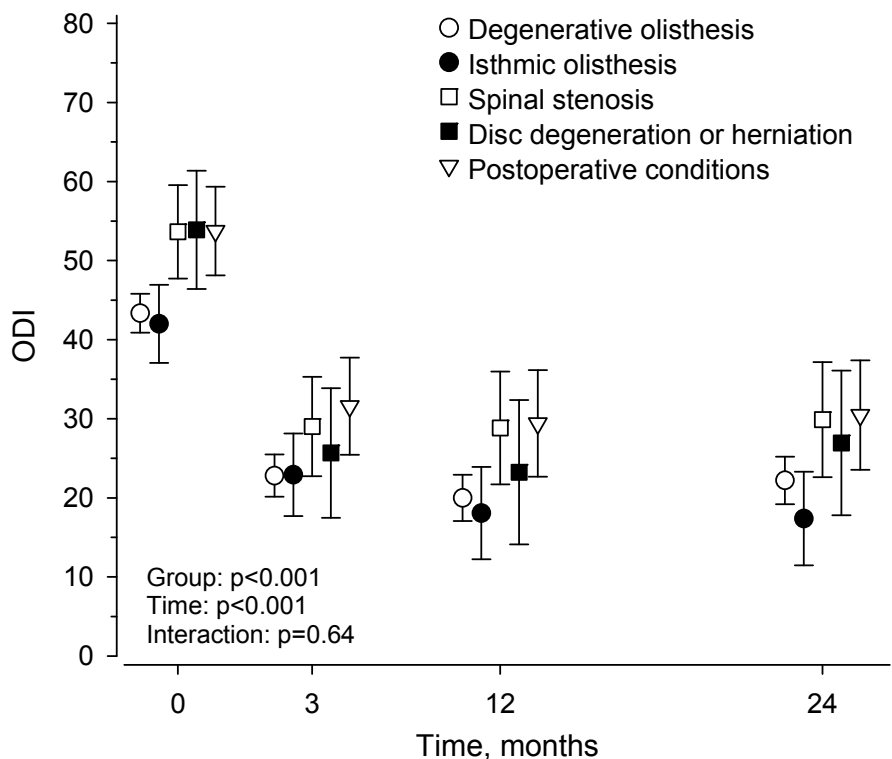

Figure 11. Oswestry Disability Index (ODI) scores stratified by surgical indication at the two-year follow-up. Values are based on linear mixed models.
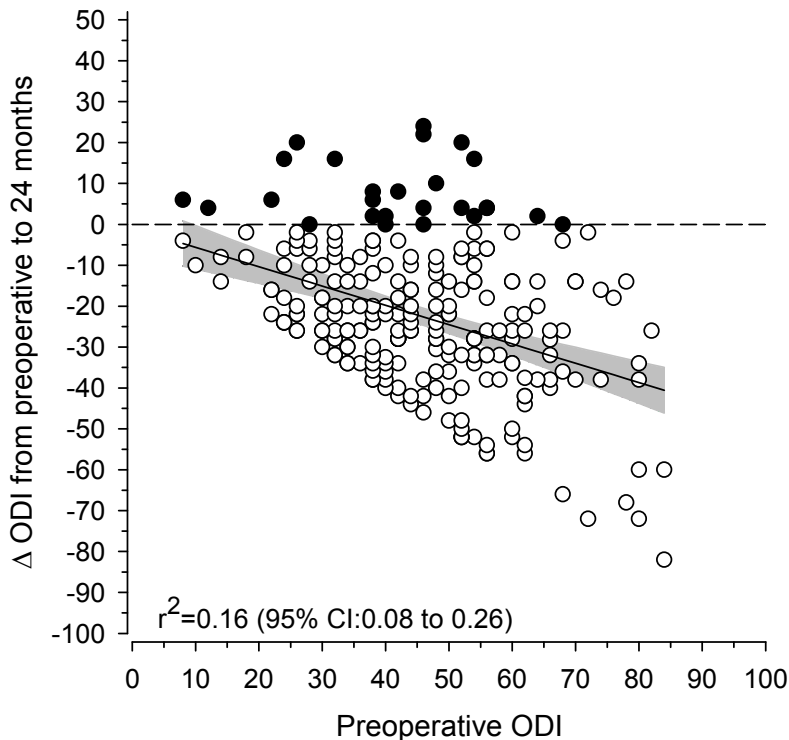
Figure 12. Change in the ODI compared to preoperative values. Black circles show worsened or stable patients and white circles improved patients. The grey area shows 95% confidence interval.

5.2 Health-Related Quality of Life (HRQoL)

5.2.1 The SF-36 in the early recovery phase after spinal fusion (II)

In the analysis of the HRQoL (II) through the SF-36 questionnaire, the poorest preoperative value was in the Role Physical dimension and highest in the Mental Health dimension. Positive changes in all eight dimensions at three months post surgery were statistically significant (Table 5).

The effect sizes as indicators of the responsiveness of the change were large in the dimensions of Bodily Pain, Physical Functioning, and Social Functioning;

medium in Vitality, Role-Physical and Mental Health and small in Role-Emotional and General Health (Table 5). When the eight scales of the SF-36 were aggregated into summary scores, preoperative PCS was 27 (95% CI: 26 to 28) and MCS was 47 (95% CI: 45 to 49). Postoperatively, the mean improvement was 9 (95% CI: 8 to 11) in PCS and 6 (95% CI: 4 to 7) in MCS; (p<0.001).

Table 5. Preoperative HRQoL dimensions (SF-36), their change and the effect-size of the change at three months.

Dimension	Preoperative, mean (SD)	Change at months 3, mean (95% CI)	P-value	Effect Size (95% CI)
Physical functioning	31 (22)	25 (22 to 29)	<0.001	1.08 (0.90 to 1.27)
Role physical	12 (24)	21 (15 to 28)	<0.001	0,69 (0.49 to 0.89)
Bodily Pain	25 (16)	30 (27 to 34)	<0.001	1,51 (1.30 to 1.76)
General health	54 (21)	4 (1 to 7)	0.005	0,20 (0.07 to 0.34)
Vitality	47 (23)	18 (15 to 21)	<0.001	0,78 (0.62 to 0.96)
Social functioning	51 (29)	22 (17 to 27)	<0.001	0,80 (0.61 to 1.02)
Role emotional	46 (43)	18 (12 to 24)	<0.001	0,44 (0.29 to 0.60)
Mental health	65 (22)	12 (9 to 15)	<0.001	0,57 (0.43 to 0.75)

5.2.2   Relationship between the ODI and the SF-36 (II)

The relationship between the changes in the total ODI score and the PCS and the MCS were assessed (Figure 13). After adjusting for age, gender, and duration of symptoms, the correlation between the change in the ODI and the PCS was -0.63 (95% CI: -0.70 to -0.54) and between the ODI and the MCS -0.35 (95% CI: -0.48 to -0.17). In 80% of patients, the change both in the total ODI and PCS scores was positive and in 64% of patients the change in both the total ODI and MCS scores was positive.
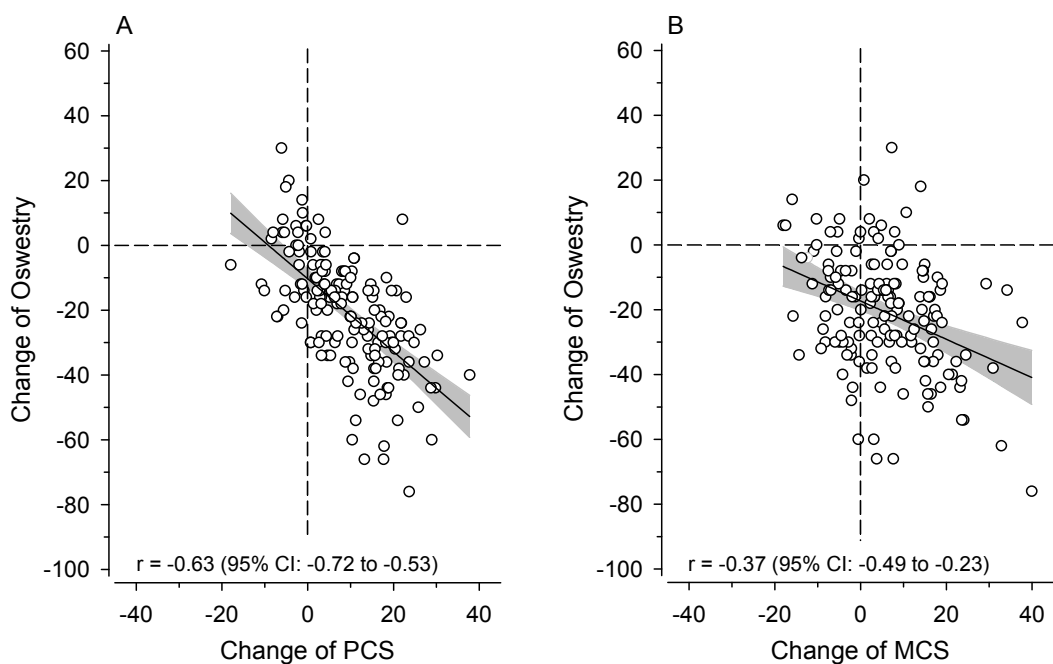
Figure 13. Relationship between the change in the Oswestry Disability Index and A) the change in the Physical Component Score (PCS) and B) the Mental Component Score (MCS) of the SF-36. Linear regression line is shown with 95% confidence intervals (grey area).

5.2.3   The SF-36 in comparison to the general population (III)

The scores for the eight dimensions of the SF-36 of the patients were compared with those of the general population (Table 6). Preoperatively in both sexes, all dimension were significantly better (p<0.001) in the general population than in the patients. The biggest difference, in both sexes, was in the dimension Role-Physical. At the one-year follow-up, the female patients reached the general population level in Vitality, Mental Health and Role-Emotional, while the male patients reached the population level only in Vitality and Mental Health.

Table 6. Health-Related Quality of Life in population and patients stratified by sex.

| SF-36 Dimensions | Population Mean (SD) | Patients | | Preoperative | Mean Ratio* (95% CI) | | |
		Preoperative Mean (SD)	12 months Mean (SD)		p-value patients vs. population preoperative	12 months	p-value patients vs. population 12 months
Female							
Physical Functioning	70 (28)	28 (19)	58 (29)	2.6 (2.3 to 3.0)	<0.001	1.3 (1.2 to 1.4)	<0.001
Role Physical	64 (42)	9 (21)	44 (43)	7.9 (2.7 to 13.0)	<0.001	1.5 (1.3 to 1.7)	<0.001
Bodily Pain	67 (27)	24 (15)	56 (25)	2.8 (2.4 to 3.3)	<0.001	1.2 (1.1 to 1.3)	<0.001
General Health	60 (22)	53 (20)	56 (21)	1.1 (1.1 to 1.2)	<0.001	1.1 (1.0 to 1.1)	0.022
Vitality	65 (23)	45 (22)	64 (23)	1.5 (1.3 to 1.6)	<0.001	1.0 (1.0 to 1.1)	0.41
Social Functioning	82 (25)	46 (28)	76 (28)	1.8 (1.7 to 2.0)	<0.001	1.1 (1.0 to 1.1)	0.004
Role Emotional	71 (39)	46 (43)	67 (41)	1.6 (1.4 to 1.8)	<0.001	1.1 (1.0 to 1.2)	0.17
Mental Health	77 (19)	63 (21)	77 (19)	1.2 (1.2 to 1.3)	<0.001	1.0 (1.0 to 1.0)	0.79
Male							
Physical Functioning	84 (22)	39 (20)	62 (26)	2.2 (1.9 to 2.4)	<0.001	1.3 (1.2 to 1.5)	<0.001
Role Physical	74 (38)	12 (21)	44 (43)	6.3 (2.1 to 10.5)	<0.001	1.7 (1.3 to 2.0)	<0.001
Bodily Pain	75 (23)	30 (16)	55 (29)	2.5 (2.1 to 3.0)	<0.001	1.4 (1.2 to 1.5)	<0.001
General Health	65 (21)	56 (21)	55 (23)	1.2 (1.1 to 1.3)	<0.001	1.2 (1.1 to 1.3)	<0.001
Vitality	71 (22)	53 (23)	66 (24)	1.4 (1.2 to 1.5)	<0.001	1.1 (1.0 to 1.2)	0.072
Social Functioning	87 (20)	62 (27)	75 (36)	1.4 (1.3 to 1.5)	<0.001	1.2 (1.1 to 1.3)	<0.001
Role Emotional	79 (35)	44 (43)	65 (42)	1.8 (1.5 to 2.2)	<0.001	1.2 (1.1 to 1.4)	0.003
Mental Health	81 (18)	68 (21)	76 (19)	1.2 (1.1 to 1.3)	<0.001	1.1 (1.0 to 1.2)	0.074

* Age adjusted

In the general population, the PCS of the SF-36 was 44 (SD 11) in females and 48 (SD 10) in males. Among the patients, the preoperative PCS was 26 (SD 7) in females and 29 (SD 6) in males. At 12 months post surgery, the improvement in the PCS was 11 (95%: CI 10 to 13) (p<0.001) in females and 10 (95% CI: 7 to 12; p<0.001) in males. In turn, the MCS of the SF-36 was 52 (SD 11) in females and 53 (SD 10) in males in the general population. The preoperative MCS was 46 (SD 13) in the female patients and 48 (SD 12) in the male patients. The positive change in the MCS from the preoperative to 12-month values was 7 (95% CI: 5 to 8; p<0.001) in females and 4 (95% CI: 1 to 6; p<0.001) in males. In the MCS, both the female (p=0.42) and male (p=0.61) patients reached the level of the general population at one year post surgery. In the PCS, however, the difference between the patients and the general population remained significant (both sexes p<0.001). See Figure 14.
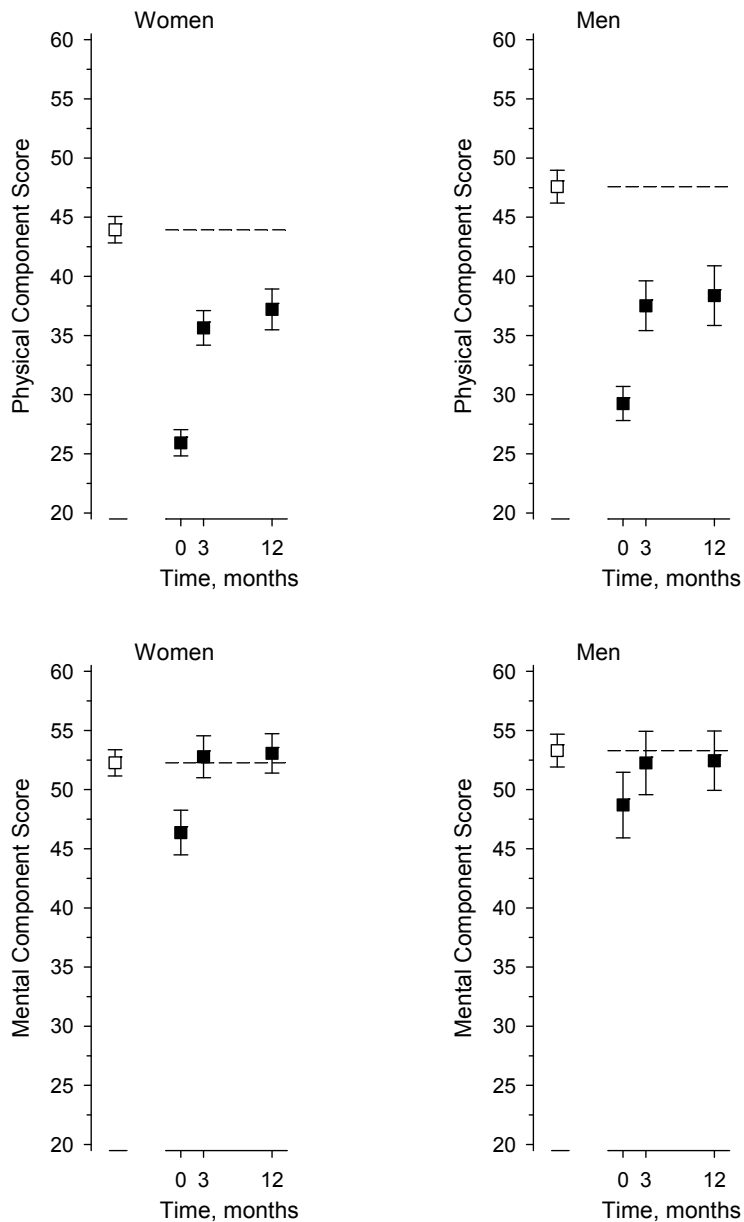
Figure 14. The change in the Physical Component Summary Score and Mental Component Summary Score of the SF-36 in patients (black square) compared with population sample (white square, dashed line). Results are expressed as means with 95% confidence interval.

5.2.4   The SF-36 stratified by surgical indication (IV)

At two years HRQoL was also assessed according to the surgical indication. The PCS was lowest preoperatively in the spinal stenosis and postoperative conditions groups, at 24 (SD 6) and 24 (SD 7), respectively, and highest in the isthmic spondylolisthesis group, at 31 (SD 8). Preoperatively, the groups differed from each other statistically significantly (p<0.001). The mean improvement in the PCS was statistically significant in all groups and varied from 9 (95% CI: 4 to 13) in the spinal stenosis group to 13 (95% CI: 7 to 18) in the degenerative disc disease or herniation group. In the PCS, the main effects of group and time were statistically significant, but their interaction was not (p=0.89). See Figure 15 A.

The preoperative MCS scores were similar between the groups (p=0.85). The changes during the 24-month follow-up varied from 3 (95% CI: -3 to 10) in the disc degeneration or herniation group to 6 (95% CI: 1 to 11) in the spinal stenosis group. The change in MCS was significant during the follow-up in the degenerative olisthesis group, spinal stenosis group and postoperative conditions group, but the change between all five groups did not differ (p=0.85). In the MCS, only the time effect was statistically significant (p<0.001). See Figure 15 B.
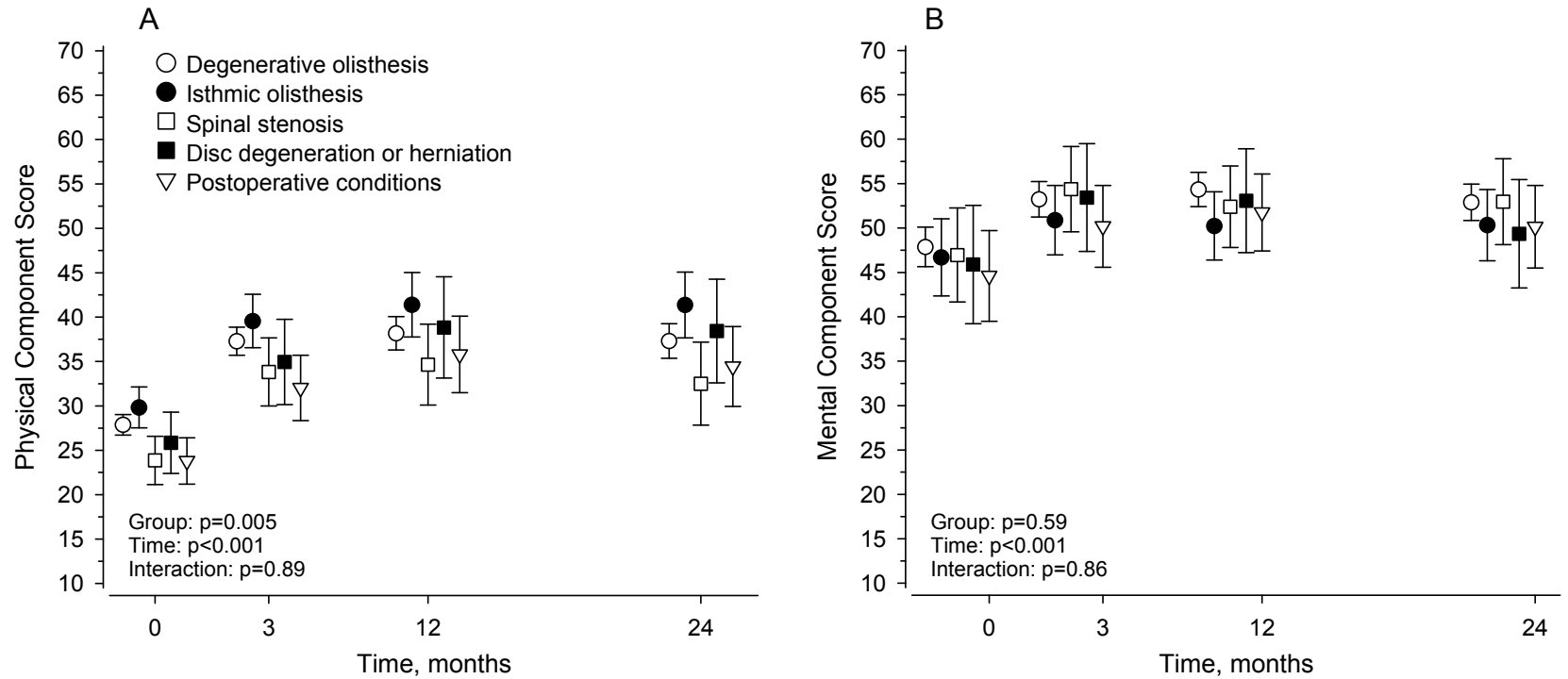
Figure 15 A, B. The Physical Component Summary Score (PCS) and the Mental Component Summary Score (MCS) of SF-36 stratified by surgical indication at two-year follow-up. Values are based on linear mixed model.

5.3 Health utility

5.3.1 The SF-6D in spinal fusion patients stratified by the surgical indication (IV)

In the health utility analysis, the patients were stratified by surgical indication. The five groups were degenerative olisthesis, isthmic olisthesis, spinal stenosis, degenerative disc disease or disc herniation and postoperative conditions. At two years the improvements in the SF-6D scores were significant in all groups. Furthermore, the changes in the SF-6D scores did not differ significantly between the groups (p=0.40). The main effects of group and time were statistically significant (p=0.023 and p<0.001, respectively) but the interaction was not significant (p=0.55). See Figure 16.
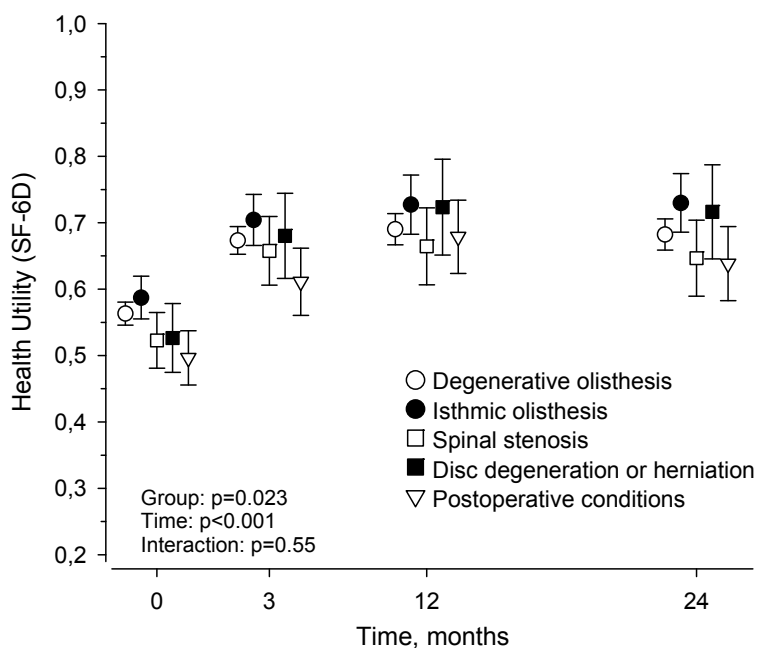

Figure 16. SF-6D scores stratified by the surgical indication in two-year follow-up. Values are based on linear mixed model.

5.3.2 Association between health utility and disability and patient´s perceived condition of back at two years (IV)

The change in the ODI explained 20% of the health utility value ($r^2=0.19$) (Figure 17). Statistically significant relationship emerged between health utility and patients´ own experience concerning the overall health status of back (Figure 18). At two years, 95% of all patients reported that they would have gone through the operation again, and there was no statistical difference observed between the indication groups ($p=0.38$). At two years, 80% of the patients indicated that the condition of their back was better than prior to the operation, 12% reported no change and 8% that the overall condition of their back was worse. There was no significant difference between the groups by surgical indication ($p=0.072$).
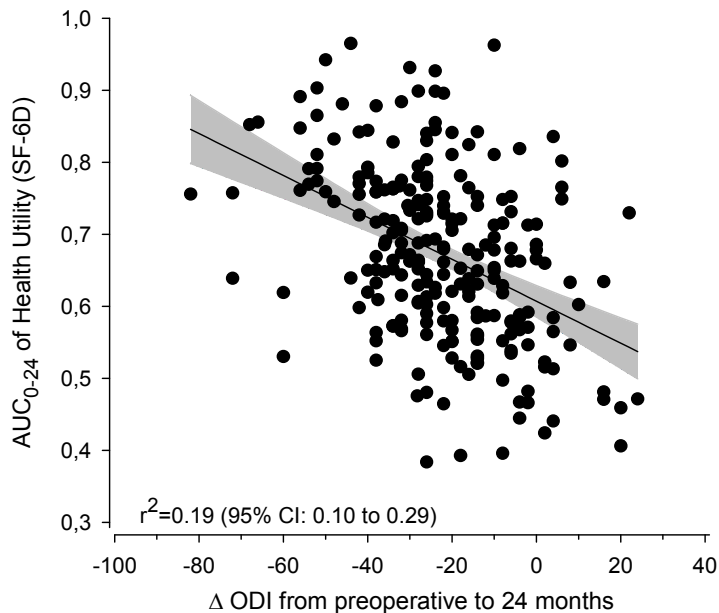

Figure 17. Relationship of change in the ODI from the preoperative to 24-month values in relation to health utility values. The grey area shows 95% Confidence Interval.
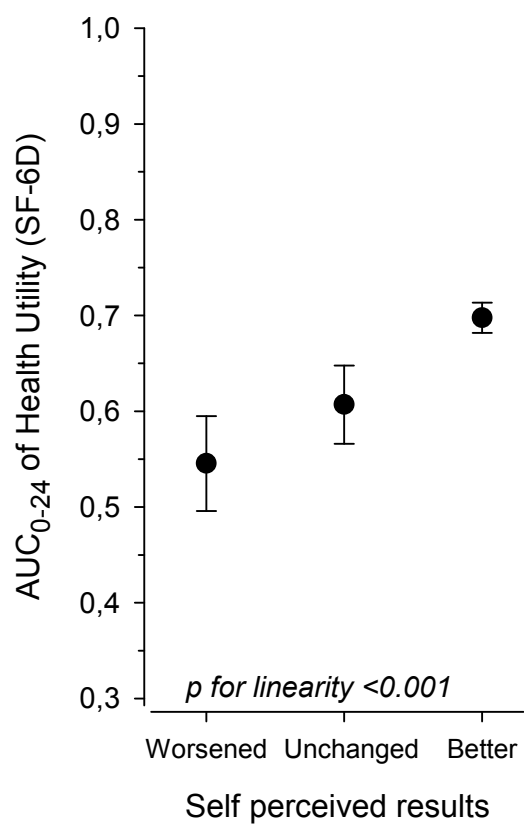
Figure 18. Self-perception of surgical outcome and health utility.

6.   Discussion

6.1   General aspects

Conservative treatment is always the primary way of treatment in back problems. Only small amount of patients end up to spinal fusion. In the health care districts of Tampere University Hospital and Jyväskylä Central Hospital the mean incidence of non-urgent adult spinal fusions has been 25 and 30 / 100 000 during the data collection for this study 2008-2012. Some form of disability also exists with the operatively treated patients and that makes it necessary to assess the outcome. The outcome assessment of spinal fusion surgery has moved towards PROs instead of judging the possible existence of solid fusion. Disability and HRQoL have established an important position as aspects of outcome. But along with limiting resources in health care and the fact that back issues cause an enormous economic burden to the community, health utility scores are also becoming more widespread. The global aim of this study was to assess the consequences of spinal fusion surgery in a prospective longitudinal spinal data base setting. Patients results were also compared to those of the general population.

6.2   Study population and study design

This study is mainly based on a uniform spinal database established in two Finnish hospitals: Tampere University Hospital and Jyväskylä Central Hospital. The collection of the data preoperatively, peroperatively and postoperatively has become a standard part of clinical practice. The data collection timepoints were strict: one to two weeks prior to the operation, and three months, one year and two years postoperatively. Co-ordinators in both hospitals collected and checked up the questionnaires from the patients. Prior to establishing the spinal database, as a part of this study, a lacking Finnish version 2.0 of the ODI was constructed for use as a tool for measuring disability.

This study is a prospective, longitudinal follow-up with very comprehensive sample of spinal fusion patients in two Finnish hospitals. The participants are highly representative of patients undergoing fusion procedures in Finland. The same group of surgeons who planned and built the spinal database originally also performed the operations. The surgical indications leading to the fusion operation have been fully discussed and have remained unchanged. This aspect of a well designed group of indications has been problematic in several previous studies. The outcome measurement tools used in this study are widely known and used in spinal surgery research. In the statistics, possible sources of errors, such as between the diagnostic indication groups, were noted. A characteristic of the Finnish health care system is national health insurance base and therefore the hospitals are covering the whole population of their catchment area and generally patients are not allowed to be referred to other hospitals. Due to this system, a hospital is responsible for the whole patient cohort and is not able to select patients or diagnoses. This results in the fact that very low number of patients was lost to follow-up. Another special feature is that the vast majority of the patients were willing to participate in the follow-up studies yielding a rate of participation almost 100%. A limitation of the present design is the proportion of patients who do not answer the questionnaires completely and therefore cause missing data.

The SF-36 version 1. was used in this study when assessing HRQoL. The questionnaire used in this study is the translated (Aalto et al 1999) and validated questionnaire RAND 36-Item Health Survey 1.0 (RAND-36). The original questionnaire is the SF-36 (Ware and Sherbourne 1992), but the Finnish translation, the RAND-36, is based on this exact replica of the SF-36 (Hays et al. 1993). However, because the RAND-36 uses different scoring algorithms for two of the 8 scales (Bodily pain, General Health), their results for those scales are not comparable with the standard SF-36. As the RAND-36 scoring is rarely used in international publications, in this thesis the original algorithms of the SF-36 were used.

In previous studies the minimal clinical important difference (MCID) has been used to assess the "smallest change that is important to patients" (Stratford et al. 1998). MCID values for the ODI have been reported in various studies. In a study of primary and revision fusion surgery the patients met a similar MCID value for the ODI of 12.4 (Carreon et al. 2013). In a study of 74 PLIF patients the MCID value of the ODI was 16 (Pollock et al. 2012). In a study of 454 patients from a large spinal database with the ODI as a disability score, 12.8 was determined as the MCID value. (Copay et al. 2008). Instead of the MCID the guidelines of Cohen for effect size were used to express the power of change in this research (Cohen 1988).

6.3   Psychometric properties of the ODI

The ODI is one of the most commonly used back-specific instruments both clinically and in the field of scientific work. It is easy and quick to complete and, for the surgeon, calculating the score is simple. The ODI is one of the measurement tools recommended by the WHO (Stucki and Sigl 2003).

The psychometric properties of the Finnish version 2.0 of the ODI were analyzed in the present study. The overall response rate for individual items was acceptable at 85%. No real floor or ceiling effect was present. Reproducibility as an indicator of reliability was assessed. The intraclass correlation coefficient (ICC) was 0.90 (95% CI: 0.85 to 0.94). The ICC was 0.83 in the earlier Finnish version of the ODI (Grönblad et al. 1993). In a recent validation study of the Hungarian version of the ODI, the ICC was 0.93 (Valasek et al. 2013). In the present study, the test-retest interval was 14 days and in the Hungarian study it was 9 days on average. A higher ICC was reported in Brazilian-Portugese (Vigatto et al. 2007) and in a Turkish (Yakut et al. 2004) ODI validation study: the ICCs were 0.99 and 0.94, respectively. The time periods in those studies were shorter, from one to seven days, and the memory-effect may have been influenced the results (Fairbank and Pynsent 2000). Cronbach´s alpha for internal consistency was 0.86, as compared to 0.76 reported in a previous study for the same version 2.0 of the ODI (Fisher and Johnson 1997). A Cronbach´s alpha higher than 0.80 is generally recommended (Streiner DL 2003), but if the value is too high, it may indicate that the items are too homogeneous.

Construct validity, in turn, was analyzed using several approaches in this study. As part of construct validity, convergent validity was assessed by testing the power of the relationship between the ODI and the Million-VAS; the correlation coefficient was acceptable 0.75 (95% CI 0.64 to 0.84). In previous reports a correlation coefficient of 0.77 was found between the ODI and the Roland Morris Questionnaire (Co et al. 1993) and a coefficient of 0.70 between the ODI and the Waddell Disability Index (Waddell and Main 1984). The correlation between the ODI and VAS_{back} and VAS_{leg} were moderate, at 0.48 (95% CI 0.32 to 0.62) and 0.41 (95% CI 0.23 to 0.57), respectively. In a study of the German version of the ODI, the correlation between the ODI and combined VAS back and leg was 0.78 (Mannion et al. 2006). Factor analysis showed that the ODI loaded on two factors, explaining 51% of the total variance.

6.4 Disability during the follow-up

The main finding was that the patients achieved a siginificant improvement in disability already in the early recovery phase three months after spinal fusion, and that this positive change was sustained up to the end of the two-year follow-up with no marked further change. A few studies have also reported early results for disability after spinal fusion operation. In a study of 30 patients undergoing a conventional or a mini-invasive fusion procedure, the preoperative mean ODI scores were 56 and 59 and at the three-month follow-up 32 and 16, respectively, and thus in favour of the mini-invasive technique (Rodriguez-Vela et al. 2009). In another study of 60 patients, an interbody fusion with either unilateral or bilateral pedicle screw fixation was performed. The preoperative mean ODI values were 49 (SD 15) and 50 (SD 13) in these two groups. After three months, the mean ODI in the unilateral group was 8 (SD 4) and in the bilateral group 5 (SD 4), and after six months the corresponding values were 9 (SD 7) and 11 (SD 9) (Mao et al. 2013). In a report on 195 fusion patients the disability outcome peaked its best at six weeks, with only a slight worsening during the rest of the two-year follow-up (Becker et al. 2010). In a study of 68 patients undergoing instrumented spinal fusion, the ODI values were reported at six months and one year (Oestergaard et al. 2012). The mean improvement in the ODI at six months were -6 and -15 in two groups, while at one year the mean ODI improvements were -5 and -20. The mean baseline ODI values were 40 and 44.

The overall trend in the present study is, that the patients benefit from the surgery concerning disability. Previous studies have also reported improvement in disability in several diagnostic entities. In a Finnish study of surgical treatment of spinal stenosis the surgical group consisting both decompressions and fusions had better ODI scores both at two years and six years postoperatively (Malmivaara et al. 2007, Slätis et al. 2011). At two years the mean ODI difference was 8 and at six years 10 in favour for surgery. In a study of both randomized and observational cohorts with degenerative olisthesis patients Weinstein et al. have reported the improvement in the ODI score of 22 in surgical and 9 in non-surgical patients at four years from the baseline score of 43 (Weinstein et al. 2009). Several studies of total disc replacement (TDR) have compared the outcome in disability with fusion. In a RCT study of 304 patients the TDR was compared to the instrumented anterior interbody fusion (Blumenthal et al. 2005). At two years the ODI score of the TDR group was 26 and of the fusion group 30 and this difference was not statistically significant. In a very recent RCT study of 152 chronic low back pain patients the ODI score of TDR patients at five years was significantly better than

that of fusion patients. The ODI scores were 17 and 23, respectively (Sköld et al. 2013). The results of the present study at the two-year follow-up were stratified by surgical indication for fusion: degenerative olisthesis, isthmic olisthesis, spinal stenosis, degenerative disc disease or disc herniation and postoperative conditions including pseudoarthrosis and postoperative instability. The groups differed significantly in disability preoperatively; the lowest disability in the ODI was measured in the isthmic olisthesis group, with a mean ODI of 40 (SD 16), and the highest in the spinal stenosis group, with a mean ODI of 53 (SD 18) and the postoperative conditions groups, also with a mean ODI of 53 (SD 15). The lowest disability in the isthmic olisthesis group was most probably caused by the younger age. At two years, biggest improvement was observed in the degenerative disc disease or disc herniation group and smallest in the degenerative olisthesis group, with ODIs varying from -27 to -21. That the change in disability was biggest in the degenerative disc disease or disc herniation group, was surprising. One explanation might be that this group was the smallest, with 15 patients, and hence further evaluation with a larger number of patients would be needed to confirm this result. In all groups, the main improvement was gained within the first three months, after which the results remained rather stable up to the end of the two-year follow-up. In a previous prospective study of Glassman et al., 428 decompression and posterolateral fusion patients were assigned to eight indication groups (Glassman et al. 2009). The number of patients in the groups ranged from 17 in the scoliosis group to 80 in the spondylolisthesis group. The number of the patients in the scoliosis group was rather small. That was also the case in the present study, and this is why this group was excluded from the two-year analysis in the present research. The biggest mean change in the study of Glassman et al. at two years was in the scoliosis group and the smallest in the nonunion group. In that study the group "spondylolisthesis" comprised both isthmic and degenerative olisthesis patients, which is possible to cause misleading since these groups are entirely different by age (Glassman et al. 2009). Specifying the surgical indication when reporting the outcome is critical in seeking to develop the evidence base for spinal fusion surgery (Glassman et al. 2009). Our study supports this view; there are major differences between the indication groups in, for example, demographic details like age. The mean age in isthmic olisthesis groups was markedly lower than, for example in degenerative olisthesis groups and therefore also the aims of surgery are entirely different between these groups. Stratifying fusion patients is a demanding task for the researcher: the number of patients in the study must be large enough and the strategy in determining the diagnostic indication must be solid, which in turn requires much planning and discussion with the surgeons involved in the patient recruitment process.

6.5 The ODI in the ICF framework

The ICF offers a unified and standard language and framework for the description of functioning, disability, health and health-related states (Andresen 2000, Andresen et al. 2000, Dijkers et al. 2000, Gray and Hendershot 2000, WHO 2001). It enables classification of the patient´s perspective at both the personal and population level (Sigl et al. 2006). Comparison of outcome instruments may be an unmanageable task, but the use of the ICF as a system linking interventions and outcome measurement tools may offer considerable benefits when comparing different investigations and their results (Cieza et al. 2005, Sigl et al. 2006). The novel contribution to the literature is that, for the first time, linkage between the ICF and the ODI has been studied in a real patient cohort three months after spinal fusion surgery.

In this study, the worst preoperative disabilities were pain (item 1) in the Sensory functions and pain domain in the Body Functions and Structures component, lifting, walking, sitting and standing (items 3,4,5,6) in the Mobility domain in the Activities component, and social function (item 9) in the Community, social, and civic life domain in the Participation component. In earlier studies, the findings concerning the most problematic sectors have been similar (Røe et al. 2008).

In the present study, the effect-size of the change in disability was largest in the Body Functions and Structures and also large in the Participation component and in the Mobility domain in the Activities components. The corresponding ODI items having the large effect-sizes are sleeping, pain, lifting, walking, sitting, standing, social life and travelling. Similar findings were reported by Djurasovic et al., according to which pain, walking and social life were the items that showed a moderate to large effect-size in change in disability after fusion surgery (Djurasovic et al. 2012). The important finding is that the ODI is comprehensive and able to distinguish different domains from each other. Previously, it has been pointed out that the ODI is comparable in physical aspects but less comparable in emotional aspects (Sigl et al. 2006). Therefore, since the ODI does not cover all the categories that exist in the core set for low back pain, clinical work requires combining clinical judgement to the data provided by the measurement tools (Røe et al. 2008).

6.6   HRQoL during the follow-up

In the literature, only a few studies have reported early recovery phase values for HRQoL. In the present study, the positive changes in all the domains of SF-36 were significant three months after spinal fusion. Bodily pain improved by 30 (95% CI: 27 to 34) from the mean preoperative value of 25 (SD 16) and Physical functioning by 25 (95% CI: 22 to 29) from the preoperative value of 31 (SD 22). Radcliff et al. reported variation in preoperative Bodily pain from 29 (SD 16) in patients with no revision surgery to 30 (SD 15) in patients with revision. The Physical functioning values of the two groups were 31 (SD 22) and 31 (SD 21), respectively. At three months, the improvement in Bodily pain was approximately 30 in patients with no revision and approximately 13 in patients with revision. In Physical functioning the corresponding changes were approximately 25 and 11 (Radcliff et al. 2013). The mean preoperative PCS was 27 (95% CI: 26 to 28) and MCS 47 (95% CI: 45 to 49). The mean improvement at three months was 9 (95% CI: 8 to 11) in PCS and 6 (95% CI: 4 to 7) in MCS. In the study of Saban et al. with 57 patients, of whom 13 had spinal fusion, the patients showed a preoperative value for the SF-12 PCS 29 (SD 8) and at three months postoperatively a value of 39 (SD 12); in the SF-12 MCS the corresponding values were 46 (SD 12) and 50 (SD11) (Saban et al. 2007). In a study of 283 patients with spinal deformity, the SF-12 PCS improved from the baseline value of 34 to 37 at 6 months and the SF-12 MCS from 51 to 52 (Glassman et al. 2009). However, in that study the changes between 6 and 12 months were not significant in either the PCS or MCS (p=0.06 and 0.26, respectively). In a study of 224 patients, in which the patients were divided into the groups of younger than 65 or 65 and older, the PCS at 6 months after single-level posterolateral fusion improved by 10 (SD 11) from the baseline score of 28 in the younger group and 14 (SD 11) from the mean baseline value of 25 in the older group (Glassman et al. 2009). Further, in a study of elderly patients undergoing spinal fusion, the best results in HRQoL were achieved already at six weeks, with only subtle deterioration observed during the rest of the two-year follow-up (Becker et al. 2010). Therefore this study confirms previous results that a significant change in HRQoL is already visible at three months postoperatively and that the change thereafter is minimal.

The relationship between HRQoL and disability was assessed in this study in the early recovery phase. There was simultaneous improvement in the ODI and the PCS more often (in 80% of cases), than in the ODI and the MCS (in 64% of the cases). The correlation between the change in the ODI and the change in the PCS was r=-0.63, while it was r=-0.37 between the change in the ODI and the change

in the MCS. This suggests that the ODI, as a measure of disability, and the physical part of the HRQoL, the PCS, express notable agreement in assessing the change in patients´ health status in the early recovery phase. Somewhat parallel findings have been reported previously in a two year follow-up study of 1,104 patients undergoing decompression and instrumented lumbar fusion (Djurasovic et al. 2012). The association between disability and HRQoL was analyzed at the end of the follow-up. In that study the items of pain, walking and social life showed a moderate correlation with the PCS, and these items were also the ones that independently predicted the physical aspect of the quality of life.

At two years after spinal fusion, the HRQoL values were also assessed after stratification by surgical indication. The preoperative PCS values differed between the groups significantly; the spinal stenosis and postoperative group showed the lowest values and the isthmic olisthesis group the highest. All groups improved during the follow-up, and there was no significant difference between the groups in the amount of change. The preoperative MCS values did not differ significantly between the groups and neither did the changes. The trend in the change in the HRQoL values in all groups was obvious, while the main improvement was achieved in the early recovery phase in all groups and this improvement seemed to be sustained through the two-year follow-up period.

6.7 Disability and HRQoL in comparison to the general population

With respect to disability, the spinal fusion patients did not reach the level of the general population despite the significant improvement from the preoperative level, in either females or males. According to the original grading of the ODI (Fairbank et al. 1980), the general population also showed minimal disability; the mean ODI of the female individuals was 15 (SD 17) and that of the male 9 (SD 13). To our knowledge, no data on the scores of disability in the general population have been reported. A possible explanation for the difference between the patients and the general population is that, most often, the longstanding back symptoms have also caused permanent changes in patients´ lives and actions. While it is unrealistic to expect the total absence of disability post surgery, the majority of the surgical patients experience reasonable improvement. It is noteworthy, however, that in some patients the disability remains at a severe level and this is already evident three months after surgery.

The HRQoL values were compared to those of the general population at one year after spinal surgery in our study. The significant improvement in the PCS from the preoperative value 26 (SD 7) in females and 29 (SD 6) in males to 37 and 39 in females and males, respectively, was not enough to match those of the general population, whose PCS was 44 (SD 11) in females and 48 (SD 10) in males. However, the MCS of the patients reached the general population values in both females and males. The MCS in the population sample was 52 (SD 11) in females and 53 (SD 10) in males and the corresponding values in the patients at one year were 53 in females and 52 in males. In an Australian study of 105 fusion patients (Mokhtar et al. 2010), the HRQoL (SF-12) values were compared to the population norms of South Australia (Avery et al. 2003). In that study the SF-12 PCS was 44 (43-46) in the population while for the patients the mean preoperative PCS 28 (27-30) improved to 39 (37-42) during the mean follow-up of 24 months. The preoperative SF-12 MCS of the patients was 47 (46-50) and after one year 52 (50-55), thereby reached the population level 54 (53-55). The authors indicate, that from substantial preoperative impairment, the patients HRQoL levels return to resemble the age-matched population norms (Mokhtar et al. 2010). In conclusion, although all the variables improve significantly after surgery in the present study, physical impairment measured by the ODI or PCS seems to remain below the level of the general population, while mental health recovery attains the level of the general population and indicates improvement in an important component of HRQoL.

6.8 The SF-6D in spinal fusion patients stratified by surgical indication

Health utility was assessed in five different patient groups, all of which irrespective their surgical indication, gained in health utility. This can be interpreted, from the social perspective to mean that surgical intervention is beneficial. The health utility values assessed with the SF-6D instrument varied preoperatively in the five different indication groups from 0.50 (SD 0.10) in the postoperative consequences group to 0.59 (SD 0.12) in the isthmic olisthesis group. The difference in the preoperative utility values between the groups was significant. The changes varied from 0.12 to 0.19 during the two year follow-up. The rates of change did not differ between the groups, and all groups showed significant improvement. The degenerative olisthesis and spinal stenosis groups showed the smallest and the disc

pathology group the biggest change. The correlation between disability and utility was assessed and the change in the ODI explained 20% of the utility ($r^2=0.19$).

Two previous studies have reported on health utility in spinal fusion patients. In a study of 1,104 patients in eight different groups, the preoperative values were at the same level as in the present study, but the changes in all groups were somewhat lower during the two-year follow-up (Carreon et al. 2012). The lowest preoperative mean SF-6D in the study by Carreon et al. was 0.492 (SD 0.075) in non-union patients, while in the present study the patients with postoperative conditions had a mean preoperative SF-6D value of 0.50 (SD 0.10). In the earlier study, the mean changes ranged from 0.050 (0.110) in the nonunion group of revision patients to 0.088 (SD 0.112) in the group of spinal stenosis patients (Carreon et al. 2012). The mean improvement in the patients with postoperative conditions in the present study was 0.15 (95% CI: 0.09 to 0.20) at two years, indicating a somewhat higher gain in health utility. In a study of 80 patients with a single-level PLF, the mean preoperative SF-6D was 0.500. Health utility showed gradual improvement during the five-year follow-up, and at two years the mean SF-6D was 0.638 (Glassman et al. 2012).

6.9   Clinical implications and future directions

The results of the present study indicate that spinal fusion is a beneficial surgical procedure for patients with various indications. The positive changes are significant in terms of disability, HRQoL and health utility regardless of surgical indication in the current practise. Moreover these positive changes are already present in the early phase of recovery three months after the operation. Nevertheless, at one year, patients show more restrictions in the physical aspect of recovery, both in terms of disability and HRQoL compared to the general population. However, in the mental aspect of HRQoL, the patients reach the level of the population.

As a part of this thesis, the ODI was translated into the Finnish language and its psychometric properties were tested. The Finnish ODI version 2.0 proved to be reliable, valid and feasible for use in both clinical work and research.

Over the last few decades, PROs have increasingly been used in evaluating the outcome of fusion surgery instead of, for example, radiological evaluation of the presence or absence of a solid fusion. The routine-like administration of some of the outcome instruments has been recommended. In clinical work, the ODI should be used before and after surgical treatment and the use of the HRQoL instrument should be used when evaluating surgical outcomes in the clinical-research setting (Chapman et al. 2011, DeVine et al. 2011).

The results of this study can be utilized at the individual level in the evaluation of treatment, and may even be helpful in treatment planning and also when giving a patient information concerning the results of surgery at the time of treatment decision. At the institutional level, the results can be utilized in the planning and development of treatment options and health care services among health care providers, and also for educational purposes. At the society level, the results provide information and understanding for policy-makers. Knowledge of the favourable outcomes of these high cost binding spinal fusion operations is an ongoing need at a time when limitations of resources are complicating the focusing of interventions in the field of health care.

7. Conclusions

The main findings of the present study can be summarized as follows:

1) Significant improvement both in disability and quality of life was seen already in the early recovery phase three months after surgery, and this improvement was maintained during the two-year follow-up period.
2) Fusion patients did not entirely achieve the level of the general population in disability or in the physical aspect of quality of life during the first year after the fusion surgery.
3) All the surgical indication groups gained in health utility after spinal fusion.
4) The Finnish Oswestry Disability Index (ODI) version 2.0 is a reliable, valid and feasible measurement tool for assessing the outcome of spinal fusion procedure.

Acknowledgements

References

Aalto A-M. Aro AR, Teperi J. RAND-36 as a measure of health-related quality of life. reliability, construct validity and reference values in the Finnish general population. 1999; STAKES Research Reports 101.

Adogwa O, Parker S L, Shau D, Mendelhall S K, Cheng J, Aaronson O, Devin C J, McGirt M J. Long-term outcomes of revision fusion for lumbar pseudarthrosis: Clinical article. J Neurosurg Spine 2011; 15 (4): 393-8.

Adult Low Back Disease (online). Current care guidelines. Working group set by the Finnish Medical Society Duodecim and the Finnish Society of Physiatrists. Finnish Medical Society Duodecim, 2008 (referred 16 june 2009). Available online at: www.kaypahoito.fi.

Agarwal V J, Golish R K,D., Alamin T F. Results of bacterial culture from surgically excised intervertebral disc in 52 patients undergoing primary lumbar disc microdiscectomy at a single level. Spine J 2010; 10: S45-S46.

Airaksinen O, Brox J I, Cedraschi C, Hildebrandt J, Klaber-Moffett J, Kovacs F, Mannion A F, Reis S, Staal J B, Ursin H, Zanoli G. COST B13 Working Group on Guidelines for Chronic Low Back Pain. Chapter 4. European guidelines for the management of chronic nonspecific low back pain. Eur Spine J 2006; 15 Suppl 2: S192-300.

Albert H B, Kjaer P, Jensen T S, Sorensen J S, Bendix T, Manniche C. Modic changes, possible causes and relation to low back pain. Med Hypotheses 2008; 70 (2): 361-8.

Albert H B, Sorensen J S, Christensen B S, Manniche C. Antibiotic treatment in patients with chronic low back pain and vertebral bone edema (Modic type 1 changes): A double-blind randomized clinical controlled trial of efficacy. Eur Spine J 2013; 22 (4): 697-707.

Anderson G B J. The epidemiology of spinal disorders. In: The adult spine: Principles and practise (Ed Frymoyer J W). Lippincott-Raven Publishers Philadelphia, PA 1997; 93-141.

Andersson G B. Epidemiological features of chronic low-back pain. Lancet 1999; 354 (9178): 581-5.

Andresen E M. Criteria for assessing the tools of disability outcomes research. Arch Phys Med Rehabil 2000; 81 (12 Suppl 2): S15-20.

Andresen E M, Lollar D J, Meyers A R. Disability outcomes research: Why this supplement, on this topic, at this time? Arch Phys Med Rehabil 2000; 81 (12 Suppl 2): S1-4.

Avery J,Grande E D, Taylor A. Quality of life in South Australia as measured by the SF-12 health status questionnanires: Population research and outcome studies unit. City of Adelaide, South Australia: Department of Human Service 2003.

Baker J D, Pynsent P B, Fairbank J C T. The Oswestry Disability Index revisited: Its reliability, repeatibility and validity, and a comparison with the St.Thomas´s disability index. In: Back pain: New approaches to rehabilitation and education. (Eds Roland M O, Jenner J R). University press Manchester, England 1989: 174-86.

Barton G R, Sach T H, Avery A J, Jenkinson C, Doherty M, Whynes D K, Muir K R. A comparison of the performance of the EQ-5D and SF-6D for individuals aged ≥ 45 years. Health Econ 2008; 17 (7): 815-32.

Beaton D E, Bombardier C, Guillemin F, Ferraz M B. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine 2000; 25 (24): 3186-91.

Becker P, Bretschneider W, Tuschel A, Ogon M. Life quality after instrumented lumbar fusion in the elderly. Spine (Phila Pa 1976) 2010; 35 (15): 1478-81.

Bener A, Verjee M, Dafeeah E E, Falah O, Al-Juhaishi T, Schlogl J, Sedeeq A, Khan S. Psychological factors: Anxiety, depression, and somatization symptoms in low back pain patients. J Pain Res 2013; 6: 95-101.

Berg S, Tullberg T, Branth B, Olerud C, Tropp H. Total disc replacement compared to lumbar fusion: A randomised controlled trial with 2-year follow-up. Eur Spine J 2009; 18 (10): 1512-9.

Bharmal M, Thomas J, 3rd. Comparing the EQ-5D and the SF-6D descriptive systems to assess their ceiling effects in the US general population. Value Health 2006; 9 (4): 262-71.

Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. J Health Econ 2002; 21 (2): 271-92.

Brazier J, Usherwood T, Harper R, Thomas K. Deriving a preference-based single index from the UK SF-36 health survey. J Clin Epidemiol 1998; 51 (11): 1115-28.

Brazier J E, Roberts J. The estimation of a preference-based measure of health from the SF-12. Med Care 2004; 42 (9): 851-9.

Brox J I, Sørensen R, Friis A, Nygaard Ø, Indahl A, Keller A, Ingebrigtsen T, Eriksen H R, Holm I, Koller A K, Riise R, Reikerås O. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. Spine (Phila Pa 1976) 2003; 28 (17): 1913-21.

Carragee E J, Cheng I. Minimum acceptable outcomes after lumbar spinal fusion. Spine J 2010; 10 (4): 313-20.

Carreon L Y, Berven S H, Djurasovic M, Bratcher K R, Glassman S D. The discriminative properties of the SF-6D compared to the SF-36 and ODI. Spine (Phila Pa 1976) 2013; 38(1): 60-64.

Carreon L Y, Bratcher K R, Canan C E, Burke L O, Djurasovic M, Glassman S D. Differentiating minimum clinically important difference for primary and revision lumbar fusion surgeries. J Neurosurg Spine 2013; 18 (1): 102-6.

Carreon L Y, Djurasovic M, Canan C E, Burke L O, Glassman S D. SF-6D values stratified by specific diagnostic indication. Spine (Phila Pa 1976) 2012; 37 (13): E804-8.

Carreon L Y, Glassman S D, Djurasovic M, Dimar J R, Johnson J R, Puno R M, Campbell M J. Are preoperative health-related quality of life scores predictive of clinical outcomes after lumbar fusion? Spine (Phila Pa 1976) 2009; 34 (7): 725-30.

Carreon L Y, Glassman S D, Kantamneni N R, Mugavin M O, Djurasovic M. Clinical outcomes after posterolateral lumbar fusion in workers' compensation patients: A case-control study. Spine (Phila Pa 1976) 2010; 35 (19): 1812-7.

Cartwright W S. Methods for the economic evaluation of health care programmes, second edition. by Michael F. Drummond, Bernie O'Brien, Greg L. Stoddart, George W.

Torrance. Oxford: Oxford university press, 1997. J Ment Health Policy Econ 1999; 2 (1): 43.

Cassisi J E, Robinson M E, O'Conner P, MacMillan M. Trunk strength and lumbar paraspinal muscle activity during isometric exercise in chronic low-back pain patients and controls. Spine (Phila Pa 1976) 1993; 18 (2): 245-51.

Cavanaugh J M, Lu Y, Chen C, Kallakuri S. Pain generation in lumbar and cervical facet joints. J Bone Joint Surg Am 2006; 88 Suppl 2: 63-7.

Cella D F, Bonomi A E. Measuring quality of life: 1995 update. Oncology (Williston Park) 1995; 9 (11 Suppl): 47-60.

Chapman J R, Norvell D C, Hermsmeyer J T, Bransford R J, DeVine J, McGirt M J, Lee M J. Evaluating common outcomes for measuring treatment success for chronic low back pain. Spine (Phila Pa 1976) 2011; 36 (21 Suppl): S54-68.

Cho W, Lenke L G, Bridwell K H, Dorward I G, Shoda N, Baldus C R, Cho S K, Kang M M, Zebala L P, Pahys J M, Koester L A. Comparison of spinal deformity surgery in patients with non-insulin-dependent diabetes mellitus (NIDDM) versus controls. Spine (Phila Pa 1976) 2012; 37 (16): E978-84.

Chou R, Baisden J, Carragee E J, Resnick D K, Shaffer W O, Loeser J D. Surgery for low back pain: A review of the evidence for an American Pain Society clinical practice guideline. Spine (Phila Pa 1976) 2009; 34 (10): 1094-109.

Chou R, Huffman L H. American Pain Society, American College of Physicians. Medications for acute and chronic low back pain: A review of the evidence for an American Pain Society/American College of Physicians clinical practice guideline. Ann Intern Med 2007; 147 (7): 505-14.

Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustun B, Stucki G. ICF linking rules: An update based on lessons learned. J Rehabil Med 2005; 37 (4): 212-8.

Cloward R B. The treatment of ruptured lumbar intervertebral discs by vertebral body fusion. I. Indications, operative technique, after care. J Neurosurg 1953; 10 (2): 154-68.

Co Y Y, Eaton S, Maxwell M W. The relationship between the St. Thomas and Oswestry disability scores and the severity of low back pain. J Manipulative Physiol Ther 1993; 16 (1): 14-8.

Cobo Soriano J, Sendino Revuelta M, Fabregate Fuente M, Cimarra Diaz I, Martinez Urena P, Deglane Meneses R. Predictors of outcome after decompressive lumbar surgery and instrumented posterolateral fusion. Eur Spine J 2010; 19 (11): 1841-8.

Cohen J. Statistical power analysis for the behavioral sciences. Rev ed. Academic Press, New York 1988.

Connor A C, Rooney J A, Carroll J P. Anterior lumbar fusion. A technique combinig intervertebral and intravertebral body fixation. Surg Clin North Am 1967; 47(1): 231-7.

Copay A G, Glassman S D, Subach B R, Berven S, Schuler T C, Carreon L Y. Minimum clinically important difference in lumbar spine surgery patients: A choice of methods using the Oswestry Disability Index, medical outcomes study questionnaire short form 36, and pain scales. Spine J 2008; 8 (6): 968-74.

Corsia M F, Wack M, Denys G. Low virulence bacterial infections of intervertebral discs and the resultant spinal disease processes. Abstract from Scoliosis Research Society (SRS) annual meeting 2003.

Coste J, Delecoeuillerie G, Cohen de Lara A, Le Parc J M, Paolaggi J B. Clinical course and prognostic factors in acute low back pain: An inception cohort study in primary care practice. BMJ 1994; 308 (6928): 577-80.

de Rooij A, Steultjens M P, Siemonsma P C, Vollebregt J A, Roorda L D, Beuving W, Dekker J. Overlap of cognitive concepts in chronic widespread pain: An exploratory study. BMC Musculoskelet Disord 2011; 12: 218,2474-12-218.

DeVine J, Norvell D C, Ecker E, Fourney D R, Vaccaro A, Wang J, Andersson G. Evaluating the correlation and responsiveness of patient-reported pain with function and quality-of-life outcomes after spine surgery. Spine (Phila Pa 1976) 2011; 36 (21 Suppl): S69-74.

Deyo R A, Battie M, Beurskens A J, Bombardier C, Croft P, Koes B, Malmivaara A, Roland M, Von Korff M, Waddell G. Outcome measures for low back pain research. A proposal for standardized use. Spine 1998; 23 (18): 2003-13.

Dijkers M P, Whiteneck G, El-Jaroudi R. Measures of social outcomes in disability research. Arch Phys Med Rehabil 2000; 81 (12 Suppl 2): S63-80.

Dimar J R,2nd, Glassman S D, Burkus J K, Pryor P W, Hardacker J W, Carreon L Y. Two-year fusion and clinical outcomes in 224 patients treated with a single-level instrumented posterolateral fusion with iliac crest bone graft. Spine J 2009; 9 (11): 880-5.

Dionne C E, Dunn K M, Croft P R, Nachemson A L, Buchbinder R, Walker B F, Wyatt M, Cassidy J D, Rossignol M, Leboeuf-Yde C, Hartvigsen J, Leino-Arjas P, Latza U, Reis S, Gil Del Real M T, Kovacs F M, Oberg B, Cedraschi C, Bouter L M, Koes B W, Picavet H S, van Tulder M W, Burton K, Foster N E, Macfarlane G J, Thomas E, Underwood M, Waddell G, Shekelle P, Volinn E, Von Korff M. A consensus approach toward the standardization of back pain definitions for use in prevalence studies. Spine (Phila Pa 1976) 2008; 33 (1): 95-103.

Djurasovic M, Glassman S D, Dimar J R,2nd, Crawford C H,3rd, Bratcher K R, Carreon L Y. Changes in the Oswestry Disability Index that predict improvement after lumbar fusion. J Neurosurg Spine 2012; 17 (5): 486-90.

Djurasovic M, Glassman S D, Howard J M, Copay A G, Carreon L Y. Health-related quality of life improvements in patients undergoing lumbar spinal fusion as a revision surgery. Spine (Phila Pa 1976) 2011; 36 (4): 269-76.

Eie N. Technique of posterior spinal fusion. Rheumatism 1966; 22: 2-6.

Ekman P, Möller H, Hedlund R. Predictive factors for the outcome of fusion in adult isthmic spondylolisthesis. Spine (Phila Pa 1976) 2009; 34 (11): 1204-10.

Ekman P, Möller H, Hedlund R. The long-term effect of posterolateral fusion in adult isthmic spondylolisthesis: A randomized controlled study. Spine J 2005; 5 (1): 36-44.

EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol group. Health Policy 1990; 16 (3): 199-208.

Fairbank J, Frost H, Wilson-MacDonald J, Yu L M, Barker K, Collins R, Spine Stabilisation Trial Group. Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: The MRC spine stabilisation trial. BMJ 2005; 330 (7502): 1233.

Fairbank J C, Couper J, Davies J B, O'Brien J P. The Oswestry low back pain disability questionnaire. Physiotherapy 1980; 66 (8): 271-3.

Fairbank J C, Pynsent P B. The Oswestry Disability Index. Spine 2000; 25 (22): 2940-52; discussion 2952.

Feeny D, Furlong W, Torrance G W, Goldsmith C H, Zhu Z, DePauw S, Denton M, Boyle M. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care 2002; 40 (2): 113-28.

Finch E, Brooks D, Stratford P W, Mayo N E. Physical rehabilitation outcome measures: A guide to enhanced clinical decision making. BC Decker Inc., Ontario 2002.

Fisher K, Johnson M. Validation of the Oswestry low back pain disability questionnaire, its sensitivity as a measure of chenge following treatment and its relationship with other aspects of the chronic pain experience. Physiother Theort Pract 1997; 13: 67-80.

Freeman B J, Davenport J. Total disc replacement in the lumbar spine: A systematic review of the literature. Eur Spine J 2006; 15 Suppl 3: S439-47.

Fritzell P, Hagg O, Wessberg P, Nordwall A, Swedish Lumbar Spine Study Group. 2001 Volvo award winner in clinical studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: A multicenter randomized controlled trial from the Swedish Lumbar Spine Study Group. Spine (Phila Pa 1976) 2001; 26 (23): 2521,32; discussion 2532-4.

Fryback D G, Dunham N C, Palta M, Hanmer J, Buechner J, Cherepanov D, Herrington S A, Hays R D, Kaplan R M, Ganiats T G, Feeny D, Kind P. US norms for six generic health-related quality-of-life indexes from the national health measurement study. Med Care 2007; 45 (12): 1162-70.

Geisler F H, Blumenthal S L, Guyer R D, McAfee P C, Regan J J, Johnson J P, Mullin B. Neurological complications of lumbar artificial disc replacement and comparison of clinical results with those related to lumbar arthrodesis in the literature: Results of a multicenter, prospective, randomized investigational device exemption study of charite intervertebral disc. Invited submission from the joint section meeting on disorders of the spine and peripheral nerves, March 2004. J Neurosurg Spine 2004; 1 (2): 143-54.

Ghiselli G, Wang J C, Bhatia N N, Hsu W K, Dawson E G. Adjacent segment degeneration in the lumbar spine. J Bone Joint Surg Am 2004; 86-A (7): 1497-503.

Gibson J N, Waddell G. Surgery for degenerative lumbar spondylosis: Updated Cochrane review. Spine (Phila Pa 1976) 2005; 30 (20): 2312-20.

Glassman S, Gornet M F, Branch C, Polly D,Jr, Peloza J, Schwender J D, Carreon L. MOS Short Form 36 and Oswestry Disability Index outcomes in lumbar fusion: A multicenter experience. Spine J 2006; 6 (1): 21-6.

Glassman S D, Carreon L, Dimar J R. Outcome of lumbar arthrodesis in patients sixty-five years of age or older. Surgical technique. J Bone Joint Surg Am 2010; 92 Suppl 1 Pt 1: 77-84.

Glassman S D, Carreon L Y, Djurasovic M, Dimar J R, Johnson J R, Puno R M, Campbell M J. Lumbar fusion outcomes stratified by specific diagnostic indication. Spine J 2009; 9 (1): 13-21.

Glassman S D, Polly D W, Bono C M, Burkus K, Dimar J R. Outcome of lumbar arthrodesis in patients sixty-five years of age or older. J Bone Joint Surg Am 2009; 91 (4): 783-90.

Glassman S D, Polly D W, Dimar J R, Carreon L Y. The cost effectiveness of single-level instrumented posterolateral lumbar fusion at 5 years after surgery. Spine (Phila Pa 1976) 2012; 37 (9): 769-74.

Glassman S D, Schwab F, Bridwell K H, Shaffrey C, Horton W, Hu S. Do 1-year outcomes predict 2-year outcomes for adult deformity surgery? Spine J 2009; 9 (4): 317-22.

Gold M, Franks P, Erickson P. Assessing the health of the nation. The predictive validity of a preference-based measure and self-rated health. Med Care 1996; 34 (2): 163-77.

Gray D B, Hendershot G E. The ICIDH-2: Developments for a new era of outcomes research. Arch Phys Med Rehabil 2000; 81(12 Suppl 2): S10-4.

Grönblad M, Hupli M, Wennerstrand P, Järvinen E, Lukinmaa A, Kouri J P, Karaharju E O. Intercorrelation and test-retest reliability of the Pain Disability Index (PDI) and the Oswestry Disability Questionnaire (ODQ) and their correlation with pain intensity in low back pain patients. Clin J Pain 1993; 9 (3): 189-95.

Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: Literature review and proposed guidelines. J Clin Epidemiol 1993; 46 (12): 1417-32.

Gum J L, Carreon L Y, Stimac J D, Glassman S D. Predictors of Oswestry Disability Index worsening after lumbar fusion. Orthopedics 2013; 36 (4): e478-83.

Gum J L, Glassman S D, Carreon L Y. Is type of compensation a predictor of outcome after lumbar fusion? Spine (Phila Pa 1976) 2013; 38 (5): 443-8.

Gumbs A A, Bloom N D, Bitan F D, Hanan S H. Open anterior approaches for lumbar spine procedures. Am J Surg 2007; 194 (1): 98-102.

Gutzman J, Esmail R, Karjalainen K. Multidisciplinary bio-psycho-social rehabilitation for chronic low back pain. (Cochrane review). The Cochrane Library 2004; Issue 3. Wiley. Chichester.

Haddad B, Makki D, Konan S, Park D, Khan W, Okafor B. Dynesys dynamic stabilization: Less good outcome than lumbar fusion at 4-year follow-up. Acta Orthop Belg 2013; 79 (1): 97-103.

Harmon P H. Anterior extraperitoneal lumbar disc excision and vertebral body fusion. Clin Orthop 1960; 18: 169-184.

Hayden J A, van Tulder M W, Malmivaara A V, Koes B W. Meta-analysis: Exercise therapy for nonspecific low back pain. Ann Intern Med 2005; 142 (9): 765-75.

Hays R D, Sherbourne C D, Mazel R M. The RAND 36-Item Health Survey 1.0. Health Econ 1993; 2 (3): 217-27.

Helenius I, Lamberg T, Österman K, Schlenzka D, Yrjönen T, Tervahartiala P, Seitsalo S, Poussa M, Remes V. Posterolateral, anterior, or circumferential fusion in situ for high-grade spondylolisthesis in young patients: A long-term evaluation using the Scoliosis Research Society questionnaire. Spine (Phila Pa 1976) 2006; 31 (2): 190-6.

Helenius I, Remes V, Lamberg T, Schlenzka D, Poussa M. Long-term health-related quality of life after surgery for adolescent idiopathic scoliosis and spondylolisthesis. J Bone Joint Surg Am 2008; 90 (6): 1231-9.

Hibbs R A. An operation for progressive spinal deformities: A preliminary report of three cases from the service of the orthopaedic hospital. 1911. Clin Orthop Relat Res 2007; 460: 17-20.

Hibbs R A, Swift W E. Developmental abnormalities at the lumbosacral juncture causing pain and disability. Surg Gynecol Obstet 1929; 48: 604-612.

Hollingworth W, Deyo R A, Sullivan S D, Emerson S S, Gray D T, Jarvik J G. The practicality and validity of directly elicited and SF-36 derived health state preferences in patients with low back pain. Health Econ 2002; 11 (1): 71-85.

Horner D, Larmer P J. Health outcome measures. N Z J Physiother 2006; 34: 17-24.

Hudson-Cook N, Tomes-Nicholson K, Breen A. A revised Oswestry Disability Questionnaire. 1989; 1: 187-204.

Hunt S M, McKenna S P, McEwen J, Williams J, Papp E. The Nottingham Health Profile: Subjective health status and medical consultations. Soc Sci Med A 1981; 15 (3 Pt 1): 221-9.

Hurri H, Karppinen J. Discogenic pain. Pain 2004; 112 (3): 225-8.

Høy K, Bunger C, Niederman B, Helmig P, Hansen E S, Li H, Andersen T. Transforaminal lumbar interbody fusion (TLIF) versus posterolateral instrumented fusion (PLF) in degenerative lumbar disorders: A randomized clinical trial with 2-year follow-up. Eur Spine J 2013; [Epub ahead of print].

Inage K, Ohtori S, Koshi T, Suzuki M, Takaso M, Yamashita M, Yamauchi K, Inoue G, Orita S, Eguchi Y, Ochiai N, Kishida S, Kuniyoshi K, Aoki Y, Nakamura J, Ishikawa T, Arai G, Miyagi M, Kamoda H, Suzuki T, Toyone T, Takahashi K. One, two-, and three-level instrumented posterolateral fusion of the lumbar spine with a local bone graft: A prospective study with a 2-year follow-up. Spine (Phila Pa 1976) 2011; 36 (17): 1392-6.

Johnsen L G, Hellum C, Nygaard O P, Storheim K, Brox J I, Rossvoll I, Leivseth G, Grotle M. Comparison of the SF-6D, the EQ-5D, and the Oswestry Disability Index in patients with chronic low back pain and degenerative disc disease. BMC Musculoskelet Disord 2013; 14: 148, 2474-14-148.

Kalichman L, Kim D H, Li L, Guermazi A, Berkin V, Hunter D J. Spondylolysis and spondylolisthesis: Prevalence and association with low back pain in the adult community-based population. Spine (Phila Pa 1976) 2009; 34 (2): 199-205.

Kasis A G, Marshman L A, Krishna M, Bhatia C K. Significantly improved outcomes with a less invasive posterior lumbar interbody fusion incorporating total facetectomy. Spine (Phila Pa 1976) 2009; 34 (6): 572-7.

Kebaish K M, Neubauer P R, Voros G D, Khoshnevisan M A, Skolasky R L. Scoliosis in adults aged forty years and older: Prevalence and relationship to age, race, and gender. Spine (Phila Pa 1976) 2011; 36 (9): 731-6.

Kim J, Henderson R A, Pocock S J, Clayton T, Sculpher M J, Fox K A, RITA-3 Trial Investigators. Health-related quality of life after interventional or conservative strategy in patients with unstable angina or non-ST-segment elevation myocardial infarction: One-year results of the third randomized intervention trial of unstable angina (RITA-3). J Am Coll Cardiol 2005; 45 (2): 221-8.

Kim Y J, Lenke L G, Bridwell K H, Cho Y S, Riew K D. Free hand pedicle screw placement in the thoracic spine: Is it safe? Spine (Phila Pa 1976) 2004; 29 (3): 333-42; discussion 342.

Koes B W, van Tulder M, Lin C W, Macedo L G, McAuley J, Maher C. An updated overview of clinical guidelines for the management of non-specific low back pain in primary care. Eur Spine J 2010; 19 (12): 2075-94.

Koes B W, van Tulder M W, Ostelo R, Kim Burton A, Waddell G. Clinical guidelines for the management of low back pain in primary care: An international comparison. Spine (Phila Pa 1976) 2001; 26 (22): 2504-13; discussion 2513-4.

Kopec J A. Measuring functional outcomes in persons with back pain: A review of back-specific questionnaires. Spine 2000; 25 (24): 3110-4.

Kopec J A, Esdaile J M, Abrahamowicz M, Abenhaim L, Wood-Dauphinee S, Lamping D L, Williams J I. The Quebec Back Pain Disability Scale: Conceptualization and development. J Clin Epidemiol 1996; 49 (2): 151-61.

Laine T, Lund T, Ylikoski M, Lohikoski J, Schlenzka D. Accuracy of pedicle screw insertion with and without computer assistance: A randomised controlled clinical study in 100 consecutive patients. Eur Spine J 2000; 9 (3): 235-40.

Lambat M P, Glassman S D, Carreon L Y. Impact of perioperative complications on clinical outcome scores in lumbar fusion surgery. J Neurosurg Spine 2013a; 18 (3): 265-8.

Lamberg T, Remes V, Helenius I, Schlenzka D, Seitsalo S, Poussa M. Uninstrumented in situ fusion for high-grade childhood and adolescent isthmic spondylolisthesis: Long-term outcome. J Bone Joint Surg Am 2007; 89 (3): 512-8.

Lane J D,Jr, Moore E S, Jr. Transperitoneal approach to the intervertebral disc in the lumbar area. Ann Surg 1948; 127 (3): 537-51.

Latimer J, Maher C G, Refshauge K, Colaco I. The reliability and validity of the Biering-Sorensen test in asymptomatic subjects and subjects reporting current or previous nonspecific low back pain. Spine (Phila Pa 1976) 1999; 24 (20): 2085-9; discussion 2090.

Lee K H, Yue W M, Yeo W, Soeharno H, Tan S B. Clinical and radiological outcomes of open versus minimally invasive transforaminal lumbar interbody fusion. Eur Spine J 2012; 21 (11): 2265-70.

Li G, Passias P, Kozanek M, Fu E, Wang S, Xia Q, Li G, Rand F E, Wood K B. Adult scoliosis in patients over sixty-five years of age: Outcomes of operative versus nonoperative treatment at a minimum two-year follow-up. Spine (Phila Pa 1976) 2009; 34 (20): 2165-70.

Liao J C, Chen W J, Chen L H, Niu C C. Outcome of the L5-S1 segment after posterior instrumented spinal surgery in degenerative lumbar diseases. Chang Gung Med J 2009; 32 (1): 81-8.

Lin P M, Cautilli R A, Joyce M F. Posterior lumbar interbody fusion. Clin Orthop Relat Res 1983; (180) (180): 154-68.

Lohr K N. Assessing health status and quality-of-life instruments: Attributes and review criteria. Qual Life Res 2002; 11 (3): 193-205.

Lowe T G, Tahernia A D, O'Brien M F, Smith D A. Unilateral transforaminal posterior lumbar interbody fusion (TLIF): Indications, technique, and 2-year results. J Spinal Disord Tech 2002; 15 (1): 31-8.

Lucas P. Surgery in the treatment of lower back pain I. Med Health R I 2012; 95 (12): 382-3.

Lund T, Nydegger T, Schlenzka D, Oxland T R. Three-dimensional motion patterns during active bending in patients with chronic low back pain. Spine (Phila Pa 1976) 2002; 27 (17): 1865-74.

Malmivaara A, Slätis P, Heliövaara M, Sainio P, Kinnunen H, Kankare J, Dalin-Hirvonen N, Seitsalo S, Herno A, Kortekangas P, Niinimäki T, Rönty H, Tallroth K, Turunen V, Knekt P, Härkänen T, Hurri H, Finnish Lumbar Spinal Research Group. Surgical or nonoperative treatment for lumbar spinal stenosis? A randomized controlled trial. Spine (Phila Pa 1976) 2007; 32 (1): 1-8.

Mannion A F, Junge A, Fairbank J C, Dvorak J, Grob D. Development of a German version of the Oswestry Disability Index. Part 1: Cross-cultural adaptation, reliability, and validity. Eur Spine J 2006; 15 (1): 55-65.

Mao L, Chen G D, Xu X M, Guo Z, Yang H L. Comparison of lumbar interbody fusion performed with unilateral or bilateral pedicle screw. Orthopedics 2013; 36 (4): e489-93.

Marshman L A, Kasis A, Krishna M, Bhatia C K. Does symptom duration correlate negatively with outcome after posterior lumbar interbody fusion for chronic low back pain? Spine (Phila Pa 1976) 2010;

Mayer T, Gatchel R, Keeley J, Mayer H, Richling D. A male incumbent worker industrial database. Part III: Lumbar/cervical functional testing. Spine (Phila Pa 1976) 1994; 19 (7): 765-70.

McCracken L M, Turk D C. Behavioral and cognitive-behavioral treatment for chronic pain: Outcome, predictors of outcome, and treatment process. Spine (Phila Pa 1976) 2002; 27 (22): 2564-73.

McDonough C M, Grove M R, Tosteson T D, Lurie J D, Hilibrand A S, Tosteson A N. Comparison of EQ-5D, HUI, and SF-36-derived societal health state values among spine patient outcomes research trial (SPORT) participants. Qual Life Res 2005; 14 (5): 1321-32.

Million R, Hall W, Nilsen K H, Baker R D, Jayson M I. Assessment of the progress of the back-pain patient 1981 Volvo award in clinical science. Spine 1982; 7 (3): 204-12.

Min S H, Yoo J S. The clinical and radiological outcomes of multilevel minimally invasive transforaminal lumbar interbody fusion. Eur Spine J 2013; 22 (5): 1164-72.

Mirza S K, Deyo R A. Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain. Spine (Phila Pa 1976) 2007; 32 (7): 816-23.

Mokhtar S A, McCombe P F, Williamson O D, Morgan M K, White G J, Sears W R. Health-related quality of life: A comparison of outcomes after lumbar fusion for degenerative spondylolisthesis with large joint replacement surgery and population norms. Spine J 2010; 10 (4): 306-12.

Möller H, Hedlund R. Surgery versus conservative management in adult isthmic spondylolisthesis-a prospective randomized study: Part 1. Spine (Phila Pa 1976) 2000; 25 (13): 1711-5.

Oestergaard L G, Nielsen C V, Bünger C E, Sogaard R, Fruensgaard S, Helmig P, Christensen F B. The effect of early initiation of rehabilitation after lumbar spinal fusion: A randomized clinical study. Spine (Phila Pa 1976) 2012; 37 (21): 1803-9.

Osthus H, Cziske R, Jacobi E. Cross-cultural adaptation of a German version of the Oswestry Disability Index and evaluation of its measurement properties. Spine 2006; 31 (14): E448-53.

O'Sullivan P. Diagnosis and classification of chronic low back pain disorders: Maladaptive movement and motor control impairments as underlying mechanism. Man Ther 2005; 10 (4): 242-55.

Paalanne N, Niinimäki J, Karppinen J, Taimela S, Mutanen P, Takatalo J, Korpelainen R, Tervonen O. Assessment of association between low back pain and paraspinal muscle atrophy using opposed-phase magnetic resonance imaging: A population-based study among young adults. Spine (Phila Pa 1976) 2011; 36 (23): 1961-8.

Park P, Garton H J, Gala V C, Hoff J T, McGillicuddy J E. Adjacent segment disease after lumbar or lumbosacral fusion: Review of the literature. Spine (Phila Pa 1976) 2004; 29 (17): 1938-44.

Parker S L, Adogwa O, Paul A R, Anderson W N, Aaronson O, Cheng J S, McGirt M J. Utility of minimum clinically important difference in assessing pain, disability, and health state after transforaminal lumbar interbody fusion for degenerative lumbar spondylolisthesis. J Neurosurg Spine 2011; 14 (5): 598-604.

Parker S L, Shau D N, Mendenhall S K, McGirt M J. Factors influencing 2-year health care costs in patients undergoing revision lumbar fusion procedures. J Neurosurg Spine 2012; 16 (4): 323-8.

Pereira-Maxwell F. A-Z of medical statistics, a companion for critical appraisal. Arnold, London, UK 1998.

Phillips F M, Slosar P J, Youssef J A, Andersson G, Papatheofanis F. Lumbar spine fusion for chronic low back pain due to degenerative disc disease: A systematic review. Spine (Phila Pa 1976) 2013; 38 (7): E409-22.

Pollock R, Lakkol S, Budithi C, Bhatia C, Krishna M. Effect of psychological status on outcome of posterior lumbar interbody fusion surgery. Asian Spine J 2012; 6 (3): 178-82.

Price D D, McGrath P A, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. Pain 1983; 17 (1): 45-56.

Radcliff K, Curry P, Hilibrand A, Kepler C, Lurie J, Zhao W, Albert T J, Weinstein J. Risk for adjacent segment and same segment reoperation after surgery for lumbar stenosis: A subgroup analysis of the spine patient outcomes research trial (SPORT). Spine (Phila Pa 1976) 2013; 38 (7): 531-9.

Radcliff K, Hwang R, Hilibrand A, Smith H E, Gruskay J, Lurie J D, Zhao W, Albert T, Weinstein J. The effect of iliac crest autograft on the outcome of fusion in the setting of degenerative spondylolisthesis: A subgroup analysis of the spine patient outcomes research trial (SPORT). J Bone Joint Surg Am 2012; 94 (18): 1685-92.

Rampersaud Y R, Wai E K, Fisher C G, Yee A J, Dvorak M F, Finkelstein J A, Gandhi R, Abraham E P, Lewis S J, Alexander D I, Oxner W M, Davey J R, Mahomed N. Postoperative improvement in health-related quality of life: A national comparison of surgical treatment for focal (one- to two-level) lumbar spinal stenosis compared with total joint arthroplasty for osteoarthritis. Spine J 2011; 11 (11): 1033-41.

Rodriguez-Vela J, Lobo-Escolar A, Joven-Aliaga E, Herrera A, Vicente J, Sunen E, Loste A, Tabuenca A. Perioperative and short-term advantages of mini-open approach for lumbar spinal fusion. Eur Spine J 2009; 18 (8): 1194-201.

Roelofs P D, Deyo R A, Koes B W, Scholten R J, van Tulder M W. Non-steroidal anti-inflammatory drugs for low back pain. Cochrane Database Syst Rev 2008; (1):CD000396. doi (1): CD000396.

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 (Phila Pa 1976) 1983; 8 (2): 141-4.

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 (Phila Pa 1976) 1983b; 8 (2): 145-50.

Roy-Camille R, Saillant G, Mazel C. Internal fixation of the lumbar spine with pedicle screw plating. Clin Orthop Relat Res 1986; (203) (203): 7-17.

Røe C, Sveen U, Bautz-Holter E. Retaining the patient perspective in the international classification of functioning, disability and health core set for low back pain. Patient Prefer Adherence 2008; 2: 337-47.

Räsänen P, Öhman J, Sintonen H, Ryynänen O P, Koivisto A M, Blom M, Roine R P. Cost-utility analysis of routine neurosurgical spinal surgery. J Neurosurg Spine 2006; 5 (3): 204-9.

Räsänen P, Roine E, Sintonen H, Semberg-Konttinen V, Ryynänen O P, Roine R. Use of quality-adjusted life years for the estimation of effectiveness of health care: A systematic literature review. Int J Technol Assess Health Care 2006; 22 (2): 235-41.

Saban K L, Penckofer S M, Androwich I, Bryant F B. Health-related quality of life of patients following selected types of lumbar spinal surgery: A pilot study. Health Qual Life Outcomes 2007; 5: 71.

Salokangas R K, Poutanen O, Stengård E. Screening for depression in primary care. development and validation of the depression scale, a screening instrument for depression. Acta Psychiatr Scand 1995; 92 (1): 10-6.

Sanden B, Forsth P, Michaelsson K. Smokers show less improvement than nonsmokers two years after surgery for lumbar spinal stenosis: A study of 4555 patients from the Swedish Spine Register. Spine (Phila Pa 1976) 2011; 36 (13): 1059-64.

Sigl T, Cieza A, Brockow T, Chatterji S, Kostanjsek N, Stucki G. Content comparison of low back pain-specific measures based on the International Classification of Functioning, Disability and Health (ICF). Clin J Pain 2006; 22 (2): 147-53.

Sintonen H. The 15D instrument of health-related quality of life: Properties and applications. Ann Med 2001; 33 (5): 328-36.

Sköld C, Tropp H, Berg S. Five-year follow-up of total disc replacement compared to fusion: A randomized controlled trial. E pub Eur Spine J 29 July 2013.

Slätis P, Malmivaara A, Heliovaara M, Sainio P, Herno A, Kankare J, Seitsalo S, Tallroth K, Turunen V, Knekt P, Hurri H. Long-term results of surgery for lumbar spinal stenosis: A randomised controlled trial. Eur Spine J 2011; 20 (7): 1174-81.

Smith-Petersen M N, Larson C B, Aufranc O E. Osteotomy of the spine for correction of flexion deformity in rheumatoid arthritis. J Bone & Joint Surg 1945; 27: 1-11.

Solberg T K, Olsen J A, Ingebrigtsen T, Hofoss D, Nygaard O P. Health-related quality of life assessment by the EuroQol-5D can provide cost-utility data in the field of low-back surgery. Eur Spine J 2005; 14 (10): 1000-7.

Steenstra I A, Verbeek J H, Heymans M W, Bongers P M. Prognostic factors for duration of sick leave in patients sick listed with acute low back pain: A systematic review of the literature. Occup Environ Med 2005; 62 (12): 851-60.

Stratford P W, Binkley J M, Riddle D L, Guyatt G H. Sensitivity to change of the Roland-Morris Back Pain Questionnaire: Part 1. Phys Ther 1998; 78 (11): 1186-96.

Streiner D L, Norman G R. Health measurement scales. A practical guide to their development and use. Oxford Medical Publications, Oxford 2003.

Strömqvist B H, Berg S, Gerdhem P, Johnsson R, Möller A, Sahlstrand T, Soliman A, Tullberg T. X-stop versus decompressive surgery for lumbar neurogenic intermittent claudication: Randomized controlled trial with 2-year follow-up. Spine (Phila Pa 1976) 2013; 38 (17): 1436-42.

Stucki G, Sigl T. Assessment of the impact of disease on the individual. Best Pract Res Clin Rheumatol 2003; 17 (3): 451-73.

Søgaard R, Christensen F B, Videbaek T S, Bunger C, Christiansen T. Interchangeability of the EQ-5D and the SF-6D in long-lasting low back pain. Value Health 2009; 12 (4): 606-12.

Takatalo J, Karppinen J, Niinimäki J, Taimela S, Mutanen P, Sequeiros R B, Näyhä S, Järvelin M R, Kyllönen E, Tervonen O. Association of Modic changes, Schmorl's nodes, spondylolytic defects, high-intensity zone lesions, disc herniations, and radial tears with low back symptom severity among young Finnish adults. Spine (Phila Pa 1976) 2012; 37 (14): 1231-9.

Takatalo J, Karppinen J, Niinimäki J, Taimela S, Näyhä S, Mutanen P, Sequeiros R B, Kyllönen E, Tervonen O. Does lumbar disc degeneration on magnetic resonance imaging associate with low back symptom severity in young Finnish adults? Spine (Phila Pa 1976) 2011; 36 (25): 2180-9.

The World Health Organization quality of life assessment (WHOQOL): Position paper from the world health organization. Soc Sci Med 1995; 41 (10): 1403-9.

Tobler W D, Gerszten P C, Bradley W D, Raley T J, Nasca R J, Block J E. Minimally invasive axial presacral L5-S1 interbody fusion: Two-year clinical and radiographic outcomes. Spine (Phila Pa 1976) 2011; 36 (20): E1296-301.

Valasek T, Varga P P, Szoverfi Z, Kumin M, Fairbank J, Lazary A. Reliability and validity study on the Hungarian versions of the Oswestry Disability Index and the Quebec Back Pain Disability Scale. Eur Spine J 2013; 22 (5): 1010-8.

van den Berg B. Sf-6d population norms. Health Econ 2012; 21 (12): 1508-12.

van Tulder M, Koes B, Bombardier C. Low back pain. Best Pract Res Clin Rheumatol 2002; 16 (5): 761-75.

van Tulder M W, Ostelo R W, Vlaeyen J. Behavioral treatment for chronic low back pain. (Cochrane review). Wiley, Chichester 2004.

van Tulder M W, Touray T, Furlan A D, Solway S, Bouter L M. Muscle relaxants for non-specific low back pain. Cochrane Database Syst Rev 2003; (2) (2): CD004252.

Vigatto R, Alexandre N M, Correa Filho H R. Development of a Brazilian Portuguese version of the Oswestry Disability Index: Cross-cultural adaptation, reliability, and validity. Spine (Phila Pa 1976) 2007; 32 (4): 481-6.

Virta L, Ronnemaa T, Österman K, Aalto T, Laakso M. Prevalence of isthmic lumbar spondylolisthesis in middle-aged subjects from eastern and western Finland. J Clin Epidemiol 1992; 45 (8): 917-22.

Waddell G. Subgroups within "nonspecific" low back pain. J Rheumatol 2005; 32 (3): 395-6.

Waddell G. Low back pain: A twentieth century health care enigma. Spine (Phila Pa 1976) 1996; 21 (24): 2820-5.

Waddell G. 1987 Volvo award in clinical sciences. A new clinical model for the treatment of low-back pain. Spine (Phila Pa 1976) 1987; 12 (7): 632-44.

Waddell G, Main C J. Assessment of severity in low-back disorders. Spine (Phila Pa 1976) 1984; 9 (2): 204-8.

Walker B F. The prevalence of low back pain: A systematic review of the literature from 1966 to 1998. J Spinal Disord 2000; 13 (3): 205-17.

Walters S J, Brazier J E. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Qual Life Res 2005; 14 (6): 1523-32.

Walters S J, Brazier J E. What is the relationship between the minimally important difference and health state utility values? The case of the SF-6D. Health Qual Life Outcomes 2003; 1: 4.

Ware J E, Kosinski M, Keller SD. SF-36 physical and mental health summary scales: A User´s manual. The Health Institute, Boston, MA 1994.

Ware J E, Kosinski M. SF-36 physical and mental health summary scales: A manual for users of version 1 (2nd ed.). QualityMetric Incorporated, Lincoln, RI, US 2001.

Ware J E, Jr, Gandek B. Overview of the SF-36 health survey and the international quality of life assessment (IQOLA) project. J Clin Epidemiol 1998; 51 (11): 903-12.

Ware J E, Jr, Sherbourne C D. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992; 30 (6): 473-83.

Weinstein J N, Lurie J D, Tosteson T D, Hanscom B, Tosteson A N, Blood E A, Birkmeyer N J, Hilibrand A S, Herkowitz H, Cammisa F P, Albert T J, Emery S E, Lenke L G, Abdu W A, Longley M, Errico T J, Hu S S. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. N Engl J Med 2007; 356 (22): 2257-70.

Weinstein J N, Lurie J D, Tosteson T D, Zhao W, Blood E A, Tosteson A N, Birkmeyer N, Herkowitz H, Longley M, Lenke L, Emery S, Hu S S. Surgical compared with nonoperative treatment for lumbar degenerative spondylolisthesis. Four-year results in the spine patient outcomes research trial (SPORT) randomized and observational cohorts. J Bone Joint Surg Am 2009; 91 (6): 1295-304.

Weinstein J N, Tosteson T D, Lurie J D, Tosteson A N, Blood E, Hanscom B, Herkowitz H, Cammisa F, Albert T, Boden S D, Hilibrand A, Goldberg H, Berven S, An H, SPORT Investigators. Surgical versus nonsurgical therapy for lumbar spinal stenosis. N Engl J Med 2008; 358 (8): 794-810.

Wessels T, van Tulder M, Sigl T, Ewert T, Limm H, Stucki G. What predicts outcome in non-operative treatments of chronic low back pain? A systematic review. Eur Spine J 2006; 15 (11): 1633-44.

WHO. International Classification of Functioning, Disability and Health. Geneva: World Health Organization; 2001.

Wood K B, Fritzell P, Dettori J R, Hashimoto R, Lund T, Shaffrey C. Effectiveness of spinal fusion versus structured rehabilitation in chronic low back pain patients with and without isthmic spondylolisthesis: A systematic review. Spine (Phila Pa 1976) 2011; 36 (21 Suppl): S110-9.

World Health Organization. Disabilities. Retrieved 11 August 2012.

Wu W J, Liang Y, Zhang X K, Cao P, Zheng T. Complications and clinical outcomes of minimally invasive transforaminal lumbar interbody fusion for the treatment of one-

Waddell G. 1987 Volvo award in clinical sciences. A new clinical model for the treatment of low-back pain. Spine (Phila Pa 1976) 1987; 12 (7): 632-44.

Waddell G, Main C J. Assessment of severity in low-back disorders. Spine (Phila Pa 1976) 1984; 9 (2): 204-8.

Walker B F. The prevalence of low back pain: A systematic review of the literature from 1966 to 1998. J Spinal Disord 2000; 13 (3): 205-17.

Walters S J, Brazier J E. Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Qual Life Res 2005; 14 (6): 1523-32.

Walters S J, Brazier J E. What is the relationship between the minimally important difference and health state utility values? The case of the SF-6D. Health Qual Life Outcomes 2003; 1: 4.

Ware J E, Kosinski M, Keller SD. SF-36 physical and mental health summary scales: A User´s manual. The Health Institute, Boston, MA 1994.

Ware J E, Kosinski M. SF-36 physical and mental health summary scales: A manual for users of version 1 (2nd ed.). QualityMetric Incorporated, Lincoln, RI, US 2001.

Ware J E, Jr, Gandek B. Overview of the SF-36 health survey and the international quality of life assessment (IQOLA) project. J Clin Epidemiol 1998; 51 (11): 903-12.

Ware J E, Jr, Sherbourne C D. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992; 30 (6): 473-83.

Weinstein J N, Lurie J D, Tosteson T D, Hanscom B, Tosteson A N, Blood E A, Birkmeyer N J, Hilibrand A S, Herkowitz H, Cammisa F P, Albert T J, Emery S E, Lenke L G, Abdu W A, Longley M, Errico T J, Hu S S. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. N Engl J Med 2007; 356 (22): 2257-70.

Weinstein J N, Lurie J D, Tosteson T D, Zhao W, Blood E A, Tosteson A N, Birkmeyer N, Herkowitz H, Longley M, Lenke L, Emery S, Hu S S. Surgical compared with nonoperative treatment for lumbar degenerative spondylolisthesis. Four-year results in the spine patient outcomes research trial (SPORT) randomized and observational cohorts. J Bone Joint Surg Am 2009; 91 (6): 1295-304.

Weinstein J N, Tosteson T D, Lurie J D, Tosteson A N, Blood E, Hanscom B, Herkowitz H, Cammisa F, Albert T, Boden S D, Hilibrand A, Goldberg H, Berven S, An H, SPORT Investigators. Surgical versus nonsurgical therapy for lumbar spinal stenosis. N Engl J Med 2008; 358 (8): 794-810.

Wessels T, van Tulder M, Sigl T, Ewert T, Limm H, Stucki G. What predicts outcome in non-operative treatments of chronic low back pain? A systematic review. Eur Spine J 2006; 15 (11): 1633-44.

WHO. International Classification of Functioning, Disability and Health. Geneva: World Health Organization; 2001.

Wood K B, Fritzell P, Dettori J R, Hashimoto R, Lund T, Shaffrey C. Effectiveness of spinal fusion versus structured rehabilitation in chronic low back pain patients with and without isthmic spondylolisthesis: A systematic review. Spine (Phila Pa 1976) 2011; 36 (21 Suppl): S110-9.

World Health Organization. Disabilities. Retrieved 11 August 2012.

Wu W J, Liang Y, Zhang X K, Cao P, Zheng T. Complications and clinical outcomes of minimally invasive transforaminal lumbar interbody fusion for the treatment of one-

or two-level degenerative disc diseases of the lumbar spine in patients older than 65 years. Chin Med J (Engl) 2012; 125 (14): 2505-10.

Wu Y, Tang H, Li Z, Zhang Q, Shi Z. Outcome of posterior lumbar interbody fusion versus posterolateral fusion in lumbar degenerative disease. J Clin Neurosci 2011; 18 (6): 780-3.

Yakut E, Duger T, Oksuz C, Yorukan S, Ureten K, Turan D, Frat T, Kiraz S, Krd N, Kayhan H, Yakut Y, Guler C. Validation of the Turkish version of the Oswestry Disability Index for patients with low back pain. Spine (Phila Pa 1976) 2004; 29 (5): 581-5; discussion 585.

Appendix

TOIMINTAKYKYINDEKSI (OSWESTRY 2.0)
Kyselyn tarkoituksena on antaa tietoa siitä, kuinka selkäkipu (alaraajakipu) on
vaikuttanut kykyynne suoriutua jokapäiväisistä toimistanne. Rastittakaa joka
kohdasta vain se ruutu, joka parhaiten kuvaa tilannettanne tänään.

1. Kivun voimakkuus
□ Minulla ei ole kipua tällä hetkellä.
□ Kipu on hyvin lievä tällä hetkellä.
□ Kipu on kohtalainen tällä hetkellä.
□ Kipu on melko voimakas tällä hetkellä.
□ Kipu on hyvin voimakas tällä hetkellä.
□ Kipu on pahin mahdollinen tällä hetkellä.

2. Itsestä huolehtiminen (peseytyminen, pukeutuminen, jne.)
□ Selviydyn näistä toimista normaalisti, eikä niistä aiheudu lisää kipua.
□ Selviydyn näistä toimista normaalisti, mutta niistä aiheutuu lisää kipua.
□ Näistä toimista selviytyminen on kivuliasta vaatien aikaa ja varovaisuutta.
□ Tarvitsen hieman apua, mutta selviydyn useimmista toimista itsenäisesti.
□ Tarvitsen apua päivittäin useimmissa näistä toimista.
□ En pukeudu, peseydyn vaivalloisesti ja pysyttelen vuoteessa.

3. Nostaminen
□ Voin nostaa raskaita taakkoja, eikä se lisää kipua.
□ Voin nostaa raskaita taakkoja, mutta se lisää kipua.
□ Kipu estää minua nostamasta raskaita taakkoja lattialta, mutta voin nostaa niitä, jos
ne on sijoitettu sopivasti, esim. pöydälle.
□ Kipu estää minua nostamasta raskaita taakkoja, mutta voin nostaa kevyitä tai
kohtalaisia taakkoja, jos ne on sijoitettu sopivasti.
□ Voin nostaa vain hyvin kevyitä taakkoja.
□ En voi nostaa tai kantaa mitään.

4. Kävely
□ Kipu ei rajoita kävelymatkaani.
□ Kipu estää minua kävelemästä yli 2 kilometriä.
□ Kipu estää minua kävelemästä yli 500 metriä.
□ Kipu estää minua kävelemästä yli 100 metriä.
□ Voin kävellä vain käyttäen keppiä tai kyynärsauvoja.
□ Olen enimmäkseen vuoteessa, ja minun on kontattava WC:hen.

5. Istuminen
□ Voin istua millaisessa tuolissa tahansa niin pitkään kuin haluan.
□ Voin istua vain määrätynlaisessa tuolissa niin pitkään kuin haluan.
□ Kipu estää minua istumasta tuntia pitempään.
□ Kipu estää minua istumasta puolta tuntia pitempään.
□ Kipu estää minua istumasta kymmentä minuuttia pitempään.
□ Kipu estää istumiseni täysin.

6. Seisominen
□ Voin seisoa niin pitkään kuin haluan, ilman että siitä aiheutuu lisää kipua.
□ Voin seisoa niin pitkään kuin haluan, mutta siitä aiheutuu lisää kipua.
□ Kipu estää minua seisomasta tuntia pitempään.
□ Kipu estää minua seisomasta puolta tuntia pitempään.
□ Kipu estää minua seisomasta kymmentä minuuttia pitempään.
□ Kipu estää seisomiseni täysin.

7. Nukkuminen
□ Kipu ei häiritse nukkumistani koskaan.
□ Kipu häiritsee nukkumistani ajoittain.
□ Kivun takia nukun alle kuusi tuntia.
□ Kivun takia nukun alle neljä tuntia.
□ Kivun takia nukun alle kaksi tuntia.
□ Kipu estää nukkumiseni täysin.

8. Sukupuolielämä
□ Sukupuolielämäni on normaalia, eikä siitä aiheudu lisää kipua.
□ Sukupuolielämäni on normaalia, mutta siitä aiheutuu hieman lisää kipua.
□ Sukupuolielämäni on lähes normaalia, mutta hyvin kivuliasta.
□ Kipu rajoittaa huomattavasti sukupuolielämääni.
□ Kivun takia sukupuolielämäni on lähes olematonta.
□ Kipu estää minulta kaiken sukupuolielämän.

9. Sosiaalinen elämä
□ Sosiaalinen elämäni on normaalia, eikä siitä aiheudu lisää kipua.
□ Sosiaalinen elämäni on normaalia, mutta siitä aiheutuu lisää kipua.
□ Kipu ei vaikuta merkittävästi sosiaaliseen elämääni, mutta se rajoittaa liikunnallisia harrastuksiani.
□ Kivun takia sosiaalinen elämäni kodin ulkopuolella on vähentynyt.
□ Kivun takia sosiaalinen elämäni on rajoittunut kotiin.
□ Kivun takia minulla ei ole mitään sosiaalista elämää.

10. Matkustaminen
□ Voin matkustaa minne tahansa, ilman että siitä aiheutuu kipua.
□ Voin matkustaa minne tahansa, mutta siitä aiheutuu kipua.
□ Kipu on voimakas, mutta voin matkustaa yli kaksi tuntia.
□ Kipu rajoittaa matkustamiseni alle tunnin kestäviin matkoihin.
□ Kipu rajoittaa matkustamiseni alle puolen tunnin kestäviin välttämättömiin matkoihin.
□ Kivun takia en voi matkustaa minnekään muualle kuin saamaan hoitoa

Original publications

HEALTH SERVICES RESEARCH

Reliability and Validity Study of the Finnish Version 2.0 of the Oswestry Disability Index

Liisa Pekkanen, MD,* Hannu Kautiainen, BA,† Jari Ylinen, MD, PhD,‡ Petri Salo, MSc,‡ and Arja Häkkinen, PhD‡§

Study Design. Prospective clinical validation study.
Objective. The aims of this study were to translate into Finnish and culturally adapt and study the psychometric properties of the Oswestry Disability Index (ODI) version 2.0.
Summary of Background Data. The ODI is one of the most commonly reported back-specific disability questionnaires. It is widely used both in clinical work and in medical studies. To date, no validated Finnish version of the ODI version 2.0 has been reported.
Methods. The ODI version 2.0 was translated into the Finnish language. A total of 115 patients with back pain, referred by the primary care physician to the outpatient clinic of the department of physical medicine and rehabilitation, were recruited for this study. The patients answered a questionnaire package that included the Finnish ODI 2.0, Back Pain Questionnaire for Visual Analogue Assessment (Million-VAS), Visual Analogue Scales of back and leg pain (VAS$_{back}$, VAS$_{leg}$), the Depressions Scale, and a question on their subjectively perceived health. The package was administered twice; 2 weeks before and at the arrival to the clinic.
Results. Reproducibility of the ODI was 0.90 (95% confidence interval [CI] = 0.85–0.94) and the internal consistency was 0.86 (95% CI = 0.81–0.90). Factor analysis showed that the ODI was loaded on 2 factors, which explained 51% of the total variance. In testing convergent validity ODI correlated with Million-VAS, $r = 0.75$ (95% CI = 0.64–0.84); VAS$_{back}$, $r = 0.48$ (95% CI = 0.32–0.62); and VAS$_{leg}$, $r = 0.41$ (95% CI = 0.23–0.57).
Conclusion. The Finnish ODI version 2.0 proved to be a valid and reliable instrument that showed psychometric properties comparable with the original English version. Therefore, it can be used in assessing the disability among Finnish-speaking patients with back pain for both clinical and scientific purposes.

Back pain accounts for a significant amount of disability in industrialized societies and can lead to absence from work or early retirement. Consequently, such costs place a burden on the health care system and on society. The strategies in the treatment of chronic low back pain (LBP) vary considerably from country to country or even between different areas within one country. In general, these patients are first treated conservatively but an overall trend is toward an increased use of operative treatment.

In treating these patients, it is crucial to get information from the patient's point of view to assess the level of severity of the symptoms and the level of disability. This information acts as an important asset in patient history and can assist in planning the most suitable treatment options. In addition, this kind of information is essential when evaluating the effects of the chosen treatment over the time course.

The use of patient-based outcome measures, are important tools to assess the patients' perceptions of their functional health status. This allows the clinician to gain a better understanding of the impact of symptoms and the effect of treatment of patients with LBP.[1,2] There are 2 types of commonly used patient-based outcome measures. First, the generic measures evaluate general health, overall disability, and quality of life. However, they are insensitive to react to clinically relevant change in single specific disease.[3,4] The Short Form Health Survey (SF-36) is used by many clinicians as a generic instrument in the lumbar pain field.[5] Second, back-specific measurements assess only aspects of patient's health that are affected by that specific disease. The main feature in these instruments is the connection of symptoms and disability to a specific disease or condition.[2] In a systematic review from 1996 to 2002 a total of 36 back-specific questionnaires were found.[6] Interestingly, only a small number of the questionnaires could be considered to be scientifically validated.[6] The most commonly used questionnaires were the Roland-Morris Disability Questionnaire (RM) and the Oswestry Disability Index (ODI). Several versions have been identified for these 2 questionnaires.[6] It has been suggested by Roland and Fairbank that RM is more suitable with patients with mild-to-moderate disability, whereas the ODI could better express disability in patients with more severe symptoms.[7]

The ODI is highly recommended and one of the most popularly used back-specific questionnaires in the field and it is also widely used in medical studies.[3,8] John O'Brien started the development of the ODI in 1976. The first ODI version 1.0 was published in 1980 with the original name "The Oswestry Low Back Pain Disability Questionnaire."[9] Since then several versions have been released and up to date information of these versions is available at the official website of the developers.[10] In addition to these there are also revised versions of the ODI, for example, by a chiropractic study group in the United Kingdom.[11] However, such revisions can lead to difficulties when comparing the results with other versions. The section concerning sex life of the patient may be omitted in some revised versions because of cultural reasons.[1,2] According to the ODI websites, the questionnaire has been translated in 29 different languages. Validation of such translated versions 2.0 or 2.1, have been made in at least 6 different countries (Germany, Italy, Japan, Korea, Poland, and Turkey).[10] The authors considered it necessary to independently validate all versions in different languages and ideally there should be a single version in use for each language.[12]

Previously, an intercorrelation analysis between ODI and Pain-Disability Index had been provided for the ODI version 1.0 in Finland.[13] In our previous study, the Finnish version of the ODI 1.0 proved to capture a wide scale of disability in lumbar disc surgery patients, thereby supporting the future use of the index. However, the "pain intensity" item concerning the use of pain killers in version 1.0 of the ODI did not support the item structure of the index. The item related to pain killers clearly did not measure pain in the same way as the other items concerning pain-related disability did.[14] The use of ODI version 2.0 might eliminate this problem, as in that version this particular question has been modified to measure pain intensity and not the use of pain killers.[7]

In this present study, the ODI questionnaire was used alongside the disability index developed by R. Million and coworkers (Back Pain Questionnaire for Visual Analogue Assessment) and published with the 1981 Volvo Award in Clinical Science.[15] According to the authors, the questionnaire was found to be satisfactory and usable in the back pain clinics.[15] Further in this manuscript the term Million-VAS is used for this index.

The primary objective of this study was to create a Finnish version of the ODI version 2.0 through proper cross-cultural adaptations and through investigation of the psychometric properties of the index. This work was accomplished by using Finnish-speaking patients with back pain.

MATERIALS AND METHODS

Patients and Data Collection

This study was approved by the local Ethics Committee and each patient signed a written consent. The patients in this study were clinically diagnosed with back problem and referred by the primary care physician to the outpatient clinic in the Department of Physical Medicine and Rehabilitation in Jyväskylä Central Hospital, Finland. A total of 115 patients were recruited for this investigation. The inclusion criteria for the patients' inclusion were: age ≥18 years, LBP with or without radicular pain in the lower extremities, and ability to communicate in written Finnish language. There were no specific exclusion criteria. The patients answered the first questionnaire package 2 weeks before arriving at the outpatient clinic and again a second time when came to the clinic.

Translation and Cross-Cultural Adaptation

The translation and cross-cultural adaptation followed the guidelines which were obtained from recently published studies.[16,17] The first step was an independent translation of the ODI version 2.0 by 2 professionals, who had Finnish as their first language. Both were professionals in the field of rehabilitation medicine and this allowed them to have a clinical perspective and to understand the practical use of this questionnaire. In the next step, we used these 2 translated versions along with the original English version and a consensus version was produced. In the third step, an individual not working in the field of medicine who had English as a first language, back-translated the version into English without any knowledge of the ODI beforehand. The Finnish version was then evaluated by a Finnish language expert in the national scientific medical association of Finland (Duodecim) and the Finnish ODI version 2.0 was introduced (Supplemental Digital Content, Appendix 1, available at: http://links.lww.com/BRS/A449). This version was tested with 30 patients before the actual patient recruitment for this trial was undertaken. The preliminary pilot testing demonstrated no concerns or reasons for any change to be made.

Questionnaire Booklets

The questionnaire booklet included the final version of the Finnish ODI 2.0 consisting of 10 items with each item having 6 statements. The first statement was linked to the lowest score of 0, whereas the last statement was linked to the highest score of 5. If more than one statement is chosen by the patient, the highest score is used. The total score is then calculated as the sum of each completed item and is expressed as a percentage of the maximum number of possible points.[12] On repeated occasions, the percentage can simply be referred to as the "ODI score" and expressed in terms of points (0–100).[18]

The Back Pain Questionnaire for Visual Analogue Assessment (Million-VAS)[15] consisted of 15 questions which were used to evaluate the severity of the symptoms, the issues that would exacerbate the symptoms and measure how much the disability affected the patients' normal activities. The scale used for each answer ranged from 0 to 100. For example, in the question "do you have discomfort when walking" the verbal scale in answering ranges from "none at all" = 0 to "intolerable" = 100.[15] The total score is the mean value of all the items completed.[15]

The questionnaire package also included the Visual Analogue Scale for back and leg pain (VAS$_{back}$, VAS$_{leg}$, scale 0-100 mm)[19] and the Depression Scale (DEPS).[20] In the 10 item DEPS, the score ranges from 0 to 30. When the score is 9 or more, the susceptibility toward depressive symptoms is increased. Only a small number of totally asymptomatic patients gain a score this high.[20] The patients' subjectively perceived health was captured by one item from the SF-36 questionnaire: "In general, would you say your health is excellent, very good, good, fair or poor."[5] In the analysis, the responses were divided into 3 categories with excellent and very good recorded as "excellent" and "fair," respectively, and poor reported as "poor." The patient was also asked to provide standard demographic data and the location and duration of the pain. The diagnosis was made on the basis of information retrieved from the medical records and if needed radiologic examinations (*e.g.*, magnetic resonance imaging). At the time of their outpatient clinic visit, in addition to the questionnaire the patients were also asked so-called transition questions to determine whether there had been any subjective change in their health condition within the last 2 weeks.

Statistical Methods

The results of the investigation are expressed as mean or median, standard deviation, or interquartile range. The 95% confidence intervals (95% CI) were obtained by bias-corrected bootstrapping (5000 replications). The "floor value" was defined in this study as the best possible value of the item or as the minimum total value of the scale, and the "ceiling value" is the worst possible value of the item or the maximum total value of the scale. The reliability of the scales was evaluated by calculating the intraclass correlation coefficient and coefficient of repeatability, with the bias corrected and accelerated bootstrapping (5000 replications) CIs. Internal consistency was estimated by calculating Cronbach's alpha with 95% CI. Construct validity was studied by maximum likelihood factor analysis with promax rotation for the ODI items matrix of polychoric correlations. Item analysis of the ODI scales was performed by analyzing item discriminating power (corrected item correlation) and item difficulty (item mean) depicted by the explanatory data analysis. Corrected item correlation was estimated using polyserial correlations. Statistical significance for hypotheses of linearity was evaluated by bootstrap-type analysis of variance, with covariates when appropriate. Correlation coefficients were calculated by the Spearman method with bootstrap-type CIs.

RESULTS

A total of 115 patients, 64 (56%) males and 51 (44%) females, participated in this study. The demographic and clinical characteristics of the study group are presented in Table 1. The ages of the patients ranged from 19 to 77 years, and the mean age was of 49 years. Of the patients, 26% were obese (body mass index ≥30). The mean VAS$_{back}$ was 60 mm and the mean VAS$_{leg}$ was 52 mm.

Table 2 shows the response rates for all ODI items as well as floor and ceiling values of the initial assessment. The floor

TABLE 1. Demographic and Clinical Characteristics of 115 Patients	
Variables	**Values**
Male, n (%)	64 (56)
Age, yr, mean (SD)	49 (13)
Body mass index, n (%)	
<25	43 (39)
25–29	39 (35)
≥30	29 (26)
Duration of symptoms, mo, n (%)	
>6	31 (27)
6–11	28 (24)
12–23	17 (15)
≥24	39 (34)
Diagnosis, n (%)	
Disc degeneration or herniation	27 (23)
Spinal stenosis	9 (8)
Isthmic spondylolysis	10 (9)
Spondylarthrosis	33 (29)
Others	36 (31)
Pain, VAS, mean (SD)	
Back	60 (25)
Leg	52 (27)
Laseq present, n (%)	11 (10)
Spinal operation(s) in history, n (%)	15 (13)
DEPS ≥9, n (%)	51 (44)

VAS indicates visual analogue scale; SD, standard deviation; DEPS, depression scale.

value was reached more often in individual items of personal care, walking, sex life, and social life. Six patients left more than 1 item unanswered, most often it was the question number 8 concerning the sex life (N = 17) together with some other item. Response rate in the other ODI 2.0 items varied between 97% and 99%. One patient got total ODI score of zero at the first assessment, whereas at the time of the repeated administration none. At the first measurement a total of 4 patients had total scores of 10 or lower and at the second assessment 5 patients had similar lower scores. No one reached the maximum total ODI score while answering the questionnaire on either the first or second assessments. The highest ODI score that was achieved was 70.

When the questionnaire was administered for the first time, the mean standard deviation ODI score was 33[14] and Million-VAS score was 53[18] for patients, who had been stable in their subjective health between the first and second survey. For these patients, reproducibility intraclass correlation

TABLE 2. Characteristics for the ODI Items and Scales of 115 Patients

Item/Scale 0–5	Response Rate, %	Floor, %*	Ceiling, %†	Mean
1. Pain	98	5	2	2.25
2. Personal care	99	43	0	0.74
3. Lifting	97	5	2	2.51
4. Walking	97	43	0	0.83
5. Sitting	98	6	1	1.91
6. Standing	97	7	0	2.07
7. Sleeping	98	8	1	1.53
8. Sex life	85	35	3	1.34
9. Social life	99	25	0	1.74
10. Travelling	98	9	1	1.52

*Best possible value of the item.
†Worst possible value of the item.

Item analysis of the ODI showed that all items had a good overall corrected item correlation (Figure 1). Item 9 had a high corrected item correlation (social life) and number 3 had low (lifting).

Items 2 (personal care) and 4 (walking) showed the lowest item mean values, whereas item 3 (lifting) showed the highest.

Factor analysis performed for construct validity showed that the ODI scale was loaded on 2 factors. Factor 1 characterizes activities of daily living as well as social life, that is, sleeping and sitting. Factor 2 characterizes activity in the upright position and pain. These factors explained 51% of the total variance (Table 4).

Correlation coefficients between ODI and Million-VAS were 0.75 (95% CI = 0.64−0.84) (Figure 2) and between the ODI and VAS_{back} or VAS_{leg} were 0.48 (95% CI = 0.32−0.62) and 0.41 (95% CI = 0.23−0.57), respectively.

The ability of the ODI to distinguish between clinical subentities at baseline is illustrated in Figure 3. After age adjustments, relationships remained linear between the ODI and self-perceived health $(P < 0.001)$ or duration of symptoms (P = 0.011) but not with the body mass index (P = 0.16).

The correlations between the ODI *versus* age and DEPS-score were 0.24 (95% CI = 0.07−0.40) and 0.52 (95% CI = 0.36−0.64), respectively.

According to the ODI score, there was no statistically significant difference in disability between the diagnostic groups. The mean ODI score ranged from 31% to 35% between the 5 diagnostic groups (*i.e.*, disc degeneration or herniation, spinal stenosis, isthmic spondylolysis, spondylarthrosis, and others like back pain symptom without a diagnosis).

coefficient was 0.90 (95% CI = 0.85–0.94) for the ODI and 0.85 (95% CI = 0.76–0.91) for the Million-VAS. In those patients who reported changes in their symptoms, both instruments showed poorer reproducibility (Table 3). When the patients were grouped according to the changes in their symptoms, the change of the ODI $(P < 0.001)$ and the change of Million-VAS (P = 0.003) decreased linearly across the groups (*i.e.*, worsened-stable-improved).

Internal cons.istency of the ODI was 0.86 (95% CI = 0.81–0.90) and that of Million-VAS was 0.90 (95% CI = 0.86–0.93).

TABLE 3. Reproducibility of the Oswestry Disability Index (ODI) and Million Visual Analoque Scale (Million-VAS) Scores

Symptoms	n	First Measurement Mean % (SD)	Change From First to Second Measurement Mean (95% CI)	Reproducibility ICC (95% CI)	CR (95% CI)
ODI					
Improved	12	29.4 (15.3)	−9.3 (−15.2 to −4.4)	0.60 (0.12–0.82)	19.4 (9.4–29.6)
Stable	86	33.1 (15.0)	−1.7 (−3.1 to −0.4)	0.90 (0.85–0.94)	12.5 (10.7–14.4)
Worsened	16	35.5 (11.4)	+ 5.1 (1.1−9.5)	0.69 (0.41–0.83)	17.6 (12.1–23.2)
Million-VAS					
Improved	12	50.9 (13.8)	−16.5 (−25.4 to −8.6)	0.17 (−0.34 to 0.54)	29.2 (20.1–38.3)
Stable	82	52.6 (18.8)	−4.8 (−6.9 to −2.8)	0.85 (0.76–0.91)	18.5 (14.6–22.5)
Worsened	13	56.9 (20.3)	+ 1.0 (−6.2 to +7.7)	0.69 (0.29–0.84)	26.4 (16.2–36.5)

Express the expected maximum size of 95% of the absolute differences between paired observations. 95% CI obtained by bias corrected and accelerated bootstrapping.
ICC indicates intraclass correlation coefficient; CR, coefficient of repeatability.
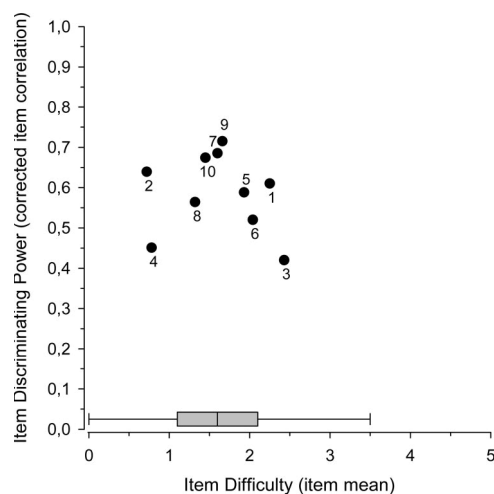
Figure 1. Item analysis for the Oswestry Disability Index (ODI) items. The bar denotes median with interquartile range. Numbers indicate corresponding items in ODI scale.

Figure 2. Correlation between ODI and Million Visual Analoque Scale (Million-VAS).

DISCUSSION

The purpose of this investigation was to cross-culturally adapt and psychometrically test the Finnish version of the ODI 2.0 to be used among Finnish-speaking patients with back pain. The ODI version 2.0 has been widely used in Finland, even though it has not been scientifically validated. According to the ODI website, this particular version has been translated or validated in many countries.[10] The general aim using this kind of standardized back-specific tool is to obtain comparable information of the disability that LBP causes to the patient as well as to be able to evaluate the effect of different kinds of treatments.

The translation process was carried out according to the accepted guidelines.[16,17] In the translation process, the only difficulties worth noticing were the concerns for the item
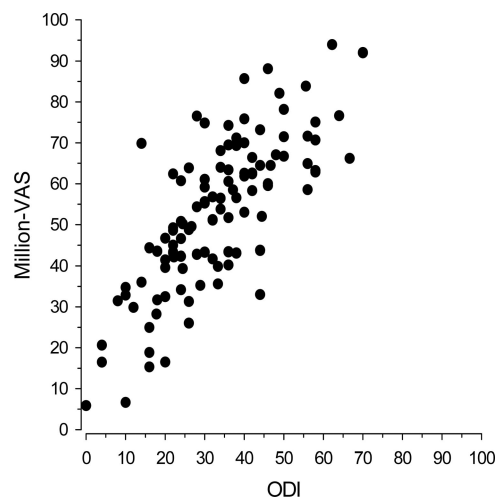
number 4, the description of the distance for walking. The consensus was to modify the original option "pain prevents me walking more than one mile" to more than 2 km, the option "a quarter of a mile" to 500 m, the option "100 yards" to 100 m. The exact distances translated from the British Imperial system into the metric system would be 1.6 km, 402 m, and 91 m. As the British Imperial System is not used in Finland and is therefore not understood, more conceptual metric distances were used for these options. For a patient it is often quite difficult to determine distances in general and to get more reliable answers the numbers in the question should be as simple as possible, and the options in the question divergent enough. Older versions of the ODI have been widely used in clinical and scientific purposes in Finland, and the second choice in the question 4 has invariably been "2 km." Therefore, it was not tempting to change the translation. In the official ODI web sites, many forms of the distances are used in different ODI translations.[10] For example, in the work of Mannion et al[18] that produced a version for use in German-speaking countries and the official German version for the Spine Society of Europe's "Spine Tango" Spine Surgery Registry of ODI, the authors modified the walking question to the distance of "1 to 2 km" instead of "1 mile."

The patients completed the ODI questionnaire twice: 2 weeks before arrival to the outpatient clinic and at the time of arrival. This interval minimizes the possibility that the quality of the answers could be affected by the memory effect.[21] Prior research has shown that the reproducibility of the functional status questionnaires is best measured within a 1–2 week time interval.[21] This time interval minimizes one's capability for recall of the previous answers as compared with the shorter time frames of same day or next day testing.[22] There was a special question added to the second administration question package to determine whether the patients' health condition had changed or whether he/she had had any kind of treatment. Using this question, we could evaluate patients who had been

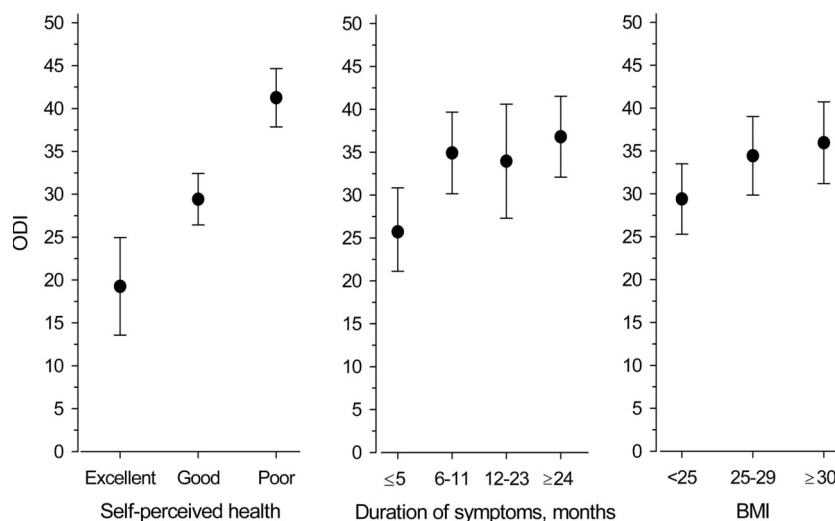| TABLE 4. Explanatory Factor Analysis With Promax-Rotated Factor Loadings of the ODI Items		
Item/Scale	Factor 1	Factor 2
7. Sleeping	0.73	
2. Personal care	0.72	
9. Social life	0.71	
8. Sex life	0.65	
3. Lifting	0.49	
5. Sitting	0.47	
4. Walking		0.88
10. Traveling		0.63
6. Standing		0.52
1. Pain		0.48
ODI indicates Oswestry Disability Index.		

Figure 3. Association between ODI and self-perceived health, duration of symptoms and body mass index.

staying stable and those patients whose disability had improved or gotten worse. On entering the clinic, patients filled out the questionnaires before seeing the physician to avoid the possibility that the treatment or discussion would affect their response.

In the original English version of the ODI, the patients completed the questionnaire with 24 hour time interval and the reproducibility was extremely high (0.99).[9] Among those patients whose symptoms remained stable between the 2-week time interval, the ODI demonstrated a solid and quite similar reproducibility with Million-VAS. ODI could also detect the change in patient's outcome. This indicates that the ODI as an instrument is able to follow the patient's own experience in their symptoms and disability.

It has been suggested that a questionnaire reaching the floor or the ceiling value of over 15% should be omitted.[23] In this evaluation, at the baseline one patient had a total score of 0 points, whereas none of the patients had a score of zero at the 2-week, that is, the floor. The ceiling was not reached, as the maximum total ODI score was 70 for one patient at the baseline testing. This indicates that a real floor or ceiling effect did not exist when using the Finnish ODI. On the other hand, for individual items expressing personal care, walking, sex life, and social life, the floor effect was recorded more frequently than the accepted amount of 15% of the patients tested. The item analysis showed that in the ODI questionnaire specifically questions 2 and 4 concerning personal care and walking were those in which the respondents expressed least disability while in question 3 concerning lifting patients experienced the most difficult form of disability. Lifting was obviously the most physically demanding task related item of the ODI. For individual items, the overall response rate was acceptable and only in the question number 8 that referred to sex life was the response rate under 85%.

The overall homogeneity of the items (*i.e.*, internal consistency) was 0.86 in this study. In previously published research using the ODI version 1.0, Cronbach's alpha has been reported to be 0.71,[24] whereas in the recently used version 2.1 of the index a score of 0.90 was reported.[18] The recommended Cronbach's alpha for group comparisons is typically higher than 0.80.[25] However, if it is "too good" the internal consistency may indicate that the items are too homogenous. From that perspective our study expressed good reliability and demonstrated that the items of the Finnish ODI were reasonably related, but still each one of them also contributed to some unique information on patient status.

The ODI presented good construct validity in this study as analyzed from several statistical approaches. Convergent validity was assessed by testing the power of the relationship between the ODI and Million-VAS. In other studies, testing the relationship between the ODI and RM, a good correlation has been reported. However, it has been noticed that the ODI was more reliable in patients with more severe disability. When the RM score is already at its maximum, the ODI is still capable to show change.[7] The ODI showed the ability to discriminate between patients that would be expected to vary in a particular characteristic. After age adjustment, there was a linearity noticed between the ODI and subjectively perceived health or duration of symptoms. In this study, there was a moderate correlation between the ODI and VAS_{back} or VAS_{leg}, age, and the DEPS. In previously published research, a correlation ($r = 0.78$) between the ODI and combined VAS for back and leg pain was also reported.[18]

Factor analysis was carried out to detect the structure in relationships between the variables. It appears that LBP is a phenomenon that cannot be explained unambiguously. The factor analysis of the ODI showed that there is loading on 2 factors and these factors explain half of the disability.

CONCLUSION

This study has produced an appropriately translated and culturally adapted ODI version 2.0. The Finnish ODI version 2.0 proved to be reliable and valid. As the ODI is easy for the patient to administer and to score both in clinical work and research, its use can now be recommended among Finnish-speaking patients.

➤ **Key Points**

❑ The ODI is one of the most recommended and popular back-specific questionnaire both in clinical use and in medical studies; therefore, there is a need to get comparable information of the disability across different countries.

❑ The purpose of this work was to translate and culturally adapt the ODI version 2.0 into Finnish language and study psychometric properties of the tool.

❑ The Finnish ODI version 2.0 seems to be valid and reliable and it shows as good psychometric properties as the original English version. The use of this instrument can be recommended among Finnish-speaking patients.

Supplemental digital content is available for this article. Direct URL citation appears in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.spinejournal.com).

References

1. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993;118:622–9.
2. Kopec JA. Measuring functional outcomes in persons with back pain: a review of back-specific questionnaires. *Spine* 2000;25:3110–4.
3. Osthus H, Cziske R, Jacobi E. Cross-cultural adaptation of a German version of the oswestry disability index and evaluation of its measurement properties. *Spine* 2006;31:E448–53.
4. Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. *Spine* 2000;25:3100–3.
5. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-item health survey 1.0. *Health Econ* 1993;2:217–27.
6. Grotle M, Brox JI, Vollestad NK. Functional status and disability questionnaires: what do they assess? A systematic review of back-specific outcome questionnaires. *Spine* 2005;30:130–40.
7. Roland M, Fairbank J. The Roland-Morris disability questionnaire and the Oswestry disability questionnaire. *Spine* 2000;25:3115–24.
8. Deyo RA, Battie M, Beurskens AJ, et al. Outcome measures for low back pain research: a proposal for standardized use. *Spine* 1998;23:2003–13.
9. Fairbank JC, Couper J, Davies JB, et al. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66:271–3.
10. Fairbank J. Oswestry Disability Index. Available at: http://www.orthosurg. org.uk. Accessed March 27, 2009.
11. Hudson-Cook N, Tomes-Nicholson K, Breen A. A revised Oswestry disability questionnaire. In: Roland M, Jenner JR, eds. *Back Pain: New Approaches to Rehabilitation and Education.* Manchester, Lancashire, United Kingdom: Manchester University Press; 1989:187–204.
12. Fairbank JC, Pynsent PB. The Oswestry disability index. *Spine* 2000;25:2940–52; discussion 2952.
13. Gronblad M, Hupli M, Wennerstrand P, et al. Intercorrelation and test-retest reliability of the pain disability index (PDI) and the oswestry disability questionnaire (ODQ) and their correlation with pain intensity in low back pain patients. *Clin J Pain* 1993;9:189–95.
14. Hakkinen A, Kautiainen H, Jarvenpaa S, et al. Changes in the total Oswestry index and its ten items in females and males pre- and post-surgery for lumbar disc herniation: a 1-year follow-up. *Eur Spine J* 2007;16:347–52.
15. Million R, Hall W, Nilsen KH, et al. Assessment of the progress of the back-pain patient 1981 volvo award in clinical science. *Spine* 1982;7:204–12.
16. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health- related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol* 1993;46:1417–32.
17. Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25:3186–91.
18. Mannion AF, Junge A, Fairbank JC, et al. Development of a German version of the Oswestry disability index. Part 1: cross-cultural adaptation, reliability, and validity. *Eur Spine J* 2006;15:55–65.
19. Price DD, McGrath PA, Rafii A, et al. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain* 1983;17:45–56.
20. Salokangas RK, Poutanen O, Stengard E. Screening for depression in primary care. Development and validation of the depression scale, a screening instrument for depression. *Acta Psychiatr Scand* 1995;92:10–6.
21. Beurskens AJ, de Vet HC, Koke AJ, et al. Measuring the functional status of patients with low back pain. Assessment of the quality of four disease- specific questionnaires. *Spine* 1995;20:1017–28.
22. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Control Clin Trials* 1991;12:142S–58S.
23. McHorney CA, Tarlov AR. Individual-patient monitoring in clinical practice: are available health status surveys adequate? *Qual Life Res* 1995;4:293–307.
24. Strong J, Ashton R, Large RG. Function and the patient with chronic low back pain. *Clin J Pain* 1994;10:191–6.
25. Streiner DL Norman GR. *Health Measurement Scales. A Practical Guide to Their Development and Use.* Vol 1. 3rd ed. Oxford, United Kingdom: Oxford Medical Publications; 2003.

informa
healthcare

RESEARCH PAPER

Decreased disability is associated with improved perceived quality of life following spinal fusion

Liisa Pekkanen[1], Marko Neva[2], Hannu Kautiainen[3,4], Kimmo Vihtonen[2], Kati Kyrölä[1], Ilkka Marttinen[2], Marko Wahlman[2] & Arja Häkkinen[5,6]

[1]*Department of Orthopaedics and Traumatology, Jyväskylä Central Hospital, Jyväskylä, Finland,* [2]*Department of Orthopaedics and Traumatology, Tampere University Hospital, Finland,* [3]*Unit of Family Practice, Central Finland Central Hospital, Jyväskylä, Finland,* [4]*Unit of Primary Health Care, Kuopio University Hospital, Finland,* [5]*Department of Physical Medicine and Rehabilitation, Jyväskylä Central Hospital, Finland, and* [6]*Department of Health Sciences, University of Jyväskylä, Jyväskylä, Finland*

Purpose: To assess the disability and relationship between functional status and health related quality of life (HRQoL) in patients in the early recovery phase following spinal fusion. *Methods*: This is a prospective cohort study. Since 2008 data of spinal fusion patients have been collected prospectively in two Finnish hospitals. In August 2009, complete data of 173 patients were available. The measurement tool of disability was the Oswestry Disability Index (ODI) and it was also examined in the framework of International Classification of Functioning, Disability and Health (ICF) using body functions and structures, activities and participation components. *Results*: Preoperatively the mean total ODI was 45 (SD17) and mean (95% confidence interval) change to 3 months postoperatively was −19 (−22 to −17). When the ODI was linked to the ICF, there was a 55% improvement in the body structure and functions component and a 44% improvement in both the activities and the participation components. However, 25% of the patients still had the total ODI score over 40 three months postoperatively. Preoperatively, the mean (95% CI) Physical Component Summary Score (PCS) of the Short Form 36-questionnaire (SF-36) was 27 (26 to 28) and the mean Mental Component Summary Score (MCS) of SF-36 was 47 (45 to 49). Postoperatively the improvement was 9 (95% CI: 8 to 11) in PCS and 6 (95% CI: 4 to 7) in MCS ($p < 0.001$). *Conclusions*: Spinal fusion is successful in the early recovery period in terms of reduction of pain and disability. The significant changes in the ODI were seen in all three components of the ICF model. In addition, improvement in functioning was significantly related to positive change in HRQoL. Still there is a subgroup

Implications for Rehabilitation

- Spinal fusion
- The focus in assessing the outcome of spinal fusion operations has shifted towards patient-based health status and quality of life scales in recent years.
- When the patient's self-rated Oswestry Disability Index (ODI) was examined in the framework of International Classification of Functioning, Disability and Health (ICF) it was competent to distinguish different domains of functioning from each other. However, the ODI does not cover earlier published comprehensive core set for low back pain completely and thus more comprehensive assessment is needed in clinical settings.
- This study showed that in the early recovery phase 3 months after the spinal fusion operation the positive changes in disability, health related quality of life and pain were significant.
- However, 25% of the patients still had the ODI score 40 or more at 3 months after the spinal fusion and these patients should be identified to be given additional attention to provide appropriate rehabilitation.

of patients having marked disability needing more intensive rehabilitation and follow-ups.

Correspondence: Liisa Pekkanen, Department of Orthopaedics and Traumatology Jyväskylä Central Hospital, Keskussairaalantie 19, 40620 Jyväskylä, Finland. Tel: +358 14 2693079. Fax: +358 14 2692786. E-mail: pekkanen.auvinen@fimnet.fi

Introduction

Lumbar spinal fusion is the most rapidly growing procedure in the USA in the field of orthopaedic surgery [1,2]. This growth has been going on steadily since 1980s [3]. Along with this growth the costs have also risen substantially over the last two decades [1,4]. It is constantly crucial to get more information about the outcome of these operations and the focus in assessing the outcome has shifted towards patient-based health status and quality of life scales in studies published in recent years [5–13].

The most commonly used instruments in these outcome studies are the Oswestry Disability Index (ODI), Short Form 36 (SF-36) and Visual Analogue Scale (VAS) [14–17]. The ODI is one of the back-specific instruments recommended by the World Health Organization (WHO) [18].

In the rehabilitation or disability perspective, patient's functioning and health are associated with their condition or disease and are not only seen as consequences of them. Moreover functioning and disability are seen in association with personal and environmental factors and not only in association with the underlying disease [18]. The basis for this association and for the inclusion of patients' personal and contextual factors is the WHO International Classification of Functioning, Disability and Health (ICF) [19,20]. Most measurements tools used in the evaluation of spinal conditions focus on impairment and activity limitations rather than participation or overall involvement in life [18,19]. No single measure can cover all the aspects of patients´ function, disability and health [21]. Comparison of measurements techniques can be complicated; this might explain why many studies compare the psychometric properties of the measurement instruments instead of the content [21,22]. Linkage to the WHO's ICF has been used as a framework in comparing how comprehensively generic and back-specific outcome measures express patient´s well-being in main health related domains [21,23,24]. In an earlier study Røe et al. reported patient perspective in the ICF core set for low back pain. The authors pointed out that all items of the ODI could be linked with the ICF. The ODI items concerning sleeping, pain, standing, sitting, lifting, walking and social function were those the patients reported to have problems with and the health care professionals' scores in the linked ICF categories captured most often [25].

The purpose of this study was to assess the relationship between changes in disability and health related quality of life (HRQoL) during the early recovery phase (3 months) after spinal fusion. The ODI as a back-specific instrument was also examined in the ICF framework through a real cohort of patients. This is a novel addition to the literature of this field.

Material and methods

This data on patients having spinal fusion surgery were collected since the beginning of the year 2008 in Tampere University Hospital and Jyväskylä Central Hospital, Finland. A total of six spine surgeons performed the operations. The data were collected prospectively as part of standard clinical practice and consisted of preoperative, perioperative, and postoperative follow up data.

The ethical committees of both hospitals approved the study plan and all patients signed a written consent prior to data collection.

By the end of August 2009, data of 220 elective patients undergoing spinal fusion were available. One hundred seventy-three of these 220 patients (79%) had completed both the preoperative and postoperative data for all outcome measures and were included in the final analysis. The patients not included did not differ according to the age, sex or level of pain from the patients included.

Data collection

Preoperatively, the diagnosis and main indication for fusion operation were specified. The main indication for fusion was not always the primary reason for surgery. For example a patient with spinal stenosis and radicular symptoms is classified as "degenerative olisthesis"; olisthesis is the indication for fusion, rather than decompression alone.

Demographic and clinical data, including gender, age, body mass index (BMI), and duration of symptoms were collected. VAS (0–100 mm) was used according to the definition to measure the average level of back and leg pain during the previous week (VAS_{back}, VAS_{leg}) [17].

The questionnaire data completed pre-and postoperatively consisted of the Oswestry disability index and SF-36 as back-specific and generic patient based assessment instruments, respectively [16,26]. The ODI is one of the most widely used back-specific questionnaires, both clinically and in research in the field of spine surgery [27,28]. The ODI contains ten items each with six statements graded from zero (lowest disability) to five (greatest disability). The total score is calculated as a sum of each completed item and expressed as a percentage of the maximum number of possible points, i.e. related to the number of items the patient has answered [15]. Scores are defined by a scale according to the original publication: 0–20 minimal, 20–40 moderate and 40–60 severe disability. A score 60–80 indicates a crippled patient and 80–100 indicates that patient is either bed-bound or exaggerating their symptoms [14]. In addition, the ODI was studied in the ICF framework. The ODI was linked to the ICF according the published guidelines [24]. The linkage has been made by separate health care professionals independently and in addition to this the professionals have discussed and made a consensus about the final linkage. The linking was made into the part functioning and disability and into the components (a) body functions and structures and (b) activities and participation. In the (a) component there are two categories: "mental function" and "sensory functions and pain". In the (b) component the categories are "mobility", "self care" and "interpersonal interactions and relationships"; and "mobility" and "community, social and civic life". The option of partial overlap between component activities and participation was chosen so, that mobility was left as a common category in both. The ODI items were linked according to the main functional issue since the pain is consisting in every item and the total score was expressed in the relative scale zero to one hundred. The linkage of the ODI to the ICF is illustrated in Figure 1.
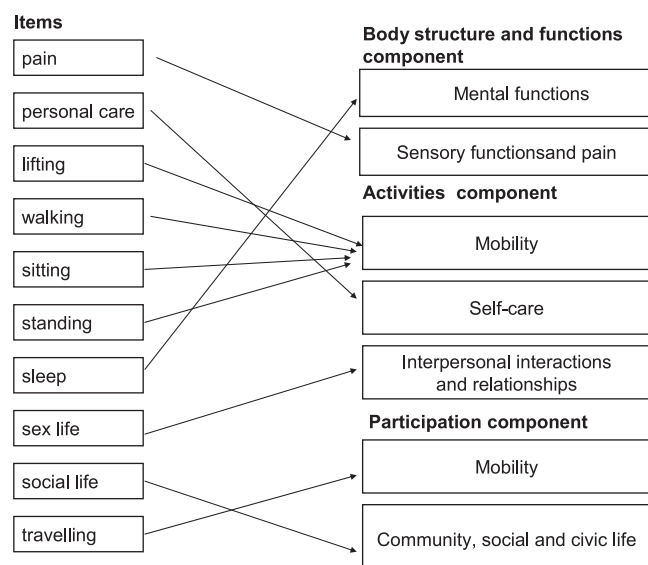
Items

Figure 1. The linkage of the Oswestry Disability Index (ODI) items to the International Classification of Functioning, Disability and Health (ICF).

SF-36 is the most widely evaluated generic patient assessed health outcome measure [29]. It reflects the patient's health condition through eight health concepts: Physical Functioning (PF), Role Physical (RP), Bodily Pain (BP), General Health (GH), Vitality (VT), Social Functioning (SF), Role Emotional (RE), and Mental Health (MH). Scores range from 0 to 100; a higher score is associated with better health. The eight scales in the SF-36 can be aggregated into two summary measures. PF, RP, BP, and GH form the physical component summary score (PCS) and VT, SF, RE, and MH form the mental component summary score (MCS) (29).

Operation technique

The fusion was performed with instrumentation in 167 of 173 patients (97%) and only six (3%) fusions were performed with non-instrumented method. In the instrumented group 138 of 167 patients had posterolateral fusion and 26 patients had 360° fusion with an additional transforaminal lumbar interbody fusion (TLIF) or posterior lumbar interbody fusion (PLIF). No minimal approaches were used in the procedures. Surgeons from two study centres have performed part of the operations together so the operation techniques are similar in these two centres.

Mobility instructions

All patients were given instructions orally and in writing during the hospital stay and 6 weeks after the operation at the outpatient clinic by a physiotherapist. Patients were recommended to sit a maximum one-half hour continuously during the first 4 weeks and avoid extreme flexion and extension of the trunk for 2 months postoperatively. Patients were encouraged to walk and perform certain light exercises and were provided with a written information leaflet containing instructions of the exercises. Use of a bicycle ergometer was allowed 1 month after the operation. Two months postoperatively, skiing, dancing, and water gymnastics were permitted. The postoperative

instructions have been developed in co-operation within the two hospitals in our study. The patients have been getting equal postoperative guidance.

Statistical analysis

Data is presented as means with standard deviations (SD) or 95% confidence intervals (95% CI) or as counts with percentages. Changes in outcomes are expressed with 95% CI and tested with paired samples t-test or permutation test. A Hotelling-type permutation test for related samples was performed simultaneously in the domains of the same ICF component.

Effect size was calculated by the method of Cohen for paired samples (mean baseline scores minus mean follow-up scores, divided by the pooled standard deviation) [30]. Effect size acts as indicator of responsiveness. Effect size standardizes mean change over time with a standard deviation and by that allows comparison of a particular intervention's different outcomes, independent of the measuring units. Effect size of >0.20 was considered small, >0.50 to medium, and >0.80 large. 95% CI for the effect sizes were obtained by bias-corrected bootstrapping (5000 replications).

Regression lines with 95% CI were used to illustrate the relationships between the changes of the ODI and Component Summary Scores of SF-36. We calculated Pearson's correlation coefficients and partial correlation coefficients adjusted for age, gender, and duration of symptoms.

Results

The study population consisted of 173 patients including 118 (68%) females and 55 (32%) males with an age range of 29–87 years. The two most common indications for the spinal fusion were degenerative olistesis (54%) and spondylolysis (19%). The mean duration of symptoms was 14 months (SD 14) preoperatively. Mean body mass index (BMI) was 28.4 (SD 4.4) at baseline. Forty percent of patients were classified as obese (BMI>30). The majority of the patients, 67% were retired and 30% were working at the time of surgery (Table I).

Preoperative mean total ODI score was 45 (SD17); at follow-up, mean change was −19 (95% CI: −22 to −17) for a mean postoperative score of 26 (SD 17). The correlation between the changes of the ODI and VAS_{back} was 0.56 (95% CI 0.55 to 0.57) and between the changes of the ODI and VAS_{leg} 0.44 (95% CI 0.44 to 0.45). At baseline 65% of patients had the total ODI score 40 or more. Three months postoperatively this value was 25%. Those patients having the ODI below 40, the mean VAS_{back} was 19 (SD 20) and patients having the ODI over 40, the mean VAS_{back} was 40 (SD 26), $p < 0.001$. The values of VAS_{leg} were 15 (SD 20) and 45 (SD 30), respectively, $p < 0.001$.

When the results were assessed with the ODI linked to the ICF, the worst disability was in the sensory functions and pain domains with the body structure and functions component; the lowest disability was reported for the self-care domain in the activities component (Table II).

Three months postoperatively the positive changes in all seven ICF domains were statistically significant. The effect

sizes of the changes were large in all domains except in the domains of self-care and interpersonal interactions and relationships. In these domains the effect sizes were moderate (Figure 2).

Mean preoperative HRQoL, through the SF-36 questionnaire, was lowest in role physical dimension and highest in mental health dimension (Table III). Three months postoperatively the changes in all eight dimensions of SF-36 were statistically significant (Table III).

The effect sizes as indicators of responsiveness of the change were largest in the dimensions of physical functioning, bodily pain, and social functioning (Figure 3).

When the eight scales of the SF-36 were aggregated into summary scores, preoperatively PCS was 27 (95% CI: 26 to 28) and MCS was 47 (95% CI: 45 to 49). Postoperatively, mean improvement was 9 (95% CI: 8 to 11) in PCS and 6 (95% CI: 4 to 7) in MCS; $p < 0.001$.

Mean preoperative VAS_{back}, and VAS_{leg} were 66 (SD27) and 67 (SD25), respectively. Mean postoperative change was −42 (95% CI: −47 to −37) for VAS_{back} and −45 (95% CI: −51 to −40) for VAS_{leg}; these changes were statistically significant ($p < 0.001$).

The relationship between the change in the total ODI score and PCS and MCS are illustrated in Figure 4A and 4B. After

adjusting for age, gender, and duration of symptoms, the correlation between the change in ODI and PCS was −0.63 (95% CI: −0.70 to −0.54) and between the ODI and MCS −0.35 (95% CI: −0.48 to −0.17).

Discussion

The main finding of this study was that patients who underwent spinal fusion had a significant decrease in disability in the early recovery phase and the decrease in disability was paralleled to corresponding improvement in perceived quality of life.

The 173 patients in this analysis having spinal fusion operation had the mean preoperative total ODI score of 45 (SD17). This score level represents severe disability according to the original interpretation of the score (14). In previous reports concerning fusion surgery the preoperative ODI scores have been ranging between 43 and 53 [5,9,10,31,32]. There have been a few earlier studies reporting short-term outcome of the ODI. In the study of Feng et al. patients were performed decompression and transforaminal lumbar interbody fusion either with unilateral (Group 1) or bilateral (Group 2) pedicle screw fixation. Preoperatively the ODI of Group 1 was 25.875 ± 12.789 and Group 2 24.083 ± 11.131 and 3 months postoperatively the ODI was 5.200 ± 11.077 and 0.000 ± 0.000, respectively [33]. In another study assessing advantages of mini-open approach the ODI changed from 55.9 ± 10.3 to 31.5 ± 14.2 in the classic approach group and from 59.1 ± 7.9 to 16.0 ± 12.1 in the mini-open approach group in 3 months [34]. Even though postoperative improvement in our material was noted, 25% of our patients had a score of over 40 postoperatively, indicating severe disability [14]. These individuals having high level of disability and pain should be identified in the early recovery phase and receive additional attention in order to provide appropriate rehabilitation. From this point of view it is also extremely important to treat the pain as well as possible in the early postoperative phase. In our study we found out that there was a correlation between the changes of the ODI and pain and the level of pain in those patients having the ODI score below or over 40 varied significantly.

One main purpose of this work was to assess disability three months after fusion through the ICF frame of reference. The ODI was studied in the ICF framework in a real patient

Table I. Baseline demographics and clinical data.

Variables	Descriptives (n = 173)
Female, n (%)	118 (68)
Age, years, mean (SD)	62 (13)
Body mass index, mean (SD)	28.4 (4.4)
Employment status	
Employed	52 (30)
Unemployed	5 (3)
Pension	116 (67)
Indication for fusion, n(%)	
Degenerative olistesis	93 (54)
Spondylolysis	32 (19)
Spinal stenosis	16 (9)
Disc herniation or degeneration	11 (6)
Postoperative conditions	16 (9)
Degenerative scoliosis	5 (3)
Duration of symptoms, months, mean (SD)	14 (14)

Table II. Oswestry Disability Index linked to the ICF components and domains.

ICF components domains	ODI item	Preoperative Mean (SD)	Change at 3 months Mean (95% CI)	*p* Value
Body structure and functions				
Mental functions	7	36 (23)	−20 (−16 to −23)	<0.001
Sensory functions and pain	1	56 (24)	−31 (−27 to −35)	<0.001
Activities				
Mobility	3,4,5,6	51 (19)	−17 (−14 to −20)	<0.001
Self-care	2	25 (23)	−11 (−8 to −15)	<0.001
Interpersonal interactions and relationships	8	36 (36)	−21 (−16 to −26)	<0.001
Participation				
Mobility	10	46 (27)	−21 (−17 to −26)	<0.001
Community, social, and civic life	9	52 (25)	−22 (−17 to −27)	<0.001

Hotelling-type permutation test for related samples performed in domains of the same ICF component simultaneously *p* < 0.001.

Body structure and functions
  Mental functions
  Sensory functions and pain

Activity
  Mobility
  Self-care
  *Interpersonal interactions
    and relationships*

Participation
  Mobility
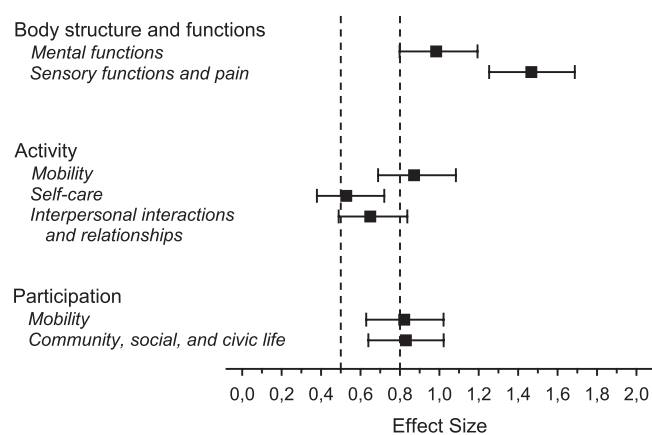 *Community, social, and civic life*

Effect Size

Figure 2. The effect-size of the change in the Oswestry Disability Index (ODI) linked to the International Classification of Functioning, Disability and Health (ICF) components and domains.

Table III. Health Related Quality of Life dimensions (SF-36).

Dimension	Preoperative Mean (SD)	Change at months 3 Mean (95% CI)	p Value
Physical functioning	31 (22)	25 (22 to 29)	<0.001
Role physical	12 (24)	21 (15 to 28)	<0.001
Bodily pain	25 (16)	30 (27 to 34)	<0.001
General health	54 (21)	4 (1 to 7)	0.005
Vitality	47 (23)	18 (15 to 21)	<0.001
Social functioning	51 (29)	22 (17 to 27)	<0.001
Role emotional	46 (43)	18 (12 to 24)	<0.001
Mental health	65 (22)	12 (9 to 15)	<0.001

Physical Functioning

Role-Physical

Bodily Pain

General Health

Vitality

Social Functioning

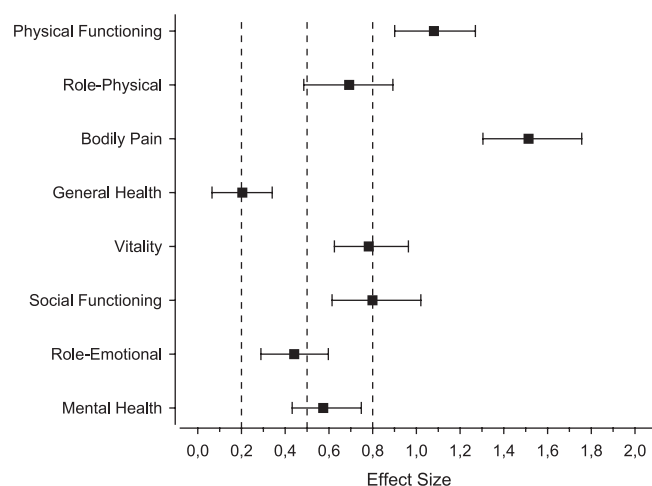Role-Emotional

Mental Health

Effect Size

Figure 3. The effect-size of the change in SF-36 dimensions.

cohort. In our study the worst preoperative disability concerned pain in sensory functions and pain domain in body structure and function component; lifting, walking, sitting and standing in mobility domain in activities component; and social function in community, social, and civic life domain in participation component. These same sectors raise up to be most problematic in an earlier study examining the health problems self-reported by patients in the ODI in relation with the health problems scored by the health professionals in the linked ICF categories in the comprehensive ICF core set for

low back pain [25]. This study also shows that the ODI does not cover all the categories that exist in the core set for low back pain defined by health professionals. This leads to the conclusion that clinical work requires wide assessment and combining clinical judgement to the data that the ODI or other measurement tools provide.

Our analysis showed that the effect size of the change in disability was largest in Body structure and functions component. The ODI is comprehensive and also competent to distinguish different domains of functioning from each other. The ICF framework gives a new dimension to the use of the ODI and helps to pay attention to different perspectives of functioning. In clinical practice it is quite adequate, however, to use the basic ODI score.

The positive changes in all SF-36 domains were significant and the effect sizes were clear in all eight domains and in both summary scores. In a study by Saban et al., the perceived HRQoL was assessed 3 months after selected types of lumbar spine surgery [35]. In their study preoperatively the PCS was 29.39 (SD 8.10) and postoperatively 38.66 (SD 11.99) while preoperatively the MCS was 46.43 (SD 11.90) and postoperatively 49.99 (SD 11.29). The improvement was significant in the PCS but not in the MCS when SF12v2 was used as a HRQoL measure. The levels of the PCS and MCS were quite the same preoperatively and postoperatively compared to our study. However, only 23% of the patients in the study of Saban et al. had spinal fusion and the sample size was 57. This may lead to contradictory results as in our study the improvement was significant.

Mean VAS_{back} and VAS_{leg} were preoperatively 66 (SD27) and 67 (SD25) respectively. In the current study performed during the rapid response phase, improvement in pain was positive; mean VAS_{back} and VAS_{leg} changed 65% and 67%, respectively. During the same time period back pain (scale 0–10) has been reported to change from 7.00 (SD 1.80) to 3.19 (SD 2.30). This change of 54% was statistically significant, $p < 0.001$ [35].

The relationship between the ODI and HRQoL was assessed in our study. There was more often simultaneous positive change in both the ODI and the PCS than in both the ODI and the MCS. In 80% of patients the changes in the total ODI and in PCS scores were positive; in 64% of patients the changes in the total ODI and in MCS scores were positive. This can be interpreted to mean that the ODI and PCS express marked agreement in evaluating the change in disability during patients' early recovery phase.

Patient's perspective has become the main focus in assessing the outcome of spinal fusion operations. A current challenge is the need for multiple measurement tools to adequately evaluate patients' function, disability, and health [21]. One enhancement is to use the WHO's ICF as a common language in describing patients' condition through different outcome measures [19].

Recovery of the patient is incomplete 3 months after spinal fusion surgery and there exists a lot of individual variety how long time the total period of recovery takes. It is possible that those individuals with poor improvement at this time may benefit from more intensive guidance and treatment. There
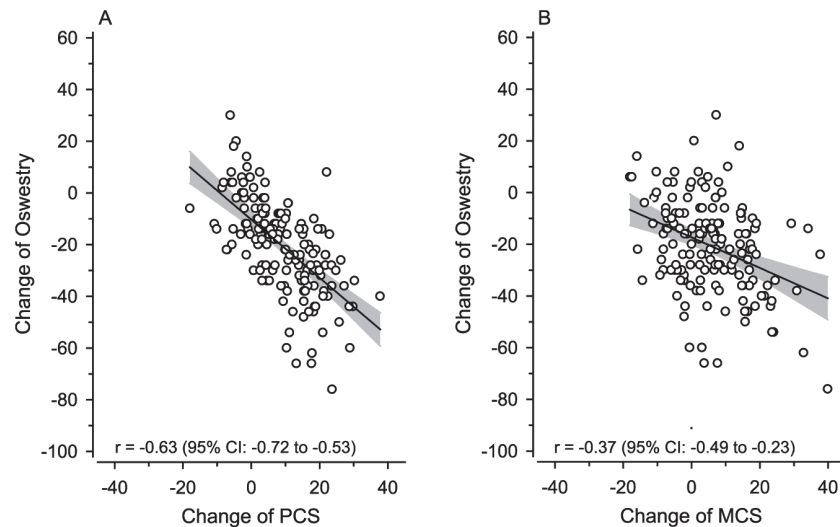
Figure 4. Relationship between the change of the Oswestry Disability Index and A) the change of Physical Component Score (PCS) and B) the change of Mental Component Score (MCS) of SF-36. Linear regression line is shown with 95% confidence intervals. Correlation is expressed with Pearson correlation coefficient.

are only a limited number of studies focusing on postoperative rehabilitation, though there is evidence that back-café concept including supplementary meetings among fusion patients during the rehabilitation period is more efficient as a rehabilitation model than physical therapy at 2 year follow-up [36]. In the future the link between ODI and ICF may be important to determine separately in fusion patients in both the good and the poor outcome sub-populations.

Study strengths and limitations

The strength of this study is the use of ICF framework in a real cohort of patients. This is the first study to our knowledge in which the analysis is performed so that the change of disability is shown via ICF domains. However, as we included only the patients with spinal fusion, generalization of these results to any back patients should be done with care. The first limitation of the study is the shortness of the follow-up. Another limitation is a lack of control group but as we are studying effectiveness of the treatment of the fusion patient cohort in usual clinical setting, the use of control group is not obligatory. One fifth of the patients were excluded because they had missing answers in the questionnaires. However, this may not lead to bias as these excluded patients did not differ from included patients according to the age, gender or level of pain (data not shown). Although extended follow-up period is needed to evaluate the long-term outcome of the surgical procedure. It is important to identify the patients at risk for poor outcome even at this early point of time to improve their recovery later. In this analysis there is no information about patients' returning to work after the surgical procedure. These data need a little longer follow-up because the usual time point of returning to work is most often three months postoperatively. A limitation is also that the female-male ratio is not equal: 68% of the patients in this material were females. There are also other studies where females are over-represented and we think that this does not distort the results [13,35].

Conclusion

Our results show that in the early recovery phase the positive changes in disability, HRQoL, and pain were significant. The improvement in functioning was significantly related to positive change in HRQoL. Still there is a subgroup of patients having marked disability possibly needing more intensive rehabilitation and follow-ups. When the ODI as a back specific measurement tool was assessed in the ICF framework the significant change at 3 months after the spinal fusion operation was seen in three components of ICF. However, the ODI was not able to cover the comprehensive core set for low back pain completely.

Acknowledgements

Declaration of Interest: The authors report no conflict of interest.

References

1. Deyo RA, Nachemson A, Mirza SK. Spinal-fusion surgery - the case for restraint. N Engl J Med 2004;350:722–726.
2. Deyo RA, Gray DT, Kreuter W, Mirza S, Martin BI. United States trends in lumbar fusion surgery for degenerative conditions. Spine (Phila Pa 1976) 2005;30:1441.
3. Andersen T, Videbaek TS, Hansen ES, Bünger C, Christensen FB. The positive effect of posterolateral lumbar spinal fusion is preserved at long-term follow-up: a RCT with 11-13 year follow-up. Eur Spine J 2008;17:272–280.
4. Haefeli M, Elfering A, McIntosh E, Gray A, Sukthankar A, Boos N. A cost-benefit analysis using contingent valuation techniques: a feasibility study in spinal surgery. Value Health 2008;11:575–588.
5. Fritzell P, Hägg O, Wessberg P, Nordwall A; Swedish Lumbar Spine Study Group. 2001 Volvo Award Winner in Clinical Studies: Lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter

randomized controlled trial from the Swedish Lumbar Spine Study Group. Spine 2001;26:2521–32; discussion 2532.

6. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY. Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the Oswestry Disability Index, Medical Outcomes Study questionnaire Short Form 36, and pain scales. Spine J 2008;8:968–974.

7. Glassman SD, Carreon LY, Djurasovic M, Dimar JR, Johnson JR, Puno RM, Campbell MJ. Lumbar fusion outcomes stratified by specific diagnostic indication. Spine J 2009;9:13–21.

8. Glassman SD, Polly DW, Bono CM, Burkus K, Dimar JR. Outcome of lumbar arthrodesis in patients sixty-five years of age or older. J Bone Joint Surg Am 2009;91:783–790.

9. Dimar JR, Glassman SD, Burkus JK, Pryor PW, Hardacker JW, Carreon LY. Two-year fusion and clinical outcomes in 224 patients treated with a single-level instrumented posterolateral fusion with iliac crest bone graft. Spine J 2009;9:880–885.

10. Becker P, Bretschneider W, Tuschel A, Ogon M. Life quality after instrumented lumbar fusion in the elderly. Spine 2010;35:1478–1481.

11. Carragee EJ, Cheng I. Minimum acceptable outcomes after lumbar spinal fusion. Spine J 2010;10:313–320.

12. Djurasovic M, Glassman SD, Howard JM, Copay AG, Carreon LY. Health-related quality of life improvements in patients undergoing lumbar spinal fusion as a revision surgery. Spine 2011;36:269–276.

13. Glassman S, Gornet MF, Branch C, Polly D, Peloza J, Schwender JD, Carreon L. MOS short form 36 and Oswestry Disability Index outcomes in lumbar fusion: a multicenter experience. Spine J 2006;6:21–26.

14. Fairbank JC, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. Physiotherapy 1980;66:271–273.

15. Fairbank JC, Pynsent PB. The Oswestry Disability Index. Spine 2000;25:2940–52; discussion 2952.

16. Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. Health Econ 1993;2:217–227.

17. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. Pain 1983;17:45–56.

18. Stucki G, Sigl T. Assessment of the impact of disease on the individual. Best Pract Res Clin Rheumatol 2003;17:451–473.

19. WHO. International classification of functioning, disability and health: ICF; 2001.

20. Cerniauskaite M, Quintas R, Boldt C, Raggi A, Cieza A, Bickenbach JE, Leonardi M. Systematic literature review on ICF from 2001 to 2009: its use, implementation and operationalisation. Disabil Rehabil 2011;33:281–309.

21. Sigl T, Cieza A, Brockow T, Chatterji S, Kostanjsek N, Stucki G. Content comparison of low back pain-specific measures based on the International Classification of Functioning, Disability and Health (ICF). Clin J Pain 2006;22:147–153.

22. Häkkinen A, Kautiainen H, Järvenpää S, Arkela-Kautiainen M, Ylinen J. Changes in the total Oswestry Index and its ten items in females and males pre- and post-surgery for lumbar disc herniation: a 1-year follow-up. Eur Spine J 2007;16:347–352.

23. Cieza A, Brockow T, Ewert T, Amman E, Kollerits B, Chatterji S, Ustün TB, Stucki G. Linking health-status measurements to the international classification of functioning, disability and health. J Rehabil Med 2002;34:205–210.

24. Cieza A, Geyh S, Chatterji S, Kostanjsek N, Ustün B, Stucki G. ICF linking rules: an update based on lessons learned. J Rehabil Med 2005;37:212–218.

25. Røe C, Sveen U, Bautz-Holter E. Retaining the patient perspective in the International Classification of Functioning, Disability and Health Core Set for low back pain. Patient Prefer Adherence 2008;2:337–347.

26. Pekkanen L, Kautiainen H, Ylinen J, Salo P, Häkkinen A. Reliability and validity study of the Finnish version 2.0 of the oswestry disability index. Spine 2011;36:332–338.

27. Osthus H, Cziske R, Jacobi E. Cross-cultural adaptation of a German version of the Oswestry Disability Index and evaluation of its measurement properties. Spine 2006;31:E448–E453.

28. Deyo RA, Battie M, Beurskens AJ, Bombardier C, Croft P, Koes B, Malmivaara A, et al. Outcome measures for low back pain research. A proposal for standardized use. Spine 1998;23:2003–2013.

29. Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. BMJ 2002;324:1417.

30. Cohen J. Statistical power analysis for the behaviral sciences. rev ed. New York: Academc Press; 1988.

31. Weinstein JN, Tosteson TD, Lurie JD, Tosteson AN, Blood E, Hanscom B, Herkowitz H, et al.; SPORT Investigators. Surgical versus nonsurgical therapy for lumbar spinal stenosis. N Engl J Med 2008;358:794–810.

32. Carreon LY, Glassman SD, Djurasovic M, Dimar JR, Johnson JR, Puno RM, Campbell MJ. Are preoperative health-related quality of life scores predictive of clinical outcomes after lumbar fusion? Spine 2009;34:725–730.

33. Feng ZZ, Cao YW, Jiang C, Jiang XX. Short-term outcome of bilateral decompression via a unilateral paramedian approach for transforaminal lumbar interbody fusion with unilateral pedicle screw fixation. Orthopedics 2011;34:364,20110317-05.

34. Rodríguez-Vela J, Lobo-Escolar A, Joven-Aliaga E, Herrera A, Vicente J, Suñén E, Loste A, Tabuenca A. Perioperative and short-term advantages of mini-open approach for lumbar spinal fusion. Eur Spine J 2009;18:1194–1201.

35. Saban KL, Penckofer SM, Androwich I, Bryant FB. Health-related quality of life of patients following selected types of lumbar spinal surgery: a pilot study. Health Qual Life Outcomes 2007;5:71.

36. Christensen FB, Laurberg I, Bünger CE. Importance of the back-café concept to rehabilitation after lumbar spinal fusion: a randomized clinical study with a 2-year follow-up. Spine 2003;28:2561–2569.

BMC
Musculoskeletal Disorders

RESEARCH ARTICLE

Open Access

Disability and health-related quality of life in patients undergoing spinal fusion: a comparison with a general population sample

Liisa Pekkanen[1*], Marko H Neva[2], Hannu Kautiainen[3,4], Joost Dekker[5], Kirsi Piitulainen[6], Marko Wahlman[2] and Arja Häkkinen[6,7]

Abstract

Background: The aim of the present study was to compare one-year-follow-up data on disability and health-related quality of life (HRQoL) between spinal fusion patients and age- and sex-matched general population.

Methods: The data on fusion patients were collected prospectively using a spinal fusion data base in two Finnish hospitals. A general population sample matched for age, sex and residential area was drawn from the Finnish Population Register. All participants completed a questionnaire and the main outcome measures were the Oswestry Disability Index (ODI) and the Short Form-36 questionnaire (SF-36).

Results: Altogether 252 (69% females) fusion patients and 682 (67% females) population sample subjects participated in the study. In general population the mean ODI was 15 (SD 17) in females and 9 (SD 13) in males. The corresponding preoperative ODI values were 47 (SD16) and 40 (SD 15) and one year follow-up values 22 (SD 17) and 23 (SD 20). In both sexes the ODI decreased significantly after surgery but remained higher than in the general population, $p < 0.001$. The physical component summary score (PCS) of the SF-36 was lower in the patients than general population sample both preoperatively and at one-year follow-up ($p < 0.001$). The mental component summary score (MCS) was lower preoperatively ($p < 0.001$), but reached the general population level after one year in both men ($p = 0.42$) and women ($p = 0.61$).

Conclusions: Disability and HRQoL improved significantly after spinal fusion surgery during a one- year follow-up. However, the patients did not reach the level of the general population in the ODI or in the physical component of HRQoL at that time, although in the mental component the difference disappeared.

Keywords: Spinal fusion, Oswestry disability index, Health-related quality of life, General population sample

Background

With the ageing of the population, an increase in degenerative spine conditions and the number of surgical patients can be expected [1]. Although instrumented spinal fusions have been performed since the early 1960s, these procedures remain controversial owing to inconsistent responses to the treatment [2,3]. In the field of spinal fusion outcome research, most earlier trials have compared surgical methods and assessed the success of the surgical procedure itself. However, over the last few decades there has been a trend towards the use of patient-reported outcomes (PROs) in evaluating the outcome of fusion surgery in addition to physical examinations, imaging or clinical outcome scales. Recently, the routine administration of certain instruments in connection with low back pain and surgical treatment has been recommended [4,5]. Condition-specific disability measures like the Oswestry Disability Index (ODI) should be used before and after surgical treatments. When evaluating surgical outcomes in the clinical-research setting, Health Related Quality of Life (HRQoL) tools, such as the Short-Form 36 (SF-36), Short Form 12 (SF-12) or EuroQol Group (EQ-5D) should be used [4].

* Correspondence: pekkanen.auvinen@fimnet.fi
[1]Department of Orthopaedics and Traumatology, Jyväskylä Central Hospital, Jyväskylä, Finland
Full list of author information is available at the end of the article

Information on whether the operation provides effective relief of symptoms and disability continues to be lacking [3]. In defining success after spinal fusion operations, calculation of the minimum clinically important difference (MCID) in PROs has been suggested. This method has limitations; for example the MCID values may differ according to multiple factors, such as original spine pathology, the method of treatment, sample size and patient-characteristics, e.g. baseline scores [6]. Another method that has been proposed is based on prospective minimum goals established by individual patients themselves. In the study by Carragee et al. isthmic spondylolisthesis patients and degenerative disc disease patients preoperatively indicated their expectations concerning level of function (ODI), work capacity, pain intensity and medication requirement [7]. One of most recent attempts to solve the difficulty in defining clinical success after spinal fusion operations is to analyze whether the patients reach the level of the general population in certain PROs. To our knowledge only one study by Mokhtar et al. [1]. has used this method. Prospective data on 100 patients undergoing spinal fusion were collected using the SF-12 questionnaire, and the results were compared to those obtained for a sample of the general population. No disease-spesific disability measurement was used in this study. As only limited amount of information exists on the use of this method, so there is a clear need for further studies.

The aim of the present study was to compare disability and HRQoL among spinal fusion patients within one-year follow-up with the values of an age- and sex-matched population resident in the same district.

Methods

Since the beginning of 2008, all patients undergoing spinal fusion surgery in Tampere University Hospital or Jyväskylä Central Hospital have been recruited to a prospective follow-up study.

In August 2010, the spinal database comprised 285 patients with the 6 most common diagnoses for elective spinal fusion. These diagnoses were degenerative spondylolisthesis, spondylolysis, spinal stenosis, disc herniation or degeneration, postoperative conditions and degenerative scoliosis. Disability and HRQoL measures were available preoperatively and 3 and 12 months postoperatively for 252 of these patients (88%) all of whom were included in this study. Six surgeons had performed the operations, and in most cases in teams of two surgeons.

The cohort of spinal fusion patients was compared to a general population sample matched according to age, sex and residential area. Four controls for each of these fusion patients was drawn from the Finnish Population Register and the sampling was performed by the Statistics Finland. A questionnaire was mailed to 1 140 controls in September 2010, and one reminder letter was sent two months later. After one reminder letter, the percentage of returned answers was 61% (n = 691) and the number of acceptable answers 682.

One to two weeks prior to the fusion operation, the patients filled in a questionnaire requesting sociodemographic and clinical information, for example weight, height, presence of co-morbidities, exercise habits, smoking and employment status. The main outcome measures were the Oswestry Disability Index (ODI) and the Short Form-36 Questionnaire (SF-36). The ODI is one of most widely used back-specific disability measurement tools in both clinical work and research.[8,9] According to the original publication, the scores are grouped into five categories: 0–20 minimal, 20–40 moderate, 40–60 severe disability; 60–80 crippled and 80–100 indicates that the patient is either bed-bound or exaggerating his or her symptoms [8]. The Finnish validated version 2.0 of the ODI was used [10]. The SF-36 is a generic patient-assessed health outcome measure for health-related quality of life with eight dimensions reflecting patients' health and welfare. The SF-36 score also divides into two summary measures: the physical component summary score (PCS) and the mental component summary score (MCS). The dimensions Physical Functioning, Role Physical, Bodily Pain and General Health form the PCS, and Mental Health, Vitality, Social Functioning and Role-Emotional the MCS. Version 1 of the SF-36 qustionnaire was used on this study.

The ethical committees in Tampere University Hospital and Jyväskylä Central Hospital approved the study plan and all the participating patients signed a written consent.

Statistics

Results are expressed as mean and standard deviation (SD). Statistical comparison between the groups was performed by t-test, bootstrap-type t-test (5000 replications), or chi-square test, where appropriate. Differences in the ODI and HRQoL between the groups were determined using generalized linear models. Repeated measures were analyzed using linear mixed models.

Results

The demographical and clinical data of the fusion patients and general population is shown in Table 1. Sixty-nine per cent of the 252 fusion patients and 67% of the 682 general population subjects were females. In the population sample, the mean age of females was higher than in the patient group: 66 (SD 11) vs. 63 (SD 12) years (p = 0.014). The mean age of males was 60 (SD 13) years in the general population and 58 years in the patients (p = 0.43). In both sexes the body mass index

Table 1 Demographical and clinical data

Variables	Female		p-value	Male		p-value
	Patients	Population		Patients	Population	
	n = 174	n = 458		n = 78	n = 224	
Body mass index, mean (SD)	28.1 (4.5)	26.9 (4.7)	0.0046	28.0 (3.8)	26.8 (3.8)	0.021
Co-morbidities, n (%)						
Cardiological	100 (59)	197 (43)	<0.001	37 (48)	81 (36)	0.065
Respiratory	24 (14)	51 (11)	0.29	4 (5)	15 (7)	0.64
Neurological	7 (4)	26 (6)	0.45	2 (3)	10 (4)	0.47
Rheumatoid	21 (12)	29 (6)	0.012	2 (3)	3 (1)	0.46
Diabetes	15 (9)	60 (13)	0.15	14 (18)	27 (12)	0.18
Psychiatric	7 (4)	18 (4)	0.90	2 (3)	7 (3)	0.82
Musculosceletal	10 (6)	48 (10)	0.080	1 (1)	7 (3)	0.39
Education, years, mean (SD)	11.3 (3.5)	11.6 (4.2)	0.33	11.5 (3.4)	11.7 (3.7)	0.78
Tobacco use, n (%)	17 (10)	46 (10)	0.99	13 (17)	42 (19)	0.72
Employment situation, n (%)			0.71			0.38
Employed	50 (29)	140 (31)		37 (47)	98 (44)	
Unemployd	4 (2)	15 (3)		2 (3)	15 (7)	
Retired	120 (69)	303 (66)		39 (50)	111 (49)	
Leisure time physical activity hours per week, mean (SD)	3.3 (3.7)	4.4 (3.9)	0.0019	4.7 (4.2)	4.6 (6.2)	0.95

(BMI) was significantly higher in the fusion patients than in the general population. The number of cardiological (p < 0.001) and rheumatoid co-morbidities (p = 0.012) was higher in the female patients than in the female population subjects and the female patients were also less physically active. In the general population, 5.7% of the females and 8.0% of the males had spinal disorders.

In the general population, females had higher ODI (15 (SD17)) than males (9 (SD 13)), (p < 0.001). In the patients, the preoperative ODI values were 47 (SD16) in females and 40 (SD 15) in males (p < 0.001) (Figure 1). One year post fusion the mean change in ODI was −25 (95% CI −28 to −22) in females and −17 (95% CI −21 to −13) in males. However, in both sexes, the age-adjusted ODI scores at baseline and at one-year were significantly higher than the mean ODI values in the general population (p < 0.001). The postoperative change in ODI between three months and one year was minor and not significant in males while in females the change was significant.

All the SF-36 dimensions in the general population were significantly better than the preoperative values of the patients, both in females and males, (p < 0.001). (Table 2). In both sexes the preoperative mean ratio between the patients and the general population subjects was biggest in the dimension Role-Physical. At the one-year follow-up the female patients reached the population level in Vitality, Mental Health and Role-Emotional,

while male patients reached the population level only in Vitality and Mental Health.

In the general population, the PCS of the SF-36 was 44 (SD 11) in females and 48 (SD 10) in males (Figure 2). Among the patients the preoperative PCS was 26 (SD 7) in females and 29 (SD 6) in males. At 12 months post surgery, the change in the PCS was 11 (95% CI 10 to 13; p < 0.001) in females and 10 (95% CI 7 to 12; p < 0.001) in males.

In turn the MCS of the SF-36 was 52 (SD 11) in females and 53 (SD 10) in males in the general population. The preoperative MCS was 46 (SD 13) in the female patients and 48 (SD 12) in the male patients. The positive change in the MCS from the preoperative to 12-month values was 7 (95% CI 5 to 8; p < 0.001) in females and 4 (95% CI 1 to 6; p < 0.001) in males (Figure 3). In the MCS, both the female (p = 0.42) and male (p = 0.61) patients had reached the level of the general population at one year post surgery, although, the difference in PCS between the patients and the general population remained significant (both sexes p < 0.001). In the patients, the changes in PCS and MCS between three months and one year after surgery, were minor and statistically non significant.

Discussion

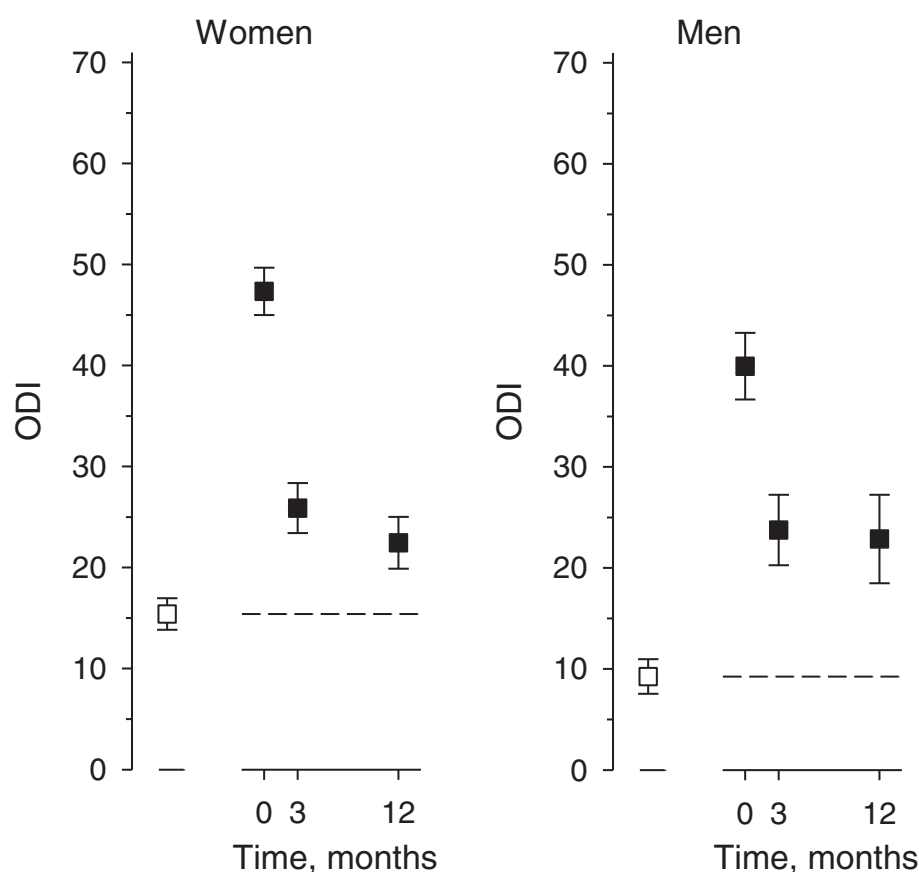Our main purpose was to study the recovery of the spinal fusion patients during a one-year follow-up and

Figure 1 The mean Oswestry Disability Index (ODI, with 95% Confidence Interval) in the patients (■) and in the population (□, dashed line).

compare our patients reported outcomes (PRO) to the values of a matched general population sample. The results showed that despite considerable improvement during the follow-up the patients did not reach the level of the matched general population in either disability or the physical component of the HRQoL.

To our knowledge, this is the first study where the PROs of spinal fusion patients have been compared to general population values. The general population subjects showed minimal disability in the mean ODI scores according to the original scoring, while the fusion patients' mean ODI scores were preoperatively severe and at one year after the spinal fusion surgery remained moderate [8]. Therefore, inspite of recovery the disability according to the ODI did not decrease to the level of general population in our follow-up in males or in females. One explanation for this might be, that the patients undergoing fusion operation have often suffered from longstanding spinal symptoms which may have caused permanent changes to their life and behavior.

Interestingly the change in the ODI between 3 months and one year was minimal. This finding suggests that

already the early recovery at three months may probably have quite high prognostic value when assessing the success of the treatment, also over a longer period. This result is supported by a finding in the earlier literature [11]. In a study of 96 patients undergoing spinal fusion, pain measurements were conducted at 6 months and then yearly over a total follow-up of 5 years. An interesting finding was that the improvement in the pain scale was biggest at 6 months and in the ODI at one year. The improvements seen in this early phase were maintained throughout the remainder of the follow-up period [11].

In the present study, one of the main findings concerning disability was that the mean levels of the ODI had not reached the values of the general population in either sex at one year post surgery. In comparison with the results in disability reported in the literature, in a trial implemented at 5 spine centers with 497 patients receiving one or two level spinal fusion, the mean ODI improved by 22 points at one year postoperatively. The preoperative level of the ODI varied in different subgroups from 48 to 56 [12]. In the

Table 2 Health-related quality of life in population and patients preoperatively stratified by sex
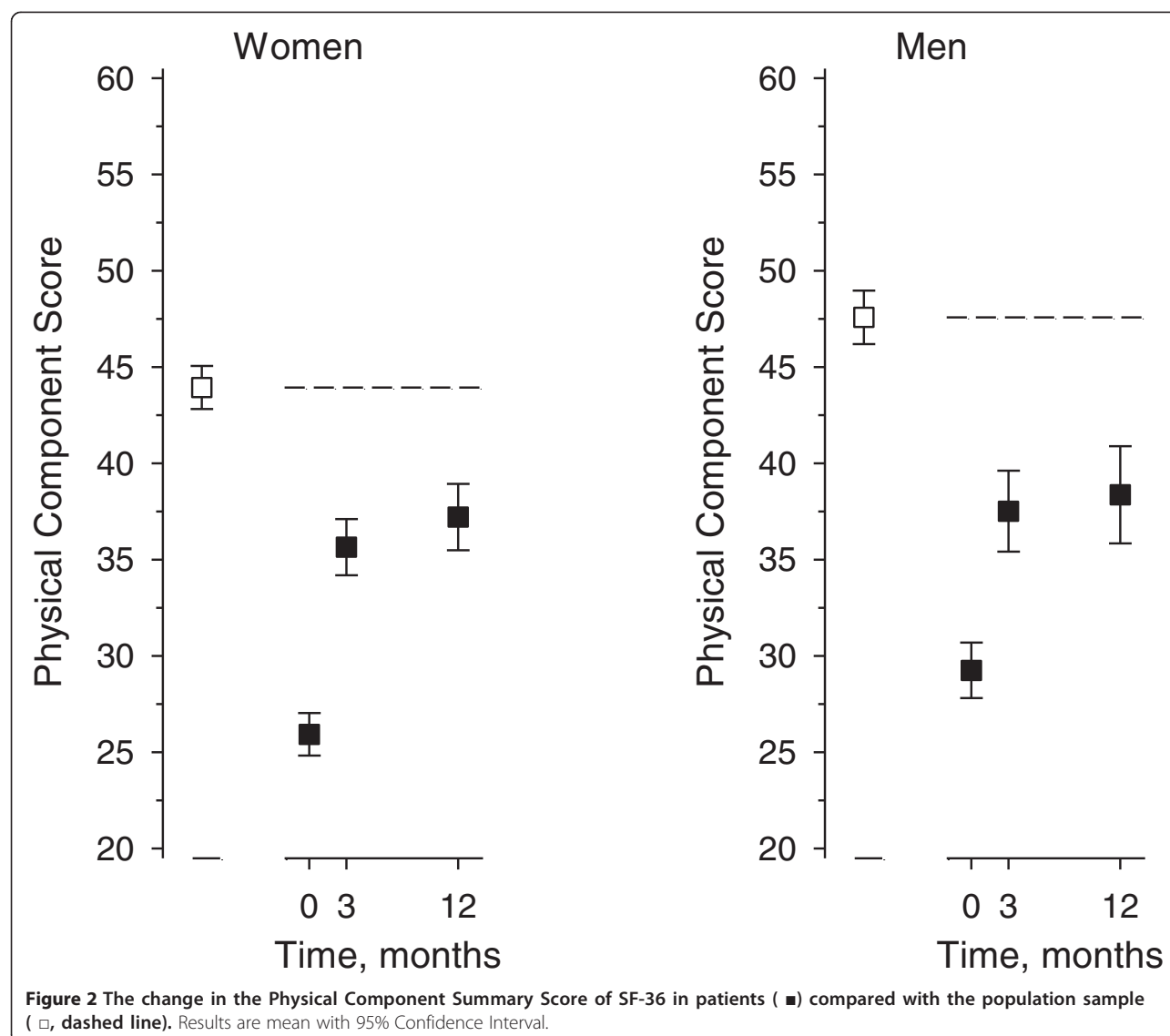
SF-36 dimensions	Population Mean (SD)	Patients Preoperative mean (SD)	Patients 12 months mean (SD)	Mean ratio* (95% CI) Preoperative	Mean ratio* (95% CI) p-value patients vs population preoperative	Mean ratio* (95% CI) 12 months	Mean ratio* (95% CI) p-value patients vs population 12 months
Female							
Physical functioning	70 (28)	28 (19)	58 (29)	2.6 (2.3 to 3.0)	<0.001	1.3 (1.2 to 1.4)	<0.001
General health	60 (22)	53 (20)	56 (21)	1.1 (1.1 to 1.2)	<0.001	1.1 (1.0 to 1.1)	0.022
Vitality	65 (23)	45 (22)	64 (23)	1.5 (1.3 to 1.6)	<0.001	1.0 (1.0 to 1.1)	0.41
Mental health	77 (19)	63 (21)	77 (19)	1.2 (1.2 to 1.3)	<0.001	1.0 (1.0 to 1.0)	0.79
Role physical	64 (42)	9 (21)	44 (43)	7.9 (2.7 to 13.0)	<0.001	1.5 (1.3 to 1.7)	<0.001
Role emotional	71 (39)	46 (43)	67 (41)	1.6 (1.4 to 1.8)	<0.001	1.1 (1.0 to 1.2)	0.17
Social functioning	82 (25)	46 (28)	76 (28)	1.8 (1.7 to 2.0)	<0.001	1.1 (1.0 to 1.1)	0.004
Bodily pain	67 (27)	24 (15)	56 (25)	2.8 (2.4 to 3.3)	<0.001	1.2 (1.1 to 1.3)	<0.001
Male							
Physical functioning	84 (22)	39 (20)	62 (26)	2.2 (1.9 to 2.4)	<0.001	1.3 (1.2 to 1.5)	<0.001
General health	65 (21)	56 (21)	55 (23)	1.2 (1.1 to 1.3)	<0.001	1.2 (1.1 to 1.3)	<0.001
Vitality	71 (22)	53 (23)	66 (24)	1.4 (1.2 to 1.5)	<0.001	1.1 (1.0 to 1.2)	0.072
Mental health	81 (18)	68 (21)	76 (19)	1.2 (1.1 to 1.3)	<0.001	1.1 (1.0 to 1.2)	0.074
Role physical	74 (38)	12 (21)	44 (43)	6.3 (2.1 to 10.5)	<0.001	1.7 (1.3 to 2.0)	<0.001
Role emotional	79 (35)	44 (43)	65 (42)	1.8 (1.5 to 2.2)	<0.001	1.2 (1.1 to 1.4)	0.003
Social functioning	87 (20)	62 (27)	75 (36)	1.4 (1.3 to 1.5)	<0.001	1.2 (1.1 to 1.3)	<0.001
Bodily pain	75 (23)	30 (16)	55 (29)	2.5 (2.1 to 3.0)	<0.001	1.4 (1.2 to 1.5)	<0.001

*Adjusted age.

Swedish Lumbar Spine Study, which was a multicenter randomized controlled trial where patients were randomized into a surgical or a control group, 222 patients received spinal fusion either by non-instrumented fusion, by instrumented posterolateral fusion or by circumferential fusion. In the surgical group, the mean ODI improved from 47 to 36 (p < 0.0001), at two-year follow up [13]. In a prospective randomized controlled study of 111 patients with adult isthmic spondylolisthesis the preoperative ODI scores were not reported but at two years in the surgical group the mean ODI score was 26 (95% CI 18.1 to 31.6) [14]. In our study at one year the mean positive change of the ODI in female patients was 25 (95% CI 22 to 28) and in male patients 17 (95% CI 13 to 21) and the corresponding mean ODI scores were 22 (SD 17) and 23 (SD 20).

In the present study, the spinal fusion patients reached the values of their matched population sample in the mental component (MCS) of SF-36 but not in the physical component (PCS). Preoperatively, the value of Role Physical was highest in patients in both sexes in the mean ratio analysis. At 12 months, Vitality, Mental Health and Role-Emotional were the only dimensions in

the female patients that reached general population values. In males, this was true only for Vitality and Mental Health. Interestingly, the Pain dimension was still significantly worse in patients at the one-year follow-up compared to the general population. This prompts the question: how should we manage the physical aspect in the long term recovery. The earlier literature includes a multicenter study with 497 patients undergoing one- or two-level spinal fusion with several techniques. The results showed an improvement in mean PCS of 9.9 points over a one-year follow up [12]. This finding is confirmed by our study in which the mean PCS improved by 11 (95% CI 10 to 13) points in females and 10 (95% CI 7 to 12) in males in one-year follow-up. Another study with 100 primary spinal fusion patients who received decompression and single-level posterior lumbar interbody fusion reported HRQoL scores in both the PCS-12 and MCS-12 that approached the Australian population norm over a follow-up varying from 12 months to 5 years [1]. The mean postoperative PCS-12 score was 39 (95% CI 37 to 42) and MCS-12 score 52 (95% CI 50 to 55) as compared with the corresponding population norm values of 44 (95% CI 43 to 46) and 54 (95% CI 53 to 55) [1]. To our knowledge no other studies have used

Figure 2 The change in the Physical Component Summary Score of SF-36 in patients (▪) compared with the population sample (□, dashed line). Results are mean with 95% Confidence Interval.

a population-based method when exploring the success of fusion operations.

After spinal fusion operations it is seldom a realistic goal to expect that all of the disability will disappear. It is also obvious that the level of disability, and hence quality of life, depends on various factors such as patient's age, possible chronic neuropathic pain, and other possible diseases in addition to spinal disorders. However, in the present study, the prevalence of diabetes or most of the other co-morbidities, was similar in patients and in the general population. Interestingly only cardiovascular and rheumatoid diseases in females were more often present in patients than in the general population. Patients who have undergone spinal fusion may also get other sources of pain like osteoarthritis of hip or knee and these reasons may confuse the answers in the

questionnaires. Furthermore, in spinal fusion surgery, complications and failed fusions may worsen the results. Finally, in the evaluation of disability of the patients it is essential to understand the level of disability in the general population of same age and sex. This data is important in evaluating the influence of surgery for the patients and also in surgical decision making in individual cases.

Study strengths and limitations

The present study includes register based, not selected, consecutive patient material. The main strength of this study is the comparison between patients and the general population in disability and HRQoL scores. To our knowledge, this is the only study in which the PROs of fusion patients have been compared to those of a
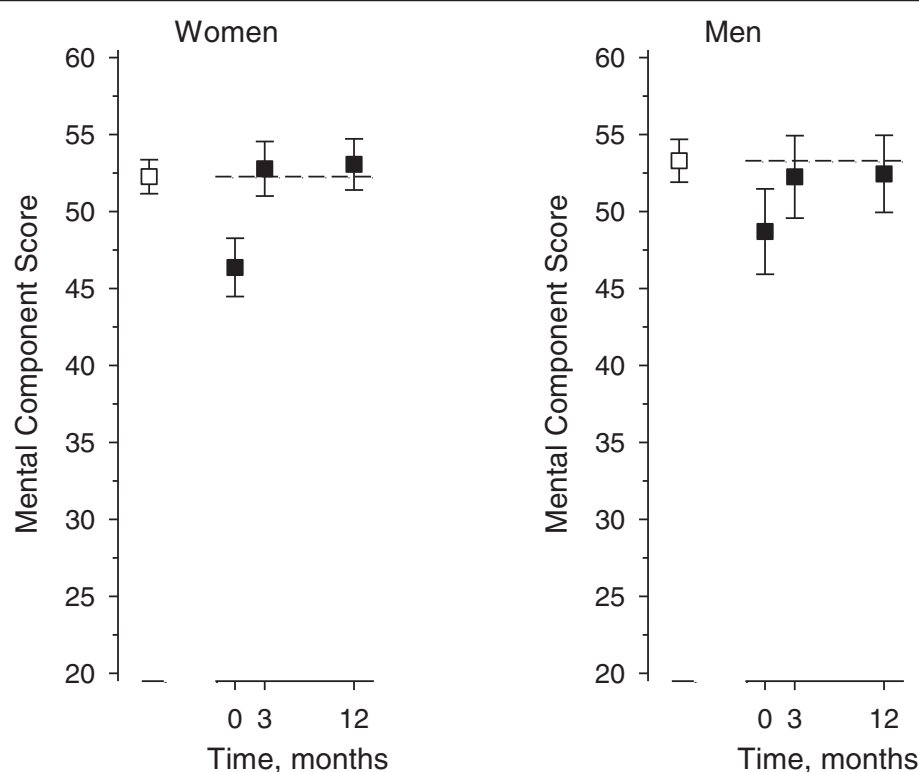
Figure 3 The change in the Mental Component Summary Score of SF-36 in patients (■) compared with the population sample (□, dashed line). Results are mean with 95% Confidence Interval.

matched general population sample. An additional strength is the accurate timing of the data collection. The preoperative data were collected one to two weeks prior to the operation and the data collection timepoints during the follow-up were strict. In addition, the population based data were collected from the same residential area compared to the patients. In the analyses, females and males have been systematically stratified. This is because the majority of the patients operated on were females and because there was a significant gender difference in the ODI in the general population between females and males. A limitation in this study is the lack of analyses stratified by surgical diagnostic indication for the fusion operation. This is due the number of patients in this material, which could have led to a too small sample size in some of the diagnostic subgroups and lack of statistical representation of the phenomenon. 3Another limitation is that as a part of the surgical procedure in our patients, also decompression through laminectomy was performed whenever appropriate. This might cause difficulty to determine how much of the total improvement of HRQoL is caused by the fusion alone and how much by the coexisting decompression procedure. Further, a limitation is also the possible bias in answering to the general population questionnaire.

Would those general population individuals who have back pain, reply more eagerly, making the observed difference between general population and patients smaller than the true value? In the literature it has been shown, that the life-time prevalence of back-pain in normal population is even 84% [15]. This leads to thinking, that even though there might be a bias in answering profile, it is not affecting the results between the patients and general population significantly. The follow-up in our study was 12 months. This period of time seemed to be sufficient to show, that results in disability and quality of life stabilized after three months. Although a one-year of follow-up indicated a trend towards recovery, further follow-ups of several years are needed to evaluate the longer term outcome.

Conclusions

In this study the data of 252 spine fusion patients was analyzed and compared with general population. Despite the significant improvement during the one-year follow-up in both disability and HRQoL, the patients did not reach the level of general population in the ODI or in the PCS. In the MCS, however, both female and male patients reached the level of general population.

Abbreviations

CI: Confidence interval; EQ-5D: The EuroQol group-5D; HRQol: Health-related quality of life; MCS: Mental component summary score; ODI: Oswestry disability index; PCS: Physical component summary score; PRO: Patient-reported outcome; SD: Standard deviation; SF-36: Short-form-36.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

LP: corresponding author,participated in the design of the study, participated in the patient collection, drafted the manuscript, finished the manuscript, read and approved the final manuscript. MHN: participated in the design of the study, participated in the patient collection, revised the manuscript critically, read and approved the final manuscript. HK: performed the statistical analysis, revised the manuscript critically, read and approved the final manuscript. JD: participated in the design of the study, read and approved the final manuscript. KP: participated in to the patient collection, read and approved the final manuscript. MW: participated in the patient collection, read and approved the final manuscrip. AH: participated in the design of the study, helped to draft the manuscript, revised the manuscript critically, read and approved the final manuscript

Author details

[1]Department of Orthopaedics and Traumatology, Jyväskylä Central Hospital, Jyväskylä, Finland. [2]Department of Orthopaedics and Traumatology, Tampere University Hospital, Tampere, Finland. [3]Unit of Family Practice, Central Finland Central Hospital, Jyväskylä, Finland. [4]Unit of Primary Health Care, Kuopio University Hospital, Kuopio, Finland. [5]VU Medical Center Amsterdam, Amsterdam, Netherlands. [6]Department of Physical Medicine and Rehabilitation, Jyväskylä Central Hospital, Jyväskylä, Finland. [7]Department of Health Sciences, University of Jyväskylä, Jyväskylä, Finland.

References

1. Mokhtar SA, McCombe PF, Williamson OD, Morgan MK, White GJ, Sears WR: **Health-related quality of life: a comparison of outcomes after lumbar fusion for degenerative spondylolisthesis with large joint replacement surgery and population norms.** *Spine J* 2010, **10**(4):306–312.
2. Roy-Camille R, Saillant G, Mazel C: **Internal fixation of the lumbar spine with pedicle screw plating.** *Clin Orthop Relat Res* 1986, **203**(203):7–17.
3. Gibson JN, Waddell G: **Surgery for degenerative lumbar spondylosis: updated cochrane review.** *Spine (Phila Pa 1976)* 2005, **30**(20):2312–2320.
4. DeVine J, Norvell DC, Ecker E, Fourney DR, Vaccaro A, Wang J, Andersson G: **Evaluating the correlation and responsiveness of patient-reported pain with function and quality-of-life outcomes after spine surgery.** *Spine (Phila Pa 1976)* 2011, **36**(21 Suppl):S69–S74.
5. Chapman JR, Norvell DC, Hermsmeyer JT, Bransford RJ, DeVine J, McGirt MJ, Lee MJ: **Evaluating common outcomes for measuring treatment success for chronic low back pain.** *Spine (Phila Pa 1976)* 2011, **36**(21 Suppl):S54–S68.
6. Copay AG, Glassman SD, Subach BR, Berven S, Schuler TC, Carreon LY: **Minimum clinically important difference in lumbar spine surgery patients: a choice of methods using the oswestry disability index, medical outcomes study questionnaire short form 36, and pain scales.** *Spine J* 2008, **8**(6):968–974.
7. Carragee EJ, Cheng I: **Minimum acceptable outcomes after lumbar spinal fusion.** *Spine J* 2010, **10**(4):313–320.
8. Fairbank JC, Couper J, Davies JB, O'Brien JP: **The oswestry low back pain disability questionnaire.** *Physiotherapy* 1980, **66**(8):271–273.
9. Fairbank JC, Pynsent PB: **The oswestry disability index.** *Spine* 2000, **25**(22):2940–2952. discussion 2952.
10. Pekkanen L, Kautiainen H, Ylinen J, Salo P, Hakkinen A: **Reliability and validity study of the finnish version 2.0 Of the oswestry disability index.** *Spine (Phila Pa 1976)* 2011, **36**(4):332–338.
11. Glassman SD, Polly DW, Dimar JR, Carreon LY: **The cost effectiveness of single-level instrumented posterolateral lumbar fusion at 5 years after surgery.** *Spine (Phila Pa 1976)* 2012, **37**(9):769–774.
12. Glassman S, Gornet MF, Branch C, Polly D Jr, Peloza J, Schwender JD, Carreon L: **MOS short form 36 and oswestry disability index outcomes in lumbar fusion: a multicenter experience.** *Spine J* 2006, **6**(1):21–26.
13. Fritzell P, Hagg O, Wessberg P, Nordwall A, Group SLSS: **Volvo award winner in clinical studies: lumbar fusion versus nonsurgical treatment for chronic low back pain: a multicenter randomized controlled trial from the swedish lumbar spine study group.** *Spine (Phila Pa 1976) 2001* 2001, **26**(23):2521–2532. discussion 2532–4.
14. Ekman P, Moller H, Hedlund R: **The long-term effect of posterolateral fusion in adult isthmic spondylolisthesis: a randomized controlled study.** *Spine J* 2005, **5**(1):36–44.
15. Walker BF: **The prevalence of low back pain: a systematic review of the literature from 1966 to 1998.** *J Spinal Disord* 2000, **13**(3):205–217.