Pain 122 (2006) 145–155

PAIN

www.elsevier.com/locate/pain

Lumbar instrumented fusion compared with cognitive intervention and exercises in patients with chronic back pain after previous surgery for disc herniation: A prospective randomized controlled study ☆

Jens Ivar Brox [a,g,*], Olav Reikerås [a], Øystein Nygaard [b], Roger Sørensen [a], Aage Indahl [c,g], Inger Holm [a], Anne Keller [f], Tor Ingebrigtsen [d], Oliver Grundnes [e], Johan Emil Lange [a], Astrid Friis [a]

[a] *Department of Orthopedics and Physiotherapy, Rikshospitalet University Hospital and Medical Faculty University in Oslo, Norway*
[b] *Department of Neurosurgery, St. Olavs Hospital, Trondheim, Norway*
[c] *Special Hospital for Rehabilitation, Stavern, Rikshospitalet University Hospital, Oslo, Norway*
[d] *Department of Neurosurgery, University North Norway, Tromsø, Norway*
[e] *Department of Orthopedics, Ullevaal University Hospital, Oslo, Norway*
[f] *Department of Physical Medicine and Rehabilitation, Ullevaal University Hospital, Oslo, Norway*
[g] *Unifob Health, University in Bergen, Norway*

Received 31 January 2005; received in revised form 9 December 2005; accepted 26 January 2006

Abstract

The effectiveness of lumbar fusion for chronic low back pain after surgery for disc herniation has not been evaluated in a randomized controlled trial. The aim of the present study was to compare the effectiveness of lumbar fusion with posterior transpedicular screws and cognitive intervention and exercises. Sixty patients aged 25–60 years with low back pain lasting longer than 1 year after previous surgery for disc herniation were randomly allocated to the two treatment groups. Experienced back surgeons performed transpedicular fusion. Cognitive intervention consisted of a lecture intended to give the patient an understanding that ordinary physical activity would not harm the disc and a recommendation to use the back and bend it. This was reinforced by three daily physical exercise sessions for 3 weeks. The primary outcome measure was the Oswestry Disability Index (ODI). Outcome data were analyzed on an intention-to-treat basis. Ninety-seven percent of the patients, including seven of eight patients who had either not attended treatment ($n = 5$) or changed groups ($n = 2$), completed 1-year follow-up. ODI was significantly improved from 47 to 38 after fusion and from 45 to 32 after cognitive intervention and exercises. The mean difference between treatments after adjustment for gender was −7.3 (95% CI −17.3 to 2.7, $p = 0.15$). The success rate was 50% in the fusion group and 48% in the cognitive intervention/exercise group. For patients with chronic low back pain after previous surgery for disc herniation, lumbar fusion failed to show any benefit over cognitive intervention and exercises.

Keywords: Randomized clinical trial; Previous discectomy; Chronic low back pain; Posterior lumbar instrumented fusion; Cognitive intervention; Exercises

1. Introduction

Success rates of 75–80% are reported after surgery for disc herniation (Asch et al., 2002). Despite these

☆ This study was supported by grants from the Norwegian Back Association and the Foundation for Health and Rehabilitation.
* Corresponding author. Tel.: +47 23076029; fax: +47 23076010.
E-mail address: jens.ivar.brox@rikshospitalet.no (J.I. Brox).

0304-3959/$32.00
doi:10.1016/j.pain.2006.01.027

findings a considerable number of patients have reported postoperative back pain with or without leg pain (Nachemson, 1993; Manniche, 1995; Berger, 2000). Manniche found that more than half of the patients suffered from considerable back pain, sciatica, and functional disabilities. These patients are often considered for repeat disc surgery or lumbar fusion. Studies on results of surgery for disc herniation report inappropriate patient selection as the leading cause of surgical failure (Nachemson, 1993; Berger, 2000). Re-operation is usually considered if the patient has a recurrent disc herniation with clinical symptoms and signs. Inferior results are reported after repeat disc surgery in patients with postoperative pain, but without recurrent disc herniation (Vik et al., 2001). According to evidence-based medicine, activity restriction is unnecessary after lumbar disc surgery (Ostelo et al., 2003).

Three randomized studies have compared lumbar fusion to non-operative treatment (Fritzell et al., 2001; Brox et al., 2003; Fairbank et al., 2005). Fritzell et al. found that, of 294 patients, lumbar fusion reduced pain and decreased disability more efficiently than usual care within the primary health care system (Fritzell et al., 2001). Only 19% of these patients had previously been operated for lumbar disc herniation. In our previous trial, which included 64 patients, we reported that lumbar fusion and cognitive intervention and exercises showed equal improvement in patients with chronic low back pain and disc degeneration (Brox et al., 2003). The study by Fairbank et al. included 349 patients and found no clear evidence that spinal fusion was more beneficial than intensive rehabilitation. In that study, patients allocated to the rehabilitation program completed exercises and cognitive behavior intervention for 75 days. Patients in the surgery group were operated on by surgeons using the surgical technique of their choice (Fairbank et al., 2005). Only 8% of the patients had post-laminectomy pain. To our knowledge, the effectiveness of lumbar fusion in patients with chronic low back pain after surgery for disc herniation has not been evaluated in a randomized study.

Posterolateral fusion with pedicle fixation is one of a variety of contemporary fusion and non-fusion techniques employed for the treatment of degenerative disc disease. Previous randomized studies consistently show no clinical advantage of any particular technique (Deyo et al., 2004),

The study design and methods, except the history of surgery for disc herniation, are similar in the present study and the previously published study (Brox et al., 2003). We have previously published results on trunk muscle strength, cross-sectional muscle area, and density for both study populations together (Keller et al., 2004). The purpose of the present study was to compare the effect of transpedicular fusion with cognitive intervention and exercise in a prospective randomized study of patients with previous surgery for disc herniation. The predefined primary outcome measure was the difference between groups in change in Oswestry Disability Index (ODI) between baseline and 1-year follow-up. Secondary outcome measures are outlined in Section 2.

2. Materials and methods

2.1. Study design

This study was a randomized, single blind, clinical trial with prospective assessment before randomization and blinded assessment of the two parallel treatment groups by two independent observers at 1-year follow-up. The study is reported according to the Consolidated Standards of Reporting Trials (CONSORT) (Moher et al., 2001).

2.2. Participants

Patients with chronic low back pain and previous surgery for disc herniation, referred from departments of orthopedic surgery, neurosurgery, and physical medicine and rehabilitation from all regions in Norway during the period 1997–2000, were eligible to participate in the study. Clinicians at the referral hospitals performed a brief examination of patients with respect to inclusion and exclusion criteria and informed eligible patients about the trial. All patients referred for inclusion were examined at the Orthopedic Department at the Rikshospitalet University Hospital. At least one spine surgeon and one specialist in physical medicine and rehabilitation examined each patient. All patients underwent plain radiography and magnetic resonance imaging of the lumbar spine. Where indicated, a gadolinium contrast magnetic resonance imaging was performed to discriminate between disc herniation and scar tissue. Discography for imaging and pain provocation testing was not used except for a few patients for whom an additional diagnostic tool was deemed necessary.

Inclusion criteria were: age 25–60 years; pain duration of at least 1 year after previous surgery for disc herniation; a score of at least 30 of 100 points on the Oswestry Disability Index (ODI); degeneration at L4–L5 and/or L5–S1 (spondylosis) on plain radiographs. Exclusion criteria were: widespread myofascial pain; spinal stenosis with reduced walking distance and neurological signs; recurrent disc herniation or lateral recess stenosis with clinical signs of radiculopathy; inflammatory disease; previous spinal fracture; previous lumbar fusion; generalized disc degeneration on plain X-ray examination; ongoing somatic or psychiatric disease that excludes either one or both treatment alternatives; registered medical abuse; reluctance to accept one or both of the treatment regimens of the study.

All eligible patients were given verbal and written information about the study and the two treatment alternatives. Each patient signed an informed consent before participating in the study. To standardize patient management, the maximum waiting time between inclusion and start of treatment was set to 3 months for both treatments. This was a possible incentive because waiting time for spinal fusion for chronic low back pain in Norway was between 6 and 12 months at that time.

The Ethics Committee for Medical Research in Health Region I of Norway approved the study.

2.3. Randomization

Participants were randomly allocated to one of two treatment groups: posterolateral fusion with pedicle fixation or cognitive intervention and exercises. Each eligible patient was assigned an identification number by the randomization central at the University of Bergen. Concealed random allocation was conducted by a computer generated random list. Blocks of 10 patients were used to ensure fairly even-numbered treatment groups.

2.4. Treatments

2.4.1. Posterolateral fusion with pedicle fixation

Patients in the fusion group were operated on by nine experienced back surgeons at four different hospital departments: Orthopedic Department, Rikshospitalet University Hospital; Neurosurgical Department, St. Olavs Hospital; Neurosurgical Department, University Hospital Northern Norway; Orthopedic Department, Ullevaal University Hospital. The fusion procedure used was posterolateral fusion with transpedicular screws of either the L4–L5 and/or the L5–S1 segment. The types of screws and instruments used were not standardized. Autologous bone was used in all cases. Physical therapists at the respective departments advised patients regarding physical activities during the first 3 months following surgery. The surgeons conducted follow-up consultations with each patient at 3 and 6 months, with the surgeon prescribing physiotherapy, including exercises, at the first follow-up.

2.4.2. Cognitive intervention and exercises

Patients allocated to the cognitive intervention and exercises group were treated at the Physiotherapy Department at the Rikshospitalet University Hospital. This rehabilitation program has been described previously (Brox et al., 1999, 2003). The average duration of the program was about 25 h per week for 3 weeks. During the first week, a specialist in physical medicine and rehabilitation gave a presentation to the patients describing pain receptors in the discs, facet joints, and muscles, the reflexive interplay between various structures, and the ability to suppress and reinforce various peripheral stimuli. Patients were assured that they could not do any harm to the disc (back) by engaging in ordinary activities of daily life and were told to use their backs and to not be too cautious (Indahl et al., 1995). They were challenged in physical activities previously labeled as not recommended, such as vacuum cleaning, jumping, lifting, and ball games. Patients were told to bend their back while lifting light objects and to bend their knees while lifting heavy objects. Endurance and coordination exercises were also recommended, not necessarily with the goal of increasing aerobic capacity and trunk muscle strength, but to gain confidence towards engaging in ordinary activities of daily life. An additional exercise that specifically trains the co-contraction of the deep abdominal muscles with lumbar multifidus was performed according to principles outlined by O'Sullivan et al. (1997).

Individual goals for the rehabilitation process were established based on the patients' answers to a comprehensive questionnaire (thoughts and feelings) and their test results (physical function and behavior). Additionally, an operant conditioning behavioral approach (Lindström et al., 1992) was used for a few patients. This approach is aimed at teaching the patients that it is safe to move, while regaining function. The original plan was that patients would fill out a training diary, but compliance was low and that was dropped.

Because the patients were recruited from all over Norway, most patients stayed at a patient hotel. Three daily workouts were performed: aerobics or outdoor activities, water gymnastics, and individual exercises. Additionally, individual consultations, group lessons, and discussions were given. Groups of patients met with a former participant in the program in order to exchange experiences.

2.5. Outcomes

The primary outcome measure was predefined in the study protocol. The Norwegian version of the original ODI (version 1.0) was used to evaluate condition-specific disability and pain (Fairbank et al., 1990). The ODI is comprised of 10 questions concerning pain and pain-related disability in activities of daily life and social participation. Each question has six different response alternatives with each response having a score. The sum of the response scores is calculated and presented as a percentage, where 0% represents no pain or disability, and 100% represents the worst possible pain and disability. In a recent study evaluating several outcome measures in the target population, the ODI showed the highest test–retest reliability of the variables used in the questionnaire (Holm et al., 2003). Changes within 12% points using the ODI could be attributed to measurement error or random variation in a single patient with chronic low back pain (Holm et al., 2003; Grotle et al., 2004a,b).

Secondary outcome measures were assessed using a standardized questionnaire, which was completed by the patients before inclusion and again at the 1-year follow-up visit. Patients ranked the intensity of their back and leg pain during the previous week on a vertical visual analog scale ranging from 0 to 100, where 100 reflected the worst pain imaginable. Maximum pain, minimum pain, and current pain were scored on three separate scales. The mean of the three measurements provided the pain index for back pain and lower limb pain (Fritzell et al., 2001). The use of daily pain medication was registered (including drug name and dosage, reported by the patient) for the week just prior to inclusion and the 1-year follow-up. The consumption of different classes of drugs was calculated with defined daily doses as a measurement unit. Drugs were classified according to therapeutic group, defined by the anatomic therapeutic chemical system (Brox et al., 1993; Rønning et al., 2000; WHO, 2002). The therapeutic groups were: analgesics; anxiolytics, hypnotics, and sedatives; antidepressants; anti-inflammatory agents; muscle relaxants. One defined daily dose equals 3 g paracetamol (acetaminophen) or 0.5 g naproxen.

The General Function Score (GFS) was used to measure back-related disability in activities of daily living (Fritzell et al., 2001). Patients answered nine questions with one of three alternatives: "can perform", "can perform with difficulty due to back complaints", and "cannot perform due to back complaints". The score is expressed as a percentage, with 100% representing maximum disability. The reliability and validity of GFS were found to be acceptable in a Swedish study (Hägg et al., 2002). The minimal clinically important difference was 12% points (Hägg et al., 2003).

Emotional distress was rated by the Hopkins Symptom Check List-25 (Derogatis et al., 1974). Patients ranked their experience of 25 symptoms from 1 (''not at all'') to 4 (''extreme''). An average score above 1.75 is a high predictor of whether a patient is currently seeking help, but seems to reflect illness or non-specific distress more than psychiatric diagnoses (Sandanger et al., 1999).

A Norwegian version of Waddell's Fear-Avoidance Belief Questionnaire (FABQ) was used to quantify fear-avoidance beliefs (Waddell et al., 1993). FABQ is comprised of two sub-scales, one for physical activity (FABQ-PA) and one for work (FABQ-Work). Each item in the questionnaire was scored from 0 to 6, with higher numbers indicating increased levels of fear-avoidance beliefs. According to Waddel et al.'s original paper seven of the 11 items in the FABQ-Work are summed to a FABQ-Work score and all five items in the FABQ-PA are summed to a FABQ-PA score. Therefore, the FABQ-Work score ranges from 0 to 42, and the FABQ-PA ranges from 0 to 24. An early Norwegian version of FABQ showed acceptable test–retest reliability and internal consistency by Chronbach's α (Holm et al., 2003). In this version, item 12 (''I should not do my normal work with my present pain'') in FABQ-Work was omitted. We adjusted for this omission by adding the mean of the other FABQ-Work items to the sum score. A new Norwegian version of the FABQ that includes all 16 items has since been validated (Grotle et al., 2004a,b).

Patients rated their overall function using the Global Back Disability Question at the 1-year follow-up visit (Holm et al., 2003). There were five response alternatives: excellent, no complaints; good, occasionally influences activity; fair, some pain, and limited function; poor, unchanged, considerable complaints, and severe disability; miserable, worse, not self-reliant in activities of daily living. The reliability and construct validity of this question are good in the target population (Holm et al., 2003).

Evaluation of work status included questions about paid work (full time, part time, and not working) and status if not working (on sick leave, vocational or medical rehabilitation, disability pension, unemployed, and other (homemaker or student)). The assessment of work status was standard and reliable in the target population (Holm et al., 2003).

Relevant information from the index operation for disc herniation was tracked on medical charts. Baseline characteristics were assessed by a standardized questionnaire. Co-morbidity was assessed by a single question: ''Do you have other diseases?'' Two visual analog scales assessed pre-treatment beliefs and expectancies about the potential benefit from surgical and non-surgical treatment, respectively. The scale ranged from 0 (no benefit) to 100 (complete recovery). The reliability and validity of these questions have not been evaluated. The number of physiotherapy sessions attended after leaving the hospital was obtained by patient self-report.

A specialist in physical medicine and rehabilitation working at another hospital and a physiotherapist not involved in either treatment and working in another building at the Rikshospitalet University Hospital evaluated the patients at 1-year follow-up. All patients wore T-shirts during the evaluation to hide any scars from surgery and were told not to disclose which treatment they had received. The physician rated the patients' overall functional and work status using the Prolo Scale (Prolo et al., 1986). This scale has been applied in large series to evaluate results after lumbar discectomy (Davis, 1994; Berger, 2000). The Prolo Scale has two parts: economic and functional status. We chose to use functional status only, which ranks pain and the effect of pain on five categories of daily living activities. The scale ranges from ''complete recovery'' to ''worse''. The physiotherapist measured fingertip–floor distance as described by Hyytiäinen with the patient standing on a platform and bending forward (Hyytiäinen et al., 1999).

In addition, experienced radiologists assessed radiographic fusion. Anteroposterior and lateral radiographs of the lumbar sacral spine were taken. Fusion was graded as fused or not fused (Christensen et al., 2001).

2.6. Sample size

The study was designed to detect the difference of as little as 10 points between groups in change on the ODI from baseline to the 1-year follow-up visit. After a pilot study, the standard deviation was estimated at 10 points (Brox et al., 1999). With α set at 0.05 (type I error) and β at 0.1 (type II error), 26 patients were required for each treatment group to complete the trial (Pocock, 1983).

2.7. Statistical analysis

The results were analyzed according to the method of intention-to-treat. The normal distribution of baseline, follow-up data, and differences was checked by histograms and discussed with a statistician. Results are presented as medians (interquartile ranges (IQR)), proportions or means (SD). Differences are presented as means (95% confidence interval (CI) or odds ratio for categories (95% CI)). Baseline categories were compared with the χ^2-test or the Fisher's exact test. Paired t-tests (2-sided) were used to assess differences within the two groups from baseline to 1-year follow-up. Correlations between improvement in ODI and 1-year results according to Global Back Disability Questionnaire and Prolo Scale for functional status were calculated (Pearson's correlation coefficient). Categorical data for work and overall ratings were dichotomized. Success was defined as the three best grades (all responses except ''unchanged'' and ''worse'') for the Prolo Scale and the Global Back Question (Gibson et al., 1999). Results were compared with the Fisher's exact test. Because of small numbers in the work categories, work was dichotomized into full time work or not working. The McNmar's test of paired proportions was used to test differences in return to full time work. Patients on disability pension and other (homemaker, unemployed, and student) at baseline were excluded from this analysis.

Multiple regression analysis predicting 1-year follow-up scores on the outcome measure adjusted for baseline scores and gender was used to measure point estimates and confidence intervals for group differences (Tables 2 and 3). We constructed a variable from baseline scores for treatment expectations: expectation of surgery for patients assigned surgery and expectation of the cognitive intervention for patients allocated to that treatment. This variable correlated −0.12 (Pearson's correlation coefficient) with change in ODI. Adjustment for this variable made small differences and is reported only for primary outcome.

A second analysis, which included only the patients who completed the study, paralleled the intention-to-treat analysis. In a third analysis, we carried the last value forward to replace post-baseline values in the few patients missing results from 1-year follow-up. Analyses were performed using SPSS software, version 12.0.

3. Results

3.1. Recruitment and baseline characteristics

Of the 113 patients referred, 53 did not fulfill the criteria for randomization (Fig. 1), and 60 gave signed, informed consent and were randomly allocated to the study groups. The baseline characteristics of the patients are shown in Table 1. The percentage of men was lower in the surgery group (38%) than in the cognitive/exercises group (64%) ($p = 0.04$). The two groups did not differ in age, duration of disease, comorbidity, the use of analgesics or smoking habits. Beliefs in surgery were considerably higher than beliefs in non-operative treatment for both groups.

The 1-year follow-up rate was 97% (Fig. 1). Six patients in the surgery group did not receive the assigned treatment because they changed their mind after having been randomized to lumbar fusion. In addition, one patient died during the follow-up period. Two patients did not receive the assigned treatment because they changed their mind after having been randomized to the cognitive/exercises group. Additionally, two patients from the cognitive/exercises group had lumbar fusion during the follow-up period. Both patients had researched the possibility of having back surgery abroad before they were randomized to the study and underwent an operation in Sweden during the follow-up period. The median number of physiotherapy sessions attended after the hospital or patient hotel stay was 32 (range 0–96) in the surgical group and 0 (range 0–40) for cognitive/exercises group ($p < 0.001$, Fisher's exact test).

3.2. Outcomes

3.2.1. Primary outcome measure: improvement in back specific pain and disability measured by ODI

The improvement in the ODI at the 1-year follow-up was significant in both groups from 47.0 (SD 9.4) to 38.1 (20.1) ($p = 0.023$) for lumbar fusion and from 45.1 (9.1) to 32.3 (19.1) ($p = 0.001$) for cognitive intervention and exercises (Table 2) The improvement in the ODI at the 1-year follow-up visit did not differ significantly between treatments (Table 2). The mean difference in change between groups was −3.7 (95% confidence interval −13.5 to 6.2), −7.3 (−17.3 to 2.7) after adjusting for gender, and −9.7 (−21.7 to 1.7) after adjusting for gender and pretreatment expectations. The mean scores at baseline, 3 and 6 months, and at the 1-year follow-up visit are shown in Fig. 2. Improvement above measurement error (12 points) in a single patient was observed in 12 of 28 surgery patients and 14 of 29 cognitive/exercise patients ($p = 0.79$, Fischer's exact test).

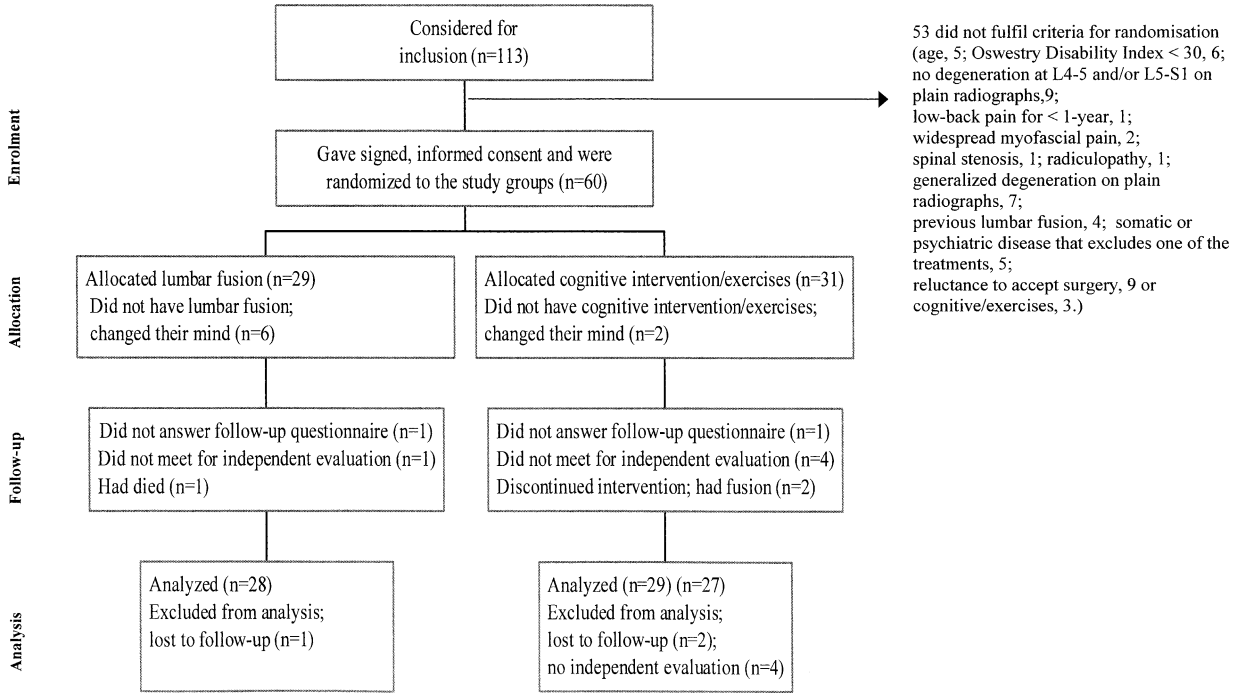

Fig. 1. Recruitment and follow-up of study participants.

Table 1
Baseline characteristics of the patients

	Lumbar fusion ($n = 29$)	Cognitive/exercises ($n = 31$)	p-value
Age (years, median, and IQR)	42 (35–45)	43 (37–50)	0.35[c]
Gender (women/total)	18/29	11/31	0.04
Duration from first onset (months, median, and IQR)	100 (39–150)	93 (36–146)	0.62
Educational level			
High school/university/college	9/29	11/31	0.72
Pre-randomization beliefs/expectancies in surgery[a] (mean (SD))	71.6 (19.3)	70.5 (16.9)	0.99[d]
Pre-randomization beliefs/expectancies in non-surgical treatment[a] (mean (SD))	40.8 (24.3)	51.8 (22.8)	0.07[d]
Comorbidity	12/29	12/31	0.83
Daily or weekly consumption of analgesics including acetaminophen	20/27	26/31	0.52[e]
Smoking	21/29	18/31	0.24
Married/living together	25/29	27/31	1.00[e]
Radiculopathy at index operation	21/27	25/29	0.50[e]
Radiculopathy improved after index operation[b]			
3 months	15/28	22/31	0.17
1 year	12/28	11/31	0.56
Second index operation (reoperation)	8/29	14/31	0.16

Numbers are proportions and Pearson's χ^2 test is used to compare groups unless stated otherwise.
[a] 0 = no effect, 100 = complete recovery.
[b] Index operation = operation for disc herniation.
[c] Mann–Whitney U-test.
[d] Independent 2-sided t-test.
[e] Fischer's exact 2-sided test.

The mean difference in change between groups for ODI from baseline to the 1-year follow-up visit in patients who adhered to their assigned treatment was −3.8 (−14.1 to 6.5), and −3.2 (−12.5 to 6.3) when the baseline values replaced missing 1-year follow-up values. ODI scores for the men in the surgery group did not improve. The mean difference in ODI for men between treatment groups was −19.8 (−32.8 to −6.8) ($p = 0.001$). The difference in women was 5.9 (−8.2 to 20.0) ($p = 0.4$). Moreover, the two men from the cognitive/exercise treatment group who had surgery during the follow-up period did not improve. The median improvement on the ODI was 16 (range −18 to 38) for single level L5–S1 fusions ($n = 8$) and 8 (range −14 to 60) for two level L4–S1 fusions ($n = 14$).

3.2.2. Secondary outcomes

Patients randomized to cognitive intervention and exercise improved significantly from baseline to 1-year follow-up in all variables except back pain ($p = 0.07$), work ($p = 0.13$), and emotional distress ($p = 0.08$). Patients in the surgery group did not improve significantly in any variable except back pain ($p = 0.02$). Differences between groups from baseline to 1-year follow-up were not significant, except for fingertip–floor distance and fear-avoidance beliefs for physical activity (Table 3). Complete results for the Global Back Disability Question and the Prolo Scale for functional status are presented in Tables 4 and 5. Correlation coefficients between improvement in ODI and Global Back Disability Question and Prolo scale were −0.75 and −0.77,

Table 2
Primary outcome

Outcome	Lumbar fusion ($n = 28$)[a]	Cognitive/exercises ($n = 29$)[a]	p-value	Mean difference (95% CI) between groups from baseline to 1-year adjusted for gender	Adjusted for gender and treatment expectations
Oswestry (0–100)					
Baseline	47.0 (9.4)	45.1 (9.1)			
1-year	38.1 (20.1)	32.3 (19.1)			
Difference	8.9 (1.3 to 18.5)[*]	12.6 (6.0–19.2)[**]	0.43	−7.3 (−17.3 to 2.7)	−9.7 (−21.7 to 1.7)[***]

Data are means (SD) for lumbar fusion and cognitive/exercises and means (95% CI) and p-values from paired t-tests for the differences between baseline and 1-year and regression analysis predicting 1-year follow-up scores on the outcome measure adjusted for baseline score and gender, and baseline score, gender and treatment expectations.
[a] Two patients randomized cognitive intervention/exercises and one patient randomized surgery did not attend 1-year follow-up. The independent observer did not examine two patients given intervention/exercises.
[*] $p = 0.023$.
[**] $p = 0.001$.
[***] $p = 0.09$.

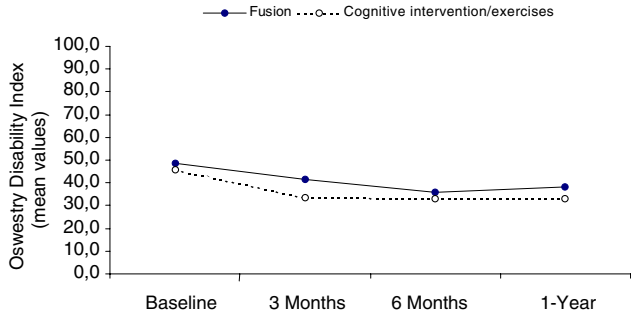

Fig. 2. Mean values for Oswestry Disability Index (ODI) in each study group. The ODI was the predefined main outcome variable and the study was designed to detect a difference in change of at least 10 points between groups. This measure consists of 10 questions about pain-related disability of activities of daily life and social participation. The total score ranges from 0 (no pain and disability) to 100 (worst possible pain and disability).

respectively, and 0.81 between the two overall rating scales. Success rate, as assessed by the patients (Global Back Disability Question), was 50% for the surgery group and 48% for the cognitive/exercise group (Table 4). Success rates according to the independent observer (functional Prolo Scale) were 57% and 67%, respectively (Table 5). The observed fusion rate was 85%.

3.2.3. Adverse effects

Early complications included two wound infections among the 23 operated patients. There were no late complications.

4. Discussion

We found no difference in primary outcome between lumbar fusion and cognitive intervention and exercises

Table 3
Secondary outcome

Outcome	Lumbar fusion ($n = 28$)[a]	Cognitive/exercises ($n = 29$)[a]	p-value	Mean difference (95% CI) between groups from baseline to 1-year adjusted for gender
General Function Score (0–100)				
Baseline (median (IQR))	40.3 (20.1)	39.1 (17.1)		
1-year	30.8 (21.6)	23.8 (21.0)	0.09	−9.5 (−20.7 to 1.6)
Back pain (0–100)				
Baseline (median (IQR))	64.6 (15.4)	64.7 (11.1)		
1-year	50.7 (27.3)	49.5 (20.0)	0.42	−5.2 (−18.0 to 7.6)
Lower limb pain (0–100)				
Baseline (median (IQR))	52.7 (20.2)	55.3 (19.4)		
1-year	45.0 (29.8)	47.7 (24.1)	0.68	−2.7 (−15.8 to 10.4)
Medication (DDD[b])				
Baseline	1.2 (1.4)	1.6 (2.0)		
1-year	0.9 (1.2)	0.9 (1.2)	0.31	−0.3 (−1.3 to 0.3)
Emotional distress (1–4)				
Baseline	2.0 (0.6)	2.0 (0.5)		
1-year	1.8 (0.6)	1.7 (0.5)	0.33	−0.1 (−0.5 to 0.2)
Fear-avoidance physical activity (0–24)				
Baseline	12.8 (5.1)	14.2 (4.4)		
1-year	11.9 (5.4)	7.9 (5.4)	0.003	−5.5 (−8.9 to −2.0)
Fear-avoidance work (0–42)				
Baseline	26.8 (10.3)	27.7 (10.1)		
1-year	24.7 (13.1)	22.5 (13.4)	0.31	−3.2 (−9.5 to 3.0)
Fingertip–floor distance (cm)				
Baseline	25.9 (20.2)	28.7 (17.1)		
1-year	19.8 (21.8)	11.7 (18.7)	0.009	−13.2 (−23.0 to −3.4)
Working full time[c]				
Baseline	1/20	3/21	n.a.	
1-year	2/20[*]	8/21[**]		

Data are means (SD) for lumbar fusion and cognitive/exercises and means (95% CI) and p-values from regression analysis predicting 1-year follow-up scores on the outcome measure adjusted for baseline score and gender unless stated otherwise. Bonferroni correction for secondary outcome measures can be obtained by multiplying p-values with 12. n.a., not applicable.

[a] Two patients randomized cognitive intervention/exercises and one patient randomized surgery did not attend 1-year follow-up. The independent observer did not examine two patients given intervention/exercises.

[b] DDD, daily defined doses based on ATC codes.

[c] Patients on disability pension and other (homemaker, student, and unemployed) are excluded.

[*] $p = 1.0$.

[**] $p = 0.13$ (McNmar's test).

Table 4
Patient overall rating by the Global Disability Question

Outcome	Lumbar fusion ($n = 28$)[a]	Cognitive/exercises ($n = 29$)[a]
Excellent	1	1
Good	5	8
Fair	8	5
Poor	12	15
Worse	2	
Proportion with success at 1 year	14/28	14/29

Data are number of patients. Odds ratio (95% CI) for success (excellent, good, or fair) of surgery at 1-year: 1.1 (0.4–3.0), $p = 0.91$ (Pearson's χ^2 test).

[a] Two patients randomized cognitive intervention/exercises and one patient randomized surgery did not attend 1-year follow-up. In addition the independent observer did not examine two patients given cognitive intervention exercieses.

after 1 year. Consistent results were observed for the primary outcome and overall evaluation at 1-year follow-up. Observed results for fear-avoidance for physical activity and fingertip–floor distance were better after non-operative treatment. The small size of the present study and the large variation between patients are reflected in the large confidence intervals. We had considered a difference between treatments of 10.0 on the Oswestry Disability Index to be clinically relevant before starting the study. After adjustment for gender and pretreatment expectations, the observed difference was 9.7 in favor of cognitive intervention and exercise. Thus, we cannot exclude the fact that cognitive intervention and exercises are more effective than lumbar fusion.

Patients with chronic low back pain after surgery for disc herniation are often considered for fusion or disc prosthesis. These patients are commonly classified as having failed back surgery syndrome and the prognosis after a second operation is generally considered poor compared with the prognosis in patients without previous surgery for disc herniation. The present patient population was more debilitated and chronically disabled than the population from our previously published study, but results were similar (Brox et al., 2003). Comparable results were reported in a sub-group of patients with the post-laminectomy syndrome in the British study (Fairbank et al., 2005). Separate results for patients who had previous surgery for disc herniation were not reported in the Swedish lumbar study (Fritzell et al., 2001).

Table 5
Independent observer overall rating by the Prolo scale for functional status

Outcome	Lumbar fusion ($n = 28$)[a]	Cognitive/exercises ($n = 29$)[a]
Complete recovery	2	6
Recurrent pain	4	7
Mild pain	9	5
Unchanged	12	9
Worse	1	
Proportion with success at 1-year	15/28	18/29

Data are number of patients. Odds ratio (95% CI) for success (excellent, good, or fair) of surgery at 1-year: 0.6 (0.2 to −1.7), $p = 0.32$. (Pearson's χ^2 test).

[a] Two patients randomized cognitive intervention/exercises and one patient randomized surgery did not attend 1-year follow-up. In addition the independent observer did not examine two patients given cognitive intervention exercieses.

The lack of a "no treatment" group is a limitation of the present study. This raises the possibility that neither of the treatments given was effective or that the modest improvements simply reflect the natural history, placebo or regression to the mean. Given the high rates of back surgery and the poorly documented results, a study including a sham operation group might be ethically acceptable, although difficult to conduct. It has been claimed that the conservative treatment given in the Swedish study resembles a non-treatment group (Mooney, 2001). A three group design including a placebo control group would have been preferable.

The greater reductions in fear-avoidance beliefs, fingertip–floor distance, and reduction in the number of physiotherapy sessions attended in the cognitive/exercise group compared with the surgery group may have occurred statistically by chance because we compared multiple outcomes, but could suggest that a change in pain behavior had occurred in the cognitive/exercise group. The reduction we found in fear-avoidance beliefs agrees with the results from a recent trial in chronic back pain patients that incorporated fear reducing and activating techniques (Von Korff et al., 2005). The avoidance of physical activity and fear of pain were the two main factors directly associated with back symptoms and changes in physical activity in a recent study (Keen et al., 1999). In the present study, the main purpose of cognitive intervention and exercises was to build confidence, via both lecture and experience, in patients that using their back would do no harm. This approach is based on the hypothesis that inaccurate and detrimental beliefs, ineffective coping behavior, negative mood states, social problems, and patho-physiological processes all interact to perpetuate the illness (Enright, 1997). Cognitive behavior therapy ascribes a central role to conscious thoughts, beliefs, and behavior in the perpetuation of disability (Vlaeyen and Linton, 2000). This therapy is a brief and problem-oriented approach. Its efficacy is documented by an increasing number of clinical studies (Vlaeyen and Linton, 2000).

Psychologists or psychiatrists were not involved in the present study. The information provided by physicians and physiotherapists and reinforced by participation in physical activities was simple and acceptable to most patients, and fear-avoidance beliefs for physical activity were significantly reduced. A recent

study suggests that high fear-avoiders benefit from a short-term exercise program (Klaber Moffett et al., 2004). Our exercise program was not strictly based on a behavioral approach. The role of exercise physiology principles and specific exercises in patients with chronic low back pain has been debated. The evidence for recommendation of exercises and/or behavioral therapy in patients with chronic low back pain is strong, but the evidence for specific exercise programs or techniques is weak (www.backpaineurope.org).

Few studies have compared surgical and non-surgical treatment, probably because such studies are demanding to conduct. In a previous study comparing shoulder surgery with a physiotherapy regimen for rotator cuff disease, we reported that patients were more reluctant to enter the study due to not wanting to be randomly assigned to surgery (Brox et al., 1993). Eligible patients in the present study were told about the study and asked if they were willing to participate at the referring hospitals. Patients reluctant to participate at this screening were not registered. Nine patients considered for inclusion were excluded because of reluctance towards undergoing a new operation. Six patients changed their mind and refused surgery. Most of the patients who withdrew from treatment attended the 1-year follow-up visit. Excluding dropouts did not change results, suggesting that dropouts did not bias our results. Men had inferior results after surgery. It is difficult to explain why surgical results were related to gender. Other factors not registered may have been associated with gender for the patients included in the present study and could have contributed to the outcome. Our findings need further exploration in future studies.

Critics may argue that posterior fusion with pedicle fixation is just one of several surgical methods and that other techniques may be preferable. Supporters of spinal fusion may even argue that the choice of posterolateral fusion is unfortunate and is not considered state of the art. However, a recent randomized study failed to show any clinical advantage of any of anterior-, posterior- or non-instrumentation fusion techniques (Fritzell et al., 2002). In another study of 140 patients, 59 of whom had previous back surgery (Sasso et al., 2004), cylindrical threaded titanium cages had a higher fusion rate, but similar improvements in pain and disability compared with femoral ring allograft.

In studies comparing fusion with and without pedicle screws, patients in whom screws were used had a higher likelihood of re-operation, a higher rate of nerve injury, greater blood loss, a longer operative time, and a higher rate of complications (Christensen et al., 2002a,b; Fritzell et al., 2003). Another previous study concluded that back surgery that included fusion was associated with a doubling of the risk of complications (Deyo et al., 1993). Common complications in instrumented spine surgery are instrument failure (7%), neural injuries (3%), infections (3%), and pulmonary embolus (2%). Therefore, we consider the adverse effects in the present study (two infections and no other complications) to be lower than expected.

Spinal-fusion surgery is undoubtedly effective for some conditions in some patients, but the rapidly rising rates of surgery and high rates of complications and re-operations raise concern that the procedure may be overused. More evidence from clinical trials using contemporary surgical methods is needed to justify recommending lumbar fusion to patients with degenerative disc disease and chronic low back pain after previous surgery for disc herniation.

Surgery carries strong implications of success (Turner et al., 1994). For example, relief of sciatica and back pain has been reported in at least one-third of back surgery patients who proved to have no disc herniation (Spongford, 1972). In the present study, the patients' pre-treatment expectancies favored surgery, but were minimally correlated with improvement in ODI.

Spinal surgeons who strongly believe in the efficacy of lumbar fusion may argue that inappropriate patient selection explains our results. To our knowledge experienced clinicians carefully evaluated the patients included in the present study. However, patients' back problems are often labeled discogenic pain although the validity of clinical signs, reduced disc height, Modic changes or high intensity zone lesions are highly questionable (Hurri et al., 2004). Recent studies indicate that invasive procedures like discography are not helpful as a diagnostic procedure in most cases (Carragee, 2000; Carragee and Hannibal, 2004). Despite the lack of validated diagnostic criteria, patients are increasingly being offered fusion surgery or disc prosthesis at spine centers in most western countries.

5. Conclusions

This randomized study did not show a significant difference in treatment effect between posterior lumbar transpedicular fusion and cognitive intervention and exercises in patients with chronic low back pain after surgery for disc herniation. Based on the results from previous studies (Fritzell et al., 2001; Fairbank et al., 2005) and the present study, our interpretation of the present evidence is that lumbar fusion should not be recommended in patients with chronic low back after surgery for disc herniation. Future studies should assess whether a recommendation is indicated in subgroups of patients with specific characteristics.

Acknowledgments

We thank the patients who participated in the trial, the nurses and nurse aids at the hospital departments

and outpatient clinics, and the referring medical doctors; H. Ursin and H. Eriksen for their work with randomization of the patients and comments on study design and the manuscript; orthopedic surgeons and neurosurgeons who assisted in the lumbar fusions; K. Frey Frøslie for statistical advice; physiotherapists A. K. Koller, M. Aarsland Fosdahl, and T. Haakenstad, for supervision of the exercise program and participation in the cognitive intervention; and the radiologists R. Gunderson and A.M. Finnanger for their assistance.

References

Asch HL, Lewis PJ, Moreland DB, Egnatchik JG, Yu YJ, Clabeau DE, et al. Prospective multiple outcomes of outpatient lumbar microdiscectomy: should 75–80% success rates be the norm? J Neurosurg 2002;96:34–44.

Berger E. Late postoperative results in 1000 work related lumbar spine conditions. Surg Neurol 2000;54:101–6.

Brox JI, Staff PH, Ljunggren AE, Brevik JI. Arthroscopic surgery compared with supervised exercises in patients with rotator cuff disease (stage II impingement syndrome). Brit Med J 1993;307:899–903.

Brox JI, Friis A, Holm I, Grundnes O, Sørensen R, Lange JE, et al. Patients with chronic degenerative spinal disease–can conservative treatment reduce the waiting list for surgery? Tidsskr Nor Lægeforen 1999;119:1784–7.

Brox JI, Sørensen R, Friis A, Nygaard Ø, Indahl A, Keller A, et al. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. Spine 2003;28:1913–21.

Carragee EJ. Is lumbar discography a determinate of discogenic low back pain: provocative discography reconsidered. Curr Rev Pain 2000;4:301–8.

Carragee EJ, Hannibal M. Discogenic evaluation of low back pain. Orthop Clin N Am 2004;35:7–16.

Christensen FB, Laursen M, Gelineck J, Eiskjær SP, Thomsen K, Bunger CE. Interobserver and intraobserver agreement of radiograph interpretation with and without pedicle screw implants: the need for a detailed classification system in posterolateral spinal fusion. Spine 2001;26:538–43.

Christensen FB, Stender Hansen E, Laursen M, Thomsen K, Bunger CE. Long-term functional outcome of pedicle screw instrumentation as a support for posterolateral spinal fusion: randomized clinical study with a 5-year follow-up. Spine 2002a;27:1269–77.

Christensen FB, Hansen ES, Eiskjær SP, Hoy K, Helmig P, Neumann P, et al. Circumferential lumbar spine fusion with Brantigan cage versus posterolateral fusion with titanium Cotrel-Dubousset instrumentation: a prospective, randomized clinical study of 146 patients. Spine 2002b;27:2674–83.

Davis RA. A long-term outcome analysis of 984 surgically treated herniated lumbar discs. J Neurosurg 1994;80:415–21.

Derogatis LR, Lipman RS, Rickels K, Uhlenhut EH, Cori L. The Hopkins symptom checklist (HSCL): a self-report inventory. Behav Sci 1974;19:1–15.

Deyo RA, Ciol MA, Cherkin DC, Loeser JD, Bigos SJ. Lumbar spinal fusion: a cohort study of complications, reoperations, and resource use in the Medicare population. Spine 1993;18:1463–70.

Deyo RA, Nachemson A, Mirza SK. Spinal fusion – the case for restraint. N Engl J Med 2004;350:722–6.

Enright S. Fortnightly review: cognitive behaviour therapy – clinical applications. Brit Med J 1997;314:1811–5.

Fairbank JCT, Davis J, Couper J, O'Brian J. The Oswestry low back pain questionnaire. Physiotherapy 1990;66:271–3.

Fairbank J, Frost H, Wilson-Mac Donald J, Yu L-M, Barker K, Collins R, Spine Stabilisation Trial Group. Randomised controlled trial to compare surgical stabilization of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC spine stabilization trial. Brit Med J 2005;330:1233–39.

Fritzell P, Hägg O, Wessberg P, Nordwall A. 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.

Fritzell P, Hägg O, Wessberg P, Nordwall A. Chronic low back pain and fusion: a comparison of three surgical techniques: a prospective multicenter randomized study from the Swedish Lumbar Spine Study Group. Spine 2002;27:1131–41.

Fritzell P, Hägg O, Nordwall A. Complications in lumbar fusion surgery for chronic low back pain: a comparison of three surgical techniques used in a prospective randomized study. Eur Spine J 2003;12:178–89.

Gibson JN, Waddell G, Grant IC. The Cocherane review for lumbar disc prolapse and degenerative lumbar spondylosis. Spine 1999;24:1820–32.

Grotle M, Vøllestad N, Veierød V, Brox JI. Fear-avoidance beliefs in relation to disability in acute and chronic low back pain. Pain 2004a;112:343–52.

Grotle M, Brox JI, Vøllestad N. Concurrent responsiveness in pain and functional status measurements used for patients with low back pain. Spine 2004b;29:E492–501.

Hägg O, Fritzell P, Oden A, Nordwell A. Simplifying outcome measurement: evaluation of instruments for measuring outcome after fusion surgery for chronic low back pain. Spine 2002;27:1805–12.

Hägg O, Fritzell P, Nordwall A. The clinical importance of changes in outcome scores after treatment for chronic low back pain. Eur Spine J 2003;12:12–20.

Holm I, Friis A, Storheim K, Brox JI. Measuring self-reported functional status and pain in patients with chronic low back pain by postal questionnaires. A reliability study. Spine 2003;28:828–33.

Hurri H, Karppinen J. Topical review. Discogenic pain. Pain 2004;112:225–8.

Hyytiäinen K, Salminen JJ, Suvitie T, Wickström G, Pentti J. Reproducibility of nine tests to measure spinal mobility and trunk muscle strength. Scand J Rehab Med 1999;23:3–10.

Indahl A, Velund L, Reikerås O. Good prognosis for low back pain when left untampered. A randomized clinical trial [see comments]. Spine 1995;20:473–7.

Keen S, Dowell AC, Hurst K, Klaber Mofett JA, Tovey P, et al. Individuals with low back pain: how do they view physical activity? Fam Pract 1999;16:39–45.

Keller A, Brox JI, Gunderson R, Holm I, Friis A, Reikerås O. Trunk muscle strength, cross-sectional area, and density in patients with chronic low back pain randomized to lumbar fusion or cognitive intervention and exercises. Spine 2004;29:3–8.

Klaber Moffett JA, Carr J, Howarth E. High fear-avoiders of physical activity benefit from an exercise program for patients with back pain. Spine 2004;29:1167–72.

Lindström I, Öhlund C, Eek C, Wallin L, Peterson LE, Nachemson A, et al. Mobility, strength, and fitness after a graded activity program for patients with subacute low back pain. Spine 1992;17:641–9.

Manniche C. Assessment and exercise in low back pain. With special reference to the management of pain and disability following first time lumbar disc surgery. Dan Med Bull 1995;42:301–13.

Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel group randomized trials. Lancet 2001;357:1191–4.

Mooney V. Point of view. Spine 2001;26:2532–3.
Nachemson AL. Lumbar disc herniation–conclusions. Acta Orthop Scand 1993;251(Suppl.):49–50.
O'Sullivan PB, Twomey LT, Allison GT. Evaluation of specific stabilizing exercise in the treatment of chronic low back pain with radiologic diagnosis of spondylolysis and spondylolisthesis. Spine 1997;22:2959–67.
Ostelo RW, de Vet HC, Waddell G, Kerckhoffs MR, Leffens P, van Tulder M, et al. Rehabilitation following first-time lumbar disc surgery: a systematic review within the framework of the Cochrane collaboration. Spine 2003;28:209–18.
Pocock SJ. Clinical trials. New York: John Wiley; 1983.
Prolo D, Oklund S, Butcher M. Toward uniformity in evaluating results of lumbar spine operations. Spine 1986;11:601–6.
Rønning M, Blix HS, Harbo BT, Strøm H. Different versions of the anatomical therapeutic chemical classification system and the defined daily dose–are drug utilisation data comparable? Eur J Clin Pharmacol 2000;56:723–7.
Sandanger I, Nygård JF, Ingebrigtsen G, Sørensen T, Dalgard OS. Prevalence, incidence and age at onset of psychiatric disorders in Norway. Soc Psych Epidemiol 1999;34:570–9.
Sasso RC, Kitchel SH, Dawson EG. A prospective randomized controlled trial of anterior lumbar interbody fusion using titanium cylindrical threaded fusion device. Spine 2004;29:113–22.
Spongford EV. The lumbar disc herniation: a computer-aided analysis of 2504 operations. Acta Orthop Scand 1972;342(Suppl.):1–95.
Turner JA, Deyo RA, Loeser JD, von Korff M, Fordyce WE. The importance of placebo effects in pain treatment and research. J Am Med Assoc 1994;271:1609–14.
Vik A, Zwart JA, Hulleberg G, Nygaard Ø. Eight year outcome after surgery for lumbar disc herniation: a comparison of reoperated and not reoperated patients. Acta Neurochir (Wien) 2001;143:607–10.
Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000;85:317–32.
Von Korff M, Balderson BHK, Saunders K, Miglioretti DL, Lin EHB, Berry S, et al. A trial of an activating intervention for chronic low back pain in primary care and physical therapy setting. Pain 2005;113:323–30.
Waddell G, Newton M, Henderson I, Sommerville D, Main C. A fear-avoidance beliefs questionnaire (FABQ) and the role of fear-avoidance beliefs in chronic low back pain and disability. Pain 1993;52:157–68.
WHO Collaborating Centre for Drug Statistics Methodology. ATC index with DDDs. Oslo, Norway: WHO Collaborating Centre; 2002.
www.backpaineurope.org. European commission research directorate general. Cost action B 13. Low back pain: guidelines for its management. Working group 2. Chronic low back pain.