Four-year follow-up of surgical versus non-surgical therapy for chronic low back pain

Jens Ivar Brox,1 Øystein P Nygaard,2 Inger Holm,3 Anne Keller,4 Tor Ingebritsen,5 Olav Reikerås6

ABSTRACT

Objectives To compare the long-term effectiveness of surgical and non-surgical treatment in patients with chronic low back pain.

Methods Two merged randomised clinical trials compared instrumented transpedicular fusion with cognitive intervention and exercises in 124 patients with disc degeneration and at least 1 year of symptoms after or without previous surgery for disc herniation. The main outcome measure was the Oswestry disability index.

Results At 4 years 14 (24%) patients randomly assigned to cognitive intervention and exercises had also undergone surgery. 15% (23%) patients assigned fusion had undergone re-surgery. The mean treatment effect for the primary outcome was 1.1; 95% CI −5.9 to 8.2, according to the intention-to-treat analysis and −1.8; 95% CI −8.9 to 5.6 in the as-treated analysis. There was no difference in return to work.

Conclusions Long-term improvement was not better after instrumented transpedicular fusion compared with cognitive intervention and exercises.

Lumbar spine fusion for chronic low back pain has increased rapidly during the past two decades.1 Four randomised studies have compared lumbar fusion and conservative treatment in patients with disc degeneration and chronic low back pain.2-5 Results up to 2 years after treatment have been published. A recent meta-analysis concluded that cumulative evidence at the present time does not support routine fusion, whereas a recent systematic review concluded that surgery may be more efficacious than unstructured care, but may not be more efficacious than structured cognitive-behavioural therapy.6-7 Methodological limitations of the randomised trials prevent firm conclusions. The Norwegian studies, published in three papers,4 5 8 were criticised for lack of power, short follow-up and a high number of withdrawals from fusion among patients with chronic low back pain after surgery for disc herniation. Because results, interventions and outcome measures were similar we merged the two Norwegian trials for long-term follow-up using a questionnaire mailed to the patients. We report the 4-year effectiveness of lumbar fusion compared with cognitive intervention and exercises in patients with chronic low back pain with and without previous surgery for disc herniation.

PATIENTS AND METHODS

Study design

The Norwegian studies were investigator initiated in 1999 and were conducted at four university hospitals. They were designed as two separate randomised trials and results were reported after 1-year follow-up.4 5 The ethics committee for medical research in health region I of Norway approved the studies.

Patients

Patients aged 25–60 years with chronic low back pain for at least 1 year, Oswestry disability index score greater than 30 and disc degeneration at L4–L5 and/or L5–S1, were eligible to participate in the study. Exclusion criteria were: widespread myofascial pain; spinal stenosis with reduced walking distance and neurological signs; disc herniation or lateral recess stenosis with clinical signs of radiculopathy; inflammatory disease; previous spinal fracture; previous fusion surgery of the spine; pelvic pain; generalised disc degeneration on plain radiographic examination; ongoing serious somatic and psychiatric disease; registered medicine abuse and reluctance to accept one of the interventions. At least one spine surgeon and one specialist in physical medicine and rehabilitation examined each patient. A research physiotherapist coordinated the study and verified eligibility. All eligible patients were given oral and written information about the study and the two interventions.

Randomisation

Patients received treatment assignments from an independent unit at Unifob Health, University of Bergen that was not involved in the treatment. Computer-generated randomly permuted blocks were used and allocation was concealed. The project coordinator telephoned the unit at Unifob Health and reported an identification number and was phoned back in order to inform the patient about the assigned intervention.4 Treatments were started within 3 months after randomisation.

Study interventions

The protocol surgery was posterolateral fusion with transpedicular screws of the L4–L5 and/or L5–S1 segment. Autologous bone was used in all cases. Postoperative rehabilitation was at the choice of the surgeon. Surgery was performed at two neurosurgical and two orthopaedic departments.4 The cognitive intervention and exercises consisted of 1 week plus 2 weeks in the outpatient clinic at the study centre interrupted by 2 weeks at home. Specialists in physical medicine and physiotherapists gave the intervention. In addition, patients met a peer for exchanging experiences. The main aim was to make the patients confident...
that they could not do any harm to the disc (back) by engaging in ordinary activities of daily life. Details of the programme have been outlined previously.4

Outcome measures
A standardised questionnaire was sent by post to all patients. The primary outcome measure was the original (version 1.0) Oswestry disability index.9 This score has 10 questions about pain and disability and ranges from 0% (no pain and disability) to 100% (worst possible disability).

Secondary outcome measures included pain,3 general function score,10 global back disability question for the assessment of patients’ overall rating,11 work and medication,11 emotional distress,12 fear-avoidance beliefs13 and life satisfaction (for details see additional supplemental file, available online only).14 The questionnaire also included questions about treatment taken after the 1-year follow-up. Additional surgery was verified from medical records.

Statistical analysis
Estimation of sample sizes in the two trials merged for 4-year follow-up has been reported previously.4 5 Results are primarily analysed with an intention-to-treat approach. Because of crossover and withdrawal, sensitivity analyses were based on the treatment actually received. Baseline characteristics in those who attended the 4-year follow-up were compared with crossover patients and withdrawals in the two treatment groups (table 1). Means (±SD) or numbers (percentages) were calculated for baseline and 4-year follow-up in those who attended, and are reported separately for intention-to-treat and as-treated analyses. The analyses of treatment effects compared differences between interventions at 4 years using linear regression with adjustments for gender, age, previous surgery for disc herniation and baseline scores. We conducted analyses with and without the most recent observed non-missing value carried forward in those who did not attend the 4-year follow-up. We used this simplistic method, being aware that more comprehensive multiple imputation techniques are available.15 The estimated treatment effects are reported as mean adjusted differences between groups (95% CI) based on analyses using the last observed value carried forward and including all patients randomly assigned (tables 2 and 3). Categorical outcomes (patients’ overall rating, medication and work) were dichotomised and logistic regression was used to calculate adjusted OR (95% CI) with adjustments for gender, age, previous surgery for disc herniation and

Table 1  Baseline characteristics of the patients*

<table>
<thead>
<tr>
<th></th>
<th>Lumbar fusion</th>
<th>Cognitive intervention and exercises</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All randomised</td>
<td>Crossover/withdrawals</td>
</tr>
<tr>
<td>Age (years)</td>
<td>42.7±8.0</td>
<td>43.9±7.3</td>
</tr>
<tr>
<td>No of men (%)</td>
<td>27 (41)</td>
<td>7 (64)</td>
</tr>
<tr>
<td>Years from first pain episode</td>
<td>8.9±7.9</td>
<td>8.1±7.9</td>
</tr>
<tr>
<td>Married/living together no (%)</td>
<td>57 (86)</td>
<td>10 (91)</td>
</tr>
<tr>
<td>Occupational education &lt;3 years no (%)</td>
<td>45 (68)</td>
<td>8 (27)</td>
</tr>
<tr>
<td>Work status no (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working</td>
<td>9 (14)</td>
<td>2 (18)</td>
</tr>
<tr>
<td>On sick leave</td>
<td>14 (21)</td>
<td>3 (27)</td>
</tr>
<tr>
<td>On rehabilitation</td>
<td>29 (44)</td>
<td>3 (27)</td>
</tr>
<tr>
<td>Disability pension</td>
<td>10 (15)</td>
<td>3 (27)</td>
</tr>
<tr>
<td>Student, homemaker, unemployed</td>
<td>3 (5)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Retirement pension</td>
<td>1 (2)</td>
<td></td>
</tr>
<tr>
<td>Back pain (0–100)†</td>
<td>63.0±14.7</td>
<td>64.2±15.5</td>
</tr>
<tr>
<td>Oswestry disability index§</td>
<td>44.5±10.7</td>
<td>45.3±10.1</td>
</tr>
<tr>
<td>Emotional distress (1–4)¶</td>
<td>1.9±0.5</td>
<td>1.9±0.7</td>
</tr>
<tr>
<td>Previous surgery for disc herniation no (%)</td>
<td>29 (44)</td>
<td>4 (36)</td>
</tr>
<tr>
<td>Beliefs in surgery**</td>
<td>69.7±18.2</td>
<td>62.1±17.7</td>
</tr>
<tr>
<td>Beliefs in non-surgical treatment**</td>
<td>40.1±25.4</td>
<td>42.1±24.5</td>
</tr>
<tr>
<td>Comorbidity no (%)</td>
<td>24 (36)</td>
<td>6 (55)</td>
</tr>
<tr>
<td>Taking analgesics daily or weekly no (%)</td>
<td>40 (61)</td>
<td>4 (36)</td>
</tr>
<tr>
<td>Smoking no (%)</td>
<td>36 (55)</td>
<td>4 (36)</td>
</tr>
</tbody>
</table>

*Plus–minus values are means±SD.
†Three patients allocated lumbar fusion died, four did not have and four had cognitive intervention and exercises. Fourteen patients allocated cognitive intervention and exercises had surgery.
‡Back pain ranges from 0 to 100, with lower scores indicating less severe symptoms.
§The Oswestry disability index ranges from 0 to 100, with lower scores indicating less severe symptoms.
¶Emotional stress ranges from 1 to 4, with lower scores indicating less severe symptoms.
**Beliefs ranges from 0 to 100, with lower scores indicating not efficient.
often men and non-smokers, had higher occupational education and higher comorbidity, but took analgesics less often at baseline. Such patients from the non-surgical group took analgesics more often at baseline.

**RESULTS**

**Patients**

A total of 124 patients was enrolled out of 234 who were eligible: 66 were assigned to the surgical group and 58 to the non-surgical group (figure 1). The 4-year follow-up rate was 92% and 86%, respectively. In the surgical group, 83% had undergone surgery at 1 year and 91% at 4 years. In the non-surgical group, 5% had undergone surgery at 1 year and 24% at 4 years.

In both groups patients had stronger beliefs in surgical compared with non-surgical treatment at baseline (table 1). Crossover patients and withdrawals from surgery were more often men and non-smokers, had higher occupational education and higher comorbidity, but took analgesics less often at baseline. Such patients from the non-surgical group took analgesics more often at baseline.

**Healthcare utilisation and return to work**

Thirty (49%) and 29 (58%) allocated surgical or non-surgical treatment, respectively, reported visits to a physician for back pain the year before the 4-year follow-up. Physiotherapy (20% vs 22%) and other treatments (16% vs 14%) were taken by a minority in both groups. More patients who had surgery (53% vs 32%) were on disability pension (adjusted OR 2.5; 95% CI 1.1 to 5.9). For the intention-to-treat analysis this difference was no longer significant (p=0.21). The number of patients working full time was not significantly different (tables 2 and 3).

**Table 2** Intention-to-treat analysis*

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Lumbar fusion (N=61)†</th>
<th>Cognitive/exercises (N=50)†</th>
<th>Adjusted treatment effect (95% CI)‡</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oswestry disability index†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>44.1±10.7</td>
<td>43.4±11.1</td>
<td>1.1 (−5.9 to 8.2)</td>
</tr>
<tr>
<td>4 Years</td>
<td>29.7±20.5</td>
<td>27.0±19.4</td>
<td></td>
</tr>
<tr>
<td><strong>Secondary</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General function score†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>37.3±19.3</td>
<td>40.0±18.9</td>
<td>−3.5 (−11.6 to 4.6)</td>
</tr>
<tr>
<td>4 Years</td>
<td>25.8±24.7</td>
<td>21.4±21.5</td>
<td></td>
</tr>
<tr>
<td>Back pain†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>62.8±14.5</td>
<td>64.2±12.5</td>
<td>2.3 (−6.4 to 10.9)</td>
</tr>
<tr>
<td>4 Years</td>
<td>42.2±23.9</td>
<td>44.7±22.8</td>
<td></td>
</tr>
<tr>
<td>Lower limb pain†</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>48.5±24.4</td>
<td>44.8±23.5</td>
<td>1.3 (−8.3 to 10.8)</td>
</tr>
<tr>
<td>4 Years</td>
<td>34.8±29.4</td>
<td>33.5±24.7</td>
<td></td>
</tr>
<tr>
<td>Emotional distress§</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>1.9±0.5</td>
<td>1.9±0.5</td>
<td>−0.1 (−0.1 to 0.3)</td>
</tr>
<tr>
<td>4 Years</td>
<td>1.7±0.6</td>
<td>1.7±0.6</td>
<td></td>
</tr>
<tr>
<td>Life satisfaction¶</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>5.0±2.2</td>
<td>4.6±1.7</td>
<td>−0.3 (−1.1 to 0.6)</td>
</tr>
<tr>
<td>4 Years</td>
<td>6.2±2.5</td>
<td>6.4±2.3</td>
<td></td>
</tr>
<tr>
<td>Fear-avoidance beliefs physical activity**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>13.0±5.0</td>
<td>15.4±5.0</td>
<td>−3.5 (−5.8 to −1.1)</td>
</tr>
<tr>
<td>4 Years</td>
<td>9.1±7.3</td>
<td>7.0±6.0</td>
<td></td>
</tr>
<tr>
<td>Fear-avoidance beliefs work**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>26.1±10.5</td>
<td>28.4±10.7</td>
<td>−4.3 (−8.3 to −0.2)</td>
</tr>
<tr>
<td>4 Years</td>
<td>23.9±13.8</td>
<td>21.1±12.5</td>
<td></td>
</tr>
<tr>
<td>Patients overall rating – no (%) success††</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Year</td>
<td>38 (62)</td>
<td>32 (64)</td>
<td>1.0 (0.8 to 1.5)</td>
</tr>
<tr>
<td>Work – no (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>9 (15)</td>
<td>8 (16)</td>
<td>0.8 (0.3 to 1.9)</td>
</tr>
<tr>
<td>1 Year</td>
<td>16 (26)</td>
<td>17 (34)</td>
<td></td>
</tr>
</tbody>
</table>

*Mean values±SD unless otherwise noted.
†The Oswestry disability index, the general function score, back and lower limb pain ranges from 0 to 100, with lower scores indicating fewer symptoms.
‡The treatment effect is the difference between patients randomly assigned to lumbar fusion and cognitive intervention and exercises at 4 years with adjustments for baseline score, age, gender and previous disc surgery. All patients randomly assigned (n=124) are included with last observed value carried forward.
††All who did not attend the 4-year follow-up are classified as non-success.
§Emotional stress ranges from 1 to 4, with lower scores indicating less severe symptoms.
¶Life satisfaction ranges from 1 to 10, with higher scores indicating better life satisfaction.
**Fear-avoidance beliefs for physical activity ranges from 0 to 24 and for work from 0 to 42, with lower scores indicating less strong beliefs for physical activity and work hurting the back.
Crossover, complications and re-operations

Non-adherence was registered in 17 (29%) patients randomly assigned to cognitive intervention and exercises, three (5%) did not have the allocated treatment and 14 (24%) patients later had surgery (figure 1). Eleven (17%) patients randomly assigned to surgery were classified as non-adherent, six (9%) did not have lumbar fusion (figure 1), two (3%) withdrew and three (5%) patients died. Deaths were not related to the surgical procedures. Four crossover patients operated (25%) in the non-surgical group and 15 (25%) in the surgical group had re-operation.

The reason was persistent complaints or deterioration of the condition. Complications have been described previously. No major complications occurred in patients operated after the 1-year follow-up.

Main treatment effects

In the intention-to-treat analysis there was no treatment effect for the Oswestry disability index. When adjusted for age, gender, baseline score and previous disc surgery the treatment effect was 1.1; 95% CI −5.9 to 8.2 (table 2). The mean adjusted treatment effect was −1.6; 95% CI −8.9 to 5.6 (table 3) according to as-treated analysis. Sensitivity analyses including only those who attended the 4-year follow-up did not alter the results.

Secondary outcome

The only treatment effect observed in the secondary outcome was a reduction of fear-avoidance beliefs favouring cognitive intervention and exercises (tables 2 and 3). The mean treatment effect for fear-avoidance beliefs for physical activity was −3.5;
Figure 1  Exclusion, enrolment, randomisation and follow-up of participants.

95% CI –5.8 to –1.1 in the intention-to-treat analysis and –2.8; 95% CI –5.3 to –0.4 in the as-treated last analysis, and –4.3; 95% CI –3.3 to –0.2 and –4.8; 95% CI –8.9 to –0.7 for fear-avoidance beliefs for work, respectively. Pain medication was taken daily or weekly by 58% treated with surgery compared with 35% not operated (adjusted OR 2.3; 95% CI 1.0 to 5.2). For the intention-to-treat analysis the difference was no longer significant (p=0.14).

DISCUSSION

In patients with chronic low back pain with and without previous surgery for disc herniation, lumbar fusion was not superior to cognitive intervention and exercises at relieving symptoms, improving function and return to work at 4-years. The results were consistent for intention-to-treat and as-treated analyses. The number of re-operations in patients randomly assigned to surgery were similar to the number patients operated in the non-surgical group.

The CI for the treatment effects were within 10 points on the Oswestry disability index that the trial was designed to detect. This indicates that lack of power is unlikely to explain the observed results.

Comparison with existing literature

The present study is the first to provide the long-term results of a randomised study comparing lumbar fusion with non-surgical treatment in patients with chronic low back pain. Results are in agreement with previously reported results at 1 and 2 years.2–5 The reported long-term results do not exclude the possibility that fusion may be indicated in carefully selected patients with chronic low back pain, but widening indications have contributed to the rise in rates of fusion surgery.16 Despite much effort to improve selection criteria, there is no agreement on providing a valid tool to diagnose discogenic pain, and even procedures such as discography and MRI are not reliable for selecting patients.17 Hägg et al16 reported that a personality characterised by low neuroticism and low disc height predicted functional improvement after surgery and that work resumption was predicted by low age and short-term sick leave.

The re-operation rate was slightly higher than previously reported after spinal surgery.19 A higher rate is not unexpected after instrumented fusion compared with laminctomy and discectomy. Although device failure and postoperative infection did not explain the cases in the present study, the outcome after surgery for back pain may be less predictable than after surgery for leg pain. Re-operation is an undesirable outcome, and the high rate observed in the present study is an argument against surgery. Preventing repeat spinal surgery is an important goal for surgeons and patients.

We observed no treatment effects in secondary outcome except for fear-avoidance beliefs. Differences in favour of non-surgical treatment for the number taking pain medication regularly or on disability pension were observed in the as-treated analysis only. The aim of the non-surgical intervention was to give patients the understanding that they could not do any harm to the disc (back) by engaging in ordinary activities of daily life. To reduce fear and avoidance and achieve confidence patients were encouraged and confronted with physical activities that were previously not recommended. Results at 4 years suggest that the reduction in avoidant behaviour observed at 1 year was maintained.

We observed that more patients used pain medication after surgery. Alternative interpretations are that these patients either experience more pain or they are habituated to pain medication. A recent study reported that surgical patients used more opiates, but that both pain medication and pain intensity were reduced after participation in a multidisciplinary pain programme.20 The reduction was attributed to a cognitive-behavioural approach to symptom management during the course of the rehabilitation programme. The possible effect of multidisciplinary pain rehabilitation on withdrawal of pain medication warrants further studies.

Most patients included were out of work at baseline. The number who had returned to work was not significantly different at 4 years, but the OR for disability pension was increased after surgery. Our interpretation is that the claim adjuster may consider that lumbar fusion represents the end stage of treatment, and consequently the claim for disability pension may be more easily accepted.
Possible confounders and weaknesses
A limitation of this study is the non-adherence to randomised treatment. Although patients consented to the protocol, some of them chose to change their consent as they are allowed to in clinical trials. The degree of non-adherence was lower than in the SPORT studies. One possible interpretation is that we aimed to conduct the interventions within 3 months after enrolment compared with 3–6 months in the SPORT studies. The consistency of results in intention-to-treat and as-treated analyses of the present study indicates that non-adherence does not play a decisive role to explain our results. Although 89% answered the follow-up questionnaire, the use of last value carried forward and not the multiple imputation technique for missing values may bias results.

Another limitation is the lack of a placebo group. Expectations are important for outcome. Sham surgery has previously shown that methods expected to be highly effective were mediated by placebo. We are unable to exclude the possibility that the observed improvements reflect the natural course, placebo or expectations and care.

Surgeons, patients and stakeholders may consider new technical surgical solutions more powerful, implying fast improvement and simple technical solutions in the hands of a skilled spinal surgeon, but postulated advantages for new procedures are based more on theories than knowledge. The introduction of new technology in clinical practice should be based on sound evidence from randomised studies. Patients allocated non-surgical treatment should be given the best evidence intervention and the same attention and care as the surgical patients.

In conclusion, patients did not have a better long-term improvement after instrumented fusion compared with cognitive intervention and exercises.

Acknowledgements
The authors would like to thank the patients who participated in the trial, the nurses and the nurse aids at the hospital departments and outpatient clinics and the referring medical doctors. A Friis for coordinating inclusion, treatments and 1-year follow-up, H Ursin and H Eriksen at Unifob Health, University of Bergen for their work with the random assignment of patients and comments on study design; AH Pipp at Rikshospitalet University Hospital for statistical advice; physiotherapists AK Keller, MFosdahl and T Haakenstad for non-surgical treatments; R Sørensen, JE Lange, R Rose and O Grundnes for lumbar fusions and the radiologists R Gunderson and AM Finnanger for their assistance.

Competing interests
None.

Patient consent
Obtained.

Ethics approval
This study was conducted with the approval of the Ethics Committee Health region I, Norway.

Provenance and peer review
Not commissioned; externally peer reviewed.

REFERENCES
Four-year follow-up of surgical versus non-surgical therapy for chronic low back pain

Jens Ivar Brox, Øystein P Nygaard, Inger Holm, Anne Keller, Tor Ingebrigtsen and Olav Reikerås

*Ann Rheum Dis* 2010 69: 1643-1648 originally published online July 26, 2009
doi: 10.1136/ard.2009.108902

Updated information and services can be found at:
http://ard.bmj.com/content/69/9/1643

These include:

**Supplementary Material**
Supplementary material can be found at:
http://ard.bmj.com/content/suppl/2010/06/06/ard.2009.108902.DC1

**References**
This article cites 23 articles, 1 of which you can access for free at:
http://ard.bmj.com/content/69/9/1643#BIBL

**Open Access**
This paper is freely available online under the BMJ Journals unlocked scheme, see http://ard.bmj.com/info/unlocked.dtl

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Topic Collections**
Articles on similar topics can be found in the following collections

- Open access (621)
- Pain (neurology) (883)
- Calcium and bone (725)
- Musculoskeletal syndromes (4951)

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/