

Lumbar Spinal Fusion

May 2013

Plain Language Summary

Chronic pain is pain that lasts more than three months. Chronic low back pain has many causes. They include wear and tear of discs and joints or muscle problems.

Treatment can be by surgery or other. Other treatments include physiotherapy, exercise, pain tablets, and mood and attitude therapy. Surgery includes spinal fusion.

Several studies have compared surgery with other treatments. They looked at pain, muscle strength, function, work and quality of life. There is insufficient evidence to determine how fusion compares with other treatments. No study has told us which patients do better with fusion surgery.

Studies found that the risk of fusion surgery was around 10% (1/10 patients). The risks included: bleeding, infection, blood clots, nerve damage, lung and heart complications.

Evidence Service

Lumbar Spinal Fusion

May 2013

Evidence Summary

Overview

Chronic low back pain with disc degeneration

We identified eight systematic reviews evaluating the effectiveness of fusion for the treatment of chronic low back pain (CLBP) with disc degeneration. Of these we assessed the most comprehensive systematic review, which included four randomised controlled trials (RCTs) comparing fusion to non-surgical treatment for chronic low back pain with disc degeneration; as well as three studies, published since the review.

Discogenic low back pain

One RCT examining the effectiveness of fusion for the treatment of discogenic low back pain without radicular (leg) pain was identified and assessed.

Isthmic spondylolisthesis

We identified and assessed one RCT, which compared the effect of fusion with non-surgical treatment in CLBP patients with isthmic spondylolisthesis.

Degenerative spondylolisthesis

One RCT comparing the effectiveness of surgical and non-surgical treatment among CLBP patients with degenerative spondylolisthesis was identified and assessed.

What is the effectiveness of spinal fusion on persistent pain, function, quality of life, return to work and medication use?

Chronic low back pain with disc degeneration

The evidence to answer this question is inconclusive.

Discogenic low back pain

The evidence to answer this question is inconclusive.

Isthmic spondylolisthesis

The evidence to answer this question is inconclusive.

Degenerative spondylolisthesis

The evidence to answer this question is inconclusive

What are the potential harms or risks of spinal fusion?

Chronic low back pain with disc degeneration

Frequency of early complications ranged from 8% to 18%, with no late complications reported. Major complications included deep wound infection, major bleeding during surgery, thrombosis, acute respiratory distress syndrome, pulmonary oedema, and heart failure.

Discogenic low back pain

Not reported.

Isthmic spondylolisthesis

Three major operative complications occurred: in two patients that had transpedicular fixation an L5 root injury with permanent sequelae occurred. Dermatomal pain developed in both patients, and one experienced permanent extension weakness of the foot. One non-instrumented fusion patient became permanently blind in one eye.

Degenerative spondylolisthesis

The most common surgical complication was dural tear (10%).

In what conditions is spinal fusion indicated?

There is insufficient evidence to conclude that the presence of disc degeneration, discogenic CLBP, degenerative spondylolisthesis, specific comorbid diseases, general health factors, psychological subpopulations, or isthmic spondylolisthesis are either indications for spinal fusion surgery, or predictors of outcomes in patients undergoing spinal fusion or non-surgical treatment.

Are there any reliable diagnostic procedures that predict the success of a fusion operation?

Not reported.

Glossary of Findings

Inconclusive evidence	Evidence exists regarding this question, but conclusions cannot be drawn from the results.
Not reported	This question was not addressed by the identified studies

Transport Accident Commission & WorkSafe Victoria

Evidence Service

Lumbar Spinal Fusion

Evidence Review

May 2013

Ornella Clavisi, Natasha Dodge, Emma Donoghue, Jason Wasiak

A joint initiative of

CONTENTS

ACKNOWLEDGEMENTS	2
BACKGROUND	3
METHODS	5
RESULTS	6
DISCLAIMER	22
CONFLICT OF INTEREST	22
REFERENCES	23

ACKNOWLEDGEMENTS

The authors would like to thank several colleagues for their assistance in preparation of this document.

Lisa Sherry from TAC/WSV for editing of Plain Language Summaries.

Anne Parkhill for her literature searching services.

Loretta Piccenna from the National Trauma Research Institute for proofreading.

BACKGROUND

Patient group

Chronic low back pain (CLBP) is characterised by pain persisting for longer than three months.¹ Some common theories as to the sources of CLBP include:¹

Herniated discs: where the outer part of the disc bulges out into the spinal canal.

Degenerative disc disease: where the discs flatten and dry out.

Isthmic spondylolisthesis: where the lowest lumbar vertebrae slips forward as a result of a stress fracture in childhood.

Degenerative spondylolisthesis: where the lowest lumbar vertebrae slips forward as a result of degenerative changes.

Lumbar spinal stenosis: when there is spinal degeneration causing the spinal canal to narrow compressing the spinal cord and nerves.

Lumbar spondylosis: where there is general degeneration of the lumbar spine, particularly the small joints (i.e. degenerative changes at L4-L5 and/or L5-S1).³

Spinal fusion

The aim of surgical fusion is to reduce pain and decrease disability associated with the above conditions. The rationale for spinal fusion is that pain arises from a degenerative motion segment and consequently, fusion surgery helps eliminate excessive motion and its subsequent pain.² There are two main types of lumbar spinal fusion, which may be used in conjunction with each other:⁴

- Posterolateral fusion: places a bone graft to form a bony bridge between the transverse processes in the back of the spine. These vertebrae are then fixed in place with screws and/or wire through the pedicles of each vertebra attaching to a metal rod, plate or cage on each side of the vertebrae (often called 'fixation').
- Interbody fusion: places a bone graft between the vertebrae in the area usually occupied by the intervertebral disc. In this instance, the disc is completely removed and replaced with a graft. This will allow the fusion to occur from one vertebral body to the other through their endplates. The surgical incision for interbody fusion will be either anterior, posterior or transforaminal.

The indication for spinal fusion surgery and how it should be performed for the relief of ongoing symptomatology however, remains unclear or controversial.²

Non-surgical treatments

The management of CLBP conditions and their symptoms remains controversial, with a variety of surgical and non-surgical treatment options available.² The traditional non-surgical approach may include a series of rehabilitative physiotherapy techniques in isolation or in combination with, epidural steroid injections, non-steroid anti-inflammatory drugs or opioid administration.⁴ Other therapies also include a series of structured clinical programs that provide intensive multidisciplinary rehabilitation encompassing a series of physical, psychological, social and occupational patient factors.⁵⁻⁷

Regulatory status

Spinal fusion surgery is listed under the Medicare Benefits Schedule (MBS) as a *Category 3 Therapeutic Procedure* and is covered by the following item numbers:

- fusion to cervical, thoracic or lumbar regions (48660-48675)
- fusion, posterior (40321,40324,40327)
- fusion, posterior interbody, with laminectomy (48654,48657)
- using segmental instrumentation (48613)

There have been no submissions to the Medical Services Advisory Committee (MSAC) or guidance from the Food and Drug Administration (FDA) in the United States regarding the safety, effectiveness, and cost considerations associated with spinal fusion surgery for the treatment of degenerative conditions of the spine that have failed to respond to conservative treatment.

Intended purpose of the review

The Transport Accident Commission (TAC) and WorkSafe Victoria (WSV) requested a review of the evidence to determine whether spinal fusion is an effective treatment compared to non-surgical treatment in patients with CLBP. This report sought to answer the following questions:

1. What is the effectiveness of spinal fusion on persistent pain?
2. What is the effectiveness of spinal fusion on function, quality of life, return to work and medication use?
3. What are the potential harms or risks of spinal fusion?
4. In what conditions is spinal fusion indicated, and are there any reliable diagnostic procedures that predict the success of a fusion operation?

METHODS

Methods are outlined briefly below. More detailed information about the methodology used to produce this report is available in Appendices 1 and 2. All appendices are located in the Technical Report accompanying this document.

Stage 1: Identify relevant research

A comprehensive search of Medline, PreMedline, EMBASE, CINAHL, the Cochrane Database of Systematic Reviews, DARE, CENTRAL, NHSEED, HTA and ACP Journal Club, and Web of Knowledge was undertaken in November 2012 to identify relevant synthesised research (i.e. evidence-based guidelines (EBGs), systematic reviews (SRs), health technology assessments (HTAs)) published from 1992 onwards. An additional search was conducted to identify any relevant randomised controlled trials (RCTs). A comprehensive search of the internet, relevant websites and electronic health databases was also undertaken (see Appendix 2, Tables A2.2-A2.4 for search details). Reference lists of included studies were also scanned to identify relevant references.

Studies identified by the searches were screened for inclusion using specific selection criteria (see Appendix 2, Table A2.1). Synthesised evidence (EBGs, SRs and HTAs) that met the selection criteria were reviewed to identify the most up-to-date and comprehensive source of evidence, which was then critically appraised to determine whether it was of high quality. This process was repeated for additional sources of evidence, if necessary, until the most recent, comprehensive and high quality source of evidence was identified for each indication. All screening and selection was conducted independently by two reviewers, results were compared and any discrepancies discussed and resolved. Findings from the best available source of evidence were compared to other evidence sources for consistency of included references and findings.

Stage 2: Address further actions identified

See algorithm in Table 1.

Table 1. Further action required to answer clinical questions.

Is there any synthesised research available? (e.g. EBGs, HTAs, SRs)				
Yes			No	
Is this good quality research?			Are RCTs available?	
Yes		No	Yes	No
Is it current (within 2 years)?				
Yes	No	Undertake new SR and/or meta-analysis	Undertake new SR and/or meta-analysis	Consider looking for lower levels of evidence
No further action	Update existing SR			

For each indication, the most recent, relevant, high quality piece of evidence was used to address the questions posed above.

RESULTS

Database searches yielded 1,719 articles, which were screened for potential relevance. Of these, 79 articles thought to be relevant were retrieved in full text and reviewed. From this review 22 articles were selected for inclusion (9 SRs and 13 RCTs). A further 3 EBGs were identified from the results of an internet search, bringing the total number of included studies to 25.

In total, 25 papers were included, consisting of:

- 12^{2, 8-18} Synthesised studies (SRs or EBGs)
- 13^{3-7, 19-26} primary study references (RCTs)

The above evidence assessed the effect of spinal fusion surgery with non-surgical treatment for the following indications: CLBP with disc degeneration, CLBP with discogenic pain, isthmic spondylolisthesis and degenerative spondylolisthesis.

Chronic low back pain with disc degeneration

Study characteristics

We identified eight SRs^{2, 8-10, 12-15} that evaluated the effectiveness of fusion for the treatment of chronic low back pain (CLBP) with disc degeneration. The most comprehensive of these was by Mirza (2007),¹⁴ which included four RCTs; Brox (2003),³ Brox (2006),⁶ Fairbank (2005)⁷ and Fritzell (2001)²⁰ comparing fusion to non-surgical treatment for chronic low back pain with disc degeneration.

Our search also identified three additional publications, that were published since Mirza (2007)¹⁴. These publications presented either long term data or secondary analyses of studies included in Mirza (2007)¹⁴. These included Brox (2010)⁵ which presented long-term pooled follow-up data for the Brox (2003)³ and Brox (2006)⁶ RCTs; and Keller (2004)²³ and Froholdt (2011),²¹ which provided a secondary analysis on 124 patients with CLBP from Brox (2003)³ and Brox (2006).⁶

Quality appraisal of the Mirza (2007)¹⁴ SR found it to have a low to moderate risk of bias (see Appendix 5 Technical report).

Participants

All of the four trials from Mirza (2007)¹⁴ included adult patients with CLBP with disc degeneration. The length of time patients had experienced CLBP varied among studies. One study included patients who had CLBP for more than two years,²⁰ two studies included patients who had CLBP for at least one year^{3, 6} and one study included patients with CLBP for more than 12 months.⁷

Three of the trials specifically recruited patients who also had spondylosis, (degenerative changes at L4-L5 and/or L5-S1).^{3, 6, 20} Two studies also required patients to have an Oswestry Disability Index (ODI) greater than 30 out of 100 points,^{3,6} and one study²⁰ required patients to have a score of at least 7 out of 10 on the “Function and Working Disability” index.

Two studies recruited patients who had undergone previous surgery: Fairbank (2005)⁷ included patients with CLBP irrespective of previous decompression or discectomy surgery, while Brox (2006)⁶ specifically recruited patients who had previously undergone disc herniation surgery.

Intervention

In all four trials included in Mirza (2007)¹⁴ the surgical intervention was fusion surgery; however the type of fusion varied. In Brox (2003)³ and Brox (2006)⁶ the type of fusion surgery was posterolateral fusion with transpedicular screws of the L4-L5 and/or L5-S1 with the use of autologous bone. The intervention in Fritzell (2001)²⁰ consisted of three surgical subgroups: Group 1a consisted of posterolateral fusion with a plastic brace post-surgery, Group 1b consisted of posterolateral fusion with the addition of an internal fixation device and a canvas corset post-surgery, and Group 1c consisted of surgery the same as Group 1b with an additional interbody bone graft either as an anterior lumbar interbody fusion or posterior lumbar interbody fusion, according to the preference of the surgeon. In the study by Fairbank (2005)⁷ the choice of fusion was made by the operating surgeon.

Comparator

The comparator in all four trials included in Mirza (2007)¹⁴ was non-surgical treatment. In three trials the non-surgical treatment consisted of intensive physical therapy with a cognitive and behavioral treatment program (75 hours over three weeks, with subsequent follow-up visits).^{3,6,7}

In the trial by Fritzell (2001)²⁰ included in Mirza (2007)¹⁴ the non-surgical treatment intervention was less intensive (70 hours of supervised physical therapy over a two-year period) and more heterogeneous (could be supplemented with other forms of treatment, such as information and education, treatments aimed at pain relief (TENS, acupuncture, injections), or cognitive and functional training and coping strategies).

Outcome

The RCTs included in Mirza (2007)¹⁴ assessed functional outcomes such as ODI^{20,3,6,7} and general function score;^{3, 6, 20} pain using a visual analogue scales (VAS);^{3, 6, 20} work status;^{3, 6, 20} quality of life using life satisfaction;³ complications;^{3, 6, 7, 20} medication use assessed as daily defined doses using Anatomical Therapeutic Chemical Classification System codes;^{3, 6} and mental health using the Medical Outcomes Study 36-Item Short Form General Health Survey (SF-36) mental component score,⁷ emotional distress using the Hopkins symptom check list-25^{3, 6} and Zung depression scale.²⁰

Keller (2004)²³ and Froholdt (2011)²¹ only assessed trunk muscle strength, cross-sectional area and density.

Follow-up

Two trials, Brox (2003)³ and (2006),⁶ measured outcomes one year after baseline assessment and two trials, Fairbank (2005)⁷ and Fritzell (2001)²⁰ measured the final outcomes at two years.

The Brox (2010)⁵ study was a four-year follow-up of patients from the Brox (2003)³ and Brox (2006)⁶ trials.

Keller (2004)²³ assessed the trunk muscle strength of patients at their one-year follow-up; these patients were again assessed by Froholdt (2004)²¹ between seven and eleven years after treatment.

Study quality

One of the main issues regarding study quality was that three of the four trials included in Mirza (2007)¹⁴ were underpowered to detect a significant difference with regards to the primary outcome.^{3, 6, 7} This particularly applies to Brox (2003)³ and (2006),⁶ which reported no significant treatment effect between fusion and non-surgical treatment. Furthermore, due to the nature of surgical trials, patients and clinicians could not be reasonably blinded to treatment. Given that surgery can be associated with important placebo effects, an overestimation of subjective outcomes such as pain cannot be ruled out.

Results

Pain

Three trials reported on the mean difference (MD) from base line for lower limb and back pain using VAS.^{3, 6, 20} The results for improvement in pain were inconsistent between trials. In the trial by Fritzell (2001)²⁰ patients randomised to surgery experienced moderately greater improvements in back pain (21.0 vs 4.3, $P = 0.0002$).²⁰ However two smaller trials, Brox (2003)³ and (2006)⁶ reported no significant difference in back pain scores from baseline, between fusion and non-surgical treatment (MD 8.6, 95%CI -3.0 to 20.1)³ and (MD -5.2, 95%CI -18.0 to 7.6).⁶

Two trials, Fritzell (2001)²⁰ and Brox (2003),³ found a significant mean difference from baseline for lower limb pain scores in favour of fusion (6.3 vs -7.0, $P=0.005$)²⁰ and (MD 17.5, 95% CI 4.3 to 30.7),³ however the third trial, Brox (2006)⁶ found no significant difference (MD -2.7, 95%CI -15.8 to 10.4).⁶ At the four-year follow-up, the pooled estimate for Brox (2003)³ and Brox (2006)⁶ showed no significant difference in pain scores between fusion and non-surgical treatment for either back or lower limb pain.⁵

Functioning

Oswestry Disability Index

All four RCTs in the Mirza (2007)¹⁴ review reported ODI as a measure of function. For Fritzell (2001),²⁰ Brox (2003)³ and Fairbank (2005)⁷ the change in ODI was greater in the surgical group than

the non-surgical group. In contrast, Brox (2006)⁶ reported a greater change in ODI in the non-surgical group compared to the surgical group.

For the ODI, we conducted a meta-analysis (Figure 1) to determine the estimates of the mean difference in ODI by combining the separate estimates of the inverse variance-weighted log mean differences from each of the studies. The results of this revealed that the mean overall difference in ODI between the fusion and non-surgical groups was -5.10 in favour of surgery (95% CI -8.19 to -2.01, $p=0.001$, $I^2=0\%$). Overall the results from each of the studies were similar with no evidence of heterogeneity.

However despite the differences in ODI being significantly in favour of surgery, the changes in scores were lower than the clinically meaningful difference. For patients following lumbar spinal surgery a difference in ODI of 12.8 points or above is considered clinically meaningful.²⁷

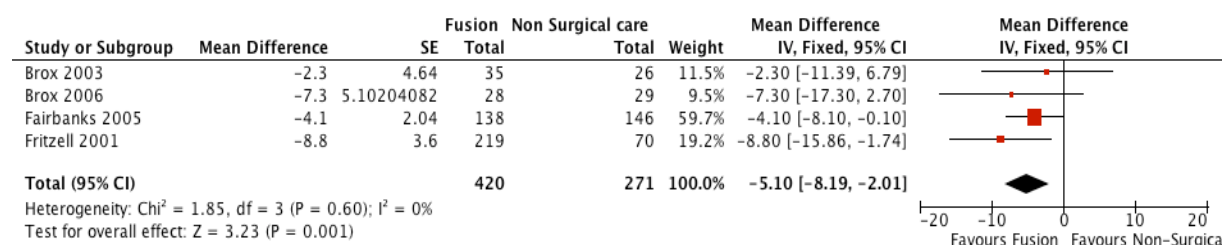


Figure 1. Meta-Analysis of ODI scores.

General function

Three trials assessed general function using a general function score.^{3, 6, 20} Only Fritzell (2001)²⁰ found that fusion significantly improved general function compared to non-surgical treatment (MD at two years, 15 vs 2.1 $P=0.005$). The two smaller trials by Brox^{3, 6} found no significant difference in general function scores for patients treated with fusion compared with non-surgical treatment (MD -4.1, 95%CI -14.9 to 6.7)³ and (MD -9.5, 95%CI -20.7 to 1.6).⁶ The trial by Fairbank (2005)⁷ reported no significant difference in the SF-36 physical and social functioning domain between the fusion and non-surgical treatment.⁷

Quality of life

One trial, Brox (2003),³ reported on quality of life, measuring “Life Satisfaction” as a secondary outcome. This study reported a significant improvement, from baseline, in life satisfaction scores for both the fusion and non-surgical groups, however, the difference between groups was not significant (MD 0.8, 95%CI: -2.1 to 0.5).³

Mental health

None of the studies found a significant difference between fusion and non-surgical treatment with regards to mental health.

Return to work

In the three trials that measured “return to work” only a small percentage of patients were working at baseline and follow-up. One study reported a significant difference in favour of fusion expressed as “net back to work” (36% fusion group vs 13% for non-surgical group $p=0.002$) and also as “back to work” (39% fusion group vs 23% non-surgical group $p=0.049$).²⁰ Two smaller trials found no significant difference in back to work rates between fusion and non-surgical treatment.^{3,6}

Brox (2010)⁵ reported an increase in the number of patients working in both the fusion and non-surgical groups at follow-up compared to baseline, however the difference between the groups was not significant.

Medication use

Two trials examined the use of pain medication and found no difference in use between fusion and non-surgical treatment (MD -0.4, 95%CI -1.1 to 0.5³ and MD -0.3, 95%CI -1.3 to 0.3⁶).

Muscle strength

Studies by Keller (2004)²³ and Froholdt (2011)²¹ conducted secondary analyses of combined data from Brox (2003),³ Brox (2006).⁶ The objective of these studies was to investigate the differences in muscle strength, cross-sectional area, and density of the back muscles in patients with CLBP with disc degeneration randomised to either lumbar fusion or cognitive intervention and exercises.

The study by Keller (2004)²³ found that at one-year follow-up, patients treated with cognitive intervention and exercise programs improved significantly in muscle strength compared with patients who underwent lumbar fusion. This effect was lost in the longer term with an analysis at seven to eleven years post intervention showing no significant difference between fusion and non-surgical treatment.²¹ The cross-sectional area remained unchanged in both treatment groups at both time points.

Complications

In the two Brox RCTs early surgical complication rates were 8%⁶ and 18%³, with no late complications reported. The study by Fritzell (2001)²⁰ reported that 17% of patients suffered early complications, the majority of which were handled with no obvious sequelae. Fairbank (2005)⁷ reported that intra-operative complications occurred in 10% of patients; with 6% requiring further operations on their lumbar spine during the two-year follow-up. Major complications included deep wound infections, major bleeding during surgery, thrombosis, acute respiratory distress syndrome, pulmonary oedema, and heart failure.

Discussion

It is unclear whether fusion is more effective than non-surgical treatment in patients CLBP with disc degeneration. Overall the body of evidence relies on a small number of trials, some of which have relatively small sample sizes and are underpowered.

Overall the results were variable between the different trials. For example the trial by Fritzell (2001)²⁰ appeared to provide strong evidence in favour of fusion, but more recent studies by Brox^{3,5,6} found no significant difference between treatment groups. A plausible explanation for the observed inconsistencies was the variability in the non-surgical therapies between the trials. For example the studies that used intensive physical therapy in combination with cognitive intervention^{3,6,7} reported no significant difference between treatment arms, whereas in the one trial that showed fusion to be more effective,²⁰ the non-surgical treatment was less structured and less intensive.

The patient groups were also different between these trials with only the Fritzell (2001)²⁰ study enrolling patients who had previously failed non-surgical treatment. Given that this trial was the only one showing a significant improvement with fusion, it may be that fusion is effective as last-line therapy in patients with CLBP, but less effective as first- or second-line therapy.

Other issues were the high crossover rate of patients from the non-surgical treatment group to fusion and high dropout rates as reported by Fairbank (2005).⁷

Conclusion

There is insufficient evidence to determine the effect of fusion compared to non-surgical treatment in patients with CLBP with disc degeneration. Although there is limited evidence to suggest that surgery could be effective as a final therapy option when compared with unstructured non-surgical treatment, it may be equally as effective when compared to intensive rehabilitation with cognitive intervention and exercises in patients who are yet to fail non-surgical therapy. Due to the methodological limitations of the current RCTs further research is required to accurately assess the effect of fusion in CLBP patients with disc degeneration.

Discogenic low back pain

Study characteristics

We identified one RCT by Ohtori (2010)²⁵ examining the effectiveness of fusion for the treatment of discogenic low back pain (DLBP) without radicular (leg) pain. This study was different from those included in the “CLBP with disc degeneration” section of the report as it specifically included CLBP patients with a diagnosis of discogenic pain (i.e. disc degeneration only at 1 level (L4/5 or L5/S1) on MRI, pain provocation on discography, and pain relief by discoblock).

Forty-one discogenic/discography positive patients were randomised into three groups; a control group and two surgery groups. The two surgery groups consisted of 15 patients who received anterior interbody fusion (ABF) and six patients who received posterolateral fusion with pedicle screws (PLF). The control group consisted of 20 patients who underwent an exercise program. The exercise program consisted of daily walking (30 minutes twice a day) and muscle stretching (body and legs) (15 minutes twice a day). The walking was performed independently by the patient at

home while the muscle stretching was performed in hospital with a physiotherapist. Patients were excluded if they did not perform the walking and stretching precisely as instructed.

This study measured functional disability using the ODI, and pain using a VAS (0, no pain; 10, worst pain) and the Japanese Orthopedic Association Score (JOAS: 0 = worst pain; 3 = no pain). These outcomes were measured at baseline and one and two years after treatment. Other outcomes such as quality of life, return to work and medication use were not reported in this trial.

Results

Pain using VAS

At two years post treatment there was an improvement in self-reported end point pain scores in those receiving fusion, with significantly lower pain VAS scores and higher pain JOAS scores in the both the ABF and PLF groups compared to the exercise group.

We calculated the mean difference between the groups at two years and found that for ABF compared to non-surgical treatment, the mean difference in VAS scores was -3.4, SE 0.35 (95% CI, -4.087, -2.71) and the mean difference for PLF compared to non-surgical treatment was -2.2, SE 0.39, (95%CI, -2.96, -1.43). For pain using JOAS the mean difference between ABF and non-surgical treatment was 1.3, SE; 0.10, (95% CI, 1.1, 1.50) and for PLF compared to non-surgical treatment 0.8, SE 0.1765, (95% CI, 0.45, 1.14).

Function

There was no difference in ODI between groups at baseline. At one and two years post treatment, ODI was significantly lower in both the ABF and PLF group compared to the exercise group ($P < 0.01$). We calculated the mean difference between the groups at two years and found the mean difference between ABF and non-surgical treatment to be -29.7, SE 2.30, (95%CI, -34.18, -25.21) and between PLF and non-surgical treatment (-18.8, SE 3.00, 95%CI, -24.69, -12.90).

Other outcomes

The study by Ohtori (2010)²⁵ did not report on quality of life, return to work, medication use or complications.

Discussion

Although the results showed a significant improvement in fusion compared with non-surgical treatment there is some uncertainty regarding the results. This study had a moderate risk of bias, its main limitation was that trial patients and clinicians could not be reasonably blinded to treatment and given that surgery can be associated with important placebo effects, an overestimation of subjective outcomes such as pain cannot be ruled out. Also the results of this study are based on a small sample of patients and given that this is the only trial investigating the effect of fusion on patients with discogenic low back pain, the generalisability of these results is unclear.

Conclusion

Given the limitations of this study and the paucity of evidence, there is insufficient evidence to confirm whether spinal fusion is as effective as non-surgical treatment in people with discogenic low back pain without radicular (leg) pain.

Isthmic spondylolisthesis

Study characteristics

We identified one RCT by Moller (2000),²⁴ which investigated the effect of fusion with non-surgical treatment in CLBP patients with isthmic spondylolisthesis. This study included adults aged 18-55 years with lumbar isthmic spondylolisthesis of any grade, with at least one year of low back pain or sciatica, and a severely restricted functional ability. In this study 111 patients were randomly allocated to an exercise program (n= 34) or posterolateral fusion with or without transpedicular fixation (n=77). All 77 patients who underwent surgery had a posterolateral fusion in situ with autologous bone transplantation harvested from the right iliac crest. Patients in the non-surgical treatment group underwent an intensive exercise program, developed by a physiotherapist with a special interest in spondylolisthesis. The exercise program was supervised by the physiotherapist and performed three times a week for the first six months, and twice a week between six and twelve months. Outcomes were assessed at one and two years post treatment and included pain using a 0-100 point VAS and functional disability using the Disability Rating Index (DRI). Nine year follow-up data were published by Ekman (2005).¹⁹ Ekman (2005)¹⁹ included an additional outcome measure, “Global Outcome”, which was assessed by patients, who classified their overall results as ‘much better,’ ‘better,’ ‘unchanged’ or ‘worse.’

Results

Pain

The surgically treated group reported a significantly lower pain index ($P = 0.002$, data not reported) at the two-year follow-up assessment than the exercise group. There were mean pain index improvements from baseline in both groups: 63 (range, 10–98) to 37 (range, 0–96) in the surgical group ($P = 0.0001$), and from 65 (range, 32–96) to 56 (range, 17–87) in the exercise group ($P = 0.024$). This difference was no longer significant at the nine-year follow-up assessment.

Function

The surgically treated group reported a significantly lower DRI ($P = 0.004$, data not reported) at the two-year follow-up assessment than the exercise group. At the two-year follow-up assessment, 11 of the 12 functional scores were significantly better in the surgical group than in the exercise group. Among all 106 patients that completed the two-year follow-up assessment, surgically treated

patients accounted for 34 of the 37 patients with a DRI lower than 20, and 28 of the 29 patients with a pain index lower than 20. This difference was no longer significant at the nine-year follow-up assessment.

Global outcome

Global outcome at nine years was significantly better in the surgical group than in the exercise group. In the surgical group 76% of patients classified their result as much better or better compared with 50% in the exercise group ($p = 0.015$). This outcome was not reported at the two-year follow-up.

Complications

In the surgical group, three major operative complications occurred. In two of the thirty-seven patients who underwent surgery with transpedicular fixation instrumentation, an L5 root injury occurred with permanent sequelae. Dermatomal pain developed in both patients, and one experienced permanent extension weakness of the foot. One non-instrumented surgical patient became permanently blind in one eye. No complications occurred in the exercise group.

Other outcome measures

The study by Moller (2000)²⁴ did not report on quality of life, return to work or medication use.

Discussion

This study had a moderate risk of bias, its main limitation was that trial patients and clinicians could not be reasonably blinded to treatment and given that surgery can be associated with important placebo effects, an overestimation of subjective outcomes such as pain cannot be ruled out. Also the results of this study are based on a small sample of patients and given that this is the only trial investigating the effect of fusion on patients with isthmic spondylolisthesis, the generalisability of these results is unclear.

Conclusions

In CLBP patients with isthmic spondylolisthesis, fusion improved the pain response and functional activity more than multidimensional supervised rehabilitation at two years. However, this is a single relatively small RCT, and although there appeared to be a relatively strong treatment effect, the results should be treated with caution.

Degenerative spondylolisthesis

Study characteristics

We identified one RCT by Weinstein (2007),⁴ which compared the effectiveness of surgical and non-surgical treatment among CLBP patients with degenerative spondylolisthesis. Patients were offered the option of enrolling into a randomised cohort (where patients were randomised to treatment) or an observational cohort (where patients were given the option to choose either fusion or non-surgical care). Treatment was standard decompressive laminectomy (with or without fusion). Non-surgical treatment included any of the following: active physical therapy, education or counseling, instructions for exercising at home, and non-steroidal anti-inflammatory agents if tolerated. Throughout the trial, patients in the observational cohort and the non-surgical treatment groups had the option of crossing over to surgery. Both the randomised and observational cohorts were followed up for two years. The primary outcome measures were the SF-36 bodily pain and physical function domains and the modified ODI.

This study compared surgical and non-surgical treatments at six weeks, three months, six months, one year, and two years using changes from baseline for SF-36 bodily pain and physical function and for the ODI. The randomised cohort was analysed on an intention-to-treat basis, however due to the high rate of crossover, an as-treated analysis was also performed.

Results

The authors enrolled 304 patients in the randomised cohort and 303 in the observational cohort. In both cohorts, >95% of patients underwent decompressive surgery with fusion. Patients in the trial underwent the following non-surgical care; physical therapy (42%), epidural steroid injections (45%), non-steroidal anti-inflammatory drugs (51%), and opioids (34%).

The rate of crossover in the randomised cohort was 49% at two years. In the observational cohort 97% of patients choosing surgery underwent surgical treatment in the first year. Of those initially choosing non-surgical treatment, 17% underwent surgery by one year and 25% by two years.

Pain

The intention-to-treat analysis of the randomised cohort showed no significant difference between surgical and non-surgical treatment groups at two years for SF-36 bodily pain (MD 1.5 95% CI: -4.2 to 7.3). In contrast the as-treated effects at two years favoured surgery, 17.8 (95% CI, 12.5 to 23.0) in the randomised cohort and 18.5 (95% CI, 13.4 to 23.6) in the observational cohort.

Function

The intention-to-treat analysis in the randomised cohort, showed no significant mean difference between fusion and non-surgical treatment at two years, 1.9 for SF-36 physical function (95% CI -3.7 to 7.5), and 2.2 for ODI (95% CI -2.3 to 6.8) while the as-treated effect significantly favoured surgery. In the randomised cohort for the as treated SF-36 physical function, the effect was 16.7 (95% CI, 11.4

to 22.1) and 19.9 (95% CI, 14.8 to 24.9) in the observational cohort; for the as treated Oswestry Disability Index, the effect was -15.9 (95% CI, -20.2 to -11.7) in the randomised cohort and -17.7 (95% CI, -21.6 to -13.7) in the observational cohort.

Other outcomes

The study by Weinstein (2007)⁴ did not report on quality of life, return to work or medication use.

Complications

There was little evidence of harm from either treatment. The most common surgical complication was dural tear (10%). The two-year re-operation rate was 12%.

Discussion

Although this study had a moderate risk of bias, a major limitation was the high degree of crossover from the non-surgical treatment group to surgery. Despite the study performing an intention-to-treat analysis the high rate of crossover would have severely diluted the effect of non-surgical treatment. Furthermore, although the as-treated effect was significant in favour of surgery, there was a high risk of selection bias as the effect of randomisation was negated.

Conclusion

Given the limitations of this study and the paucity of evidence, there is insufficient evidence to confirm whether spinal fusion is as effective as non-surgical treatment in people with CLBP with degenerative spondylolisthesis.

In what conditions or patient groups is spinal fusion indicated?

Three SRs^{9, 12, 18} were identified that examined whether certain factors (e.g. psychological subpopulations, smoking status, etc.) were treatment effect modifiers in trials comparing fusion with non-surgical treatment in patients with CLBP.

Comorbid disease

A SR by Choma (2011)⁹ examined whether comorbid disease or general health factors such as obesity, smoking, and alcohol and/or drug use were treatment effect modifiers of fusion versus non-surgical treatment in CLBP patients. This review included two RCTs that compared spinal fusion with non-surgical treatment in patients with CLBP,^{7, 22} see Table 2 for further detail on the Hagg (2003)²² and Fairbank (2005)⁷ studies. In this review, analyses were performed on a study level; data between studies were not pooled as the studies looked at different subgroups and outcomes were too heterogeneous. Forest plots for standardised mean differences and risk differences with their 95% confidence intervals were constructed comparing fusion to conservative management by subgroup to evaluate whether a treatment worked better in some subgroups than others.

The authors found that non-smokers and patients with no additional comorbidities may respond better to surgical fusion than non-surgical treatment, and recommended optimising the management of medical comorbidities and smoking cessation before considering surgical fusion in CLBP patients. This review had a low to moderate risk of bias due to the lack of information about the search strategy used to identify included studies. In addition to this, the authors of the study noted that their recommendations were based on 'weak' evidence, due to a significant lack of research in this area. The lack of evidence did not allow the authors to draw conclusions on the effectiveness of spinal fusion in comparison to conservative management, but did allow for hypotheses to be generated and considered for clinical decision making and future research planning.

Psychological indications

A SR by Daubs (2011)¹² examined whether fusion was superior to conservative management in certain psychological subpopulations (depression, stress/anxiety, personality disorders) and to determine the most common psychological screening tests and their ability to predict outcomes after treatment in patients with CLBP. This review included one RCT²² (see Table 2 for further information). In this review, all analyses were performed on a study level; data were not pooled as only one article was identified. For binary outcomes, risk differences and 95% confidence intervals were calculated and a forest plot was constructed comparing fusion to conservative management by subgroup to evaluate whether treatment worked better in some subgroups than others. For studies that reported continuous scores of a particular subgroup at baseline, paired t tests were used to compare the differences in baseline scores between fusion and conservative groups by outcome.

The authors found that patients with depression, neuroticism, and certain personality disorders appeared to respond more favourably to non-surgical treatment and those without a personality disorder more favourably to fusion. This SR had a low to moderate risk of bias due to the lack of information about the search strategy used to identify included studies. In addition, the recommendations of the review are noted to be based on insufficient evidence, meaning that evidence is either unavailable or does not permit a conclusion. The authors of the review note that one study is not enough to make treatment recommendations, but is sufficient to generate hypotheses to be considered in clinical decision making and future research.

Isthmic spondylolisthesis

A systematic review by Wood (2011)¹⁸ examined whether the presence of isthmic spondylolisthesis (IS) modified the effect of treatment (fusion vs. supervised rehabilitation) in patients with CLBP. This review failed to find any studies that compared outcomes between CLBP patients with and without IS, and instead included three RCTs of patients without IS,^{3, 7, 20} and two publications relating to one RCT of CLBP patients with IS.^{19, 24}

In this review, the focus of the analysis was to evaluate subgroups within larger comparative trials. All analyses were performed on a study level. Data were not pooled because of potentially important differences in patient populations among studies. The standardized mean differences, risk differences and 95% confidence intervals comparing fusion versus multidimensional supervised

rehabilitation were calculated as effect estimates where appropriate. Effect estimates were qualitatively compared visually with forest plots to evaluate whether a treatment worked better in one subgroup compared with the other.

This review found that the presence of IS in patients with CLBP may positively modify the treatment effect of fusion vs. multidimensional supervised rehabilitation with respect to pain and function, however, they classed the overall strength of evidence behind these findings as 'low', meaning that the authors have low confidence that the evidence reflects the true effect, and that further research is likely to change these findings. This review had a low to moderate risk of bias. Variations in the patient populations, interventions, comparators, outcomes length of follow-up for the included studies make generalisation of the findings difficult.

Conclusion

Overall, the three SRs identified do not provide sufficient evidence to conclude that the presence of specific comorbid diseases, general health factors, psychological subpopulations, or isthmia spondylolisthesis are either indications for spinal fusion surgery, or predictors of outcomes in patients undergoing spinal fusion or non-surgical treatment. Neither do they provide sufficient evidence to conclude that spinal fusion is not effective or not indicated in patients with specific comorbid diseases, general health factors, psychological indications, or isthmia spondylolisthesis in addition to CLBP.

Table 2. Characteristics of RCTs of indications/effect modifiers.

Study	Patient population	Intervention	Comparator	Outcomes & other variables
Fairbank 2005⁷	N=349 Age 18-55 years No more than 12 month history of chronic low back pain, with or without referred pain and irrespective of previous decompression or discectomy surgery.	Spinal stabilisation surgery- Particular technique of fusion was left up to the surgeon.	Intensive rehabilitation: daily outpatient program of education and exercise 5 days per week. For 3 weeks continuously. Exercises individually tailored aimed to build upon patients baseline ability. Exercises included stretching of major muscle groups, spinal flexibility exercises, cardiovascular exercises. Principles of cognitive behaviour therapy were used to identify and overcome fears and unhelpful beliefs.	<p><i>Outcome measures:</i></p> <ul style="list-style-type: none"> -Back specific pain questionnaire (ODI) -Standardized walking test <p>Secondary outcome measures</p> <ul style="list-style-type: none"> - SF-36 <p><i>Psychological assessment</i></p> <ul style="list-style-type: none"> - DRAM <p>Complications of surgery</p> <p>Work Status</p> <p><i>Other variables:</i></p> <p><u>Sociodemographic data:</u> age, gender, duration of back pain, smoking status, litigation, work status, effect of back pain on work</p> <p><u>Clinical classification:</u> spondylolisthesis, post-laminectomy, CLBP</p>
Hagg 2003²²	N=264 3 surgical groups: (n=201) 1 Non-surgical group (n= 63) Age 25–65 years Severe CLBP of at least 2 years duration, with no signs of root compression. Patients must have been on sick leave (or have had equivalent disability) for at least 1 year, and non-surgical treatment efforts should	<p>Surgery (fusion): patients allocated to surgery were operated according to one of three commonly used surgical techniques at one or both of the two lower lumbar levels.</p> <p>Group1: non-instrumented posterolateral fusion</p> <p>Group 2: instrumented posterolateral fusion</p> <p>Group 3: instrumented posterolateral fusion +</p>	Group 4: control group, “treated with commonly used non-surgical treatments outlined in a study protocol and executed according to local preferences.”	<p><i>Outcome measures:</i></p> <ul style="list-style-type: none"> - Patient global assessment of treatment effect - Change of disability (>=50% reduction in ODI score) - Work status <p><i>Other variables:</i></p> <p><u>Sociodemographic data:</u> age, sex, occupation, work status, marital status, co-morbidity, workers’ compensation, duration of CLBP, duration of sick leave, previous surgery</p>

	have been unsuccessful.	interbody fusion		and smoking.
	A score of at least 7/10 points on the Function and Working disability Score.			<p><u>Psychological assessment:</u> personality traits (KSP), personality disorders (SCID II), depressive symptoms (ZDS), pain behaviour (Waddell inappropriate signs and symptoms test & UAB Pain Behaviour Scale)</p> <p><u>Pain assessment:</u> VAS</p> <p><u>Disability assessment:</u> ODI & GFS</p> <p><u>Clinical findings:</u> Pain, motor, reflex and sensation</p> <p><u>Radiography:</u> to identify traction spurs, loss of disc height, vertebral slip, scoliosis and lordosis</p>
Moller 2000²⁴/ Ekman 2005¹⁹	<p>N = 111</p> <p>Aged 18-55yr</p> <p>Lumbar isthmic spondylolisthesis of any grade, no previous spine surgery, CLBP or sciatica of at least 1 year duration, severely restricted functional ability</p>	<p>Fusion (n = 77)</p> <ul style="list-style-type: none"> - PLF in situ with autologous bone transplantation harvested from the right iliac crest; without instrumentation (n = 40) and with rigid pedicle screw fixation (n = 37) - Non instrumented patients wore a daytime brace for 6 mo after surgery - No postoperative exercise or physiotherapy program was given 	<p>Non operative (n = 34)</p> <ul style="list-style-type: none"> - Exercise program based on strength and postural training; overseen by a physiotherapist with special interest in spondylolisthesis - Patients exercised 3x wk the first 6 mo, and 2x wk between 6 and 12 mo (duration approximately 45 min); after 1 yr patients were instructed to continue with a home program consisting of the 8 exercises that did not require special equipment 	<p><i>Outcome measures:</i></p> <p>DRI</p> <p>Pain index</p> <p>Patient and observer perceived improvement</p> <p>Patient also answered the question: "Would you go through the treatment again now that you know the result?"</p> <p>Global assessment</p>

Are there any reliable diagnostic procedures that predict the success of a fusion operation?

No evidence was identified to answer this question.

DISCLAIMER

The information in this report is a summary of that available and is primarily designed to give readers a starting point to consider currently available research evidence. Whilst appreciable care has been taken in the preparation of the materials included in this publication, the authors and the National Trauma Research Institute do not warrant the accuracy of this document and deny any representation, implied or expressed, concerning the efficacy, appropriateness or suitability of any treatment or product. In view of the possibility of human error or advances of medical knowledge the authors and the National Trauma Research Institute cannot and do not warrant that the information contained in these pages is in every aspect accurate or complete. Accordingly, they are not and will not be held responsible or liable for any errors or omissions that may be found in this publication. You are therefore encouraged to consult other sources in order to confirm the information contained in this publication and, in the event that medical treatment is required, to take professional expert advice from a legally qualified and appropriately experienced medical practitioner.

CONFLICT OF INTEREST

The TAC/WSV Evidence Service is provided by the National Trauma Research Institute. The NTRI does not accept funding from pharmaceutical or biotechnology companies or other commercial entities with potential vested interest in the outcomes of systematic reviews.

The TAC/WSV Health Services Group has engaged the NTRI for their objectivity and independence and recognise that any materials developed must be free of influence from parties with vested interests. The Evidence Service has full editorial control.

REFERENCES

1. Standifer S, Griggs M, Keenan D, Crouch D. Chronic Back Pain. Handbook of Disabilities: University of Missouri. 2001; Available from: <http://dps.missouri.edu/resources/Handbook/back.pdf>.
2. Ibrahim T, Tleyjeh IM, Gabbar O. Surgical versus non-surgical treatment of chronic low back pain: a meta-analysis of randomised trials.[Erratum appears in Int Orthop. 2009 Apr;33(2):589-90]. Int Orthop. 2008;32(1):107-13.
3. Brox JI, Sorensen R, Friis A, Nygaard O, 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(17):1913-21.
4. Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA, et al. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. NEJM. 2007;356(22):2257-70.
5. Brox JI, Nygaard OP, Holm I, Keller A, Ingebrigtsen T, Reikeras O. Four-year follow-up of surgical versus non-surgical therapy for chronic low back pain. Ann Rheumatic Dis. 2010;69(9):1643-8.
6. Brox JI, Reikeras O, Nygaard O, Sorensen R, Indahl A, Holm I, et al. 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. Pain. 2006;122(1-2):145-55.
7. Fairbank J, Frost H, Wilson-MacDonald J, Yu LM, Barker K, Collins R, et al. 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.
8. Carreon LY, Glassman SD, Howard J. Fusion and nonsurgical treatment for symptomatic lumbar degenerative disease: a systematic review of Oswestry Disability Index and MOS Short Form-36 outcomes. Spine J. 2008;8(5):747-55.
9. Choma TJ, Schuster JM, Norvell DC, Dettori JR, Chutkan NB. Fusion versus nonoperative management for chronic low back pain do comorbid diseases or general health factors affect outcome? Spine. 2011;36(21):S87-S95.
10. Chou R, Baisden J, Carragee EJ, Resnick DK, Shaffer WO, Loeser JD. Surgery for Low back pain a review of the evidence for an American Pain Society clinical practice guideline. Spine. 2009;34(10):1094-109.
11. Chou R, Loeser JD, Owens DK, Rosenquist RW, Atlas SJ, Baisden J, et al. Interventional therapies, surgery, and interdisciplinary rehabilitation for low back pain an evidence-based clinical practice guideline from the American Pain Society. Spine. 2009;34(10):1066-77.
12. Daubs MD, Norvell DC, McGuire R, Molinari R, Hermsmeyer JT, Fournery DR, et al. Fusion versus nonoperative care for chronic low back pain do psychological factors affect outcomes? Spine. 2011;36(21):S96-S109.
13. Gibson JN, Waddell G. Surgery for degenerative lumbar spondylosis. Cochrane Database Syst Rev. 2008(2):CD001352.
14. Mirza SK, Deyo RA. Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain. Spine. 2007;32(7):816-23.

15. Nordin M, Balague F, Cedraschi C. Nonspecific lower-back pain: Surgical versus nonsurgical treatment. *Clin Orthop Relat Res.* 2006(443):156-67.
16. Resnick DK, Choudhri TF, Dailey AT, Groff MW, Khoo L, Matz PG, et al. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 7: intractable low-back pain without stenosis or spondylolisthesis. *J Neurosurg.* 2005;Spine. 2(6):670-2.
17. Savigny P KS, Watson P, Underwood M, Ritchie G, Cotterell M, Hill D, Browne N, Buchanan E, Coffey P, Dixon P, Drummond C, Flanagan M, Greenough, C, Griffiths M, Halliday-Bell J, Hettinga D, Vogel S, Walsh D. *Low Back Pain: early management of persistent non-specific low back pain (NICE CG88).* London: National Collaborating Centre for Primary Care and Royal College of General Practitioners; 2009.
18. Wood KB, Fritzell P, Dettori JR, 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.* 2011;36(21):S110-S9.
19. 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.
20. Fritzell P, Hagg O, Wessberg P, Nordwall A, Swedish Lumbar Spine Study G. 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(23):2521-32; discussion 32-4.
21. Froholdt A, Holm I, Keller A, Gunderson RB, Reikeraas O, Brox JI. No difference in long-term trunk muscle strength, cross-sectional area, and density in patients with chronic low back pain 7 to 11 years after lumbar fusion versus cognitive intervention and exercises. *Spine J.* 2011;11(8):718-25.
22. Hagg O, Fritzell P, Ekselius L, Nordwall A, Swedish Lumbar Spine S. Predictors of outcome in fusion surgery for chronic low back pain. A report from the Swedish Lumbar Spine Study. *Eur Spine J.* 2003;12(1):22-33.
23. Keller A, Brox JI, Gunderson R, Holm I, Friis A, Reikeras 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(1):3-8.
24. Moller H, Hedlund R. Surgery versus conservative management in adult isthmic spondylolisthesis--a prospective randomized study: part 1. *Spine.* 2000;25(13):1711-5.
25. Ohtori S, Koshi T, Yamashita M, Yamauchi K, Inoue G, Suzuki M, et al. Surgical versus nonsurgical treatment of selected patients with discogenic low back pain a small-sized randomized trial. *Spine.* 2011;36(5):347-54.
26. Weinstein JN, Lurie JD, Tosteson TD, Zhao W, Blood EA, Tosteson AN, et al. 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.* 2009;American volume. 91(6):1295-304.
27. 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):968-74.

Evidence Service

Lumbar Spinal Fusion

Technical Report: Appendices 1-5

May 2013

Ornella Clavisi, Natasha Dodge, Emma Donoghue, Jason Wasiak

INTRODUCTION

This technical report is a companion document to “Lumbar Spinal Fusion: Evidence Review”. It contains detailed information about the methods used in the development of the Evidence Review, summaries of the studies included in the review, and quality appraisal results for the most recent and/or most relevant included studies.

CONTENTS

APPENDIX 1: REVIEW PROCESS.....	3
APPENDIX 2: METHODS	4
APPENDIX 3: LIST OF INCLUDED STUDIES	13
APPENDIX 4: SUMMARY OF KEY STUDIES.....	15
APPENDIX 5: APPRAISAL TABLES FOR KEY STUDIES.....	25

APPENDIX 1: REVIEW PROCESS

A two-staged approach was undertaken.

STAGE 1

Identify evidence available for each intervention

- Run search in health databases, websites and on the internet, limit to Evidence-Based Guidelines (EBGs), Health Technology Assessments (HTAs), Systematic Reviews (SRs) and Randomised Controlled Trials (RCTs)
- Apply inclusion and exclusion criteria

Critically appraise synthesised research

- Start with most recent review, apply standard appraisal criteria
- If found to be of high quality, cross check to ensure references from all other synthesised research are included and check for consistency of findings
- If not high quality, appraise next most recent and repeat process
- If there are inconsistent findings across the existing reviews, investigate the possibility of synthesis of this information or whether a new SR is required

Decide on actions for Stage 2

- Identify whether sufficient high level evidence exists to answer questions or identify what further action needs to be taken (see algorithm in Table A1.1)

STAGE 2

Address further actions identified.

Table A1.1. Further action required to answer clinical questions

Is there any synthesised research available? (e.g. EBGs, HTAs, SRs)				
Yes		No		
Is this good quality research?		Are RCTs available?		
Yes	No	Yes	No	
Is it current (within 2 years)?				
Yes	No			
No further action	Update existing SR	Undertake new SR and/or meta-analysis	Undertake new SR and/or meta-analysis	Consider looking for lower levels of evidence

APPENDIX 2: METHODS

TAC/WSV staff assisted in the development of search terms and inclusion and exclusion criteria.

Inclusion and exclusion criteria

Inclusion and exclusion criteria were established *a priori* (Table A2.1). These criteria were applied by two reviewers independently and any discrepancies were discussed and resolved.

Table A2.1 Inclusion and Exclusion criteria

Patient/ population	Inclusion: Adults with persistent spinal pain. Post-traumatic and degenerative spinal conditions
	Exclusion: Paediatrics. Acute trauma related, fracture-stabilisation surgery. Tumour related. Pregnancy related
Intervention/ indicator	Inclusion: Lumbar spinal Fusion, vertebral arthrodesis lumbar, single or multilevel, first fusion, as a first or second line treatment for persistent spinal pain
	Exclusion: laminectomy, artificial disc replacement, subsequent fusions
Comparison/ control	Inclusion: Non-operative care (usual care, rehab, physio, medications, standard care)
	Exclusion: Not compared to placebo or sham operations
Outcomes	Inclusion: Pain reduction, functional improvement, quality of life, return to work, medication use, harmful effects or adverse events
	Exclusion: N/A
Setting	Inclusion: Inpatient
	Exclusion: outpatient, Long term care
Study Design	Inclusion: SRs, HTA and EBGs, recent RCTs
	Exclusion: Low level evidence
Publication details	Inclusion: English language studies on humans
	Exclusion: Non English studies, animal studies
Time period	Inclusion: Studies with in the last 20 years
	Exclusion: Studies earlier than 1992

Searches undertaken

Search strategy development

The breadth of this topic required reduction of possible search terms so that the resulting reference list was a manageable number for manual reviewing. To assist with this, ten previously known references were searched prospectively in the Web of Knowledge database.

After discussion, the decision was made to undertake a preliminary search to identify high-level evidence for spinal fusion without specifying to each of the individual questions, and an overview Medline strategy was developed (see Table A2.2). This Medline strategy was translated into other databases (see Table A2.3). A further search specifically for RCTs was conducted (see Table A2.2). None of the databases were searched before 1992.

Internet searches to identify relevant websites and online evidence

Guidelines are generally published as electronic 'stand alone' documents on the internet rather than papers in peer reviewed journals. The search to identify guidelines is therefore quite different to searches usually undertaken for SRs

of primary research in the health literature. Methods to identify guidelines include identification of relevant websites followed by searches within them, direct searches of the internet and searches within the electronic health databases.

Unlike SRs of the primary research literature, which usually involve a single clinical question, a guideline is a compilation of SRs arising from numerous clinical questions. These reviews may be grouped within the guideline into sections or chapters based on patient characteristics, health care settings, domains of clinical care or other categories related to the particular condition or patient population. Due to this complexity and comprehensiveness, the information contained in a particular section or chapter may be found in more than one type of guideline.

The reviewers were aware of websites of guideline clearinghouses, guideline developers, centres of evidence-based practice and Australian government health services known to contain evidence-based resources. The 43 websites listed below were searched for relevant EBGs (see Table A2.4).

An additional internet search strategy was conducted using the Google 'Advanced Search' function. The search strings were limited to documents in English and were used to identify any information in the public domain that could further inform the Spinal Fusion research. The first 100 hits were screened for inclusion.

Table A2.2 Medline search strategy

Strategy for synthesized research	
1	spinal fusion/
2	((spinal or spine* or lumbar) adj2 (fusion* or fusing or fuse)).ti,ab.
3	or/1-2
4	"review"/ or review.pt. or review.ti.
5	(systematic or evidence\$ or methodol\$ or quantitativ\$ or analys\$ or assessment\$).ti,sh,ab.
6	4 and 5
7	meta-analysis.pt.
8	Meta-Analysis/
9	"systematic review*".ti,ab.
10	(meta-analy\$ or metanaly\$ or metaanaly\$ or meta analy\$).mp.
11	((systematic\$ or evidence\$ or methodol\$ or quantitativ\$) adj5 (review\$ or survey\$ or overview\$)).ti,ab,sh.
12	((pool\$ or combined or combining) adj2 (data or trials or studies or results)).ti,ab.
13	practice guideline/
14	(clinical adj3 guideline*).ti,ab.
15	or/6-14
16	and/3,15
17	limit 16 to yr="1992 -Current"
Strategy for RCTs	
1	spinal fusion/
2	((spinal or spine* or lumbar) adj2 (fusion* or fusing or fuse)).ti,ab.
3	or/1-2
4	randomized controlled trial.pt. or randomized.mp. or placebo.mp.
5	and/3,4
6	limit 5 to yr="1992 -Current"

** Adapted for use in other databases

Table A2.3 Databases accessed

Database name	Dates covered	Date searched	Refs
Medline (Ovid)	1946 to November Week 3 2012	27 th Nov 2012	542
PreMedline (Ovid)	November 27, 2012	27 th Nov 2012	25
All EBM (Ovid) *	Complete databases – July 2012	28 th Nov 2012	402
CINAHL (Ovid)	1992 - date	28 th Nov 2012	63
EMBASE	1980 to 2012 Week 47	28 th Nov 2012	1082
WoK	Complete databases – July 2012	5 th Nov 2012	194

*including The Cochrane Database of Systematic Reviews, DARE, CENTRAL, NHSEED, HTA and ACP Journal Club

Table A2.4 Website searches to identify relevant EBGs

Search 1: Identification of relevant guidelines for Spinal Fusion using specific guideline-related websites		
Guideline services	Results	Search
National Health and Medical Research Council (NHMRC)	www.nhmrc.gov.au	Terms used: Spinal fusion N/A
National Institute for Health and Clinical Excellence UK (NICE)	www.nice.org.uk	Terms used: Spinal fusion Transaxial interbody lumbosacral fusion (IPG387) Lateral (including extreme, extra and direct lateral) interbody fusion in the lumbar spine (IPG321) Low back pain (CG88)
New Zealand Guideline Group (NZGG)	www.nzgg.org.nz	Terms used: Spinal fusion N/A
Scottish Intercollegiate Guidelines Network (SIGN)	www.sign.ac.uk	Terms used: Spinal fusion N/A
Guidelines Advisory Committee	www.gacguidelines.ca	Scanned their list of Endorsed guidelines. N/A
Centre for Effective Practice	www.effectivepractice.org	Terms used: Spinal fusion Low Back Pain strategy health.gov.on.ca/en/pro/programs/ecfa/action/primary/lower_back.aspx
National Guideline Clearinghouse US (NGC)	guideline.gov	Terms used: Spinal fusion <u>Cervical and thoracic spine disorders</u> . 2011. NGC:008890 American College of Occupational and Environmental Medicine - Medical Specialty Society. <u>Low back pain. Early management of persistent non-specific low back pain</u> . 2009 May. NGC:007269 National Collaborating Centre for Primary Care - National Government Agency [Non-U.S.]. <u>Guideline for the evidence-informed primary care management of low back pain</u> . 2009 Mar (revised 2011). NGC:009259 Institute of Health Economics - Nonprofit Research Organization; Toward Optimized Practice - State/Local Government Agency [Non-U.S.]. <u>Diagnosis and treatment of degenerative lumbar spondylolisthesis</u> . 2008. NGC:006568 North American Spine Society - Medical Specialty Society <u>Low back - lumbar & thoracic (acute & chronic)</u> . 2003 (revised 2011 Mar 14). NGC:008517 Work Loss Data Institute - For Profit Organization. <u>Laminectomy and fusion for the treatment of cervical degenerative myelopathy</u> . 2009 Aug. NGC:008134 American Association of Neurological Surgeons - Medical Specialty Society; Congress of Neurological Surgeons - Professional Association <u>Cervical surgical techniques for the treatment of cervical spondylotic myelopathy</u> . 2009 Aug. NGC:008132 American Association of Neurological Surgeons - Medical Specialty Society; Congress of Neurological Surgeons - Professional Association. <u>ACR Appropriateness Criteria® low back pain</u> . 1996 (revised 2011). NGC:008863 American College of Radiology - Medical Specialty Society

		<p><u>Thoracolumbar spine surgery: a guide to preoperative and postoperative patient care</u>. 2012. NGC:008893 American Association of Neuroscience Nurses - Professional Association</p> <p><u>Neck and upper back (acute & chronic)</u>. 2003 (revised 2011 Apr 7). NGC:008518 Work Loss Data Institute - For Profit Organization.</p> <p><u>Diagnosis and treatment of degenerative lumbar spinal stenosis</u>. 2002 (revised 2011). NGC:008766 North American Spine Society - Medical Specialty Society.</p> <p><u>Radiographic assessment of cervical subaxial fusion</u>. 2009 Aug. NGC:008141 American Association of Neurological Surgeons - Medical Specialty Society; Congress of Neurological Surgeons - Professional Association.</p> <p><u>Cervical laminectomy for the treatment of cervical degenerative myelopathy</u>. 2009 Aug. NGC:008133 American Association of Neurological Surgeons - Medical Specialty Society; Congress of Neurological Surgeons - Professional Association</p> <p><u>Cervical laminoplasty for the treatment of cervical degenerative myelopathy</u>. 2009 Aug. NGC:008135 American Association of Neurological Surgeons - Medical Specialty Society; Congress of Neurological Surgeons - Professional Association.</p> <p><u>Low back disorders</u>. 1997 (revised 2007). NGC:006456 American College of Occupational and Environmental Medicine - Medical Specialty Society.</p>
TRIP Database	www.tripdatabase.com	<p>Terms used: Spinal Fusion</p> <p>12 references downloaded to Endnote library</p>
Australian Government websites containing guidelines		
Australian Government Department of Health & Ageing	www.health.gov.au	<p>Terms used: Spinal Fusion</p> <p>N/A</p>
Australian Institute of Health and Welfare	www.aihw.gov.au	<p>Terms used: Spinal Fusion</p> <p>N/A</p>
Health Insite	www.healthinsite.gov.au	<p>Terms used: Spinal Fusion</p> <p>N/A</p>
ACT Health	www.health.act.gov.au	<p>Terms used: Spinal Fusion</p> <p>N/A</p>
NSW Health	www.health.nsw.gov.au	<p>Terms used: Spinal Fusion</p> <p>N/A</p>
NT Department of Health and Community Services	www.health.nt.gov.au	<p>Terms used: Spinal Fusion</p> <p>N/A</p>
Queensland Health	www.health.qld.gov.au	<p>Terms used: Spinal Fusion</p> <p>N/A</p>
SA Department of Health and Human Services	www.health.sa.gov.au	<p>Terms used: Spinal Fusion</p> <p>N/A</p>
Tasmanian Department of Health and Human Services	www.dhhs.tas.gov.au	<p>Terms used: Spinal Fusion</p> <p>N/A</p>

Victorian Department of Human Services	www.dhs.vic.gov.au	Terms used: Spinal Fusion N/A
WA Department of Health	www.health.wa.gov.au	Terms used: Spinal Fusion Spinal Pain Model of Care www.healthnetworks.health.wa.gov.au/modelsofcare/docs/Spinal_Pain_Model_of_Care.pdf
Centres of evidence based practice websites		
Western Australian Centre for Evidence Informed Healthcare Practice	wacebnm.curtin.edu.au	Terms used: Spinal Fusion N/A
Other accident commissions		
Transport Accident Commission	www.tac.vic.gov.au	Terms used: Spinal Fusion N/A
Australian Transport Safety Bureau	www.atsb.gov.au	Terms used: Spinal Fusion N/A
Road Safety Victoria (TAC)	www.tacsafety.com.au	Terms used: Spinal Fusion N/A
WorkSafe Victoria	www.worksafe.vic.gov.au	Terms used: Spinal Fusion N/A
Traffic Injury Research Foundation	www.trafficinjuryresearch.com	Terms used: Spinal Fusion N/A
Motor Accidents Authority NSW	www.maa.nsw.gov.au	Terms used: Spinal Fusion N/A
WorkSafe British Columbia	www.worksafebc.com	Terms used: Spinal Fusion 1. Rehabilitation Services and Claims Manual Volume II, WorkSafeBC www.worksafebc.com/publications/policy_manuals/rehabilitation_... 2. Rehabilitation Services and Claims Manual Volume I, WorkSafeBC www.worksafebc.com/publications/policy_manuals/rehabilitation_... 3. Rehabilitation Services and Claims Manual Volume II, WCB of BC www.worksafebc.com/publications/policy_manuals/Rehabilitation_... 4. Rehabilitation Services and Claims Manual Summary of Amendments www.worksafebc.com/publications/policy_manuals/Rehabilitation_... 5. Workers' Compensation Reporter 19-2, WCB of BC www.worksafebc.com/publications/newsletters/wc_reporter/volume_19/... 6. Workers' Compensation Reporter 19-1, WCB of BC www.worksafebc.com/publications/newsletters/wc_reporter/volume_19/... 7. Workers Compensation Reporter 18-2, WCB of BC www.worksafebc.com/publications/newsletters/wc_reporter/volume_18/... 8. Workers' Compensation Reporter Volume 13, Number 2 www.worksafebc.com/publications/newsletters/wc_reporter/volume_13/...

		<p>9. Workers' Compensation Reporter Volume 10, Number 4 www.worksafebc.com/publications/newsletters/wc_reporter/volume_10/...</p> <p>10. Workers' Compensation Reporter Volume 9, Number 1 www.worksafebc.com/publications/newsletters/wc_reporter/volume_09/...</p> <p>11. Workers' Compensation Reporter Volume 8, Number 6 www.worksafebc.com/publications/newsletters/wc_reporter/volume_08/...</p> <p>12. Workers' Compensation Reporter Volume 8, Number 3 www.worksafebc.com/publications/newsletters/wc_reporter/volume_08/...</p> <p>13. Workers' Compensation Reporter Volume 8, Number 2 www.worksafebc.com/publications/newsletters/wc_reporter/volume_08/...</p> <p>14. Decision of the Review Division 2523, WCB of BC www.worksafebc.com/publications/newsletters/assets/pdf_reviewdiv/2523.</p> <p>15. Resolution of the Board of Directors 2003/06/17-06, WCB of BC www.worksafebc.com/publications/newsletters/assets/pdf_bod/20030617_...</p>
Accident Compensation Corporation	www.acc.co.nz	Terms used: Spinal Fusion N/A
Pain Treatment Topics	pain-topics.org	Terms used: Spinal Fusion N/A
The George Institute	www.georgeinstitute.org.au	Terms used: Spinal Fusion N/A
Injury Research and Prevention Unit	www.injuryresearch.bc.ca	Terms used: Spinal Fusion N/A
The Brain Trauma Foundation	www.braintrauma.org	Terms used: Spinal Fusion N/A
Safer Roads	www.rta.nsw.gov.au	Terms used: Spinal Fusion N/A
Rail Accident Investigation Branch	www.raib.gov.uk	Terms used: Spinal Fusion N/A
Oslo Sports Trauma Research Centre	www.klokavskade.no/en	Terms used: Spinal Fusion N/A
Oregon Evidence-Based Practice Centre	www.ohsu.edu/epc	Terms used: Spinal Fusion N/A
Injury Prevention Network of Aotearoa New Zealand	ipnanz.org.nz	Terms used: Spinal Fusion N/A
Trauma Centre at Justice Resource Centre	www.traumacenter.org	Terms used: Spinal Fusion N/A
The DANA Foundation	www.dana.org	Terms used: General introduction. Back Pain and Disk Disease — The Dana Guide

		www.dana.org/news/brainhealth/detail.aspx?id=9782
European Association for Injury Prevention and Safety Promotion	www.eurosafe.eu.com	Terms used: Spinal Fusion N/A
New Zealand Injury Prevention strategy	www.nzips.govt.nz	Terms used: Spinal Fusion N/A
NHS Health at Work	www.nhshealthatwork.co.uk	Terms used: Spinal Fusion General introduction Treating back pain www.nhs.uk/Conditions/Back-pain/Pages/Treatment.aspx
The Canadian Association of Road Safety Professionals	www.carsp.ca	Terms used: Spinal Fusion N/A
Search 2: Identification of relevant studies for spinal fusion using Google		
Find web pages that have all these words		
Find web pages that have this exact wording or phrase	Spinal Fusion	
Find web pages that have any of these words	guideline random	
Find web pages that are in the site or domain	.edu; .org; .gov; .net	
Limits	English	
Results 28/11/2012 completed search		

Appraisal

Appraisal was undertaken in steps:

1. The most recent review (EBG, SR or HTA) for each indication was assessed for quality using standard appraisal criteria.
2. If found to be of high quality, it was cross checked against the other available reviews to compare scope and consistency of findings.
3. If found not to be of high quality, the next most recent was appraised and the above process repeated.

Quality

EBGs and SRs were appraised using standard criteria by a single reviewer in consultation with colleagues as required. RCTs were also appraised using standard criteria by a single reviewer in consultation with colleagues as required. Details of quality appraisals are included in Appendix 5.

Data extraction

Data on characteristics of the studies were extracted and summarised.

APPENDIX 3: LIST OF INCLUDED STUDIES

Primary studies

Randomised controlled trials

1. Brox JI, Nygaard OP, Holm I, Keller A, Ingebrigtsen T, Reikeras O. Four-year follow-up of surgical versus non-surgical therapy for chronic low back pain. *Ann Rheumatic Dis.* 2010;69(9):1643-8.
2. Brox JI, Reikeras O, Nygaard O, Sorensen R, Indahl A, Holm I, et al. 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. *Pain.* 2006;122(1-2):145-55.
3. Brox JI, Sorensen R, Friis A, Nygaard O, 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(17):1913-21.
4. 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.
5. Fairbank J, Frost H, Wilson-MacDonald J, Yu LM, Barker K, Collins R, et al. 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.
6. Fritzell P, Hagg O, Wessberg P, Nordwall A, Swedish Lumbar Spine Study G. 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(23):2521-32; discussion 32-4.
7. Froholdt A, Holm I, Keller A, Gunderson RB, Reikeraas O, Brox JI. No difference in long-term trunk muscle strength, cross-sectional area, and density in patients with chronic low back pain 7 to 11 years after lumbar fusion versus cognitive intervention and exercises. *Spine J.* 2011;11(8):718-25.
8. Hagg O, Fritzell P, Ekselius L, Nordwall A, Swedish Lumbar Spine S. Predictors of outcome in fusion surgery for chronic low back pain. A report from the Swedish Lumbar Spine Study. *Eur Spine J.* 2003;12(1):22-33.
9. Keller A, Brox JI, Gunderson R, Holm I, Friis A, Reikeras 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(1):3-8.
10. Moller H, Hedlund R. Surgery versus conservative management in adult isthmic spondylolisthesis--a prospective randomized study: part 1. *Spine.* 2000;25(13):1711-5.
11. Ohtori S, Koshi T, Yamashita M, Yamauchi K, Inoue G, Suzuki M, et al. Surgical versus nonsurgical treatment of selected patients with discogenic low back pain: a small-sized randomized trial. *Spine.* 2011;36(5):347-54.
12. Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA, et al. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. *NEJM.* 2007;356(22):2257-70.
13. Weinstein JN, Lurie JD, Tosteson TD, Zhao W, Blood EA, Tosteson AN, et al. 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.* 2009;American volume. 91(6):1295-304.

Synthesised studies

Systematic reviews

1. Carreon LY, Glassman SD, Howard J. Fusion and nonsurgical treatment for symptomatic lumbar degenerative disease: a systematic review of Oswestry Disability Index and MOS Short Form-36 outcomes. *Spine J.* 2008;8(5):747-55.
2. Choma TJ, Schuster JM, Norvell DC, Dettori JR, Chutkan NB. Fusion versus nonoperative management for chronic low back pain do comorbid diseases or general health factors affect outcome? *Spine.* 2011;36(21):S87-S95.
3. Chou R, Baisden J, Carragee EJ, Resnick DK, Shaffer WO, Loeser JD. Surgery for low back pain a review of the evidence for an American Pain Society clinical practice guideline. *Spine.* 2009;34(10):1094-109.
4. Daubs MD, Norvell DC, McGuire R, Molinari R, Hermsmeyer JT, Fournay DR, et al. Fusion versus nonoperative care for chronic low back pain do psychological factors affect outcomes? *Spine.* 2011;36(21):S96-S109.
5. Gibson JN, Waddell G. Surgery for degenerative lumbar spondylosis. *Cochrane Database of Systematic Reviews* (Online). 2008(2):CD001352.
6. Ibrahim T, Tleyjeh IM, Gabbar O. Surgical versus non-surgical treatment of chronic low back pain: a meta-analysis of randomised trials.[Erratum appears in *Int Orthop.* 2009 Apr;33(2):589-90]. *Int Orthop.* 2008;32(1):107-13.
7. Mirza SK, Deyo RA. Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain. *Spine.* 2007;32(7):816-23.
8. Nordin M, Balague F, Cedraschi C. Nonspecific lower-back pain: Surgical versus nonsurgical treatment. *Clin Orthop Relat Res.* 2006(443):156-67.
9. Wood KB, Fritzell P, Dettori JR, 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.* 2011;36(21):S110-S9.

Evidence-based guidelines

1. Chou R, Loeser JD, Owens DK, Rosenquist RW, Atlas SJ, Baisden J, et al. Interventional therapies, surgery, and interdisciplinary rehabilitation for low back pain an evidence-based clinical practice guideline from the American Pain Society. *Spine.* 2009;34(10):1066-77.
2. Resnick DK, Choudhri TF, Dailey AT, Groff MW, Khoo L, Matz PG, et al. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 7: intractable low-back pain without stenosis or spondylolisthesis. *J Neurosurg: Spine.* 2005;2(6):670-2.
3. Savigny P KS, Watson P, Underwood M, Ritchie G, Cotterell M, Hill D, Browne N, Buchanan E, Coffey P, Dixon P, Drummond C, Flanagan M, Greenough C, Griffiths M, Halliday-Bell J, Hettinga D, Vogel S, Walsh D. Low back pain: early management of persistent non-specific low back pain (NICE CG88). London: National Collaborating Centre for Primary Care and Royal College of General Practitioners; 2009.

APPENDIX 4: SUMMARY OF KEY STUDIES

Table A4.1 Summary of studies.

1 st author, year, title	Inclusion, exclusion criteria (for P.I.C.O)	Study design	Conclusion/recommendation	Recommendation category	Other comments
SYSTEMATIC REVIEWS					
Carreon 2008 Fusion and nonsurgical treatment for symptomatic lumbar degenerative disease: a systematic review of Oswestry Disability Index and MOS Short Form-36 outcomes.	<p>POPULATION/CLINICAL INDICATION</p> <p>Included: Patients with low back pain of at least 12 weeks duration and older than 18 years, with prospectively collected ODI scores and at least a 12-month follow-up</p> <p>Excluded: nonfusion surgery, such as a disc replacement, intradiscal electrothermal therapy, decompression alone</p> <p>INTERVENTION</p> <p>surgical fusion</p> <p>COMPARATOR</p> <p>Nonsurgical treatment</p> <p>OUTCOMES:</p> <p>ODI and Short Form-36 (SF-36)</p>	SR	<p>Conclusions: Substantial improvement can be expected in patients treated with fusion, regardless of technique, when an established indication such as spondylolisthesis or degenerative disc disease exists. CLBP patients are less disabled and experience less improvement.</p> <p>In conclusion, the number of well-designed and adequately reported studies on interventions used to treat symptomatic lumbar degenerative disease is limited. Definite proof of treatment efficacy for both fusion and nonsurgical treatment of symptomatic lumbar degenerative disease remains unclear.</p>	<p><i>Positive (fusion more effective than non-intensive nonsurgical therapy)</i></p> <p><i>Insufficient evidence to draw conclusions</i></p>	
Choma 2011 Fusion Versus nonoperative management for chronic low back pain do comorbid diseases or general health factors affect outcome?	<p>POPULATION/CLINICAL INDICATION</p> <p>Included: Adults with chronic low back pain (CLBP), centralised or radiating</p> <p>Excluded: <18 years old; predominant neurological involvement; predominant spondylolisthesis or stenosis; cancer, deformity, instability, infection, inflammatory arthritis, trauma</p> <p>INTERVENTION</p> <p>surgical fusion</p> <p>COMPARATOR</p> <p>Nonsurgical treatment</p> <p>OUTCOMES:</p> <p>pain, physical function, quality of life, return to work</p>	SR	<p>Evidence Summary: The overall strength of the evidence evaluating whether specific disease or general health subpopulations modify the effect of fusion <i>versus</i> conservative management in the treatment of CLBP is “insufficient,” that is, evidence either is unavailable or does not permit a conclusion; however, some hypotheses can be generated and considered in clinical decision making and in future research planning.</p> <p>Conclusion. It is unclear from the literature which patients are the best candidates for fusion <i>versus</i> conservative management when experiencing CLBP without significant neurological impairment. Nonsmokers may be more likely to have a favorable surgical fusion outcome in CLBP patients. Comorbid disease presence has not been shown to definitively modify the effect of fusion. Further prospective studies that are designed to evaluate these and other subgroup effects are encouraged to confirm these findings.</p> <p>Clinical Recommendations. We recommend optimizing the management of medical co-morbidities and smoking</p>	<p><i>Insufficient evidence to draw conclusions</i></p>	

1 st author, year, title	Inclusion, exclusion criteria (for P.I.C.O)	Study design	Conclusion/recommendation	Recommendation category	Other comments
			cessation before considering surgical fusion in CLBP patients. Strength of recommendation: Weak		
Chou 2009 Surgery for low back pain a review of the evidence for an American Pain Society clinical practice guideline.	POPULATION/CLINICAL INDICATION Included: Adult patients with nonradicular back pain with common degenerative changes. Excluded: pregnant patients, acute major trauma, cancer, infection, caudaequina syndrome, and osteoporosis or vertebral compression fracture INTERVENTION surgical fusion COMPARATOR Nonsurgical treatment OUTCOMES: ODI, back pain, VAS, SF-36	SR	For nonradicular low back pain with common degenerative changes, there is fair evidence from randomised trials that fusion is no more effective than intensive rehabilitation with a cognitive behavioral emphasis, but slightly to moderately more effective than standard (non-intensive) nonsurgical therapy for improvement in pain and function.	<i>Neutral (no difference in effect between fusion and intensive rehab with cognitive behavioural element)</i> <i>Positive (fusion more effective than non-intensive nonsurgical therapy)</i>	
Daubs 2011 Fusion versus nonoperative care for chronic low back pain do psychological factors affect outcomes?	POPULATION/CLINICAL INDICATION Included: Adults with CLBP, centralised or radiating Excluded: <18 years old; predominant neurological involvement; predominant spondylolisthesis or stenosis; cancer, deformity, instability, infection, trauma INTERVENTION spinal fusion COMPARATOR conservative treatment OUTCOMES: Included: Pain, physical function, quality of life Excluded: cost effectiveness	SR	Psychological disorders affect chronic lower back pain treatment outcomes. Patients with a personality disorder appear to respond more favorably to conservative management and those without a personality disorder more favorably to fusion. Patients with higher depression and neuroticism scores may also respond more favorably to conservative management. Clinical Recommendations. Recommendation 1: Chronic LBP patients with depression, neuroticism, and certain personality disorders should preferentially be treated nonoperatively. Strength of recommendation: Weak.	<i>Weak recommendation regarding which patient subgroups should not be treated operatively</i>	The authors of the study note that: “despite finding that patients with a personality disorder may respond more favourably to conservative management and those without such a disorder more favorably to fusion, one study is not enough to make treatment recommendations, especially since subgroup analyses of secondary data are more appropriately considered hypothesis generating.”
Gibson 2008 Surgery for degenerative lumbar spondylosis.	POPULATION/CLINICAL INDICATION Patients over 18 years of age with degenerative lumbar spondylosis treated by surgery. INTERVENTION Spinal fusion COMPARATOR	SR	Fusion is more effective than continued, failed, standard 1990s ‘usual care’; it does not appear to be any more effective than a modern rehabilitation program	<i>Positive (fusion more effective than fusion and standard 1990’s usual care)</i> <i>Neutral (no</i>	

1 st author, year, title	Inclusion, exclusion criteria (for P.I.C.O)	Study design	Conclusion/recommendation	Recommendation category	Other comments
	Nonsurgical treatment OUTCOMES: ODI, back pain, work status and patient rating			<i>difference in effect between fusion and a modern rehab program)</i>	
Ibrahim 2008 Surgical versus non-surgical treatment of chronic low back pain: a meta-analysis of randomised trials.	POPULATION/CLINICAL INDICATION Patients undergoing surgical and non-surgical treatment for low back pain. INTERVENTION spinal fusion COMPARATOR nonsurgical treatment OUTCOMES: ODI	SR	We found that surgical fusion may improve the ODI compared to non-surgical intervention at the two-years follow-up for chronic low back pain. This improvement in the ODI compared to the original article (-4.87 compared to -4.13) was statistically significant and is of minimal clinical importance; consequently, surgeons should recommend spinal fusion cautiously to patients with chronic low back pain.	<i>Positive/Neutral (fusion more effective than nonsurgical treatment, but difference of minimal clinical importance)</i>	
Mirza 2007 Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain.	POPULATION/CLINICAL INDICATION <u>Included:</u> Adult patients with symptoms of low back pain for 12 months or longer without a specific diagnosis <u>Excluded:</u> Studies focused on patients with isthmic spondylolisthesis INTERVENTION spinal fusion COMPARATOR nonsurgical treatment OUTCOMES: Back-specific disability, pain, general function, psychological function, work status, radiographic results, complications and patient satisfaction. Oswestry Disability Index (ODI).	SR	These trials do not allow a general statement regarding the efficacy of fusion over nonoperative care for discogenic back pain. All 4 trials suggest that any advantage of surgery over nonsurgical care is modest, on average near or below the minimally important change in the disability score. The difference in the magnitude of non-surgical improvement suggests that the nature of non-surgical treatment may be critical.	<i>Positive/Neutral (fusion more effective than nonsurgical treatment, but difference of minimal clinical importance)</i>	
Nordin 2006 Nonspecific lower-back pain: surgical versus nonsurgical treatment.	POPULATION/CLINICAL INDICATION patients diagnosed with nonspecific lower-back pain INTERVENTION Surgical treatment (including fusion) COMPARATOR nonsurgical treatments OUTCOMES: Pain, disability, mood, return to work	SR	In cases of chronic nonspecific lower-back pain (> 12 weeks duration of pain) a variety of treatments are available with limited and similar efficacy on pain and disability reduction. There is moderate evidence that surgery in chronic nonspecific lower-back pain is as effective as cognitive behavioral treatment with regard to pain, function, mood and return to work. Surgical indications for chronic nonspecific lower-back pain remain ill defined.	<i>Neutral (no difference in effect between surgery and non-surgical treatment)</i>	
Wood 2011 Effectiveness of spinal fusion versus structured rehabilitation in chronic	POPULATION/CLINICAL INDICATION <u>Included:</u> Adults with Chronic Low Back Pain (CLBP) with or without isthmic spondylolisthesis <u>Excluded:</u> patients <18 years of age; LBP with predominantly neurological	SR	Conclusion. The overall strength of evidence evaluating whether the presence of isthmic spondylolisthesis modifies the effect of fusion compared with rehabilitation patients with CLBP is "low." Fusion should be considered for patients	<i>Weak recommendation regarding which patient subgroups</i>	

1 st author, year, title	Inclusion, exclusion criteria (for P.I.C.O)	Study design	Conclusion/recommendation	Recommendation category	Other comments
low back pain patients with and without isthmic spondylolisthesis a systematic review.	involvement; spinal stenosis; tumor; moderate to severe osteoporosis; trauma; osteomyelitis; infection of soft tissue adjacent to spine; systemic infection INTERVENTION <u>Included:</u> Fusion with or without instrumentation or decompression <u>Excluded:</u> surgical procedures other than spinal fusion (± decompression) due to CLBP COMPARATOR <u>Included:</u> Supervised rehabilitation <u>Excluded:</u> other lumbar surgeries; non multidimensional supervised rehabilitation OUTCOMES: <u>Included:</u> Physical function; pain (VAS back and leg); health related quality of life; patient satisfaction <u>Excluded:</u> nonclinical outcomes		with low back pain and isthmic spondylolisthesis who have failed nonoperative treatment. Clinical Recommendations. We recommend considering fusion for patients with isthmic spondylolisthesis and lower back pain who have failed nonoperative treatment. Recommendation: Weak.	<i>should be considered for spinal fusion</i>	
EVIDENCE-BASED GUIDELINES					
Chou 2009 Interventional therapies, surgery, and interdisciplinary rehabilitation for low back pain. An evidence-based clinical practice guideline from the American Pain Society.	<i>This EBG is based on the systematic review by Chou 2009. Please see Chou 2009 in the systematic review section for details.</i>	EBG	Recommendations on use of interventional diagnostic tests and therapies, surgery, and interdisciplinary rehabilitation are presented. Due to important trade-offs between potential benefits, harms, costs, and burdens of alternative therapies, shared decisionmaking is an important component of a number of the recommendations.	<i>see Chou 2009 in the systematic review section for details</i>	
Resnick 2005 Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 7: intractable low-back pain without stenosis or	POPULATION/CLINICAL INDICATION <u>Included:</u> patients with intractable low-back pain without stenosis or spondylolisthesis INTERVENTION Lumbar fusion COMPARATOR any	EBG	<i>Standards.</i> Lumbar fusion is recommended as a treatment for carefully selected patients with disabling low-back pain due to one- or two-level degenerative disease without stenosis or spondylolisthesis. <i>Guidelines.</i> There is insufficient evidence available to support a treatment guideline. <i>Options.</i> An intensive course of physical therapy and cognitive therapy is recommended as a treatment option for patients	<i>Both fusion and non-surgical treatments recommended in different circumstances</i>	

1 st author, year, title	Inclusion, exclusion criteria (for P.I.C.O)	Study design	Conclusion/recommendation	Recommendation category	Other comments
spondylolisthesis.	OUTCOMES: Not specified		with low-back pain in whom conventional medical management has failed.” “Class I medical evidence exists in support of the use of lumbar fusion as a treatment standard for carefully selected patients with low-back pain intractable to the best medical management. There is Class III medical evidence that suggests that a course of intensive cognitive and physical therapy may be an efficacious treatment option for the treatment of patients with chronic disabling low-back pain.		
Savigny 2009 Low back pain: early management of persistent non-specific low back pain	POPULATION/CLINICAL INDICATION Included: patients with non-specific low back pain INTERVENTIONS & COMPARATORS Pharmacological treatments, non-pharmacological treatments and referral to surgery OUTCOMES: Not specified	EBG	Consider referral for an opinion on spinal fusion for people who: • have completed an optimal package of care, including a combined physical and psychological treatment programme and • still have severe non-specific low back pain for which they would consider surgery.	<i>Positive (in favour of referral for an opinion on fusion for those failing optimal non-surgical treatment)</i>	
RANDOMISED CONTROLLED TRIALS					
Brox 2010 Four-year follow-up of surgical versus non-surgical therapy for chronic low back pain.	POPULATION/CLINICAL INDICATION Included: Adults aged 25-60 yrs, with chronic low back pain for at least 1 year. ODI greater than 30 and disc degeneration at L4-L5 and/or L5-S1. Excluded: 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 of the interventions. INTERVENTION Posterolateral fusion with transpedicular screws of the L4-L5 and or L5-S1 segment. COMPARATOR Cognitive intervention and exercises. OUTCOMES Primary outcome measure Oswestry Disability Index. Secondary outcome measure: pain, general function score, global back disability question for the assessment of patients overall rating, work and medication, emotional distress, fear-avoidance beliefs, and life satisfaction.	RCT	In conclusion, patients did not have better long-term improvement after instrumented fusion compared with cognitive intervention and exercises.	<i>Neutral (no difference in long-term effect between instrumented fusion and cognitive intervention and exercises)</i>	

1 st author, year, title	Inclusion, exclusion criteria (for P.I.C.O)	Study design	Conclusion/recommendation	Recommendation category	Other comments
Brox 2006 Lumbar instrumented fusion compared with cognitive intervention and exercises in patients with chronic low back pain after previous surgery for disc herniation: A prospective randomised controlled trial.	<p>POPULATION/CLINICAL INDICATION</p> <p>Included: Adults aged 25-60 yrs, with chronic low back pain for at least 1 year after previous disc herniation surgery, ODI greater than 30 and disc degeneration at L4-L5 and/or L5-S1.</p> <p>Excluded: 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 of the interventions.</p> <p>INTERVENTION</p> <p>Posterolateral fusion with transpedicular screws of the L4-L5 and or L5-S1 segment.</p> <p>COMPARATOR</p> <p>Cognitive intervention and exercises.</p> <p>OUTCOMES</p> <p>Primary outcome measure Oswestry Disability Index. Secondary outcome measure: pain, general function score, global back disability question for the assessment of patients overall rating, work and medication, emotional distress, fear-avoidance beliefs, and life satisfaction.</p>	RCT	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.	<i>Neutral (no difference in long-term effect between instrumented fusion and cognitive intervention and exercises)</i>	
Brox 2003 Randomised clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration.	<p>POPULATION/CLINICAL INDICATION</p> <p>Included: Adults aged 25-60 yrs, with chronic low back pain for at least 1 year. ODI greater than 30 and disc degeneration at L4-L5 and/or L5-S1.</p> <p>Excluded: 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 of the interventions.</p> <p>INTERVENTION</p> <p>Posterolateral fusion with transpedicular screws of the L4-L5 and or L5-S1 segment.</p> <p>COMPARATOR</p> <p>Cognitive intervention and exercises.</p>		It was concluded that after 1 year of follow-up, the difference between the groups given lumbar instrumented fusion and cognitive intervention and exercise was neither clinically important nor significant.	<i>Neutral (no difference in long-term effect between instrumented fusion and cognitive intervention and exercises)</i>	

1 st author, year, title	Inclusion, exclusion criteria (for P.I.C.O)	Study design	Conclusion/recommendation	Recommendation category	Other comments
	OUTCOMES Primary outcome measure Oswestry Disability Index. Secondary outcome measure: pain, general function score, global back disability question for the assessment of patients overall rating, work and medication, emotional distress, fear-avoidance beliefs, and life satisfaction.				
Ekman 2005 The long-term effect of posterolateral fusion in adult isthmic spondylolisthesis: a randomized controlled study.	POPULATION/CLINICAL INDICATION <u>Included:</u> patients aged 18 to 55 years with adult lumbar isthmic spondylolisthesis at L5 or L4 level of all degrees, and at least 1-year's duration of severe lumbar pain with or without sciatica. <u>Excluded:</u> INTERVENTION Posterolateral fusion without pedicle screw instrumentation, posterolateral fusion with pedicle screw instrumentation. COMPARATOR 1 year exercise program OUTCOMES Pain and functional disability was quantified by pain (VAS), the Disability Rating Index (DRI), the Oswestry Disability Index (ODI) work status, and global assessment of outcome by the patient into much better, better, unchanged or worse. Quality of life was assessed by the SF-36.	RCT	Posterolateral fusion in adult lumbar isthmic spondylolisthesis results in a modestly improved long-term outcome compared with a 1-year exercise program.	Positive/Neutral (fusion more effective than nonsurgical treatment, but difference of minimal clinical importance)	
Fairbank 2005 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.	POPULATION/CLINICAL INDICATION <u>Included:</u> Patients aged 18-55 years with chronic low back pain of at least one year duration <u>Excluded:</u> Infection, or other comorbidities (inflammatory disease, tumours, fractures), psychiatric disease, inability or unwillingness to complete the trial questionnaires or pregnancy. INTERVENTION Lumbar spine Fusion COMPARATOR Intensive rehabilitation programme based on principles of cognitive behavior therapy. OUTCOMES ODI, SF-36, Shuttle walking test.	RCT	No clear evidence emerged that primary spinal fusion surgery was any more beneficial than intensive rehabilitation.	Neutral (no difference in long-term effect between instrumented fusion and cognitive intervention and exercises)	

1 st author, year, title	Inclusion, exclusion criteria (for P.I.C.O)	Study design	Conclusion/recommendation	Recommendation category	Other comments
Fritzell 2001 2001 Volvo award winner in clinical studies: lumbar fusion versus nonsurgical treatment for chronic low back pain.	POPULATION/CLINICAL INDICATION <u>Included:</u> Aged 25-65 years with CLBP for at least 2 years and with radiological evidence of disc degeneration at L4-L5, L5-S1 or both. <u>Excluded:</u> Psychiatric illness, previous spine surgery except for successful removal of a herniated disc more than 2 years before entering the study and with no persistent nerve root symptoms Specific radiological findings such as spondylolisthesis, new or old fractures, infection, inflammatory process, or neoplasm. INTERVENTION Different methods of Fusion COMPARATOR Physical Therapy OUTCOMES Essential improvement, ODI, use of pain meds, VAS, General Function Score, Life satisfaction, Zung depression scale, Work status, independent observer, x rays, group changers, complications	RCT	Lumbar fusion in a well-informed and selected group of patients with severe CLBP can diminish pain and decrease disability more efficiently than commonly used non-surgical treatment.	<i>Positive/Neutral (fusion more effective than nonsurgical treatment)</i>	
Froholdt 2011 No difference in the long-term trunk muscle strength, cross-sectional area and density in patients with chronic low back pain 7 to 11 years after lumbar fusion versus cognitive intervention and exercises.	<i>Long term follow-up of patients from Keller 2004. Please see Keller 2004 for details.</i>	RCT	Although this study did not directly assess muscle morphology of muscles likely damaged by surgery, gross muscle strength, cross-sectional area, and density above the lesion are not different between those who have had lumbar fusion or cognitive intervention and exercises at 7- to 11-years after lumbar fusion.	<i>Neutral (no difference in long-term effect between instrumented fusion and cognitive intervention and exercises)</i>	
Hagg 2003 Predictors of outcome in fusion surgery for chronic low back pain. a report from the Swedish Lumbar Spine Study.	<i>Outcome predictor paper based on Fritzell study 2001. See Fritzell 2001 for details of study criteria.</i>	RCT	Depressive symptoms predicted functional improvement after non-surgical treatment. Work resumption was predicted by low age and short sick leave in the surgical group, and by short sick leave in the non-surgical group. We conclude that improved selection of successful surgical candidates with CLBP seems to be promoted by attention to severe disc degeneration, evaluation of personality traits and shortening of preoperative sick leave.	NA	

1 st author, year, title	Inclusion, exclusion criteria (for P.I.C.O)	Study design	Conclusion/recommendation	Recommendation category	Other comments
Keller 2004 Trunk muscle strength, cross sectional area, and density in patients with chronic low back pain randomized to lumbar fusion or cognitive intervention and exercises.	POPULATION/CLINICAL INDICATION <u>Included:</u> Age 25–60 years; reported low back pain for at least 1 year; a score of 30 of 100 points on the Oswestry Disability Index (ODI); and degenerative changes at the L4–L5 and/or L5–S1 on plain radiographs or previously performed surgery for disc herniation with laminectomy. <u>Excluded:</u> widespread myofascial pain, spinal stenosis with reduced walking distance and neurologic signs, disc herniation or lateral recess stenosis with clinical signs of radiculopathy, inflammatory disease, previous spinal fracture, the pelvic girdle syndrome, generalized degenerative changes on plain radiograph examination, serious somatic or psychiatric disease that excluded either one or both treatment alternatives, registered medical abuse, or reluctance to accept one or both the treatment regimens of the study. INTERVENTION Posterolateral fusion with transpedicular screws of the L4-L5 and or L5-S1 segment. COMPARATOR Cognitive intervention and exercises OUTCOMES Trunk Muscle strength, cross sectional area and density of back muscles	RCT	Patients with chronic low back pain who followed cognitive intervention and exercise programs improved significantly in muscle strength compared with patients who underwent lumbar fusion. In the lumbar fusion group, density decreased significantly at L3–L4 compared with the exercise group.	<i>Negative (exercise better than fusion for improving muscle strength)</i>	
Moller 2000 Surgery versus conservative management in adult isthmic spondylolisthesis.	POPULATION/CLINICAL INDICATION <u>Included:</u> lumbar isthmic spondylolisthesis of any grade, at least 1 year of low back pain or sciatica, and a severely restricted functional ability in individuals 18 to 55 years of age. <u>Excluded:</u> not specified INTERVENTION Posterolateral fusion with or without transpedicular fixation COMPARATOR Exercise Program	RCT	Surgical management of adult isthmic spondylolisthesis improves function and relieves pain more efficiently than an exercise program.	<i>Positive (fusion more effective than nonsurgical treatment, but results should be accepted with caution.</i>	
Ohtori 2011 Surgical versus nonsurgical treatment of selected patients with discogenic low back pain.	POPULATION/CLINICAL INDICATION <u>Included:</u> Patients with LBP , continuing for at least 2 years, with no accompanying radicular pain. <u>Excluded:</u> Patients who had previously undergone spinal surgery, with LBP after traffic accidents, recipients of workers compensation. INTERVENTION Anterior interbody Fusion and posterolateral fusion with pedicle screws.	RCT	If DLBP is strictly diagnosed, surgical therapy is suitable for its treatment. ABF gives good results, but PLF is an option for patients for anterior surgery is unsuitable.	<i>Positive/Neutral (fusion more effective than nonsurgical treatment, but difference of minimal clinical</i>	

1 st author, year, title	Inclusion, exclusion criteria (for P.I.C.O)	Study design	Conclusion/recommendation	Recommendation category	Other comments
	COMPARATOR Conservative (exercise) treatment OUTCOMES Pain Visual Analog Scale (VAS) Score, Japanese Orthopedic Association Score and Oswestry Disability Index (ODI)			<i>importance)</i>	
Weinstein 2007 Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis.	POPULATION/CLINICAL INDICATION <u>Included:</u> All patients had neurogenic claudication or radicular leg pain with associated neurologic signs, spinal stenosis shown on cross-sectional imaging, and degenerative spondylolisthesis shown on lateral radiographs obtained with the patient in a standing position. The patients had had persistent symptoms for at least 12 weeks and had been confirmed as surgical candidates by their physicians. <u>Excluded:</u> Patients with spondylolysis and isthmic spondylolisthesis INTERVENTION Standard decompressive laminectomy (with or without fusion) COMPARATOR Usual nonsurgical care OUTCOMES Medical Outcomes Study 36-Item Short- Form General Health Survey (SF-36) bodily pain and physical function scores and the modified Oswestry Disability Index.	RCT	In nonrandomised as-treated comparisons with careful control for potentially confounding baseline factors, patients with degenerative spondylolisthesis and spinal stenosis treated surgically showed substantially greater improvement in pain and function during a period of 2 years than patients treated non-surgically.	<i>Positive/Neutral (fusion more effective than nonsurgical treatment, weak evidence)</i>	
Weinstein 2009 Surgical compared with nonoperative treatment for lumbar degenerative spondylolisthesis. Four-year results in the spine patient outcomes research trial (SPORT) randomized and observational cohorts.	<i>Long term follow-up of patients from Weinstein 2007, see Weinstein 2007 for details.</i>	RCT	Conclusions: Compared with patients who are treated nonoperatively, patients in whom degenerative spondylolisthesis and associated spinal stenosis are treated surgically maintain substantially greater pain relief and improvement in function for four years. Overview: In the as-treated analysis, combining the randomised and observational cohorts of patients with spinal stenosis secondary to degenerative spondylolisthesis, those treated surgically were found to have significantly greater improvement in scores for pain, function, satisfaction, and self-rated progress over four years compared with patients treated non-operatively. The results in both groups were stable between two and four years.		

APPENDIX 5: APPRAISAL TABLES FOR KEY STUDIES

Synthesised studies

Table A5.1 Critical appraisal table (Choma 2011).

Study: Choma TJ, Schuster JM, Norvell DC, Dettori JR, Chutkan NB. Fusion Versus Nonoperative Management for Chronic Low Back Pain Do Comorbid Diseases or General Health Factors Affect Outcome? *Spine*. 2011;36(21):S87-S95.

Description of Study: systematic review of 2 randomised controlled trials.

Patient/population	Adults with chronic low back pain (CLBP), centralised or radiating	
N	2 RCTs	
Setting	Not specified	
Intervention/indicator	Reference	Intervention
	Hagg 2003	Spinal fusion
	Fairbank 2005	Spinal fusion
Comparison/control	Reference	Comparison
	Hagg 2003	Nonsurgical care
	Fairbank 2005	Intensive rehabilitation
Outcomes	Subgroup analysis of ODI and patient global assessment scores for smokers vs nonsmokers, and comorbidities vs no comorbidities	
Inclusion criteria	Patients: adults with CLBP, centralised or radiating	
	Interventions: surgical fusion, conservative management	
	Prognostic factors: diseases or symptoms associated with primary diagnosis – comorbidities; general health conditions such as – obesity, smoking, alcohol use, drug use	
	Outcomes: pain, physical function, quality of life, return to work	
Exclusion criteria	Study design: meta-analyses, RCTs, comparative observational studies, registry studies, studies including sub-analyses of subgroups	
	Patients: <18 years old; predominant neurological involvement; predominant spondylolisthesis or stenosis; cancer, deformity, instability, infection, inflammatory arthritis, trauma	
	Interventions: other surgical management strategies	
	Study design: no separate treatment effect for each subgroup of interest; included risk factor regression analysis, but did not test for interaction; case reports; nonclinical studies; case series.	

Study Validity.

Is it clear that there were no conflicts of interest in the writing or funding of this review?	Yes	"AOSpine of North America and Foundation funds were received to support this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript."
Does the review have a clearly- focused question?	Yes	"The objectives of this systematic review were to determine if comorbid disease and general health factors modify the effect of fusion versus nonoperative management in chronic low back pain (CLBP) patients?"
Is a systematic review the appropriate method to answer the question?	Yes	
Does the review have specified inclusion/exclusion criteria?	Yes	<i>See above</i>
If there were specified inclusion/ exclusion criteria, were these appropriate?	Yes	
Does the review document a comprehensive search strategy?	Not reported	Search strategy/terms not reported.
Were reviewers blind to authors, institutions and affiliations?	Not reported	
Were 2 or more independent reviewers used for: 1. application of inclusion criteria to assess eligibility of studies?	Not reported	
2. extraction of data from study reports?	Yes	"Each retrieved citation was reviewed by two independently working reviewers"
3. appraisal of study quality?	Yes	"Level of evidence ratings were assigned to each article independently by two reviewers using criteria set by The Journal of Bone and Joint Surgery, American Volume (J Bone Joint Surg Am) ²⁷ for therapeutic studies and modified to delineate criteria associated with methodological quality and described elsewhere. ²⁸ "
Were the strengths and limitations of included studies and potential impact on the results discussed?	Yes	
Was the validity of included trials appraised using appropriate criteria?	Yes	
Is there a summary of the results of individual studies?	Yes	
If meta-analyses were conducted, was it reasonable to do so?	N/A	Forest plots were generated separately for each of the two included studies, however, "data between studies were not pooled for two primary reasons: (1) we did not identify multiple studies of the same subgroup or (2) outcomes were too heterogeneous to standardize for pooling purposes."

If meta-analyses were conducted, was it done appropriately?	N/A	
What is the overall risk of bias?	Low to Moderate	Low to Moderate - Most of the criteria have been fulfilled and those criteria that have not been fulfilled may affect the conclusions of the study.

Results.

“Study Selection

We identified 127 total citations from our search strategy (Supplemental Digital Content 1, <http://links.lww.com/BRS/A546>). Of these, 93 were excluded by abstract and 34 full text articles were retrieved to determine if they met criteria. From these 34, 10 reported subgroup effects; however, only 5 reported treatment effects (fusion vs. nonoperative management) separately by disease and general health subgroups of interest. Three of these were excluded because they included patients with predominantly neurological involvement, spondylolisthesis, or stenosis (Figure 2).

Do Comorbid Diseases Modify the Treatment Effect of Fusion Versus Nonoperative Management in CLBP Patients?

Only one study was identified in the literature comparing fusion to nonoperative management that met our subject criteria and reported results separately by presence of comorbid disease subgroups for this study question.³⁰ This both highlights the gaps in the literature comparing fusion to nonoperative management in subgroups with CLBP and can only serve to provide hypotheses regarding the possibility of HTE by disease and general health subgroups. In the RCT by Hägg (n = 264 patients; 91 with additional comorbidities) comparing fusion to nonsurgical care, the authors did not specify the additional comorbid diseases. Among the 157 subjects without additional comorbidities, 61% improved (“better or “much better” using the Patient Global Assessment) with fusion 2 years after surgery and 23% “improved” with nonoperative care (Table 1). Among patients with additional comorbidities, 66% improved with fusion and 40% improved with nonoperative care. The RD comparing fusion to nonoperative management in those without additional comorbidities was 38% in favor of fusion and in those with additional comorbidities 26% in favor of fusion. The RD favouring those without additional comorbidities is explained by the difference in the nonoperative group primarily, as the improvement rates in the surgical group are similar. It is unclear why patient with comorbidities would do better with nonoperative care compared with those without comorbidities. Furthermore, the nonoperative group intervention was not well defined. It is possible that this group had more room for improvement, which was appreciated more in the nonoperative group than the fusion group. The author did not report a test for interaction on these treatment effect differences so it is not clear if there is statistical effect modification. Further the CIs overlapped suggesting the difference in treatment effects were not statistically significant (Figure 3).

Do General Health Risk Factors Modify the Effect of Fusion Versus Nonoperative Management in Patients With Chronic Low Back Pain?

Fairbank³¹ and Hägg³⁰ evaluated the effect of smoking on the comparison of fusion to nonoperative management (Table 1). Neither gave details as to what constituted a “smoker” other than patient self-report. In the RCT by Hägg (n = 264), nonsmokers (n = 144) were considered “improved” (“better or “much better” using the Patient Global Assessment) in 66% of the fusion and 26% of the nonoperative groups, 2 years after surgery. The rates in smokers (n = 112) were 58% and 32%, respectively. The RD comparing fusion to conservative management in nonsmokers was 41% in favor of fusion and in smokers 26% in favor of fusion (Table 1). The authors did not report a test for interaction on these treatment effect differences, but in examining raw scores, nonsmokers benefited more from fusion than smokers; however, the confidence interval overlapped suggesting this difference was not statistically significant (Figure 3). In the RCT by Fairbank, (n = 349 patients with CLBP with or without referred pain), comparing fusion to intensive rehabilitation, the ODI (the lower score the greater the function) change scores (from baseline to 2 years follow-up) for nonsmokers (n = 199) were –16.0 and –8.5, respectively (treatment effect = 7.5 in favor of fusion; no P value or standard deviations reported) (Table 1). The change scores for smokers (n = 150) were –7.2 and –8.5, respectively (treatment effect = –1.3 in favour of conservative management; no P value or standard deviations reported) (Table 1). The SMDs comparing surgery to conservative management were 1.7 ± 0.18 and 0.2 ± 0.15 in nonsmokers and smokers, respectively (Table 1). The authors did not report a test for interaction on these treatment effect differences; however, in our calculations of SMDs, the CIs did not overlap suggesting a statistically significant difference that nonsmokers benefited more from fusion than smokers (Figure 4).

Evidence Summary

The overall strength of the evidence evaluating whether specific disease or general health subpopulations modify the effect of fusion *versus* conservative management in the treatment of CLBP is “insufficient,” that is, evidence either is unavailable or does not permit a conclusion; however, some hypotheses can be generated and considered in clinical decision making and in future research planning (Table 2). Detailed data from individual articles evaluated for this manuscript are available in Table 3.”

See original article for Tables and Figures

Author’s Conclusions.

“Evidence Summary

The overall strength of the evidence evaluating whether specific disease or general health subpopulations modify the effect of fusion *versus* conservative management in the treatment of CLBP is “insufficient,” that is, evidence either is unavailable or does not permit a conclusion; however, some hypotheses can be generated and considered in clinical decision making and in future research planning (Table 2). Detailed data from individual articles evaluated for this manuscript are available in Table 3.”

“Key Points

- When comparing surgical fusion to nonoperative management for CLBP, the treatment benefit favouring fusion is greater in nonsmokers than smokers.
- When comparing surgical fusion to nonoperative management for CLBP, the treatment benefit favouring fusion may be slightly larger for those patients with no additional comorbidities.
- Future research designed to determine if comorbid disease and general health subpopulations modify the effect of fusion *versus* conservative management is needed.”

“Conclusion. It is unclear from the literature which patients are the best candidates for fusion *versus* conservative management when experiencing CLBP without significant neurological impairment.

Nonsmokers may be more likely to have a favorable surgical fusion outcome in CLBP patients. Comorbid disease presence has not been shown to definitively modify the effect of fusion. Further prospective studies that are designed to evaluate these and other subgroup effects are encouraged to confirm these findings.

Clinical Recommendations. We recommend optimizing the management of medical co-morbidities and smoking cessation before considering surgical fusion in CLBP patients. Strength of recommendation: Weak”

Our Comments/Summary.

This systematic review has a low to moderate risk of bias. In addition to this, the recommendations of the study are noted to be based on weak evidence, therefore, findings should not be generalised.

The authors of the study note the following: “the limitations in this article include the small number of studies identified meeting our study criteria, which limited our study power. However since subgroup analyses of secondary data are more appropriately considered hypothesis generating, we erred on the more focused side with respect to treatment and patient populations. We feel our findings are more generalizable and provide evidence that the literature is significantly limited in this area.”

Table A5.2 Critical appraisal table (Chou 2009).

Study: Chou R, Baisden J, Carragee EJ, Resnick DK, Shaffer WO, Loeser JD. Surgery for Low Back Pain A Review of the Evidence for an American Pain Society Clinical Practice Guideline. *Spine*. 2009; 34(10):1094-109

Description of Study: systematic review of 4 randomised controlled trials.

Patient/population	Adult patients with nonradicular back pain with common degenerative changes.	
N	4 RCTs Please note all these studies may not have assessed our outcomes of interest.	
Setting	Multicenter trials	
Intervention/indicator	Reference	Intervention
	Brox 2003	Instrumented Posterolateral Fusion
	Brox 2006	Instrumented Posterolateral Fusion
	Fairbank 2005	Graf ligamentoplasty (15%) or fusion with technique left to the discretion of the surgeon.
	Fritzell 2001	Noninstrumented posterolateral fusion (1/3), instrumented posterolateral fusion (1/3) or instrumented circumferential fusion (1/3).
Comparison/control	Reference	Comparison
	Brox 2003	Intensive rehabilitation with a cognitive-behavioural component.
	Brox 2006	Intensive rehabilitation with a cognitive behavioural component.
	Fairbank 2005	Intensive rehabilitation with a cognitive behavioural component.
	Fritzell 2001	Nonintensive physical therapy.
Outcomes	ODI, back pain, VAS, SF-36,	
Inclusion Criteria	RCT's and SR's written in English, evaluating nonpregnant patients, evaluating surgery for nonradicular low back pain reporting on at least one of the following outcomes; back specific function, generic health status, pain, work disability or patient satisfaction.	
Exclusion Criteria	Trials of surgery for low back pain associated with acute major trauma, cancer, infection, caudaequina syndrome, and osteoporosis or vertebral compression fracture. Outdated systematic reviews (published before 2000).	

Study Validity.

Is it clear that there were no conflicts of interest in the writing or funding of this review?	Yes	.
Does the review have a clearly- focused question?	Yes	
Is a systematic review the appropriate method to answer the question?	Yes	

Does the review have specified inclusion/exclusion criteria?	Yes	"Inclusion/exclusion criteria listed above"
If there were specified inclusion/ exclusion criteria, were these appropriate?	Yes	
Does the review document a comprehensive search strategy?	Yes	
Were reviewers blind to authors, institutions and affiliations?	Not reported	
Were 2 or more independent reviewers used for:	Not reported	
1. application of inclusion criteria to assess eligibility of studies?		
2. extraction of data from study reports?	Yes	
3. appraisal of study quality?	Yes	
Were the strengths and limitations of included studies and potential impact on the results discussed?	Yes	
Was the validity of included trials appraised using appropriate criteria?	Yes	
Is there a summary of the results of individual studies?	Yes	
If meta-analyses were conducted, was it reasonable to do so?	N/A	
If meta-analyses were conducted, was it done appropriately?	N/A	
What is the overall risk of bias?	Low	

Results.

"Four higher-quality trials of fusion surgery *versus* nonsurgical therapy evaluated patients with moderately severe pain (mean score: 63–65 on a 0–100 scale^{54,55,57}) or disability (mean ODI score: 4556) for at least 1 year, unresponsive to standard nonsurgical therapy. Positive results on provocative discography were not required for enrollment in any trial. Exclusion criteria included significant psychiatric or somatic illness, ongoing compensation issues or presence of other chronic pain conditions. Surgical techniques involved some type of fusion procedure, though specific methods varied (Table 1).

The trials reported inconsistent results.^{54–57} In the Swedish Lumbar Spine Study (N = 294), independent assessors rated outcomes as "excellent" or "good" (no more than sporadic pain, slight restriction of function, and occasional analgesics) in 46% of those randomized to surgery *versus* 18% randomized to nonsurgical therapy after 2 years ($P < 0.0001$).⁵⁷ More of the surgical patients rated results as "better" or "much better" (63% vs. 29%, $P < 0.0001$). Patients randomized to surgery also experienced moderately greater improvements in pain (mean change from baseline on 0 to 100 VAS pain score 21.0 vs. 4.3, $P = 0.0002$) and slightly greater improvements in ODI scores (mean change from baseline 11.6 vs. 2.8, $P = 0.015$), and a higher proportion returned to work (36% vs. 13%, $P = 0.002$). Two smaller (N=60 and 64) trials conducted by the same Norwegian investigators found no statistically significant differences between surgery *versus* nonsurgical therapy on any of the main outcomes after 1 year among patients either with⁵⁴ or without⁵⁵ prior discectomy. However, in the latter trial surgery was associated with a trend toward slightly superior outcomes on the ODI (mean difference = -7.3, 95% CI = -17.3 to +2.7) and back pain scores (mean difference = -5.2, 95% CI = -18.0- +7.6).⁵⁴ The Medical Research Council Spine Stabilization Trial (N = 349) found surgery associated with statistically significant improvements in ODI scores after 24 months compared to nonsurgical therapy, but the difference did not reach clinical significance (mean difference: -4.1, 95% CI: -9.1 to -0.1, $P = 0.045$).⁵⁶ There were no differences in other outcomes, including short-form 36 (SF-36) scores and the shuttle walking test.

The inconsistent results between trials could be related to differences in nonsurgical comparator treatments. In the 3 trials that found clinically or statistically insignificant benefits following surgery,

nonsurgical treatment consisted of intensive rehabilitation incorporating cognitive behavioral therapy (75 hours over 3 weeks, with subsequent follow-up visits).^{54–56} In the 1 trial showing surgery associated with clinically and statistically significant benefits, the nonsurgical treatment intervention was less intensive (70 hours of supervised physical therapy over a 2-year-period) and more heterogeneous (could be supplemented by other interventions such as transcutaneous electrical nerve stimulation, acupuncture, injections, advice, and cognitive therapy).⁵⁷ In addition, one of the criteria for enrolment in this trial was inadequate response to nonsurgical treatment, but patients randomized to the nonsurgical arm may have continued to receive previously ineffective interventions.

Two higher-quality systematic reviews also found inconsistent results for surgery *versus* no surgery that could be explained by the nonsurgical comparator intervention.^{22,23,28} Another higher-quality, quantitative systematic review found no difference between surgery and nonsurgical therapy when data from 3 trials^{55–57} were pooled (-4.13 , 95% CI = -9.08 – 0.82), but heterogeneity was present, in part because trials of intensive and standard rehabilitation were combined.²⁶ Two lower quality systematic reviews estimated success rates of 67% to 79% following fusion, but pooled data across primarily uncontrolled observational studies.^{16,17} A third systematic review³⁴ postulated that lack of efficacy observed in smaller ($N < 100$) trials could have been due to small sample sizes and insufficient power to detect differences. However, even if statistically significant, results in the smaller trials would either favor nonsurgical therapy (2.3 points on the ODI55) or would only slightly favor surgery (7.3 points on the ODI54).

There is insufficient evidence to determine optimal fusion methods. Instrumentation and electrical stimulation appear to enhance fusion rates, but effects on clinical outcomes are not established.^{22,23} Although pooled estimates in a higher-quality Cochrane review found instrumentation superior to no instrumentation (OR = 0.49 for poor clinical outcome, 95% CI = 0.28 – 0.84), results are sensitive to inclusion of 2 older, lower-quality outlier trials (1 nonrandomized⁹⁶) reporting unusually favourable results (83%⁹⁶ and 93%⁹⁴ success with instrumented fusion). After limiting the pooled analysis to higher-quality trials published since 1997, effects of fusion were marginal (74% vs. 68% rates of clinical success). There are conflicting results from head-to-head trials regarding relative effectiveness of various types of fusion (anterior, posterior, or combined).^{22,23}

No operative deaths were reported in randomized trials of fusion *versus* nonsurgical therapy.^{54–57} The pooled rate of early surgical complications from 3 trials^{55–57} was 16% (95% CI = 12%–20%).²⁶ Major complications included deep wound infections, major bleeding during surgery, thrombosis, acute respiratory distress syndrome, pulmonary edema, and heart failure. One trial, which evaluated different fusion techniques, found higher risks of complications with more technically difficult procedures.⁹⁷ The total complication rate after 2 years was 12% with noninstrumented posterolateral fusion, 22% with instrumented posterolateral fusion, and 40% with circumferential fusion. A recent, large observational study based on the Nationwide Inpatient Sample reported <1% in-hospital mortality for all fusion procedures.⁹⁸ In systematic reviews including observational studies, complication rates following fusion vary widely and are difficult to interpret due to differences in techniques, study populations, and methodologic shortcomings.^{16,95}

The review also reported on artificial disc replacement versus fusion and disectomy versus nonsurgical therapy; however we are only comparing fusion to non-surgical interventions. In addition to nonradicular low back pain the study also assessed the benefit of surgery for treatment of Symptomatic spinal stenosis, with or without degenerative spondylolisthesis.

See original article for Tables and Figures

Author's Conclusions.

“For nonradicular low back pain with common degenerative changes, there is fair evidence from randomized trials that fusion is no more effective than intensive rehabilitation with a cognitive behavioural emphasis, but slightly to moderately more effective than standard (nonintensive) nonsurgical therapy for improvement in pain and function.”

Our Comments/Summary.

Low Risk of Bias

Table A5.3 Critical appraisal table (Daubs 2011).

Study: Daubs MD, Norvell DC, McGuire R, Molinari R, Hermsmeyer JT, Fourney DR, et al. Fusion Versus Nonoperative Care for Chronic Low Back Pain Do Psychological Factors Affect Outcomes? *Spine*. 2011;36(21):S96-S109.

Description of Study: systematic review of 1 randomised controlled trial.

Patient/population	Adults with chronic low back pain (CLBP), centralised or radiating	
N	1 RCT	
Setting	Not specified	
Intervention/indicator	Reference	Intervention
	Hagg	Spinal fusion
Comparison/control	Reference	Comparison
	Hagg	Conservative treatment
Outcomes	Pain, physical function, quality of life	
Inclusion Criteria	Patients: adults with CLBP, centralised or radiating	
	Prognostic factors: adverse mental health predictors: depression; stress/anxiety; personality disorders; screening tests	
	Outcomes: pain, physical function, quality of life	
Exclusion Criteria	Study design: meta-analyses, RCTs, comparative observational studies, registry studies, studies including sub-analyses of risk factors	
	Patients: <18 years old; predominant neurological involvement; predominant spondylolisthesis or stenosis; cancer, deformity, instability, infection, trauma, Outcomes: cost effectiveness	
	Study design: no separate treatment effect for each subgroup of interest; included risk factor regression analysis, but did not test for interaction; case reports; non-clinical studies; case series.	

Study Validity.

Is it clear that there were no conflicts of interest in the writing or funding of this review?	Yes	"Professional Organization and Foundation funds were received to support this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript."
Does the review have a clearly- focused question?	Yes	"The objectives of this systematic review were to determine whether fusion is superior to conservative management in certain psychological subpopulations and to determine the most common psychological screening tests and their ability to predict outcome after treatment in patients with chronic lower back pain."
Is a systematic review the appropriate method to answer the question?	Yes	
Does the review have specified inclusion/exclusion criteria?	Yes	See above

If there were specified inclusion/ exclusion criteria, were these appropriate?	Yes	No mention of intervention or comparators in the inclusion criteria, but these are mentioned in the title and objectives
Does the review document a comprehensive search strategy?	Not reported	Search strategy/terms not reported
Were reviewers blind to authors, institutions and affiliations?	Not reported	
Were 2 or more independent reviewers used for: 1. application of inclusion criteria to assess eligibility of studies?	Yes	"Each retrieved citation was reviewed by two independently working reviewers (D.C.N. and E.E.). Some articles were excluded on the basis of information provided by the title or abstract if they clearly were not appropriate. Citations that appeared to be appropriate or those that could not be excluded unequivocally from the title and abstract were identified, and the corresponding full-text reports were reviewed by the two reviewers. Any disagreement between them was resolved by consensus."
2. extraction of data from study reports?	Not reported	
3. appraisal of study quality?	Yes	"level of evidence ratings were assigned to each article independently by two reviewers using criteria set by The Journal of Bone and Joint Surgery, American Volume ¹⁹ to delineate criteria associated with risk of bias and methodological quality described elsewhere." ²⁰
Were the strengths and limitations of included studies and potential impact on the results discussed?	Yes	
Was the validity of included trials appraised using appropriate criteria?	Yes	
Is there a summary of the results of individual studies?	Yes	
If meta-analyses were conducted, was it reasonable to do so?	N/A	A forest plot was generated, but this was for the purpose of subgroup analysis rather than metaanalysis
If meta-analyses were conducted, was it done appropriately?	N/A	
What is the overall risk of bias?	Low to Moderate	Low to Moderate - Most of the criteria have been fulfilled and those criteria that have not been fulfilled may affect the conclusions of the study.

Results.

"Is Fusion Superior to Conservative Management in Certain Psychological Subpopulations?" Only one study was identified meeting our study criteria that compared outcomes by treatment group stratified by a psychological subgroup. This highlights the limitations of the literature comparing fusion to conservative management in psychosocial subgroups with chronic LBP and can only serve to provide hypotheses regarding the possibility of treatment effect heterogeneity by psychological factors. In the RCT by Hägg (n = 264 patients with severe chronic LBP) comparing fusion to nonsurgical care (Table 1), several psychological assessments were performed on patients undergoing both treatments to include personality traits (neuroticism, aggressiveness, social introversion, and impulsiveness) using the Karolinska Scales of Personality, existence of a personality disorder (cluster A, B, or C) using the Swedish version of the Structured Clinical Interview for *Diagnostic and Statistical Manual of Mental Disorders* (Third Edition

Revised), and depressive symptoms using the Zung Depressive Scale. Eleven percent of fusion patients *with a personality disorder* were considered “improved” (“better” or “much better” using the Patient Global Assessment) *versus* 17% of conservatively managed patients 2 years after surgery (Table 2). In contrast, 18% of fusion patients *without a personality disorder* and 8% who were conservatively managed were considered “improved” 2 years after surgery. The risk difference comparing fusion to conservative management in those with a *personality disorder* was – 6% (RD = –0.06; 95% CI, –0.39 to 0.27) in favor of conservative management. The risk difference for patients *without a personality disorder* was 10% (RD = 0.10; 95% CI, –0.003 to 0.21) in favor of fusion (Table 2). The authors did not report a test for interaction on these treatment effect differences, but in examining raw scores, those *with a personality disorder* benefited more from conservative management and those *without a personality disorder* benefited more from fusion; however, the confidence intervals overlapped likely because of the small sample size in the personality disorder group (Figure 3).

The same authors also reported baseline psychological measure scores for each treatment arm by those who “improved” and those who “did not improve” using the Patient Global Assessment as the outcome. The mean baseline depressive symptom score (Zung Depression Scale) (the higher the score the higher level of depression; major depression > 58) for those designated “improved” was 39.0 ± 13.4 points for the fusion group and 48.0 ± 11.3 points for the nonoperative group (Table 3). This difference was statistically significant ($P = 0.009$). The baseline differences for those “not improved” were similar between fusion and nonsurgical groups (Table 3). The mean difference comparing fusion to nonoperative groups in baseline depression scores among those who were designated “improved” was 9.0 points (95% CI: 2.3–5.7) and in those designated “not improved” 1.0 point (95% CI: –4.0 to 6.0) (Table 3). In other words, patients who improved with nonoperative treatment had higher levels of depression at baseline than those that improved with fusion. This may suggest that patients with higher levels of depression have better outcomes with nonoperative treatment.

There were similar results with the presence of a neurotic personality trait as determined by the Karolinska Scales of Personality and translated by the authors as a person who is “tense and stiff, restless, uneasy, panicky, easily fatigued, remorseful, experiencing tremor and palpitations under stress.” The mean difference comparing nonoperative management to fusion in baseline neuroticism scores among those who were designated “improved” was 6.4 points (95% CI: 2.1–10.7) and in those designated “not improved” –0.9 points (95% CI: – 4.5 to 2.7) (Table 3). This may suggest again that patients with higher baseline neuroticism scores benefit more from nonoperative management than fusion. There was little difference in baseline scores between nonoperative and fusion groups in those who did not improve.

This paper goes on to discuss various psychological screening tests and measures

Evidence Summary: The overall strength of the evidence evaluating whether specific *psychological* subpopulations modify the effect of fusion *versus* conservative management in the treatment of chronic LBP is “insufficient,” that is, evidence either is unavailable or does not permit a conclusion; however, some hypotheses can be generated and considered in clinical decision making and in future research planning (Table 6).”

See original article for Tables and Figures

Author’s Conclusions.

“Key Points: Very few randomized control trials evaluating the treatment of chronic lower back pain have evaluated the effects of psychological factors on outcomes by treatment intervention.”

“Conclusion. Psychological disorders affect chronic lower back pain treatment outcomes. Patients with a personality disorder appear to respond more favorably to conservative management and those without a personality disorder more favorably to fusion. Patients with higher depression and neuroticism scores may also respond more favorably to conservative management.

Clinical Recommendations.

Recommendation 1: Chronic LBP patients with depression, neuroticism, and certain personality disorders should preferentially be treated nonoperatively. Strength of recommendation: Weak.

Recommendation 2: Consider the use of a validated psychological screening questionnaire such as the BDI, FABQ, DRAM, ZDI or STAI, when treating patients with CLBP. Strength of recommendation: Weak.”

Our Comments/Summary.

This systematic review has a low to moderate risk of bias. In addition to this, the recommendations of the study are noted to be based on weak evidence, therefore, findings should not be generalised.

The authors of the study note the following: “despite finding that patients with a personality disorder may respond more favourably to conservative management and those without such a disorder more favorably to fusion, one study is not enough to make treatment recommendations, especially since subgroup analyses of secondary data are more appropriately considered hypothesis generating.”

Table A5.4 Critical appraisal table (Gibson 2008).

Study: Gibson JN, Waddell G. Surgery for degenerative lumbar spondylosis. *Cochrane Database Syst Rev.* 2008(2):CD001352.

Description of Study: systematic review of 3 randomised controlled trials.

Patient/population	Patients over 18 years of age with degenerative lumbar spondylosis treated by surgery.	
N	31 RCTs were included in this review but not all were relevant to our intervention.	
Setting	Trials conducted with inpatients and outpatients.	
Intervention/indicator	Reference	Intervention
	Fritzell 2001	Surgical a) Posterolateral fusion b) Instrumented posterolateral fusion (Steffee system) c) Interbody (ALIF or PLIF (autogenous graft) + b)
	Brox 2003	Posterolateral instrumented fusion (pedicle systems)
	Brox 2004	Posterolateral instrumented fusion (pedicle systems)
Comparison/control	Reference	Comparison
	Fritzell 2001	Non-surgical treatment
	Brox 2003	Cognitive intervention / exercises
	Brox 2004	Modern rehabilitation programme
Outcomes	ODI, back pain, work status and patient rating.	
Inclusion Criteria	Randomised (RCTs) or quasi-randomised trials of surgical treatment of lumbar spondylosis.	
Exclusion Criteria	Not specified.	

Study Validity.

Is it clear that there were no conflicts of interest in the writing or funding of this review?	Yes	There were no conflicts of interest declared.
Does the review have a clearly- focused question?	Yes	To assess the current scientific evidence on the effectiveness of surgical interventions for degenerative lumbar spondylosis.
Is a systematic review the appropriate method to answer the question?	Yes	
Does the review have specified inclusion/exclusion criteria?	Partial	Exclusion criteria were not defined.
If there were specified inclusion/ exclusion criteria, were these appropriate?	Partial	
Does the review document a comprehensive search strategy?	Yes	.
Were reviewers blind to authors, institutions and affiliations?	Not reported	

Were 2 or more independent reviewers used for: 1. application of inclusion criteria to assess eligibility of studies?	Yes	
2. extraction of data from study reports?	Not reported	
3. appraisal of study quality?	Yes	
Were the strengths and limitations of included studies and potential impact on the results discussed?	Yes	
Was the validity of included trials appraised using appropriate criteria?	Yes	
Is there a summary of the results of individual studies?	Yes	
If meta-analyses were conducted, was it reasonable to do so?	N/A	
If meta-analyses were conducted, was it done appropriately?	N/A	
What is the overall risk of bias?	Low	

Results.

“Effects of interventions: Data from thirty-one RCTs of all forms of surgical treatment for degenerative lumbar spondylosis are included in this updated review. In the first edition of this review nine of the 16 trials identified were found on MEDLINE, four from personal bibliographies and four from abstracts of meeting proceedings. The new trials were mainly collected by the authors from personal literature review or after notification by colleagues of the Cochrane Back Review Group. Three trials originally included have now been deleted from the review (see Characteristics of Excluded Trials table) as originally, they were abstracts of work in progress and no data have been published over the intervening years (Emery 1995; Rogozinski 1995; Zdeblick 1996). Six further trials are included as ongoing studies. The majority of the trials compared two or more surgical techniques. From a surgical perspective, the trials now fall into three broad sections: 1) surgical treatment (decompression with or without fusion) for spinal stenosis and / or nerve root compression 2) surgical treatment (fusion, intra-discal electrotherapy or disc arthroplasty) for back pain 3) comparison of different techniques of spinal fusion. In the first section, one trial compared surgical treatment with conservative therapy and one compared different techniques of decompression for spinal stenosis. Three trials compared decompression alone with decompression and some form of fusion. One trial compared outcomes following use of an interspinous spacer with those after a non-operative regime, including epidural injection. A further two trials of surgery for isthmic spondylolisthesis were included. The second section included two trials of fusion to relieve discogenic back pain compared with different forms of conservative treatment, and preliminary results from three small trials of intra-discal electrotherapy (IDET) and two trials of disc arthroplasty. In the third section, 15 trials considered the role of instrumentation in fusion and four trials that of electrical stimulation (direct current and pulsed electromagnetic stimulation) in postero-lateral fusion. Five trials included sub-groups of participants and are included in more than one section. Analysis of the included trials is complicated by the inclusion of participants with varied pathology and a lack of consistency in treatment methods. Only five of the trials (Moller 2000; Amundsen 2000; Fritzell 2001; Brox 2003; Brox 2004) had a conservative treatment arm. It was not possible to analyze participants according to duration of their symptoms, type of previous conservative treatment, or indications for surgery, as few of the trials provided these data in usable form. Although many trials provided limited information on selected complications, these were not comparable between trials. Three trials provided comparative information on operating time and blood loss, and three trials provided information on progression of spondylolisthesis. No other adverse effects could be reviewed. A cost analysis was performed in one trial (Fritzell 2001), although the methodological criticisms by Goosens (Goosens 1998) should be noted. ...

2. Surgery for back pain without neurological compromise: At the time of the original Cochrane review of degenerative lumbar spondylosis (1999) there were no published RCTs on the effectiveness of fusion for chronic back pain, compared with natural history, conservative treatment or placebo. There are now two new trials. The Swedish trial of lumbar fusion versus physiotherapy treatment for chronic low back pain (Fritzell 2001) included 294 individuals presenting at 19 spinal centres over a six-year period. Strict inclusion criteria limited trial entry to those who had low back pain more pronounced than leg pain, lasting longer than two years, and no evidence of nerve root compression. Each patient had to have completed a course of conservative treatment that had failed to produce relief. Nineteen per cent had

previous surgery. Individuals were randomised into four treatment groups. Seventy-two patients had conservative treatment and 222 had one of three different fusion techniques. There was a 98% follow-up at two years. Twenty-five subjects did not complete treatment according to random allocation, but these 'group changers' were included in the original 'intention-to-treat' analysis. At two years, independent assessors rated 46% of the surgical group as 'excellent' or 'good', compared with 18% of the conservative group ($P < 0.0001$). More surgical patients rated their results as 'better' or 'much better' (63% versus 29%, $P < 0.0001$). The surgical patients had significantly greater improvement in pain (visual analogue scale) and disability (Oswestry scale). The "net back to work rate" was significantly in favour of surgical treatment (36% versus 13%, P value 0.002). There were no significant differences in any of these outcomes between the three surgical groups. The Swedish trial also provided one of the few cost-effective analyses of spinal surgical treatment. The cost differences between the surgical and conservative groups were significant, mainly because more individuals went back to work in the surgical group (Fritzell 2001).

The major question about the Swedish trial was the nature of the conservative treatment used as the control intervention (Mooney 1990). The investigators tried to ensure that each patient understood that "no treatment method, as far as was known, was superior to any other". Nevertheless, the control group essentially received more of the same 'usual non-surgical treatment' that had already failed, and the failure of which was one of the indications leading to consideration of surgery. In view of the likely negative patient expectations, it is hardly surprising that the results in the control group appear to have been poorer than most epidemiological studies of natural history. Strictly speaking, this trial provided the first substantive evidence that fusion is more effective than continued, standard 1990s, 'usual care'.

The Norwegian trial (Brox 2003; Brox 2004) compared posterolateral fusion with transpedicular screws and post-operative physiotherapy versus a modern 'rehabilitation' type of programme, consisting of an educational intervention (Indahl 1995) and a three-week course of intensive exercise sessions, based on cognitive-behavioural principles. Sixty-four patients with low back pain lasting longer than one year plus disc degeneration at L4/5, L5/S1 or both (Brox 2003), and a further 60 patients with chronic low back pain more than one year after previous discectomy (Brox 2004) were randomised and reported on separately. There was a 97% follow-up at one year and intention-to-treat analysis. In both series, there were no significant differences in any of the main outcomes of independent observer rating, patient rating, pain, disability or return to work. Radiating leg pain improved significantly more after surgery, whereas fear avoidance beliefs and forward flexion improved significantly more after conservative management. At one-year follow-up, the conservative groups had significantly better muscle strength and endurance (Keller 2004). Despite the relatively small size of these trials (though the number randomized to conservative treatment is comparable to the Swedish trial, 57 compared to 72), the consistent results in both first time and previously failed surgical patients and lack of any trends make a Type II error unlikely. In contrast to the Swedish trial, these results suggest that fusion and a modern rehabilitation approach can produce comparable outcomes. ...

There are now two trials on the effectiveness of fusion compared with conservative treatment. The first (Swedish) trial (Fritzell 2001) appeared to provide strong evidence in favour of fusion, but the more recent (Norwegian) trial (Brox 2003; Brox 2004) refutes this. The difference may lie in the treatment given to the control group. Fusion is more effective than continued, failed, standard 1990s, 'usual care'; it does not appear to be anymore effective than a modern rehabilitation programme. Clearly, there are still open questions about the scientific evidence on the clinical effectiveness of fusion. Further evidence is required, which hopefully will be provided by the multi-centred RCTs of fusion that are presently underway in the US and the UK."

See original publication for tables and figures.

Author's Conclusions.

Fusion is more effective than continued, failed, standard 1990s 'usual care'; it does not appear to be any more effective than a modern rehabilitation program.

Our Comments/Summary.

Low risk of bias.

Table A5.5 Critical appraisal table (Ibrahim 2008).

Study: Ibrahim T, Tleyjeh IM, Gabbar O. Surgical versus non-surgical treatment of chronic low back pain: a meta-analysis of randomised trials. *International Orthopaedics*. 2008; 32(1): 107-13 & erratum
Ibrahim T, Tleyjeh IM, Gabbar O. Surgical versus non-surgical treatment of chronic low back pain: a meta-analysis of randomised trials. *International Orthopaedics*. 2009; 33: 589-90

Description of Study: systematic review of 4 randomised controlled trials.

Patient/population	Patients undergoing surgical and non-surgical treatment for low back pain.	
N	4 studies Please note all these studies may not have assessed our outcomes of interest.	
Setting	Not Specified	
Intervention/indicator	Reference	Intervention
	Fairbank et al 2005	Spinal Fusion with some cases of flexible stabilisation
	Fritzell et al 2001	Posterolateral fusion (PLF) or PLF with screws or circumferential fusion
	Brox et al 2003	PLF with pedicle screws
	Ekman et al 2005	PLF with or without pedicle screws.
Comparison/control	Reference	Comparison
	Fairbank et al 2005	Spinal Fusion with some cases of flexible stabilisation
	Fritzell et al 2001	Posterolateral fusion (PLF) or PLF with screws or circumferential fusion
	Brox et al 2003	PLF with pedicle screws
	Ekman et al 2005	PLF with or without pedicle screws.
Outcomes	Oswestry Disability Index (ODI)	
Inclusion Criteria	Only RCT's comparing surgical and non-surgical treatments for chronic low back pain were considered for inclusion. The trials must have reported ODI as an outcome measure.	
Exclusion Criteria	RCTs with unique primary outcomes where the pooling of results was not possible.	

Study Validity.

Is it clear that there were no conflicts of interest in the writing or funding of this review?	Not reported	
Does the review have a clearly- focused question?	Yes	To compare the effectiveness of surgical fusion for the treatment of chronic low back pain compared to non-surgical intervention.
Is a systematic review the appropriate method to answer the question?	Yes	
Does the review have specified inclusion/exclusion criteria?	Yes	

If there were specified inclusion/ exclusion criteria, were these appropriate?	Yes	
Does the review document a comprehensive search strategy?	Yes	
Were reviewers blind to authors, institutions and affiliations?	Not reported	
Were 2 or more independent reviewers used for:	Yes	
1. application of inclusion criteria to assess eligibility of studies?		
2. extraction of data from study reports?	Yes	
3. appraisal of study quality?	Yes	
Were the strengths and limitations of included studies and potential impact on the results discussed?	Yes	
Was the validity of included trials appraised using appropriate criteria?	Yes	
Is there a summary of the results of individual studies?	Yes	
If meta-analyses were conducted, was it reasonable to do so?	Yes	
If meta-analyses were conducted, was it done appropriately?	Yes	
What is the overall risk of bias?	Low	

Results.

“Quantitative results of the meta-analysis: Figure 1 displays the cumulative meta-analytic comparison. The mean overall difference in the ODI between the surgical and non-surgical groups was –4.87 in favour of surgery and statistically significant (95% CI: –1.62 to –8.12, $p=0.0003$). There was no evidence of heterogeneity between the studies, and the I^2 was 0%.

Results of sensitivity analyses: Sensitivity analyses were performed that included the study by Ekman et al. [2]. This study recruited patients with chronic low back pain caused by isthmic spondylolisthesis and was the only study to do so. Analyses including this study, showed a mean overall difference in the ODI of –4.53 (95% CI: –1.48 to –7.59, $p=0.004$, $I^2=0\%$) in favour of surgery (Fig. 2).”

See original publications for tables and figures

Author’s Conclusions.

“We found that surgical fusion may improve the ODI compared to non-surgical intervention at the two-years follow-up for chronic low back pain. This improvement in the ODI compared to the original article (–4.87 compared to –4.13) was statistically significant and is of minimal clinical importance; consequently, surgeons should recommend spinal fusion cautiously to patients with chronic low back pain.”

Our Comments/Summary.

Low risk of Bias.

Table A5.6 Critical appraisal table (Mirza 2007).

Study: Mirza SK, Deyo RA. Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain. *Spine*. 2007;32(7):816-23.

Description of Study: systematic review of 4 randomised controlled trials.

Patient/population	Adult patients with symptoms of low back pain for 12 months or longer without a specific diagnosis.		
N	4 RCTs Please note all these studies may not have assessed our outcomes of interest.		
Setting	Multicenter trials conducted in Europe.		
Intervention/indicator	Reference	Intervention	
	Fritzell	Fusion(posterolateral fusion with iliac crest autograft and no fixation)	
		Fusion (posterolateral fusion with iliac crest autograft and pedicle screw fixation)	
		Fusion (circumfrential fusion consisting of posterolateral fusion and fixation supplemented by interbody fusion using autogenous iliac crest bone block inserted anteriorly or posteriorly)	
	Brox 2003	Fusion (specifically Posterolateral fusion using iliac crest autograft)	
	Brox 2006	Fusion (specifically Posterolateral fusion using iliac crest autograft)	
	Fairbank	Fusion (surgical approach, impant, interbody cages and bone graft material chosen by surgeon)	
Comparison/control	Reference	Comparison	
	Fritzell	Routine non operative care (main component physical therapy, supplemented with education, information and treatment aimed at pain relief)	
	Brox 2003	Structured nonsurgical intervention (cognitive and behavioural treatment program)	
	Brox 2006	Structured nonsurgical intervention (cognitive and behavioural treatment program)	
	Fairbank	Structured nonsurgical intervention (cognitive and behavioural treatment program)	
Outcomes	Back-specific disability, pain, general function, psychological function, work status, radiographic results, complications and patient satisfaction, ODI		
Inclusion Criteria	Trials comparing surgical to nonsurgical treatment for discogenic back pain.		
Exclusion Criteria	Studies focussed on patients with isthmic spondylolisthesis.		
Study Validity.			
Is it clear that there were no conflicts of interest in the writing or funding of this review?	Unsure	“Although one or more of the author(s)has/have received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this manuscript, benefits will be directed solely to a research fund, foundation, educational institution, or other nonprofit organization which the author(s) have been associated.	

Does the review have a clearly- focused question?	Partial	Study and design objective “comparing surgical to non-surgical treatment for chronic low back pain associated with disc degeneration”
Is a systematic review the appropriate method to answer the question?	Yes	
Does the review have specified inclusion/exclusion criteria?	Partial	It is not obvious from the study what the exclusion criteria is.
If there were specified inclusion/ exclusion criteria, were these appropriate?	Partial	
Does the review document a comprehensive search strategy?	Yes	
Were reviewers blind to authors, institutions and affiliations?	Not reported	
Were 2 or more independent reviewers used for:	Not reported	
1. application of inclusion criteria to assess eligibility of studies?		
2. extraction of data from study reports?	Not reported	
3. appraisal of study quality?	Not reported	
Were the strengths and limitations of included studies and potential impact on the results discussed?	Yes	
Was the validity of included trials appraised using appropriate criteria?	Yes	“we also assessed quality using a checklist for the Consolidated Standards of Reporting Trials (CONSORT)” It was not clear however, if more than one assessor was used.
Is there a summary of the results of individual studies?	Yes	
If meta-analyses were conducted, was it reasonable to do so?	N/A	
If meta-analyses were conducted, was it done appropriately?	N/A	
What is the overall risk of bias?	Low/Moderate	

Results.

“Outcomes

All 4 studies showed similar functional improvement for the surgical arm: 8.9–15.6 points in the ODI score; 19% to 37% of the baseline score (Table 4). Although the improvement in the ODI score with surgery in all trials was greater than the 4.0-point threshold for the minimal clinically important difference, the magnitude of improvement exceeded the FDA threshold of 15 points only in the Brox et al¹ 2003 study. Improvement in the nonsurgical arm ranged from 2.8 to 13.3 points in the ODI score, 5.8% to 30.1% of the baseline score. This change was below the minimal clinically important change threshold of 4.0 points in the Fritzell et al³ trial, which did not specify nonsurgical treatment. In all 3 trials that used a structured rehabilitation program incorporating cognitive-behavior therapy, improvement in the ODI score was

similar in magnitude to the improvement seen with surgery (Table 4).

The Fritzell et al³ study showed an advantage for fusion in the primary outcome and all secondary outcomes at 2 years (Tables 4, 5). When comparing the magnitude of improvement in the surgical and nonsurgical groups, the difference between the mean changes in the surgical and nonsurgical groups (the difference of the differences:

Δ ODIsurgery - Δ ODInonop) at 2 years was 8.8 points (Table 4). The Fairbank et al² study also observed slightly greater improvement for surgery at 2 years: Δ ODIsurgery - Δ ODInonop was 4.1 (95% confidence interval [CI] 0.1– 8.1); after imputation for missing data, this difference was 4.5 (95% CI 0.8–8.2). The authors did not consider this difference clinically important, and no significant difference was observed across all other outcomes.²

Both Brox et al^{1,4} studies observed no clinically meaningful difference between surgical to nonsurgical treatment at 1 year for all outcomes listed in Table 3.

In patients with no prior surgery, Δ ODIsurgery - Δ ODInonop was 2.3 (95% CI - 6.8 to 11.4), 2.6 (- 6.5 to 11.7) after adjustment for gender, and 2.7 (-6.8 to 12.2) after adjustment for gender and pretreatment beliefs.¹ In patients with prior discectomy, Δ ODIsurgery - Δ ODInonop was -3.9 (greater improvement in the nonsurgical group); it was -7.3 (95% CI -17.3 to 2.7) after adjustment for gender and - 9.7 (95% CI - 21.7 to 1.7) after adjustment for gender and treatment expectations.⁴

Complications from surgery ranged from 9% to 18% (17%,³ 18%,¹ 11%,² and 9%⁴) (Table 5).

Study Power

All 4 studies presented sample size calculations. The Fritzell et al³ study calculated sample size based on “essential improvement,” defined as the proportion of patients expected to rate themselves as “much better” or “better” at 2 years. They set a sample size at 300 for 80% power to detect a 13% difference between groups (18% of the surgical patients rating themselves in the “essential improvement” category vs. 5% of the nonsurgical group). The difference they actually observed in this outcome was much larger: 62.6% of the surgical versus 29.0% of the nonsurgical group. A difference of this magnitude would have been detectable in a much smaller study.

Both Brox et al^{1,4} studies calculated sample size for a group difference in the ODI (Δ ODI) of greater than 10 points, with standard deviation (SD) of 10, and determined that 26 subjects in each group were sufficient for 95% power. The reference both studies cite for the SD is a Norwegian publication not accessible through MEDLINE. Since the CI of the mean group difference observed in both studies included 10 (absolute value), neither study had sufficient power to exclude a difference of this magnitude.

The Fairbank et al² study estimated a sample size for Δ ODI > 4.0, but the SD they used was not specified. The CI for the observed group difference was 0.1– 8.1, indicating insufficient power to exclude a value of 4.0 but sufficient power for excluding a difference greater than 8.1.”

see original publication for Tables and Figures

Author’s Conclusions.

“These trials do not allow a general statement regarding the efficacy of fusion over nonoperative care for discogenic back pain. All 4 trials suggest that any advantage of surgery over nonsurgical care is modest, on average near or below the minimally important change in the disability score. The difference in the magnitude of non surgical improvement suggests that the nature of non surgical treatment may be critical.”

Our Comments/Summary.

Low to moderate risk of bias.

Table A5.7 Critical appraisal table (Wood 2011).

Study: Wood KB, Fritzell P, Dettori JR, 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*. 2011;36(21):S110-S9.

Description of Study: systematic review of 4 randomised controlled trials.

Patient/population	Adults with Chronic Low Back Pain (CLBP) with or without isthmic spondylolisthesis	
N	3 RCTs (fusion vs nonoperative care in CLBP without isthmic spondylolisthesis) 1 RCTs (fusion vs nonoperative care in CLBP with isthmic spondylolisthesis)	
Setting	Not specified	
Intervention/indicator	Reference	Intervention
	Brox 2003 (CLBP <i>without</i> isthmic spondylolisthesis)	Fusion (n=37) • PLF (posterolateral fusion) with transpedicular screws and autologous bone at L4–L5 and/or L5–S1 • Postoperative rehabilitation: standardized advice given for first 3 mo; otherwise, postoperative rehabilitation was the surgeon's choice and not standardized
	Fairbank 2005 (CLBP <i>without</i> isthmic spondylolisthesis)	Fusion (n = 176) • Fusion: 85% (149/176) • Flexible stabilization (Graf technique): 15% (27/176) • Fusion technique left to the discretion of the operating surgeon (including surgical approach, implant, if any, interbody cages, and bone graft material; NR) • Postoperative rehabilitation: NR
	Fritzell 2001 (CLBP <i>without</i> isthmic spondylolisthesis)	Fusion (n = 222) • PLF using autologous bone: 33% (73/222) • PLF + internal fixation device [VSP with pedicle screws and plates (DePuy Acromed, Raynham, MA)]: 33% (74/222) • PLF + VSP (as earlier) + interbody bone graft (ALIF or PLIF) (surgeon preference): 34% (75/222) • All surgical patients wore a brace (PLF only group) or corset for 5 mo postoperatively • Surgical technique was assigned by randomization • Only L4–L5 and/or L5–S1 fused
	Moller 2005 (CLBP <i>with</i> isthmic spondylolisthesis)	Fusion (n = 77) • PLF <i>in situ</i> with autologous bone transplantation harvested from the right iliac crest; without instrumentation (n = 40) and with rigid pedicle screw fixation (n = 37) • Noninstrumented patients wore a daytime brace for 6 mo after surgery • No postoperative exercise or physiotherapy program was given

Comparison/control	Reference	Comparison
	Brox 2003 (CLBP <i>without</i> isthmic spondylolisthesis)	Nonoperative treatment (n = 27) • Cognitive therapy + exercises as well as education and encouragement to engage in normal activities and “not be too cautious” • Physical exercises aimed to increase endurance and coordination • Supervised for 1 wk in facility, then 2 wk at home, then 2 wk in facility
	Fairbank 2005 (CLBP without isthmic spondylolisthesis)	Nonoperative treatment (n = 173) • Intensive rehabilitation program of education and exercise running on 5 d/ wk for 3 wk
	Fritzell 2001 (CLBP <i>without</i> isthmic spondylolisthesis)	Nonoperative treatment (n = 72) • Physical therapy supplemented with other types of care, such as TENS, acupuncture, injections for pain relief, cognitive and functional therapy, education, and coping strategies. • Length/structure of rehabilitation: NR
	Moller 2005 (CLBP <i>with</i> isthmic spondylolisthesis)	Nonoperative (n = 34)
Outcomes	Physical function Pain (VAS back and leg) Health related quality of life Patient satisfaction	
Inclusion Criteria	Participants: Adults with CLBP with or without isthmic spondylolisthesis Intervention: Fusion with or without instrumentation or decompression Comparators: Supervised rehabilitation Outcomes: Physical function; pain (VAS back and leg); health related quality of life; patient satisfaction Study Design: RCTs “To evaluate whether the effects of treatment in LBP patients were modified by the presence of isthmic spondylolisthesis, we first sought RCTs evaluating surgical fusion versus supervised rehabilitation for patients with CLBP that contained subpopulations of patients with and without isthmic spondylolisthesis. Having found none, we then searched the literature for RCTs that compared spine fusion versus supervised rehabilitation among those within a specific subgroup of patients with CLBP patients (i.e. , among those with isthmic spondylolisthesis only) to compare with other RCTs that were conducted among patients in the other subgroup (i.e. , among those without isthmic spondylolisthesis).”	

Exclusion Criteria	<p>Participants: patients <18 years of age; LBP with predominantly neurological involvement; spinal stenosis; tumor; moderate to severe osteoporosis; trauma; osteomyelitis; infection of soft tissue adjacent to spine; systemic infection</p> <p>Intervention: surgical procedures other than spinal fusion (± decompression) due to CLBP</p> <p>Comparators: other lumbar surgeries; non multidimensional supervised rehabilitation</p> <p>Outcomes: nonclinical outcomes</p> <p>Study Design: nonrandomized controlled trials; reviews; editorials; studies not written in English</p> <p>“We excluded studies comparing surgery other than fusion to supervised rehabilitation, surgery versus surgery; case series (a series of patients all receiving the same treatment) were also excluded. In addition, articles were excluded if they had either (a) significant dropout of eligible participants; (b) excluded eligible patients for criteria other than standard acceptable entrance characteristics; or (c) failed to describe a supervised nonoperative program. Articles were excluded if they were pediatric studies (< 18 years of age), or if they included patients with predominantly neurological involvement, stenosis, lumbar tumors, osteomyelitis, systemic infection, infection of soft tissue adjacent to the spine, trauma, or moderate to severe osteoporosis. Other exclusions included reviews, editorials, case reports, and non-English language studies”</p>
---------------------------	--

Study Validity.		
Is it clear that there were no conflicts of interest in the writing or funding of this review?	Yes	“AOSpine North America and Foundation funds were received to support this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript.”
Does the review have a clearly- focused question?	Yes	“Objective. To determine if the presence of isthmic spondylolisthesis modifies the effect of treatment (fusion vs. multidimensional supervised rehabilitation) in patients with chronic low back pain (CLBP).”
Is a systematic review the appropriate method to answer the question?	Yes	
Does the review have specified inclusion/exclusion criteria?	Yes	See above.
If there were specified inclusion/ exclusion criteria, were these appropriate?	Yes	
Does the review document a comprehensive search strategy?	Not reported	Search terms not specified
Were reviewers blind to authors, institutions and affiliations?	Not reported	
Were 2 or more independent reviewers used for: 1. application of inclusion criteria to assess eligibility of studies?	Yes	“Each retrieved citation was reviewed by two independently working reviewers (J.R.D. and R.H.). Some articles were excluded on the basis of information provided by the title or abstract if they clearly were not appropriate. Citations that appeared to be appropriate or those that could not be excluded unequivocally from the title and abstract were identified, and the corresponding full text was reviewed by the two reviewers. Any disagreement between them was resolved by consensus.”
2. extraction of data from study reports?	Not reported	

3. appraisal of study quality?	Yes	“Level-of-evidence ratings were assigned to each article independently by two reviewers (R.H. and J.R.D.) using criteria set by The Journal of Bone and Joint Surgery, American Volume, ²⁴ for therapeutic studies and modified to delineate criteria associated with methodological quality and described elsewhere”
Were the strengths and limitations of included studies and potential impact on the results discussed?	Partial	There was discussion around the inconsistencies between studies in terms of patients, treatments, comparators and follow-up, but not discussion of the strengths and limitations of individual studies
Was the validity of included trials appraised using appropriate criteria?	Yes	“Level-of-evidence ratings were assigned to each article independently by two reviewers (R.H. and J.R.D.) using criteria set by The Journal of Bone and Joint Surgery, American Volume, ²⁴ for therapeutic studies and modified to delineate criteria associated with methodological quality and described elsewhere (see Supplemental Digital Content 1, http://links.lww.com/BRS/A550).”
Is there a summary of the results of individual studies?	Yes	
If meta-analyses were conducted, was it reasonable to do so?	N/A	Forest plots were generated, but studies were presented separately in these, the results were not synthesized
If meta-analyses were conducted, was it done appropriately?	N/A	
What is the overall risk of bias?	Low to Moderate	Low to Moderate - Most of the criteria have been fulfilled and those criteria that have not been fulfilled may affect the conclusions of the study.

Results.

“Study Selection

We identified 228 total citations from our electronic search strategy. Of these, 212 were excluded by title/abstract and 16 full text articles were evaluated to determine if they met the inclusion criteria. From these 16 studies, 12 were excluded on the following reasons: eight included patients with symptoms of spinal stenosis, one included only patients who had a prior surgery, one had both subgroups but did not present results by subgroup, one compared fusion with laminectomy or discectomy, and one had no multidimensional supervised rehabilitation as a control. The remaining four RCTs met our inclusion criteria and are summarized in this report (Figure 2). Details with respect to the articles excluded and the critical appraisal summary for included articles can be found in the supplementary tables (see Table 3, Supplemental Digital Content 1, <http://links.lww.com/BRS/A550>)

Study Characteristics

We found no studies meeting our inclusion criteria that compared the two subgroups of patients (those with and without isthmic spondylolisthesis); therefore, we assessed studies that compared fusion with multidimensional supervised rehabilitation in patients with or without isthmic spondylolisthesis. To this end, three RCTs were identified that compared fusion with multidimensional supervised rehabilitation in CLBP patients without isthmic spondylolisthesis^{19–21} and one RCT was found that evaluated these treatments in CLBP patients with isthmic spondylolisthesis.²² These four studies varied in patient characteristics, type of fusion and multidimensional supervised rehabilitation, outcomes assessed, and in length of follow-up (Table 1). In Fairbank *et al*,²⁰ at baseline, 8% of the patients had received previous lumbar surgery, 11% had spondylolisthesis (degenerative or isthmic, not stated), an unknown proportion had leg pain, and 43% were on sick leave. Because only 11% of patients had an unspecified type of spondylolisthesis, we considered this to be a study primarily on patients without spondylolisthesis. In Fritzell *et al*,²¹ 18% of patients had undergone previous lumbar surgery, none had spondylolisthesis, an unknown proportion had leg pain, and 57% were on sick leave. In Brox *et al*,¹⁹ no patient had prior surgery, an unknown proportion had leg pain, and 28% were on sick leave. There was no mention of spondylolisthesis in this study. These three studies can be compared with Möller and Hedlund study,²² in which all of the patients had isthmic spondylolisthesis: none had undergone prior surgery, 67% reported leg pain, and 69% were on sick leave. There were also differences in the types of fusion and the control treatments used across studies. Fairbank *et al*²⁰ used an unspecified variety of fusion techniques (the details of which were left to the discretion of the operating surgeon) while Fritzell *et al*²¹ randomly assigned a third of the patients to receive PLF, PLF plus variable screw placement, or PLF plus variable screw placement plus interbody bone graft (either anterior or posterior lumbar interbody fusion). Brox *et al*¹⁹ exclusively used PLF with pedicle screws. Möller and Hedlund²² treated the patients with

isthmus spondylolisthesis with PLF with (52%) or without (48%) instrumentation. Although all studies used multidimensional supervised rehabilitation, the details varied by study. Treatments included the following: cognitive therapy, exercise, education, encouraging patients to engage in normal activities, transcutaneous electrical nerve stimulation, acupuncture, and injections for pain relief (see Table 1 for details).

Pain

Pain was reported by two^{19,21} of the three studies on patients without isthmus spondylolisthesis and one 22 study on patients with isthmus spondylolisthesis. The former study^{19,21} reported back and leg pain separately; both studies had patients score the intensity of their back and lower limb pain on vertical visual analog scales scores that ranged from 0 to 100. Maximum pain, minimum pain, and current pain were scored on three different scales, and the mean of the three measurements provided the pain index for back pain and lower limb pain, respectively. In the study on patients with isthmus spondylolisthesis, 22 an overall pain index was reported and did not distinguish between back and leg pain. The pain index was calculated by taking the mean of the visual analog scale scores for “pain right now” and that for “worst pain last week.” The SMDs for pain in favor of fusion were modest among those without isthmus spondylolisthesis as reported by Brox *et al*¹⁹ at 1 year and Fritzell *et al*²¹ at 2 years for back pain: 0.29 (95% CI = -0.22, 0.80) and 0.70 (95% CI = 0.41, 0.99); and leg pain: 0.63 (95% CI = 0.11, 1.15) and 0.50 (95% CI = 0.21, 0.80), respectively. In contrast, there was a large effect in favor of fusion for isthmus spondylolisthesis patients as reported by Möller and Hedlund²² at 2 years, SMD: 2.31 (95% CI = 1.79, 2.84) (Figure 3).

Function

Function was assessed in all four studies. All three studies^{19–21} on patients without isthmus spondylolisthesis evaluated function using the Oswestry Disability Index; two of these three studies^{19,21} also utilized the General Function Score. The study on patients with isthmus spondylolisthesis measured function using the Disability Rating Index.²² In patients without isthmus spondylolisthesis, the SMDs for function as measured by both the Oswestry Disability Index and the General Function Score were relatively small and in favor of fusion: Oswestry Disability Index SMDs = 0.13 (95% CI = -0.38, 0.64) at 1 year, 0.18 (95% CI = -0.05, 0.42) at 2 years and 0.50 (95% CI = 0.30, 0.88) at 2 years as reported by Brox *et al*,¹⁹ Fairbank *et al*,²⁰ and Fritzell *et al*,²¹ respectively; General Function Score SMDs = 0.25 (95% CI = -0.27, 0.75) at 1 year and 0.59 (95% CI = 0.30, 0.88) at 2 years for Brox *et al*¹⁹ and Fritzell *et al*,²¹ respectively. In contrast, the SMD for function as measured by the Disability Rating Index was appreciably higher in favor of fusion in patients with isthmus spondylolisthesis as reported by Möller and Hedlund²² at 2 years: SMD = 3.03 (95% CI = 2.45, 3.62) (Figure 3).

Patient Improvement Rating

Two studies reported whether patients were improved (“better” or “much better”) as assessed by patient perception and whether patients were willing to go through the treatment again. In both studies, the proportion of patients that reported improvement was higher in the fusion group than in the multidimensional supervised rehabilitation group. However, the proportions favoring fusion were similar at 2 years comparing groups without²¹ and with²² isthmus spondylolisthesis, with the following RDs: 34% (95% CI = 21%, 48%) compared with 33% (95% CI = 13%, 53%), as reported by Fritzell *et al*²¹ and Möller and Hedlund,²² respectively (Figure 4). In addition, a slightly higher proportion of patients without 21 isthmus spondylolisthesis receiving fusion reported a willingness to go through the same treatment again compared with patients with 22 isthmus spondylolisthesis, RDs = 22% (95% CI = 8%, 37%) compared with 11% (95% CI = -8%, 30%), as reported by Fritzell *et al*²¹ and Möller and Hedlund²² at 2 years, respectively (Figure 4).

Complications

The complication rates reported by the studies varied considerably and, given the small number of studies cited, whether they are truly clinically significant remains uncertain. The study by Fritzell *et al*²¹ as well as that by Brox *et al*¹⁹ both reported an 18% complication rate, whereas that of Fairbank *et al*²⁰ was 11%. These included wound infections, dural tears, vascular injuries (anterior approaches), excessive bleeding, implant problems, thrombosis, and iatrogenic lumbar radiculopathies. The study by Möller and Hedlund²² of patients with isthmus spondylolisthesis, only described three complications in 77 patients (4%), although they were serious; two permanent nerve root injuries and one patient with permanent blindness.

Evidence Summary

The overall strength of evidence evaluating whether the presence of isthmus spondylolisthesis modifies the effect of fusion compared with comprehensive rehabilitation in patients with CLBP is “low.” A low strength of evidence suggests that further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. (Table 2)”

From abstract: “Results. No studies were found that directly compared the two subgroups. Three RCTs compared fusion with supervised nonoperative care in patients with CLBP without isthmus

spondylolisthesis; one RCT evaluated these treatments in patients with isthmic spondylolisthesis. There were study differences in patient characteristics, type of fusion, the nature of the rehabilitation, outcomes assessed, and length of follow-up. The SMDs for pain in favor of fusion were modest at 2 years for those without isthmic spondylolisthesis, but large in favor of fusion for those with isthmic spondylolisthesis compared with rehabilitation. Similarly, the SMDs for function in patients without isthmic spondylolisthesis compared with rehabilitation was small at 2 years, but appreciably higher in favor of fusion in patients with isthmic spondylolisthesis.”

See original article for Tables and Figures

Author’s Conclusions.

“Nonetheless, on the basis of our well-defined review of the available literature, it does appear that patients with CLBP and isthmic spondylolisthesis, as compared with those without, can be treated differently when considering care, especially surgery: They are different and they respond differently.”

“Key Points

- Three RCTs compared fusion with multidimensional supervised rehabilitation in patients with CLBP without isthmic spondylolisthesis and one RCT evaluated these treatments in patients with isthmic spondylolisthesis.
- There were study differences in patient characteristics, type of fusion, the nature of the rehabilitation, outcomes assessed, and length of follow-up.
- The SMDs for pain and function in favor of fusion were modest at 2 years among those without isthmic spondylolisthesis, but large in favor of fusion among those with isthmic spondylolisthesis compared with rehabilitation.
- Fusion should be considered for patients with LBP and isthmic spondylolisthesis who have failed nonoperative treatment.”

From abstract: **“Conclusion.** The overall strength of evidence evaluating whether the presence of isthmic spondylolisthesis modifies the effect of fusion compared with rehabilitation patients with CLBP is “low.” Fusion should be considered for patients with low back pain and isthmic spondylolisthesis who have failed nonoperative treatment. Clinical Recommendations. We recommend considering fusion for patients with isthmic spondylolisthesis and lower back pain who have failed nonoperative treatment. Recommendation: Weak.”

Our Comments/Summary.

This review has a low to moderate risk of bias due to its failure to provide details of the search strategy used, and the lack of discussion around the quality of the individual studies and the impact this may have on the results.

It should be noted that the recommendations made by this study are reported to be ‘weak’ as they are based on a low strength of evidence.

Variations in the patient populations, interventions, comparators, outcomes length of follow-up for the included studies make generalisation of the findings difficult.

Key primary studies not included in synthesised studies

Table A5.8 Critical appraisal table (Brox 2010).

Study: Brox JI, Nygaard OP, Holm I, Keller A, Ingebrigtsen T, Reikeras O. Four-year follow-up of surgical versus non-surgical therapy for chronic low back pain. *Annals of the Rheumatic Diseases*. 2010;69(9):1643-8.

Description of study: *randomised controlled trial.*

Patient/population	Adults aged (25-60yrs) with chronic low back pain for at least 1 year.
N	N=124. Surgical group N=66. Non surgical group N=58.
Setting	Study conducted at four university hospitals.
Intervention/indicator	Posterolateral fusion with transpedicular screws of the L4-L5 and or L5-S1 segment.
Comparison/control	Cognitive intervention and exercises.
Outcomes	Primary outcome measure Oswestry Disability Index. Secondary outcome measure: pain, general function score, global back disability question for the assessment of patients overall rating, work and medication, emotional distress, fear-avoidance beliefs, and life satisfaction.
Inclusion Criteria	Adults aged 25-60 yrs, with chronic low back pain for at least 1 year. ODI greater than 30 and disc degeneration at L4-L5 and/or L5-S1.
Exclusion Criteria	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 of the interventions.

Study Validity.

Is it clear that there are no conflicts of interest in the writing or funding of this study?	Yes	The study documents that there are no competing interests.
Does the study have a clearly focused question?	Yes	"to compare the long-term effectiveness of surgical and non surgical treatment in patients with chronic low back pain."
Is a RCT the appropriate method to answer this question?	Yes	
Does the study have specified inclusion/exclusion criteria?	Yes	The exclusion criteria are extremely widespread.
If there were specified inclusion/ exclusion criteria, were these appropriate?	Yes	
Did the study have an adequate method of randomisation?	Yes	"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."
Was allocation to intervention group concealed?	Yes	
Were patients blind to intervention group?	No	Blinding is not possible for this type of study.

Were investigators and care providers blind to intervention group?	No	Blinding is not possible for this type of study.
Were outcome assessors blind to intervention group?	Not reported	
All outcomes were measured in a standard, valid and reliable way?	Yes	Outcomes were measured using the Oswestry Disability Index standardised questionnaire.
Were outcomes assessed objectively?	Yes	
Were outcomes assessed independently?	Partial	Primary and secondary measures were assessed using the same questionnaire.
Were the groups similar at baseline with regards to key prognostic variables?	Yes	
Aside from the experimental intervention, were the groups treated the same?	Yes	
Were the outcomes measured appropriate?	Yes	
Was there sufficient duration of follow-up?	Yes	The aim of this study is to assess patients at four-year follow-up.
Was there ≤20% drop-out?	Yes	
Was the study sufficiently powered to detect any differences between the groups?	Yes	
If statistical analysis was undertaken, was this appropriate?	Yes	
Were all the subjects analysed in the groups to which they were randomly allocated (ie intention to treat analysis)?	Yes	7% crossovers/withdrawals were reported in the Lumbar fusion group. 8% crossovers/withdrawals were reported in the comparator group.
Is the paper free of selective outcome reporting?	Yes	
What is the overall risk of bias?	Low	

Results.

“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; 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 –8.3 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 ($p=0.14$)."

see original publication for Tables and Figures

Author's Conclusions.

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

Our Comments/Summary.

Low risk of bias.

Table A5.9 Critical appraisal table (Froholdt 2011).

Study: Froholdt A, Holm I, Keller A, Gunderson RB, Reikeraas O, Brox JI. No difference in long-term trunk muscle strength, cross-sectional area, and density in patients with chronic low back pain 7 to 11 years after lumbar fusion versus cognitive intervention and exercises. *Spine J.* 2011;11(8):718-25.

Description of Study: randomised controlled trial.

Patient/population	Patients with CLBP and disc degeneration
N	N total = 124. N (Lumbar Fusion) = 66. N(Cognitive exercise therapy)=58
Setting	Surgery; Inpatient. Cognitive exercise therapy; patient hotel.
Intervention/indicator	Lumbar fusion (posterolateral autologous bone transplantation and transpedicular screw fixation of the L4-L5 and /or L5-S1 segments)
Comparison/control	Cognitive and exercise therapy.
Outcomes	Trunk muscle strength, cross-sectional area and density.
Inclusion Criteria	Patients aged 25 to 60 years, with reported low back pain for at least one year, a score of 30 out of 100 points on the ODI and degenerative changes at L4-L5 and/or L5-S1 on plain radiographs, or previously performed surgery for disc herniation with laminectomy.
Exclusion Criteria	Exclusion criteria is widespread including, myofascial pain, spinal stenosis with reduced walking distance and neurologi signs, disc herniation, or lateral recess stenosis with clinical signs of radiculopathy, inflammatory disease, previous spinal fracture, the pelvic girdle syndrome, generalised degenerative changes on plain radiograph examination, serious somatic or psychiatric disease that excluded either one or both treatment alternatives, registered medical abuse, or reluctance to accept one or both of the treatment regimens of the study.

Study Validity.

Is it clear that there are no conflicts of interest in the writing or funding of this study?	Yes	All authors report that they have no conflicts of interest.
Does the study have a clearly focused question?	Yes	To compare the long term effect of lumbar fusion and cognitive intervention and exercises on muscle strength, cross sectional area, density and self rated function in patients with chronic low back pain and disc degeneration.
Is a RCT the appropriate method to answer this question?	Yes	
Does the study have specified inclusion/exclusion criteria?	Yes	Inclusion/Exclusion criteria listed above.
If there were specified inclusion/ exclusion criteria, were these appropriate?	Yes	

Did the study have an adequate method of randomisation?	Not reported	Subjects were selected from a previous cohort of subjects in another study. (unsure how subjects in the previous study were selected). Although the study says that the subjects were randomised, the method of randomisation was not reported.
Was allocation to intervention group concealed?	Not reported	
Were patients blind to intervention group?	No	Blinding is not possible for this study due to the nature of the intervention and comparator.
Were investigators and care providers blind to intervention group?	No	Blinding is not possible for this study due to the nature of the intervention and comparator.
Were outcome assessors blind to intervention group?	Not reported	
All outcomes were measured in a standard, valid and reliable way?	Yes	
Were outcomes assessed objectively?	Yes	
Were outcomes assessed independently?	Yes	
Were the groups similar at baseline with regards to key prognostic variables?	Yes	
Aside from the experimental intervention, were the groups treated the same?	Yes	
Were the outcomes measured appropriate?	Yes	
Was there sufficient duration of follow-up?	Yes	This study focused on the long term follow up of patients between 7 and 11 years after lumbar fusion versus cognitive intervention and exercises.
Was there ≤20% drop-out?	Yes	
Was the study sufficiently powered to detect any differences between the groups?	No	The study was not sufficiently powered to measure the difference in muscle strength.
If statistical analysis was undertaken, was this appropriate?	Yes	

Were all the subjects analysed in the groups to which they were randomly allocated (ie intention to treat analysis)?	Yes	
Is the paper free of selective outcome reporting?	Yes	
What is the overall risk of bias?	Low	

Results.

“Sociodemographic characteristics are reported in [Table 1](#). There were no significant differences at baseline between the two study groups or between the 55 patients included in the present study compared with the 124 patients included in the original studies ([Table 1](#)).

Fifty-five of 61 patients (90%) attended the long-term follow-up examination. One patient in the randomized lumbar fusion group and five patients in the cognitive intervention and exercises group did not attend ([Figure](#)). Two patients were seriously ill, three were not willing to participate, and one did not respond to the inquiry.

Twenty-four patients (44%) were originally allocated lumbar fusion, and 31 patients (56%) were allocated cognitive intervention and exercises. Nine patients who were allocated cognitive intervention and exercises crossed over after the 1-year follow-up and had surgery, whereas one patient who was allocated fusion refused surgery. At the long-term follow-up, 32 patients (58%) had lumbar fusion, whereas 23 patients (42%) had cognitive intervention and exercises and no later surgery.

Trunk muscle flexion and self-rated function significantly improved, whereas cross-sectional area and muscle density significantly decreased within groups at the long-term follow-up examination ([Table 2](#)).

The adjusted mean difference between the groups for trunk extension was 31 joules (J) (95% CI, -188 to 250, $p=0.78$). The adjusted mean difference between the groups for trunk flexion was -88 J (95% CI, -320 to 145, $p=0.45$). The adjusted mean difference in the E/F ratio at 60°/s was significantly higher in the lumbar fusion group (0.24; 95% CI, 0.02–0.47; $p=0.04$).

Based on the as-treated analyses, there were no differences between the two treatment groups, except for density, which was significantly lower in patients who had lumbar fusion (mean difference, -7.5; 95% CI, -16.6 to -1.5; $p=0.015$).

The correlation coefficient (Spearman R) between cross-sectional area and density was 0.42, and the values were 0.24 and 0.44 between density or cross-sectional area and trunk muscle strength, respectively.”

See original article for Tables and Figures.

Author’s Conclusions.

“Although this study did not directly assess muscle morphology of muscles likely damaged by surgery, gross muscle strength, cross-sectional area, and density above the lesion are not different between those who have had lumbar fusion or cognitive intervention and exercises at 7- to 11-years after lumbar fusion.”

Our Comments/Summary.

Low risk of bias

Table A5.10 Critical appraisal table (Keller 2004).

Study: Keller A, Brox JJ, Gunderson R, Holm I, Friis A, Reikeras 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(1):3-8.

Description of Study: *secondary analysis of 2 randomised controlled trials.*

Patient/population	Patients with chronic low back pain + disc degeneration (CLBP + DD) or chronic low back pain + post-laminectomy syndrome (CLBP + PLS). This study is a secondary analysis of data from 2 RCTs that were identical in design and treatment, but differed in patient groups: one had patients with CLBP + DD, and the other had patients with CLBP + PLS
N	Of the 124 patients enrolled, 112 were tested for muscle strength: <ul style="list-style-type: none"> • fusion group n=60 (CLBP + DD n=33, CLBP + PLS n=27) • cognitive intervention and exercise group n=52 (CLBP + DD n=25, CLBP + PLS n=27) 61 of the 112 tested for muscle strength were tested for cross sectional area and density of back muscles (the number of patients in each treatment group was not reported).
Setting	Inpatient and outpatient settings, departments of orthopaedics and physiotherapy, Rikshospitalet University Hospital, Oslo, Norway
Intervention/indicator	Spinal fusion surgery (lumbar)
Comparison/control	Cognitive intervention + exercises
Outcomes	Trunk muscle strength, cross sectional area and density of back muscles
Inclusion Criteria	"The criteria for inclusions were as follows: age 25–60 years; reported low back pain for at least 1 year; a score of 30 of 100 points on the Oswestry Disability Index (ODI); and degenerative changes at the L4–L5 and/or L5–S1 on plain radiographs or previously performed surgery for disc herniation with laminectomy."
Exclusion Criteria	Patients were excluded if they had widespread myofascial pain, spinal stenosis with reduced walking distance and neurologic signs, disc herniation or lateral recess stenosis with clinical signs of radiculopathy, inflammatory disease, previous spinal fracture, the pelvic girdle syndrome, generalized degenerative changes on plain radiograph examination, serious somatic or psychiatric disease that excluded either one or both treatment alternatives, registered medical abuse, or reluctance to accept one or both the treatment regimens of the study.

Study Validity.

Is it clear that there are no conflicts of interest in the writing or funding of this study?	Yes	"No funds were received in support of this work. No benefits in any form have been or will be received from a commercial party related directly or indirectly to the subject of this manuscript."
Does the study have a clearly focused question?	Yes	"The purpose of the present study was to investigate the differences in muscle strength, cross-sectional area, and density of the back muscles in patients with chronic low back pain and disc degeneration and in patients with postlaminectomy syndrome, randomized to either lumbar fusion or cognitive intervention and exercises."
Is a RCT the appropriate method to answer this question?	Yes	
Does the study have specified inclusion/exclusion criteria?	Yes	See above

If there were specified inclusion/ exclusion criteria, were these appropriate?	Yes	
Did the study have an adequate method of randomisation?	Yes	<p>“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.” Brox 2006</p> <p>“When a patient was found eligible and had given an informed, signed consent, the project coordinator (A.F.) telephoned the randomization central at the University of Bergen and reported an identification number. Within an hour the patient was allocated to one of the intervention groups, the project coordinator was phoned back, and the patient was informed. The method of concealed random allocation was used. Simple randomization was conducted by a computer-generated random list. Blocks of patients were used to ensure fairly equal treatment numbers. The project coordinator was not aware of the block size and could not predict the group assignments.” Brox 2003</p>
Was allocation to intervention group concealed?	Yes	See above
Were patients blind to intervention group?	No	Not possible to blind patients when comparing surgery against non-surgical interventions
Were investigators and care providers blind to intervention group?	No	Not possible to blind care providers when comparing surgery against non-surgical interventions
Were outcome assessors blind to intervention group?	Yes	<p>“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.” Brox 2006</p> <p>“A physical therapist and a specialist in physical medicine and rehabilitation carried out blind follow-up measurements 1 year after the first day of treatment. A nurse in the outpatient clinic always told the patients not to mention anything about their treatment to the independent observers.” Brox 2003</p>
All outcomes were measured in a standard, valid and reliable way?	Yes	<p>“Trunk muscle strength was measured on a Cybex Isokinetic Trunk Extension Flexion Device (model 6000; Cybex-Lumex Inc., Ronkonkoma, NY) at the inclusion and at the 1-year follow-up examination... In addition, muscle strength was evaluated by the Biering-Sørensen Test, which measures how many seconds the participant is able to keep the unsupported upper part of the body in a horizontal position... CT (Toshiba XPEED, Tokyo, Japan) was used to measure the cross-sectional area and density of the back muscles.”</p>
Were outcomes assessed objectively?	Yes	See above
Were outcomes assessed independently?	Not reported	However, the outcomes were objectively assessed, meaning that their independence from each other is less important.
Were the groups similar at baseline with regards to key prognostic variables?	Yes	
Aside from the experimental intervention, were the groups treated the same?	Not reported	
Were the outcomes measured appropriate?	Yes	
Was there sufficient duration of follow-up?	Yes	1 year follow-up

Was there ≤20% drop-out?	Partial	Of the 124 patients enrolled, 112 were tested for muscle strength: <ul style="list-style-type: none"> fusion group n=60 (CLBP + DD n=33, CLBP + PLS n=27) cognitive intervention and exercise group n=52 (CLBP + DD n=25, CLBP + PLS n=27) 61 of the 112 tested for muscle strength were tested for cross sectional area and density of back muscles (the number of patients in each treatment group was not reported). (>20% not evaluated for cross sectional area and density)
Was the study sufficiently powered to detect any differences between the groups?	Not reported	Not reported for Keller 2004 or Brox 2006. However, Brox 2003 was reported to be underpowered (the confidence interval for the main outcome included a clinically important difference).
If statistical analysis was undertaken, was this appropriate?	Yes	However, it is unclear if the statistical analysis in this secondary analysis was planned a priori
Were all the subjects analysed in the groups to which they were randomly allocated (ie intention to treat analysis)?	Yes	
Is the paper free of selective outcome reporting?	Yes	The Keller paper focused on the specific outcomes of back muscle strength, cross-sectional area and density, however other outcomes were reported in Brox 2003 and Brox 2006
What is the overall risk of bias?	Low to Moderate	Most of the criteria have been fulfilled and those criteria that have not been fulfilled may affect the conclusions of the study.

Results.

"The sociodemographic characteristics at inclusion are given in Table 1. The patients who had CT scans did not differ significantly from those who were not examined by this method, and the patients who performed the muscle strength tests were not different from those who did not (Table 1).

At the inclusion and at the 1-year follow-up, muscle strength, cross-sectional area, and density of back muscles did not differ significantly between patients with chronic low back pain/disc degeneration and postlaminectomy syndrome. At the inclusion and at the 1-year follow-up examination, 14 and 21 patients, respectively, were not able to perform the isokinetic muscle strength test, and 8 and 18 patients, respectively, were not able to carry out the Biering-Sørensen Test. Moreover, at the inclusion there was no significant difference in muscle strength, cross-sectional area, and density for patients randomized to either lumbar fusion or cognitive intervention and exercises. This was also the case for the ODI.

At the 1-year follow-up examination, both patients randomized to lumbar fusion and cognitive intervention and exercises improved significantly in disability ratings (both $P < 0.001$), but there was no significant difference between the groups (Table 2).

At the 1-year follow-up examination, muscle strength had increased significantly in the exercise group but not in the lumbar fusion group, with the mean difference between the groups of 184 Nm (95% confidence interval [CI], 64–303 Nm; $P = 0.003$). For the Biering-Sørensen Test, there was a significant decrease in the muscle strength in the lumbar fusion group with the mean difference between the treatment groups being 21 seconds (95% CI, 6–36 seconds; $P = 0.006$) (Table 2). Limiting the analyses to include only the 61 patients who had CT scans performed gave similar results: the mean difference between the groups was 193 Nm (95% CI, 32–353 Nm; $P = 0.02$) for the isokinetic test and 24 seconds (95% CI, 5–43 seconds; $P = 0.01$) for the Biering-Sørensen Test.

There were no significant changes in the crosssectional area from the inclusion to the 1-year follow-up at either L3–L4 and T12–L1, in the two treatment groups (Table 2). There was a significant increase in density at level T12–L1 in the exercise group and a nonsignificant increase in the lumbar fusion group. The mean difference in changes between the groups was 2.4 HU (95% CI, -4.7 to 9.6 HU, not significant). The density at L3–L4 decreased significantly in the lumbar fusion group and was unchanged in the exercise group. The mean difference in changes between the groups was 5.3 HU (95% CI, 1.1–9.5 HU; $P = 0.01$) (Table 2).

There were no significant correlations between changes in muscle performance and density and changes in cross-sectional area, from the inclusion to the 1-year follow-up examination. The mean number of

physiotherapy sessions given after the hospital stay was 31 in the lumbar fusion group and 4 in the group who followed the exercise program.”

See original article for Tables and Figures

Author’s Conclusions.

“Conclusions. Patients with chronic low back pain who followed cognitive intervention and exercise programs improved significantly in muscle strength compared with patients who underwent lumbar fusion. In the lumbar fusion group, density decreased significantly at L3–L4 compared with the exercise group.”

“There was a significant difference in muscle performance between patients randomized to lumbar fusion and cognitive intervention and exercises. Density at L3–L4 decreased significantly in the lumbar fusion group but remained unchanged in the exercise group. The difference between the treatments was significant. The cross-sectional area remained unchanged at the two spinal levels in both treatment groups. There was no correlation between change in muscle strength and muscle morphology.”

Our Comments/Summary.

This study has a moderate risk of bias due to: the discrepancies between the number of patients enrolled in the study, and the number assessed for the outcomes of cross-sectional area and muscle density (only around 50%) and the lack of blinding for patients and caregivers. It should be noted however that blinding may not be possible in a study comparing surgical and non-surgical interventions.

The likelihood that the 2 original studies were underpowered mean that the results may not be generalizable.

Table A5.11 Critical appraisal table (Ohtori 2011).

Study: Ohtori S, Koshi T, Yamashita M, Yamauchi K, Inoue G, Suzuki M, et al. Surgical Versus Nonsurgical Treatment of Selected Patients With Discogenic Low Back Pain A Small-Sized Randomized Trial. *Spine*. 2011;36(5):347-54.

Description of Study: randomised controlled trial.

Patient/population	Patients with low back pain (LBP) for at least 2 years
N	$N_{total} = 41$, $N_{Exercise} = 20$, $N_{ABF} = 15$, $N_{PLF} = 6$
Setting	Not specified
Intervention/indicator	Lumbar Fusion; Anterior Interbody Fusion (ABF) and Posterolateral fusion (PLF).
Comparison/control	Exercise
Outcomes	Pain Visual Analog Scale (VAS) Score, Japanese Orthopedic Association Score and Oswestry Disability Index (ODI).
Inclusion Criteria	LBP for at least 2 years with no accompanying radicular pain.
Exclusion Criteria	Patients who had severe spondylosis or disc degeneration

Study Validity.

Is it clear that there are no conflicts of interest in the writing or funding of this study?	Yes	
Does the study have a clearly focused question?	Yes	To examine the difference in surgical versus nonsurgical treatment of selected patients with discogenic low back pain (DLBP).
Is a RCT the appropriate method to answer this question?	Yes	
Does the study have specified inclusion/exclusion criteria?	Yes	Inclusion/exclusion criteria outlined above.
If there were specified inclusion/ exclusion criteria, were these appropriate?	Yes	
Did the study have an adequate method of randomisation?	Yes	"The ethics committee of our institution randomized the patients according to the minimization method for invasive surgery or minimal treatment control"
Was allocation to intervention group concealed?	Not reported	

Were patients blind to intervention group?	No	Difficult to blind patients for this type of study.
Were investigators and care providers blind to intervention group?	No	Difficult to blind care providers for this type of study.
Were outcome assessors blind to intervention group?	Not reported	
All outcomes were measured in a standard, valid and reliable way?	Partial	
Were outcomes assessed objectively?	Partial	One of the outcomes was subjective with patients answering questions about their how they felt/feel prior to and after surgery or exercise.
Were outcomes assessed independently?	Not reported	
Were the groups similar at baseline with regards to key prognostic variables?	Yes	
Aside from the experimental intervention, were the groups treated the same?	Yes	
Were the outcomes measured appropriate?	Yes	
Was there sufficient duration of follow-up?	Yes	Patients were followed-up at one and two years post treatment.
Was there ≤20% drop-out?	Yes	
Was the study sufficiently powered to detect any differences between the groups?	Not reported	This was a small study of only 41 patients. Power was not reported.
If statistical analysis was undertaken, was this appropriate?	Yes	
Were all the subjects analysed in the groups to which they were randomly allocated (ie intention to treat analysis)?	Yes	
Is the paper free of selective outcome reporting?	Yes	

What is the overall risk of bias?	Moderate	
-----------------------------------	----------	--

Results.

“No significant differences in pain scores before treatment among the three groups ($P>0.05$). In the minimal treatment control, these 3 scores were slightly improved at two years compared with before treatment. VAS and ODI were significantly improved during 2 years ($p<0.05$); however improvement of JOAS was not significant. In comparison of the scores at 1 and 2 years, there was a small, but significant difference in ODI score in the minimally treated control group ($P<0.05$). In the surgical groups VAS, JOAS and ODI at 2 years after treatment were significantly improved compared with those before treatment. At 1 year after treatment, the scores were significantly improved in the 2 surgical groups compared with the minimally treated group ($p<0.01$). At 2 years after treatment, VAS scores, JOAS and ODI were significantly improved in the 2 surgical groups compared with the minimally treated group ($p<0.01$).”

Please refer to original text for figures.

Author's Conclusions.

“In conclusion, if DLBP is strictly diagnosed, surgical therapy is suitable for its treatment. ABF gives good results, but PLF is also an option for patients who do not want anterior surgery or who present a difficult approach because of anterior vessels.”

Our Comments/Summary.

Moderate risk of bias.

Small sample size means it is unlikely that the results of the study are generalisable.

Table A5.12 Critical appraisal table (Moller 2000)

Study: Moller H, Hedlund R. Surgery versus conservative management in adult isthmic spondylolisthesis--a prospective randomized study: part 1. *Spine*. 2000;25(13):1711-5.

Description of Study: *randomised controlled trial.*

Patient/population	Isthmic spondylolisthesis
N	111 patients were randomly allocated to an exercise program (n =34) or posterolateral fusion with or without transpedicular fixation (n = 77).
Setting	Inpatient setting
Intervention/indicator	Posterolateral fusion with or without transpedicular fixation
Comparison/control	Exercise program
Outcomes	All patients completed a questionnaire concerning their symptoms, functional disability, and pain before treatment, then at the 1- and 2-year follow-up assessments. Functional disability was quantified by the Disability Rating Index (DRI).
Inclusion Criteria	Lumbar isthmic spondylolisthesis of any grade, at least 1 year of low back pain or sciatica, and a severely restricted functional ability in individuals 18 to 55 years of age
Exclusion Criteria	Patients with mild symptoms, previous spine surgery, or alcohol/drug abuse were excluded

Study Validity.

Is it clear that there are no conflicts of interest in the writing or funding of this study?	Yes	Supported by research grants from the Karolinska Institute and the King Oscar II and Queen Sofia's Golden Anniversary Foundation.
Does the study have a clearly focused question?	Yes	
Is a RCT the appropriate method to answer this question?	Yes	
Does the study have specified inclusion/exclusion criteria?	Yes	See above inclusion/exclusion criteria
If there were specified inclusion/ exclusion criteria, were these appropriate?	Yes	
Did the study have an adequate method of randomisation?	Partial	Randomization without stratification was used.
Was allocation to intervention group concealed?	Yes	For each patient, three different notes were each marked with one of the three different treatment methods, one note was blindly chosen by the attending nurse in the outpatient ward. This procedure gave each patient the same chance to become part of each treatment group
Were patients blind to intervention group?	No	The type of treatment was unknown to the patient and the physician until after the patient had given consent.
Were investigators and care providers blind to intervention group?	No	

Were outcome assessors blind to intervention group?	Not reported	
All outcomes were measured in a standard, valid and reliable way?	Yes	
Were outcomes assessed objectively?	Yes	Objective tools and measures were used. See outcome instruments
Were outcomes assessed independently?	Yes	At follow-up evaluation, the patient was not seen by the operating surgeon, but by an independent orthopedic surgeon, the main surgeon's assistant, or the physiotherapist leading the exercise program. Some patients who underwent surgery were seen at follow-up evaluation by the physiotherapist, whereas other patients who received the exercise program were at seen at follow-up assessment by a surgeon to avoid bias.
Were the groups similar at baseline with regards to key prognostic variables?	Yes	
Aside from the experimental intervention, were the groups treated the same?	Yes	
Were the outcomes measured appropriate?	Yes	
Was there sufficient duration of follow-up?	Yes	.
Was there ≤20% drop-out?	Yes	Five patients were excluded from the study and accounted for by the study authors
Was the study sufficiently powered to detect any differences between the groups?	Yes	
If statistical analysis was undertaken, was this appropriate?	Yes	
Were all the subjects analysed in the groups to which they were randomly allocated (ie intention to treat analysis)?	Yes	
Is the paper free of selective outcome reporting?	Yes	
Was this intervention suitable for a cross-over study?	Unclear	
Was the washout period adequate?	Yes	

Other		
What is the overall risk of bias?	Moderate	<i>Moderate - Some of the criteria have been fulfilled and those criteria that have not been fulfilled may affect the conclusions of the study.</i>

Results.

"The surgically treated group reported a significantly lower DRI ($P=0.004$) and pain index ($P=0.002$) at the 2-year follow-up assessment than the exercise group. At the 2-year follow-up assessment, 11 of the 12 functional scores were significantly better in the surgical group than in the exercise group. The score for running was the only score that was not significantly lower in the surgical group (44) than in the exercise group (55) ($P = 0.13$). Among all 106 patients that completed the 2-year follow-up assessment, 34 of the 37 patients with a DRI lower than 20, and 28 of the 29 patients with a pain index lower than 20 were treated surgically

In the longitudinal analysis, the pain index and DRI before treatment and at the 2-year follow-up assessment were compared. The mean DRI improved from 48 (range, 7–83) to 29 (range, 0–79) in the surgical group ($P = 0.0001$), but remained unchanged in the exercise group, which had a DRI of 44 (range, 6–75) before treatment and 44 (range, 15–84) at the 2-year follow-up evaluation. The mean pain index improved in both groups, from 63 (range, 10–98) to 37 (range, 0–96) in the surgical group ($P = 0.0001$), and from 65 (range, 32–96) to 56 (range, 17–87) in the exercise group ($P = 0.024$)."

Please refer to original article for Tables and Figures.

Author's Conclusions.

"Surgical management of adult isthmic spondylolisthesis improves function and relieves pain more efficiently than an exercise program"

Our Comments/Summary.

The study confirmed the impression of other studies that fusion improved the pain response and functional activity better than multidimensional supervised rehabilitation in the subgroup of CLBP patients with isthmic spondylolisthesis. However, this is a one single relatively small RCT, and although there appeared to be a relatively strong treatment effect, their results should be treated with caution.

Table A5.13 Critical appraisal table (Weinstein 2007)

Study: Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA, et al. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. *NEJM*. 2007;356(22):2257-70.

Description of Study: combined cross-over randomised controlled trial and cohort study.

Patient/population	Degenerative spondylolisthesis
N	304 patients in the randomized cohort and 303 in the observational cohort.
Setting	13 United States (US) surgical centres
Intervention/indicator	Standard decompressive laminectomy (with or without fusion)
Comparison/control	Usual nonsurgical care
Outcomes	Medical Outcomes Study 36-Item Short- Form General Health Survey (SF-36) bodily pain and physical function scores and the modified Oswestry Disability Index.
Inclusion Criteria	All patients had neurogenic claudication or radicular leg pain with associated neurologic signs, spinal stenosis shown on cross-sectional imaging, and degenerative spondylolisthesis shown on lateral radiographs obtained with the patient in a standing position. The patients had had persistent symptoms for at least 12 weeks and had been confirmed as surgical candidates by their physicians.
Exclusion Criteria	Patients with spondylolysis and isthmic spondylolisthesis

Study Validity.

Is it clear that there are no conflicts of interest in the writing or funding of this study?	Yes	Supported by a grant from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) and the Office of Research on Women's Health; by the National Institutes of Health; by the National Institute of Occupational Safety and Health and the Centers for Disease Control and Prevention; and by a grant to the Multidisciplinary Clinical Research Center in Musculoskeletal Diseases) and a Research Career Award from NIAMS.
Does the study have a clearly focused question?	Yes	
Is a RCT the appropriate method to answer this question?	Yes	This is a combined RCT and cohort study
Does the study have specified inclusion/exclusion criteria?	Yes	See above
If there were specified inclusion/ exclusion criteria, were these appropriate?	Yes	
Did the study have an adequate method of randomisation?	Yes	Participants in the randomized cohort received computer generated random treatment assignments blocked according to center. The patients were offered enrolment in either cohort and gave written informed consent after viewing videotapes explaining the expected benefits, risks, and uncertainties of the treatments

Was allocation to intervention group concealed?	Not reported	
Were patients blind to intervention group?	Not reported	
Were investigators and care providers blind to intervention group?	Not reported	
Was this intervention suitable for a cross-over study?	Yes	
Was the washout period adequate?	Yes	
Were outcome assessors blind to intervention group?	Not reported	
All outcomes were measured in a standard, valid and reliable way?	Yes	
Were outcomes assessed objectively?	Yes	
Were outcomes assessed independently?	Not reported	
Were the groups similar at baseline with regards to key prognostic variables?	Yes	In addition the randomized and observational cohorts were similar at baseline.
Aside from the experimental intervention, were the groups treated the same?	Yes	
Were the outcomes measured appropriate?	Yes	
Was there sufficient duration of follow-up?	Yes	A total of 601 patients (99%) completed at least one follow-up visit and were included in the analysis; between 83% and 95% of patient supplied data at each follow-up visit.
Was there $\leq 20\%$ drop-out?	Yes	
Was the study sufficiently powered to detect any differences between the groups?	Yes	
If statistical analysis was undertaken, was this appropriate?	Yes	The randomized cohort was initially analysed on an intention-to-treat basis. However, because of the extent of crossover, subsequent analyses combined the randomized cohort and the observational cohort and were based on treatments actually received.
Were all the subjects analysed in the groups to which they were randomly allocated (ie intention to treat analysis)?	Yes	The one-year crossover rates were high in the randomized cohort (approximately 40% in each direction) but moderate in the observational cohort (17% crossover to surgery and 3% crossover to nonsurgical care).
Is the paper free of selective outcome reporting?	Yes	

What is the overall risk of bias?	Moderate	Moderate - Some of the criteria have been fulfilled and those criteria that have not been fulfilled may affect the conclusions of the study.
--	----------	--

Results.

"The authors enrolled 304 patients in the randomized cohort and 303 in the observational cohort. In both cohort studies >95% of patients underwent decompressive surgery with fusion. Patients in the trial underwent the following nonsurgical care physical therapy (42%), epidural steroid injections (45%), non-steroidal anti-inflammatory drugs (51%), and opioids (34%). The rate of crossover in the randomised cohort was 49% at two years. In the observational cohort 97% of patients choosing surgery, underwent surgical treatment in the first year. Of those initially choosing non-surgical treatment, 17% underwent surgery by 1 year and 25% by 2 years

Pain

For the randomised cohort, the treatment effects at 2 years were 1.5 for SF-36 bodily pain (95% confidence interval [CI], -4.2 to 7.3; P = 0.52), whilst the combined cohort treatment effects at 2 years was 18.1 for SF-36 bodily pain (95% CI, 14.5 to 21.7). When comparing between the two groups, the effect was 17.8 (95% CI, 12.5 to 23.0) in the randomised cohort as compared with 18.5 (95% CI, 13.4 to 23.6) in the observational cohort.

Function

For the randomised cohort, there was no significant difference between fusion and conservative management, 1.9 for physical function (95% CI, -3.7 to 7.5; P = 0.71) and 2.2 for the Oswestry Disability Index (95% CI, -2.3 to 6.8; P = 0.68). The combined cohort treatment effect at 2 years was 18.3 for physical function (95% CI, 14.6 to 21.9), and -16.7 for the Oswestry Disability Index (95% CI, -19.5 to -13.9).

When comparing between the two groups for the SF-36 physical function, the effect was 16.7 (95% CI, 11.4 to 22.1) in the randomized cohort as compared with 19.9 (95% CI, 14.8 to 24.9) in the observational cohort; and for the Oswestry Disability Index, the effect was -15.9 (95% CI, -20.2 to -11.7) in the randomized cohort as compared with -17.7 (95% CI, -21.6 to -13.7) in the observational cohort.

Overall, the as-treated analysis for both cohorts combined showed a significant advantage for surgery at 3 months that increased at 1 year and diminished only slightly at 2 years. The as treated analysis showed treatment effects at 2 years were 18.1 for bodily pain (95% CI, 14.5 to 21.7), 18.3 for physical function (95% CI, 14.6 to 21.9), and -16.7 for the Oswestry Disability Index (95% CI, -19.5 to -13.9)."

See original article for Tables and Figures.

Author's Conclusions.

"In nonrandomised as-treated comparisons with careful control for potentially confounding baseline factors, patients with degenerative spondylolisthesis and spinal stenosis treated surgically showed substantially greater improvement in pain and function during a period of 2 years than patients treated non-surgically"

Our Comments/Summary.

In patients with degenerative spondylolisthesis, the intention-to-treat analysis found no significant advantage for surgery over non-surgical care, but the analysis was severely limited by treatment crossover. Given the limitations of this study and the paucity of evidence there is insufficient evidence to confirm whether spinal fusion is as effective as conservative management in people with chronic low back pain with degenerative spondylolisthesis.