Health Technology Assessment

Spinal Fusion and Discography

In Chronic Uncomplicated Lumbar Degenerative Disc Disease

Date: September 18, 2007

Presented by:

ECRI Institute
The Discipline of Science. The Integrity of Independence.

Health Technology Assessment Program
P.O. Box 42712
Olympia, WA 98504-2712
http://www.hta.hca.wa.gov

ECRI Institute
5200 Butler Pike
Plymouth Meeting, PA 19462
Phone: (610) 825-6000
Fax: (610) 834-1275

Spinal Fusion and Discography in Chronic Uncomplicated Lumbar Degenerative Disc Disease

Contract No. 6020-001788

Task Order No. 2

September 18, 2007
DRAFT (Not for Distribution)

Prepared for:
Washington State Health Care Authority
676 Woodland Square Loop, SE
Olympia, Washington 98504-2712

Spinal Fusion and Discography in Chronic Uncomplicated Lumbar Degenerative Disc Disease

Policy Statement

This systematic review is distributed solely for the purpose of pre-release peer review. It has not been otherwise disseminated by Washington State Health Care Authority.

Prepared by:
Karen Schoelles, M.D., S.M.
James Reston, Ph.D., M.P.H.
Jonathan Treadwell, Ph.D.
Meredith Noble, M.S.
David Snyder, Ph.D.

Table of Contents

Tables

Figures

Executive Summary

This technology assessment was commissioned by the Washington State Health Technology Assessment Program for use by the Health Technology Clinical Committee (HTCC). The HTCC uses evidence, primarily as assessed in this report to determine whether health technologies are safe, effective, and cost effective, and therefore should be covered by state programs that pay for health care.

This report evaluates relevant published research describing use of lumbar fusion and discography in patients with chronic uncomplicated degenerative disc disease (DDD). ECRI Institute's technology assessment provides an independent, in-depth, formal evaluation of the strength of evidence for the safety and efficacy of lumbar fusion for the treatment of DDD associated with chronic low back pain. This report also evaluates the role of discography prior to lumbar fusion in this patient population. It is based on systematic review of the published, peer-reviewed scientific literature and methodological precepts described in Appendix C.

The degeneration of intervertebral discs is associated with altered biomechanics of adjacent vertebrae, musculature, and connective tissue, and with back pain and sciatica.(1) Discs are present between lower cervical (neck) vertebrae, thoracic (mid-back) vertebrae, and low back (lumbar) vertebrae. Discs at any level can degenerate and cause pain, but this most often occurs at cervical and lumbar levels, where there is the greatest amount of mobility. Patients with DDD in the absence of chronic low back pain would not be considered candidates for lumbar fusion.

The clinical presentation of low back pain may prompt the clinician to order diagnostic imaging. Since disc degeneration does not cause pain in all individuals, imaging alone cannot be considered diagnostic. However, a clinical diagnosis of discogenic back pain can be confirmed with radiological imaging. Both plain films and magnetic resonance imaging (MRI) can aid the clinician in confirming their diagnosis. Typical findings include disc space collapse, endplate sclerosis, and vacuum disc phenomenon.(2) On a MRI, disc dehydration, high intensity zones, and endplate edema may also be evident.(2) Discography has also been used to support a diagnosis of discogenic back pain (see below for background on discography).

Low back pain has been called "the leading cause of pain and disability in adults in North America."(3) It was the most common cause of disability in persons younger than 45 in the U.S. in 2005.(4) It causes the most loss of productivity of any medical condition.(4) Only upper respiratory complaints cause people to miss more days of work annually.(4) In the United States, an estimated range of 8-56% of the population (the reason for this variation is unclear, but may be due to differences in diagnostic criteria or definition) experiences lower back pain every year, and the lifetime incidence rate is reportedly between 65% and 80%.(5) 2.4 million people are disabled because of low back pain, 1.2 million of them chronically.(4) Most patients improve within weeks; only 5-10% of people with low back pain develop chronic back pain.(5,6) Among U.S. physician office visits for low back pain, nonspecific backache accounts for 57% (more than 17 million visits), degenerative changes account for 12.5% (3.7 million visits), and herniated discs account for 11.1% (3.3 million visits).(7)

Low back pain is typically managed conservatively for at least six months before surgery is considered. Rest is usually only recommended for the first couple days of onset.(4) A variety of conservative treatments can be tried, including back education, cognitive behavioral therapy,

physical therapy, exercise, weight reduction, and alternative therapies (e.g., chiropractic manipulation), medications, and epidural injections.(2,8)

When conservative treatments fail after at least six months, spinal fusion may be considered. Between 150,000 and 250,000 people in the U.S. undergo lumbar spinal fusion surgery annually to treat discogenic back pain.(9) The goal of spinal fusion (also known as spinal arthrodesis) is to permanently immobilize the spinal column vertebrae surrounding the disc(s) that is (are) diagnosed as causing discogenic low back pain. Immobilizing the vertebrae is believed to reduce pain by limiting painful movement (e.g., pinching of nerves or rubbing of bone on bone) that may occur as degenerated discs subside. Spinal fusion is also used to treat other painful conditions, including spondylolisthesis (forward displacement of one of the lower lumbar vertebrae over the vertebra below it or on the sacrum), trauma resulting in spinal nerve compression, abnormal spinal curvatures (scoliosis or kyphosis), and vertebral instability caused by infections or tumors. Several surgical procedures may be used to achieve spinal fusion in patients with discogenic low back pain. They differ by surgical approach and instrumentation used. Some fusion strategies may be particularly appropriate for certain patient populations. All methods have advantages and disadvantages.

The role of lumbar discography in selection of patients as surgical candidates is controversial. Discography is a diagnostic procedure in which contrast material is injected into the nucleus pulposus of a lumbar disc. The general intent is to determine whether the disc itself is the source of pain (i.e., a diagnosis of discogenic pain). This diagnosis has been used to justify the need for surgical intervention involving discectomy and lumbar fusion. Thus, discography may influence important decisions about the appropriateness of surgical intervention.

Discography yields two types of results: pain provocation (whether the patient's typical pain was reproduced by the injection), and morphology (whether the dye images an abnormal pattern in the disc, often based on CT scan). Controversy exists about the relative importance of these two test results. Some authors(4,10) assign much greater importance to pain provocation; for example, Bogduk (1996)).(10) stated that "the morphology of the disc as revealed by discography is essentially irrelevant." By contrast, Buenaventura et al. (2007) cited disc morphology as the gold standard for discogenic pain, stating that "the imaging information is important since treating an anatomically normal disc, irrespective of its ability to cause pain, seems unethical."(11) Walsh (1990) proposed that a discography result should only be considered positive if the patient's typical pain was reproduced *and* the morphology was abnormal.(12) The extent of spread of the contrast material from the nucleus pulposus determines disc morphology. The Dallas Discogram Description categorizes several levels of disruption of the disc annulus, ranging from Grade 0 (normal) to Grade 5 (highest level of disruption).(13,14)

One major concern about discography is the rate of false positive results. Several authors have found that among people with no previous pain, the discography result can be positive.(15-22) Also, discography in lumbar discs has been reported to reproduce pain known to originate elsewhere in the body.(23) Various solutions have been proposed for these phenomena, including a more stringent definition of a positive test to require both typical pain provocation and abnormal morphology (Walsh definition),(12) the requirement that adjacent discs test negative,(24,25) and the avoidance of high pressure (≥22 pounds per square inch).(16) Carragee et al. (2006) found, however, that even when all of these conditions were met, the rate of false positives was still 25%.(26) Many have suggested that the origin of many false positives

lies with the psychological status of some patients; a positive discography may be more likely in patients with psychological comorbidities who are predisposed to report pain.(17-20,22,23,27,28)

The analysis of evidence in this assessment is divided into two sections: Part I evaluates evidence comparing outcomes of lumbar fusion and nonsurgical treatments, while Part II evaluates evidence concerning the role of discography prior to lumbar fusion. We examined the evidence in the context of six clinical questions (three for Part I and three for Part II). Our strength of evidence ratings take into consideration not only the individual study quality for relevant outcomes, but also the quantity, consistency, and robustness of the evidence, in addition to the magnitude of observed effects. The instruments used to rate individual study quality appear in Appendix C, along with our system for rating the strength of evidence.

Part I – Lumbar fusion surgery and nonsurgical treatments for chronic lumbar back pain

1) Does lumbar fusion surgery reduce pain and improve functional status/quality of life more effectively than nonsurgical treatments?
2) What are the rates of adverse events (perioperative, long-term events, and reoperations) for lumbar fusion surgery and nonsurgical treatments?
3) What patient characteristics (i.e., workers' compensation population, patients with chronic pain, psychological distress, and age-groups) are associated with differences in the benefits and adverse events of lumbar fusion surgery?

Part II – Role of discography prior to lumbar fusion surgery

4) In patients being considered for lumbar fusion surgery, what is the reliability of discography?
    a. Test-retest reliability
    b. Inter-reader reliability
5) In patients undergoing lumbar fusion surgery, do the results of pre-surgical discography predict the degree of pain reduction or improvement in functional status/quality of life after lumbar fusion surgery?
6) In patients being considered for lumbar fusion surgery, do patients who receive discography that influences the treatment choice have better treatment outcomes than patients who do not receive discography?

Part I - Lumbar fusion surgery and nonsurgical treatments for chronic lumbar back pain

Overall, 30 articles reporting on 27 studies were included to address the clinical questions in Part I. Four randomized controlled trials (RCTs) that enrolled a total of 767 patients met the inclusion criteria for Key Question 1, which required a comparison of lumbar fusion to non-operative treatment in patients with DDD. These same RCTs also reported treatment complications and therefore also met the inclusion criteria for Key Question 2. In addition to the four RCTs described above, 23 studies with a total of 5,639 patients also met the inclusion criteria for Key Question 2. These studies were either case series of lumbar fusion or controlled studies (some randomized) that compared different lumbar fusion procedures. Data from one

separate publication of one RCT (also included in Key Question 1 and 2) that enrolled 294 patients met the inclusion criteria for Key Question 3.

The primary outcomes of interest addressing Key Question 1 are functional status measured by the Oswestry Disability Index (ODI), back pain measured by a visual analog scale (VAS), and quality of life measured by a previously validated instrument; the only instrument used to measure quality of life in the available evidence base was the short-form (SF)-36 questionnaire. The ODI is comprised of 10 questions on pain and pain-related disability in activities of daily life and social participation. Each question has six response alternatives, and the overall score ranges from 0 (no disability) to 100 (totally disabled or bedridden). The VAS for back pain is also scored from 0 (no pain) to 100 (worst pain imaginable). A recent study calculated the minimal clinically important difference for the ODI and VAS of back pain using linear regression analysis of score change compared to pre-treatment scores. The authors determined that the minimal clinically important difference for the ODI was 10, and for the VAS of back pain it was 18-19.(29) Accordingly, we used a difference of 10 for the ODI and a difference of 20 for the VAS as the minimal clinically important difference in our assessment of these outcomes. The SF-36 is scored from 0 (worst health state) to 100 (best health state); we used a difference of 5 in the SF-36 as the minimal clinically important difference based on data from an earlier study that investigated this issue.(30)

A quality rating (and strength of evidence rating) was applied only to studies comparing lumbar fusion to non-operative therapy in Key Question 1. The remaining studies addressing Key Question 2 were not used to address comparative event rates of fusion and non-operative care; they were used only to provide additional data on adverse events and adverse event rates for lumbar fusion. Due to variability in the way complications are reported among different studies, lists of complications do not lend themselves to evidence ratings.

Our detailed assessments of the quality of the RCTs addressing Key Question 1 appear in Table 13 of Appendix D. The average quality of the studies was moderate due to several limitations, most notably lack of blinding of patients, providers, and outcome assessors (for the majority of outcomes) in all studies. This could lead to biased interpretation or reporting of outcomes, particularly of subjective outcomes. Two of the studies were further limited because more than 15% of patients did not receive their assigned treatment, either because they crossed over to the alternative treatment group or did not receive any of the trial treatments. Crossover to alternative treatments would tend to diminish a between-group difference in treatment outcome if it exists. Another potential limitation was differences between groups in additional treatments received during the trials (most trials did not record this information).

The average age of patients in all four RCTs was about 40-45 years, and the average age of patients in the additional 23 studies that addressed Key Question 2 ranged from 39 to 54 years, which is representative of the age at which most patients with degenerative disease undergo surgery in clinical practice. The proportion of patients receiving workers' compensation varied considerably (ranging from 21% to 94%) in the 12 studies that reported this information. Although the types of fusion procedures varied among different studies, all studies used fusion procedures that are currently employed in clinical practice.

Results and conclusions (Part I)

1. Does lumbar fusion surgery reduce pain and improve functional status/quality of life more effectively than nonsurgical treatments?

   ECRI Institute evidence assessments:

   We did not find sufficient evidence that lumbar fusion surgery is more effective to a clinically meaningful degree than nonsurgical treatments for any of the following patient populations, comparisons and outcomes:

   - Meta-analysis of postoperative changes in Oswestry disability scores from two moderate quality RCTs (n = 413 patients) revealed no clinically meaningful difference between fusion and intensive exercise/rehabilitation plus cognitive behavioral therapy (CBT) in patients without prior back surgery (95% CI 0.2 to 7.5, *a priori* 10 point difference defined as clinically meaningful), although the difference slightly favored fusion. Strength of evidence: Weak.
   - The evidence was insufficient to determine whether lumbar fusion provides a greater improvement in back pain (one moderate-quality RCT, n = 64 patients) or quality of life (no acceptable evidence) compared to intensive exercise/rehabilitation plus CBT in patients without prior back surgery.
   - The evidence from one moderate quality RCT (n = 60 patients) was insufficient to determine the relative benefits of lumbar fusion compared to intensive exercise/rehabilitation in patients with prior back surgery.
   - The evidence from one moderate quality RCT (n = 294 patients) was insufficient to determine the relative benefits of lumbar fusion compared to conventional physical therapy in patients with or without prior back surgery.

The four trials that met our inclusion criteria for this question differed in potentially important ways. Based upon independent assessment by two methodologists, we assumed that one difference that was likely to create variation in the effect size among trials was the intensity of non-operative therapy in the control groups. Three trials (Brox et al. 2003; Brox et al. 2006; Fairbank et al. 2005) used more intensive exercise/rehabilitation with cognitive behavioral strategies, while the remaining trial (Fritzell et al. 2001) used non-intensive physical therapy as the main component of an unstructured nonsurgical treatment program. The more intensive therapy seems more likely to benefit patients than the less intensive treatment (which patients had undergone without improvement prior to enrollment). If the amount of patient benefit from surgery is assumed to be the same in all studies, then one would expect a greater difference in patient benefit between patients treated surgically and patients treated with conventional physical therapy compared with patients treated surgically and patients treated with multidisciplinary and intensive exercise/rehabilitation. This is important to our analysis because the mean difference measures the difference between treatment and control groups. Therefore, the mean difference would vary depending on the control selected, causing heterogeneity (differences) in study findings. For this reason, the data from Fritzell et al. were not combined with data from the other three trials.

Another factor that might create heterogeneity among effect sizes is whether the patients had back surgery before enrolling in the studies in question. Patients with prior back surgery may be

less likely to benefit from further surgery than patients who have never had back surgery. One of the three trials that used intensive exercise/rehabilitation (Brox et al. 2006) included only patients who had undergone prior surgery for disc herniation (most likely discectomy or laminectomy, as none of the patients had undergone prior lumbar fusion). The authors mentioned that "the prognosis after a second operation is generally considered poor compared with the prognosis in patients without previous surgery for disc herniation".(31) Of the remaining two trials, Brox et al. (2003) included no patients with prior back surgery, while Fairbank et al. (2005) had a small proportion of patients (8%) who had undergone prior laminectomy. Based upon the differences in the patient populations, we determined that the data from Brox et al. (2006) should not be combined with data from the remaining two trials.

Although the control therapies and patient characteristics were similar in the trials by Brox et al. (2003) and Fairbank et al. (2005), the two trials differed in the types of fusion performed and the length of followup. Brox et al. (2003) exclusively used posterolateral fusion (PLF) with pedicle screws, while Fairbank et al. (2005) used an unspecified variety of fusion procedures. Also, Brox et al. reported treatment outcomes at one year of followup, while Fairbank et al. reported treatment outcomes at two years of followup. However, we considered differences in the fusion procedure and length of followup less likely to create heterogeneity in effect sizes than the other factors described above. Therefore, we determined that combining the data from these two trials was appropriate.

The four RCTs were therefore analyzed in three separate groups: fusion versus intensive exercise/rehabilitation plus CBT – divided into patients without prior back surgery (Brox et al. 2003, Fairbank et al. 2005) and patients with prior back surgery (Brox et al. 2006) – and fusion versus non-intensive physical therapy (Fritzell et al. 2006).

Fusion versus Intensive Exercise/Rehabilitation Plus CBT in Patients without Prior Back Surgery

Two multicenter RCTs with a total of 413 patients compared intensive exercise/rehabilitation with cognitive behavioral therapy to fusion in patients who had not undergone back surgery before. Both studies reported the between group difference in the pre-post change in ODI score (see Brox et al. 2003 and Fairbank et al. 2005 in Table 13, Appendix D). Both studies also reported the change scores adjusted for baseline values by analysis of covariance (ANCOVA); this is the best method for adjusting for imbalances in patient characteristics.(32) Thus, our analysis is based on the adjusted change scores.

As described above, these studies were considered suitable for a combined data analysis (meta-analysis), so the change score data were combined in a random effects meta-analysis. As shown in Figure 3, fusion led to a small but statistically significant increase in ODI change scores compared to intensive exercise/rehabilitation plus CBT; however, the upper 95% confidence limit (7.5) was below the minimum level that is considered clinically significant (ODI = 10). We therefore conclude that changes in ODI scores did not show a clinically meaningful difference between fusion and intensive exercise/rehabilitation plus CBT in patients without prior back surgery, although the difference slightly favored fusion (95% CI 0.2 to 7.5). Because the evidence base is of moderate quality and limited quantity, the strength of evidence supporting this conclusion is weak.

Only one of these studies (Brox et al. 2003) evaluated VAS back pain. This study reported no statistically significant difference in change in VAS scores between patients undergoing fusion

and patients undergoing intensive exercise/rehabilitation plus CBT. Because the 95% CI overlapped with zero and the boundary of minimum clinical significance, the evidence is insufficient to allow a conclusion for this outcome.

Although one of these studies measured quality of life using the SF-36 instrument, this outcome was excluded from analysis because <80% of patients completed the instrument.

Fusion versus Intensive Exercise/Rehabilitation Plus CBT in Patients with Prior Back Surgery

One RCT (Brox et al. 2006) with 60 patients studied the efficacy of exercise/rehabilitation plus cognitive behavioral therapy to fusion in patients who had previously undergone back surgery. This study reported the between-group difference in the pre-post change in ODI score, using ANCOVA to adjust for baseline between-group differences in gender and treatment expectations (see data in Table 13, Appendix D). The adjusted comparison showed a trend favoring a larger change in ODI in the control group. However, the results were inconclusive because the 95% CI overlapped with zero (not statistically significant) as well as the boundary of clinical significance (ODI = -10), meaning the true difference (if one exists) could favor either treatment. Thus, the evidence is insufficient for a conclusion regarding the relative benefit of fusion versus intensive exercise/rehabilitation plus CBT in patients with prior back surgery.

This same study reported no statistically significant difference in change in VAS scores between patients undergoing fusion and patients undergoing intensive exercise/rehabilitation plus CBT. Although the 95% CI did not overlap with the boundary of minimum clinical significance, the results of a single moderate quality study are insufficient to allow a conclusion for this outcome.

Fusion versus Non-intensive Physical Therapy in Patients without Prior Back Surgery

One RCT (Fritzell et al. 2001) with 294 patients addressed this comparison; however, a minority of patients (18.7%) had prior discectomy. This study reported ODI pre-post change scores for each comparison group (see data in Table 13, Appendix D). A significantly larger improvement in ODI was observed in the fusion group compared to the physical therapy group (11.6 vs 2.8, $p = 0.015$); group changers were included in the analysis of difference (although not in their tabled data). However, although the difference in change is statistically significant, the mean difference in change between groups (ODI = 8.8) is below the level of clinical significance (ODI = 10). Because this is a single trial of moderate quality, the evidence is insufficient to allow a conclusion for this comparison.

This same study reported a statistically significant difference in the change in VAS score favoring fusion when compared to non-intensive physical therapy. However, the mean difference between groups (16.7) did not exceed the boundary of minimum clinical significance for VAS back pain (difference = 20). Because this study did not include group changers in their tabled data, we cannot be certain of the difference if group changers had been included. In any event, because this is a single study of moderate quality without a large effect, the evidence is inconclusive for this outcome.

2. What are the rates of adverse events (perioperative, long-term events, and reoperations) for lumbar fusion surgery and nonsurgical treatments?

- Lumbar fusion leads to higher rates of both early and late adverse events compared to non-intensive physical therapy or intensive exercise/rehabilitation plus CBT.
- None of the four RCTs comparing fusion to non-intensive physical therapy or intensive exercise/rehabilitation plus CBT reported any adverse events occurring in patients who only received non-operative care. Most of the reported adverse events for patients in the surgical group could not have occurred in patients who did not undergo surgery (e.g., surgical complications).
- Categories of adverse events most frequently reported in fusion studies include reoperation (18/27 studies), infection (14/27 studies), various device-related complications (13/27 studies), neurologic complications (12/27 studies), thrombosis (11/27 studies), bleeding/vascular complications (10/27 studies), and dural injury (10/27 studies).
- The ranges of rates of the most frequently reported complications in fusion studies were: reoperation (0% to 46.1%), infection (0% to 9%), device-related complications (0% to 17.8%), neurologic complications (0.7% to 25.8%), thrombosis (0% to 4%), bleeding/vascular complications (0% to 12.8%), and dural injury (0.5% to 29%).

All four RCTs with 767 patients that met our inclusion criteria for Key Question 1 compared adverse event rates for lumbar fusion surgery and nonsurgical treatments. None of the trials reported the rate of total adverse events (from intraoperative to last followup). Instead, they generally divided complication rates by time of occurrence.

Two trials (Brox et al. 2003, Fritzell et al. 2001) separately reported "early" (usually meaning perioperative) and "late" complications (which either occur at a later time or are persistent or permanent). Fritzell et al. defined early as within the first two weeks post-treatment, while Brox et al. did not report the cutoff time for early complications (although it likely did not exceed one month). Another trial by Brox et al. (2006) appeared not to report all early complications; the authors stated that "early complications included two wound infections among the 23 operated patients", but no other early complications are mentioned. Thus, we cannot be certain that these were the only early complications. However, the authors stated that no late complications occurred. The remaining trial (Fairbank et al. 2005) divided adverse events into intraoperative (during surgery) and post-operative (any time after surgery) categories, which is a somewhat different division than early and late. The only postoperative complications mentioned were need for reoperation; we cannot be certain that there were no late complications that did not require reoperation.

All trials calculated adverse event rates on a per protocol basis, meaning only patients who actually received surgery were included in calculations of surgical adverse events. This is the most conservative approach for analysis of adverse events; calculations on an intent-to-treat basis would underestimate the surgical complication rate, as some patients assigned to surgery never received it.

The results for overall early adverse events appear in Table 15, Appendix D. Despite variation in types of fusion and nonsurgical therapies used in these studies, the four trials had one factor in common; none of them identified any adverse event resulting from nonsurgical treatment

(intensive exercise/rehabilitation plus CBT in three trials, non-intensive physical therapy in one trial). The three trials that reported overall intraoperative or early adverse event rates found similar rates (range 12.7% to 18%) despite differences in the time period observed (intraoperative to one month). The differences between early adverse events in the surgical versus physical therapy groups was statistically significant in all three of these trials. The reported early adverse events in the surgical groups included bleeding, thrombosis, wound infection (deep and superficial), neurological (pain, sympathetic cord damage) complications, device-related (problems with screws or implants) complications, reoperations for various causes, and others (dural tears, peritoneal tears). A complete list of reported early complications and their occurrence rates in these trials appears in Table 17, Appendix D (note: some complications in this table may not be early; most studies did not report time cutoffs for the complications). Most of these complications could not have occurred in the absence of surgery.

Overall late adverse event rates showed more variation among studies, ranging from 0% to 7.4% (Table 15, Appendix D). A number of factors might account for this variation. It could have resulted from differences in the length of followup; the two trials with only one-year followup reported no late events, while the two trials with two-year followup reported that 6.2% and 7.4% of patients who underwent fusion had late events (in both trials, the difference in event rates between surgical and nonsurgical patients was statistically significant). The size of the trials may also have influenced these differences, as the two trials with one-year followup were also much smaller than the other trials, and therefore less likely to detect less common adverse events. A third factor is that the authors of these trials may have had different definitions of what constitutes an adverse event. Reported late adverse events most frequently included reoperations for various problems (mostly infections and pseudoarthroses) and continuing pain at the donor site from bone graft harvesting. Specific causes of reoperations and other late complications and their rates are listed in Table 18, Appendix D. Again, these events could not have occurred in the absence of surgery.

We examined additional studies of lumbar fusion that lacked a non-operative control group to determine whether these studies report adverse events not reported in the four RCTs described above, and also to determine if the adverse event rates differed from those reported in the RCTs. We selected studies with at least 100 patients total that received any type of lumbar fusion procedure and met all of our other inclusion criteria.

Twenty-three studies with a total of 5,639 enrolled patients met our criteria for this question. Fourteen of these studies were prospective studies; of these 14, six were randomized trials comparing different fusion procedures (a comparison not addressed in this report). The remaining studies were retrospective. Some studies focused only on specific adverse events such as need for reoperation, while others reported all adverse events that occurred during the course of the study. Only eight studies reported any type of overall adverse event rates (operative, postoperative, total, etc.), and the studies varied considerably in the manner in which these events were summarized (Table 16, Appendix D). Because a patient may experience more than one adverse event, we could not calculate the percent of patients experiencing any adverse event when studies only reported rates for specific adverse events. These studies also showed considerable variation in the types of fusion procedures performed, which may contribute to variation in the types of adverse events that occurred in different studies.

A concise summary of reported ranges of specific adverse event rates appears in Table 4. These ranges combine data from the four RCTs described earlier with data from the 23 additional

studies. In this table, we do not attempt to separate early from late events, as several studies did not report the specific time of occurrence for each event. Categories of adverse events most frequently reported in fusion studies include reoperation (18/27 studies), infection (14/27 studies), neurologic complications (12/27 studies), thrombosis (11/27 studies), bleeding/vascular complications (10/27 studies), and dural injury (10/27 studies). Death related to surgery was relatively rare, occurring only in 4/27 studies with a maximum reported rate of 2% (we assumed no deaths related to surgery occurred in the other 23 studies). Certain adverse events showed substantial variation in reported rates: these include reoperation (0% to 46.1%), dural injury (0.5% to 29%), neurologic complications (0.7% to 25.8%), and device-related complications (0% to 17.8%). Reported rates in the four RCTs comparing fusion to non-operative care were either at the low end (0% for death) or within the indicated ranges but below the maximum reported rate.

Complete information on the rates of all adverse events reported in these studies is summarized in Tables 19 and 20, Appendix D.

3. What patient characteristics (i.e., workers' compensation population, patients with chronic pain, psychological distress, and age-groups) are associated with differences in the benefits and adverse events of lumbar fusion surgery?

   ECRI Institute Evidence Assessment:

   - The evidence from one moderate-quality RCT (n = 294 patients) is insufficient to determine what patient characteristics are associated with differences in the benefits and adverse events of lumbar fusion surgery.

One RCT (Hagg et al. 2003) with 294 patients met the inclusion criteria for this question. This was another publication derived from the Swedish Lumbar Spine Study originally described in Fritzell et al. (2001). The efficacy and safety findings of Fritzell et al. were discussed under Key Questions 1 and 2. In their subsequent publication, Hagg et al. presented data concerning prognostic factors that was not included in Fritzell et al. Hagg et al. conducted a multivariate analysis to identify factors that predicted various outcomes of treatment in the surgical and nonsurgical (non-intensive physical therapy) patient groups. The main outcome measures in their analysis included change of disability (measured as ≥50% reduction of the ODI score), patient global assessment of treatment effect (improvement/no improvement), and work status at followup. Stepwise, forward multiple logistic regression analyses were performed within each treatment group, with the outcomes as dependent variables.

As shown in Table 21 (Appendix D), only one patient characteristic (neurotic personality) showed a statistically significant association with change in disability in the surgical group; patients with neurotic personalities were less likely to show improvement in the ODI score. No patient characteristic was significantly associated with improvement in ODI score in the nonsurgical group.

The study also identified patient characteristics significantly associated with the patient global assessment (improved or not improved). In the surgical group, neurotic personality was again associated with poor outcome (less likely to be improved), while disc height <50% was significantly associated with improvement. In the nonsurgical group, one patient characteristic (depressive symptoms) was significantly associated with poor outcome. No other factors were significantly associated with patient global assessment in either group.

Certain patient characteristics were significantly associated with work status at followup in both groups. Among surgical patients, older age and longer period of current sick leave were significant predictors of not working at followup. Among nonsurgical patients, only longer period of current sick leave was significantly associated with not working at followup. No positive predictors of working at followup were identified for either patient group.

The following variables did not show significant associations with any of the three outcomes at followup: pain (multiple measures), clinical findings (multiple measures), sociodemographics (disability pension, workers' compensation, unemployment, heavy job, comorbidity, smoking, prior surgery, gender, or marital status), other psychological measures (pain behavior, personality disorders), or radiographic indicators.

Although not specifically stated in the text of the study, it appears that patients who changed treatment groups after enrollment were not included in the analyses described above. The effect this might have on the observed associations is unknown.

Although multicenter, this was a single study of moderate quality; furthermore, none of the observed associations were large effects. Therefore, the evidence is insufficient to allow a conclusion regarding patient characteristics associated with differences in the benefits and adverse events of lumbar fusion surgery.

Part II – Role of discography prior to lumbar fusion surgery

Overall, six studies were included to address the clinical questions in Part II.

Results and Conclusions – Part II

4. In patients being considered for lumbar fusion surgery, what is the reliability of discography?
    a. Test-retest reliability
    b. Inter-reader reliability

    ECRI Institute Evidence Assessment:

    - The evidence was insufficient to permit conclusions about the reliability of discography for patients with chronic uncomplicated lumbar degenerative disc disease.

Two studies met the inclusion criteria for this Key Question[1].(33,34) Agorastides (2002)(33) reported data on both test-retest reliability and inter-rater reliability (133 discs in 72 patients), whereas Milette (1999)(34) only reported data on inter-rater reliability (132 discs in 45 patients).

Both studies investigated at least one specific type of reliability: whether a given discogram is judged to have the same morphology grade by the same reader at different times (i.e., test-retest) or by different readers (i.e., inter-rater). Notably, neither study performed two discography exams on the same disc to determine whether the results were consistent between discography injections. Also, neither study investigated the reliability of patients' reports of pain provocation or similarity to their typical pain. These types of reliability represent additional potential sources

[1] After finding only two studies, we removed the date requirement (that studies must have been published in 1990 or later), but when we examined earlier studies, none of them met the other inclusion criteria.

of variability in discography examinations that have not been assessed in patients with chronic uncomplicated lumbar degenerative disc disease.

We rated the quality of both studies as moderate (quality scores of 7.1 and 7.9). Both studies used consecutive enrollment, reported data on all or almost all enrolled patients, and the discograms were read without consultation of prior discograms or other clinical information about the patient. However, both were retrospective studies that did not report the funding source, and also the Agorastides study did not report whether patient inclusion/exclusion criteria were applied consistently to all patients.

For test-retest reliability, the Agorastides study observed good reliability (values for kappa ranging from 0.80 to 0.85 for the three raters),[2] but because it was a single moderate-quality study at a single center, we deemed this evidence insufficient quantity to permit conclusions. For inter-rater reliability, neither study observed large reliability (values for kappa ranging from 0.66 to 0.77), and neither study was multicenter. These factors, considered together with the moderate quality and limited quantity, mean that the evidence base was insufficient to permit conclusions.

5. In patients undergoing lumbar fusion surgery, do the results of pre-surgical discography predict the degree of pain reduction or improvement in functional status/quality of life after lumbar fusion surgery?

   ECRI Institute Evidence Assessment:

   - Because of low quality and heterogeneous results from three studies (n = 330 patients), the evidence was insufficient to permit conclusions about the use of discography to predict fusion outcomes in patients with chronic uncomplicated lumbar degenerative disc disease.

This question involves a comparison in surgical outcomes between those who had a positive discography before surgery and those who had a negative discography before surgery. Three studies met the inclusion criteria.(36-38) Willems (2007)(36) included 82 patients, Gill (1992)(37) included 53 patients, and Colhoun (1988)(38) included 195 patients.

Importantly, the three studies each used a different definition of a "positive" discography test:

- Willems (2007)(36) categorized two groups of patients based on *typical pain provocation* in *adjacent-disc(s)*: 1) patients whose adjacent lumbar disc(s) provoked typical pain on discography (N = 22); and 2) patients whose adjacent lumbar disc(s) did not provoke typical pain (or no pain) on discography (N = 60).
- Gill (1992)(37) categorized three groups of patients based on the *morphology* of the *suspected disc*: 1) annular tear beyond the periphery (N = 20); 2) annular tear and contrast extension to the periphery, but not beyond (N = 19); and 3) small annular tear that did not extend to the periphery (N = 14).

[2] Kappa measures chance-corrected agreement. 0 represents chance, and 1 represents perfect agreement. The standard interpretation of kappa values is that Below 0.0 is Poor agreement; 0.00-0.20 is Slight agreement; 0.21-0.40 is Fair agreement; 0.41-0.60 is Moderate agreement; 0.61-0.80 is Substantial agreement; 0.81-1.00 is Almost Perfect agreement.(35)

- Colhoun (1988)(38) categorized four groups of patients based on both *typical pain provocation and morphology* of the *suspected disc:*1) typical pain provocation and abnormal morphology (N = 137); no pain provocation and abnormal morphology (N = 25); 3) neither pain provocation nor abnormal morphology (N = 6); and 4) total disc resorption of contrast material thus morphology not assessable and pain provocation not reported (N = 27).

Also, the three studies assessed different surgical outcomes:

- Willems (2007)(36) reported mean VAS pain scores at followup as well as the percentage of patients who experienced at least 30% pain relief (at a mean followup 6.7 years)
- Gill (1992)(37) reported a composite outcome involving the percentage of patients showing "improvement on functional testing and pain report", which was based on three items (Oswestry Pain Questionnaire, VAS, and pain drawing) (at a mean followup 3 years)
- Colhoun (1988)(38) reported a composite outcome involving the percentage of patients who were considered a "success", which was defined as meeting all three conditions: 1) complete relief or significant subjective improvement in symptoms; 2) resumption of work and/or normal duties; 3) no intake of analgesics (at a mean followup 3.6 years).

Furthermore, the three studies reported qualitatively different results (the data appear in Table 33 of Appendix E):

- Willems (2007)(36) found evidence of no statistical difference in VAS pain scores at followup between the two groups, suggesting that discography results do not predict surgical outcomes.
- Gill (1992)(37) did not enroll enough patients to determine whether their data demonstrated a difference or no difference, leaving open the question of whether discography results predict surgical outcomes.
- Colhoun (1988)(38) found evidence of a difference in success rates, suggesting that discography results do predict surgical outcomes. Specifically, "success" was found to be more likely among patients with positive pain provocation and abnormal morphology (88%) than for other groups (52% to 85%).

We rated the quality of all three studies as low (with scores ranging from 4.1 to 4.3). All three were retrospective, non-randomized, unblinded studies. Only one of the three studies (Willems) reported baseline data to assess comparability of patient groups at baseline or attempted to enhance comparability using statistical methods.

Given the low quality, the different definitions of a positive discography, the different outcomes examined, and the qualitatively different results reported, we drew no conclusions about whether discography results predict surgical outcomes.

6. In patients being considered for lumbar fusion surgery, do patients who receive discography that influences the treatment choice have better treatment outcomes than patients who do not receive discography?

   ECRI Institute Evidence Assessment:

- No evidence of acceptable quality was available to address this question; thus, the evidence was insufficient to permit conclusions about the influence of discography on fusion outcomes in patients with chronic uncomplicated lumbar degenerative disc disease.

This question involves comparison of treatment outcomes between patients who had received discography before treatment and patients who had not received discography before treatment. Only one study met the inclusion criteria. Madan (2002)(39) retrospectively compared the surgical outcomes of two groups of patients at a single center: 32 patients who were seen between January 1998 and January 1999 and had a positive discography result; and 2) 41 patients who were seen prior to 1998 and had not received discography. All patients underwent the same surgical procedure (instrumented PLIF with posterolateral fusion).

Our quality assessment indicated that the study was very low quality (score 3.4), therefore we excluded the study from further consideration. The primary factors influencing this quality rating were a retrospective, non-concurrent, non-randomized, unblinded design in which the groups were not well-matched at baseline and authors had not attempted statistical methods that may have enhanced group comparability. Due to the lack of evidence of sufficient quality, we drew no conclusions about whether performing discography influences surgical outcomes.

Introduction

Degenerative Disc Disease and Low Back Pain

Description of Healthy Discs

Intervertebral discs form resilient fibrocartilaginous joints between vertebral bodies of the spinal column. These discs begin between the second and third cervical vertebrae and are present through the end of the lumbar portion of the vertebral column.

A disc is composed of two parts, the annulus fibrosus and the nucleus pulposus. The annulus fibrosus is the outer portion of the disc, and is composed of concentric fibrous rings of collagen fibers; it allows motion between the vertebral bodies and acts as a shock absorber. The annulus fibrosus encircles the nucleus pulposus, which is the gelatinous center of the disc, composed of radially arranged collagen and elastin fibers. The nucleus pulposus functions as an incompressible ball bearing that allows the vertebral bodies to roll forward and backward.(40)

Discs are the "joints" of the spinal column, primarily playing a mechanical role by transmitting forces resulting from body weight and activity. Their flexibility allows for multidirectional movement of the spine.

Degenerative Disc Disease (DDD)

The degeneration of intervertebral discs is associated with altered biomechanics of adjacent vertebrae, musculature, and connective tissue, and with back pain and sciatica.(1) Discs are present between lower cervical (neck) vertebrae, thoracic (mid-back) vertebrae, and low back (lumbar) vertebrae. Discs at any level can degenerate and cause pain, but this most often occurs at cervical and lumbar levels, where there is the greatest amount of mobility. However, degenerative disc disease (DDD) is not always associated with low back pain. This report focuses on DDD associated with chronic low back pain, as patients with DDD in the absence of chronic low back pain would not be considered candidates for lumbar fusion.

Low Back Pain Associated with DDD

Discogenic low back pain results when discs in the lumbar spine degenerate and cause pain. The pain may have a deep boring sensation, often with a distribution to the upper thighs and buttocks.(2) Patients with discogenic low back pain typically suffer from pain that is worsened with bending, twisting, squatting, or stooping, and possibly accompanied by leg pain or numbness(41), which may be relieved somewhat with a reclined position, such as with the legs elevated.(2)

Diagnosis

The clinical presentation of discogenic back pain (described above) may prompt the clinician to order diagnostic imaging. Imaging alone cannot be considered diagnostic, because disc degeneration does not cause pain in all individuals. However, a clinical diagnosis of discogenic back pain can be confirmed with radiological imaging. Both plain films and magnetic resonance imaging (MRI) can aid the clinician in confirming their diagnosis. Typical findings include disc space collapse, endplate sclerosis, and vacuum disc phenomenon.(2) On a MRI, disc

dehydration, high intensity zones, and endplate edema may also be evident.(2) Discography has also been used to support a diagnosis of discogenic back pain (see below for background on discography).(42)

Etiology of Low Back Pain Associated with DDD

Discogenic back pain is caused by degeneration of intervertebral disks and its sequelae. However, according to Urban and Roberts, the etiology of disc degeneration is "a difficult entity to study" because "its definition is vague, with diffuse parameters that are not always easy to quantify."(1) Other researchers have noted that DDD is a "nonspecific pathologic diagnosis" that is "very poorly defined" in the literature.(43)

The disc nuclei distribute forces equally throughout the annulus, transmitting a greater portion of loads to, and contributing to tears in, the annulus. Aging, decreased nutrition to the disc, genetic predisposition, and trauma may all play a role in the degeneration of vertebral discs. Age-related decreases in water content in the nucleus, along with changes in the structure of the collagen, make the disc more rigid, contributing to tearing. As degeneration progresses, collagen continues to break down, and larger tears form in the annulus. Material may fragment from these tears and bulge from the disc causing a "herniation."(41) Trauma may also contribute to herniation. Although commonly thought to cause pain by impinging on spinal nerves, disc herniation is now thought to possibly cause pain by activating an inflammatory cascade of irritating biochemical processes.(1)

When disc degeneration is present, the spinal column becomes unstable because the disc can no longer hold the vertebral bodies in their proper positions. Dysfunction results, which can lead to outer annular tears, separation of the endplate, cartilage destruction, and facet synovial reaction.(8) Pain may be the result because nearby paraspinal muscles, facet joint capsules, periosteum, intraspinal ligaments and tendons, and sacroiliac joints are innervated with nociceptive fibers.(2) If the articular cartilage between the discs erodes, it can also lead to damage of the joint and nearby ligaments.(1,41,44,45)

Epidemiology of Low Back Pain

Low back pain has been called "the leading cause of pain and disability in adults in North America."(3) It was the most common cause of disability in persons younger than 45 in the U.S. in 2005.(4) It causes the most loss of productivity of any medical condition.(4) Only upper respiratory complaints cause people to miss more days of work annually.(4)

In the United States, an estimated range of 8-56% of the population (the reason for this variation is unclear, but may be due to differences in diagnostic criteria or definition) experiences lower back pain every year, and the lifetime incidence rate is reportedly between 65% and 80%.(5) 2.4 million people are disabled because of low back pain, 1.2 million of them chronically.(4) Most patients improve within weeks; only 5-10% of people with low back pain develop chronic back pain.(5,6)

Among U.S. physician office visits for low back pain, nonspecific backache accounts for 57% (more than 17 million visits), degenerative changes account for 12.5% (3.7 million visits), and herniated discs account for 11.1% (3.3 million visits).(7)

Natural History of Low Back Pain Associated with DDD

About 90% of patients with low back pain due to any etiology experience pain relief, regardless of treatment, in about six weeks.(4)

Progressive loss of disc height and tension characterizes degenerative disc disease. For this reason, patients typically experience different signs and symptoms as their condition progresses. Following the painful dysfunction phase (described above), progressive degeneration may lead to an instability phase. This is typified by disc resorption, loss of disc height, and facet capsular laxity. The instability phase poses a higher likelihood of disc tears and herniation and an increased risk of vertebral subluxations, and may result in continued back or leg pain.(3,8)

Continued loss of disc height and the resulting settling of the vertebrae together contributes to a stiffer motion segment. Finally, the motion segment(s) begin to restabilize. Osteophyte formation may alleviate the severity of back pain, but muscle tenderness, stiffness, reduced movement, and scoliosis may remain.(8) This may result in less segmental backache, but is associated with an increased potential for radicular pain resulting from stenosis (narrowing of the spinal canal) and may speed degeneration in adjacent discs due to their relative hypermobility.(3)

Studies examining the natural history of discogenic back pain are scarce. A retrospective study of 25 patients provides some information on the natural course of the disorder.(46) This study examined patients who had discogenic low back pain confirmed with positive discography. Six months of conservative treatment had failed to relieve their pain, and spinal surgery was their next treatment option, but they refused the surgery. At an average of 4.9 years of followup, improvements had occurred in 68% (17/25) of patients, no change in 8% (2/25), and worsening in 24% (6/25). These results suggest that, in some patients, discogenic back pain improves without surgery. One problem with interpreting these results, mentioned in an editorial by Deyo,(47) is that patients who refuse surgery may be fundamentally different from patients who would accept surgery: refusers may be more likely to improve without surgery. Deyo thus commented that the study "cannot be taken to establish that natural history," but that it still motivates the need for additional research.(47) Although this is a flawed study (retrospective, small size, possibly with sampling bias), it was the only one identified in our literature search that examined the natural course of DDD. This issue has not yet been adequately addressed by larger or better-quality studies.

Treatment of Low Back Pain Associated with DDD

Discogenic back pain is typically managed conservatively for at least six months before surgery is considered. Rest is usually only recommended for the first couple days of onset.(4) A variety of conservative treatments can be tried:(2,8)

- Back education: To relieve pain and improve function by adapting body mechanics for everyday activities
- Cognitive behavioral therapy (CBT): To relieve pain by improving coping and adaptive skills
- Physical therapy: To decrease inflammation and alleviate pain, and, once there is sufficient improvement, to strengthen and stabilize the lumbar area of the back.
- Exercise: To achieve lumbar stabilization, and to relax tense muscles.

- Weight reduction: In overweight patients, to relieve mechanical pressure on discs and surrounding structures
- Alternative therapies, such as chiropractic manipulation, acupuncture, therapeutic massage: To relieve pain
- Medications: To reduce inflammation (antiinflammatories), relax muscles (muscle relaxants), and relieve pain (analgesics, antiinflammatories, muscle relaxants)
- Epidural injections: To reduce inflammation (steroids) and pain (steroids or morphine)

Spinal Fusion

When conservative treatments fail after at least six months, spinal fusion may be considered. Between 150,000 and 250,000 people in the U.S. undergo lumbar spinal fusion surgery annually to treat discogenic back pain.(9)

Underlying Theory

The goal of spinal fusion (also known as spinal arthrodesis) is to permanently immobilize the spinal column vertebrae surrounding the disc(s) that is (are) diagnosed as causing discogenic low back pain. Immobilizing the vertebrae is believed to reduce pain by limiting painful movement (e.g., pinching of nerves or rubbing of bone on bone) that may occur as degenerated discs subside. Spinal fusion is also used to treat other painful conditions, including spondylolisthesis (forward displacement of one of the lower lumbar vertebrae over the vertebra below it or on the sacrum), trauma resulting in spinal nerve compression, abnormal spinal curvatures (scoliosis or kyphosis), and vertebral instability caused by infections or tumors.

Basic Procedure

Several surgical procedures may be used to achieve spinal fusion in patients with discogenic low back pain. They differ by surgical approach and instrumentation used. Some fusion strategies may be particularly appropriate for certain patient populations. All methods have advantages and disadvantages.

For any fusion procedure, surgeons may or may not elect to use instrumentation. Many types are commercially available, including pedicle and facet screws, rods, and cages. Cages are manmade implants intended to stabilize the motion segments. Several types of cages are available, and they can be divided into three groups: cylindrical threaded titanium interbody cages (tubes to be implanted and packed with bone graft); cylindrical threaded cortical bone dowels (disc-shaped dowels to be implanted in the center of the disc space); vertical interbody rings or boxes (which are also implanted in the center of the disc space).(48)Premade femoral ring allograft implants, which may be packed with allograft or demineralized bone matrix, as well as bone spacers, are also commercially available.

Metal or polytheretherketone (PEEK) cages to be packed with autograft are in wide diffusion. Disadvantages include subsidence and complication of radiological assessment. Titanium and carbon fibre cages may more closely approximate bone.(49) BAK cages are stand-alone threaded cages, but these have become unpopular in Europe due to instability with spinal extension leading to pseudarthrosis and poor clinical outcomes.(49) Newer cages are typically made of

PEEK or metal and include locking screws and/or plates. These are suitable for patients with end-stage degenerative disc disease and very rigid facet joints.(49)

Pedicle screw fixation with adjoining rods provides immediate immobilization, theoretically improving the odds of successful fusion and enabling earlier mobilization of the patient.(50) However, the use of pedicle screw fixation may be associated with an increase in vascular, neurological, and soft tissue complications, and the metals may affect MRI.(49) Facet screw fixation may be used instead when there are one or two levels to fuse, the facets are intact, and the disc segment of interest is collapsed.(49) Potential advantages over pedicle screws include less soft tissue dissection, more space for bone graft, lower rate of neurological complications, and substantially lower cost.(49)

Packing material is always used to stabilize the fusion by promoting new bone growth. Autografts harvested from the patient's own iliac crest are commonly used. Allografts, bone harvested from another human, are also in use. Another type of packing material is demineralized bone matrix (DBM). DBM contains osteoinductive proteins that improve new bone formation by inducing the production of chondrocytes and new cartilage; the resulting cartilage is resorbed and replace by bone.(51) If proven effective, DBM could reduce the reliance on bone autografts, which are associated with short-term and long-term patient morbidity (e.g., harvest site pain). The use of DBM is under study and not widely diffused.

There are five main types of fusion surgeries: posterolateral fusion, posterior lumbar interbody lumbar fusion, transforaminal lumbar interbody fusion, anterior lumbar interbody fusion, and circumferential fusion. These surgeries are discussed in the text to follow.

Posterolateral fusion (PLF)

Posterolateral fusion (PLF) involves a dorsal surgical approach that joins vertebrae by the space just outside the spine. Bone is grafted onto decorticated laminae and spinous processes.(49) It is the most commonly used method of spinal fusion.(52) Potential advantages of PLF include ease of approach, low complication rates, and familiarity with the approach among spine surgeons.(2) In addition, there is less soft tissue disruption, and theoretically, a lower chance of infection.(49) Less bone graft may be needed for this type of fusion because the parts of the vertebrae that are closest together are fused.(49) Instrumentation may or may not be used in PLF.

Drawbacks of PLF compared with interbody techniques include the mechanically disadvantageous position for the bone graft, and the need to strip muscles to gain adequate exposure of the area to be treated.(2) The rate of pseudarthrosis (unsuccessful fusion) may also be higher compared with intertransverse fusion.(52) Also, discogenic pain may continue after successful PLF, because small amounts of motion still occur in the pedicles.(52) PLF may not enable a biomechanically ideal placement of the bone as interbody techniques (described in the sections to follow).(2)

PLF may be most appropriate for older patients who cannot undergo interbody techniques due to osteoporosis or medical comorbidity, for patients with three or more intervertebral discs involved, or for whom translational instability is thought to be causing the back pain.(53)

Posterior lumbar interbody fusion (PLIF)

Interbody fusion techniques join adjacent vertebrae body-to-body, utilizing the disc space. In posterior lumbar interbody fusion (PLIF), posterior fixation and the use of interbody grafts are employed.(2)Interbody grafts are materials placed between the bodies of the vertebrae to provide

structural support and to facilitate fusion. If used, instrumentation helps to restore the dorsal tension band and maintain lordosis.(2) Instrumentation, including cages and pedicle screws, is commonly used with PLIF. PLIF is considered appropriate for patients who need concomitant posterior decompression, and patients with spondylolisthesis with retained disc space height.(2)

Advantages of PLIF include improved maintenance of sagittal balance, opportunity for nerve root decompression, and construction of an environment conducive to fusion.(49) Disadvantages of PLIF include manipulation of nerve roots, and limitation of the size of the interbody graft that can be used.(2)

Transforaminal lumbar interbody fusion (TLIF)

Transforaminal lumbar interbody fusion (TLIF) is a variation of PLIF which involves a unilateral (rather than bilateral) fusion.(2) The main advantages of TLIF are that little or no retraction of the thecal sac is required to gain access to the interbody space, and that contralateral nerve roots do not need to be exposed or manipulated.(2) TLIF is more appropriate for patients requiring fusion at the upper lumbar spine, as anatomy at L5-S1 makes this approach more complicated, though not impossible.(2)

Anterior lumbar interbody fusion (ALIF)

Anterior lumbar interbody fusion (ALIF) involves approaching the spine from the patient's front (supine side). Approaches include open transperitoneal (through the peritoneum) or retroperitoneal (from behind the abdominal cavity) and mini-open or laparoscopic techniques, with retroperitoneal being most commonly advocated.(52)

Advantages include avoiding paraspinal muscle dissection (as with PLF) or dissection with the neural elements (such as with PLIF).(2) Furthermore, ALIF is thought by some experts to have a lower failure rate than uninstrumented posterior fusion, as well as reducing the risk of canal stenosis and reducing movement across the disc at the fused level.(49) Disadvantages include the frequent need for an access surgeon in addition to the spine surgeon, because careful retraction of great vessels, reliable identification of the midline, and avoidance of monopolar cautery around the sympathetic plexus, are required.(2) Most complications associated with ALIF are associated with the surgical approach.

ALIF is typically used in patients with collapsed disc space height, since greater distraction of the disc space is required to maintain tension of the annular ligament to help stabilize the spine.(2) It may be particularly useful for patients with abnormalities in the anterior or middle column, after failed posterior surgical procedures, and when the motion segments of interest are kyphotic.(52) Usually patients with only 1 or 2 levels needing surgery are selected, due to the difficulties of accessing the spine anteriorly. ALIF is not appropriate for patients with substantial translational deformity.(52)

Circumferential fusion

Circumferential, or three hundred sixty-degree fusion, is typically comprised of interbody plus PLF. ALIF and PLF are the traditional combination. Circumferential results in high fusion rates.(49) However, because this technique is more expensive, time-consuming, and risky for the patient, it is usually not used for patients with discogenic back pain, but reserved for patients with pseudarthrosis.(2)

Patient Indications/Contraindications

Indications

According to clinical practice guidelines (for more information and citations, refer to the *Clinical Practice Guidelines* section of this report), the main indication for spinal fusion for discogenic back pain is moderate to severe chronic back pain (lasting at least six months) that is inadequately relieved by conservative treatment with degenerative disc disease confirmed by radiologic images. A discogram (described in Part II) may or may not be considered, at the surgeon's discretion.

Contraindications

Absolute contraindications are conditions under which the treatment must never be administered. One absolute contraindication to lumbar fusion was identified. Initial laminectomy/discectomy related to unilateral compression of a lumbar nerve root was cited as an absolute contraindication by The Washington State Department of Labor and Industries.(54) Some exclusion criteria in studies in the evidence-based portion of this report include metabolic bone disease, spondylolisthesis greater than grade I, multi-level degeneration, and significant endplate sclerosis.

Relative contraindications are conditions under which additional consideration is required before treatment is administered, because those conditions provide a less ideal situation for treatment success. Relative contraindications to fusion include(8,54): smoking, morbid obesity, active infection, multiple level degenerative disease, severe physical deconditioning, disability for one year or longer prior to consideration of fusion, absence of evidence of functional recovery following most recent spine surgery, and severe medical or psychological problems. Psychological factors correlated with poor outcomes include: history of drug or alcohol abuse, high degrees of somatization on clinical or psychological evaluation, presence of a personality disorder or major psychiatric illness, and/or current evidence of a factitious disorder.(54)

Procedure Charges and Cost Considerations

We searched for reports on the cost or charge for lumbar fusion in the United States with data not more than five years old. One cost report of the retail price of spinal fusion in the United States was identified. According to the active Web site of a commercial carrier, the average billed charge (retail price) for an inpatient spinal fusion surgery cost $62,982 (range of 25th to 75th percentile $42,447-$76,794). This does not reflect any payer negotiation. The average amount of time spent in hospital was 3.0 days (range 1 to 4 days).(55)

One benefit-cost ratio using the Centers for Medicare and Medicaid Services (CMS) reimbursement rate from 2005 (which ranged from $9,900-11,300) was identified. Cost estimates were calculated for a minimal clinically important improvement in quality of life measured by the SF-36 Physical Component Score (a difference of 5.42 was considered clinically important). In this study, the cost per benefit achieved, in terms of quality of life, of lumbar fusion was comparable to that seen with knee and hip replacement.(56)

Discography Prior to Lumbar Fusion

Lumbar discography is a diagnostic procedure in which contrast material is injected into the nucleus pulposus of a lumbar disc. The general intent is to determine whether the disc itself is the source of pain (i.e., a diagnosis of discogenic pain). This diagnosis has been used to justify the need for surgical intervention involving discectomy and lumbar fusion. Thus, discography may influence important decisions about the appropriateness of surgical intervention.

This section provides background on numerous aspects of discography, including patient indications, injection techniques, testing on adjacent discs, the types of discography results and their interpretation, concerns about false positives, and the usage rates of discography.

Indications

Patient indications for discography prior to lumbar fusion are not well-defined in the literature. A 2005 guideline published by the American Association of Neurological Surgeons (AANS) recommended that "discography be reserved for use in patients with equivocal MR imaging findings, especially at levels adjacent to clearly pathological levels."(24) Guyer and Ohnmeiss (2003), however, suggested additional indications prior to fusion surgery such as "further evaluation of demonstrably abnormal discs to help assess the extent of abnormality" and "assessment of discs before fusion to determine if discs within the proposed fusion segment are symptomatic and to determine if discs adjacent to this segment are normal."(57) A 2007 guideline from the American Society of Interventional Pain Physicians echoed these additional indications, but also stated that "Generally, discography should be viewed as an invasive test to be used to seek abnormalities when results from other tests are equivocal or inconsistent, in a patient with symptoms severe enough to require further evaluation."(58) For additional information on indications for discography, refer to the section on *Clinical Practice Guidelines* (within the *Discussion* section).

Basic Procedure

Some variability exists in how different practitioners perform discography injections. In an attempt to reduce this variability, Sachs et al. (1989) provided detailed guidelines.(59) General recommendations included placing the patient in the left decubitus position (lying on the left side), a two-needle technique (i.e., one needle through the other) to avoid infection of the disc, no transdural puncture, use of a full syringe with 3 cubic centimeters of contrast material, injection with maximal force at a steady rate, and CT scanning within seven hours comprising 3-4 slices per disc at 5 mm thickness and 4 mm intervals. Specific needle recommendations for L3-4 or L4-5 discography included: one 3.5 inch 18-gauge guide needle, and one 6-inch 22-gauge insertion needle. Specific needle recommendations for L5-S1 discography included: one 6 inch 18-gauge guide needle, and one 8-inch 22-gauge insertion needle. Reitman et al. (2001)(60) emphasized the need to standardize patient positioning after observing that different spinal positions resulted in different levels of strains and bulges in the annulus. Guyer and Ohnmeiss (2003)(57) recommended the use of a water-soluble radiopaque contrast.

Some authors advocate that discography be performed not only in the suspected disc, but also in discs adjacent to the suspected disc. One reason to test these "control" discs is that if they test positive, this raises a suspicion that the positive test in the suspected disc was actually a false positive (we discuss false positives in more detail below).(24,25) Another proposed justification for adjacent-disc discography concerns the eventual efficacy of fusion surgery. If adjacent discs

are abnormal but are not included in the levels undergoing fusion, the rigidity imposed by fusion may accelerate disc degeneration in the adjacent levels above and below, thereby compromising the efficacy of surgery.(36,57,61)

Discography yields two types of results: pain provocation (whether the patient's typical pain was reproduced by the injection), and morphology (whether the dye images an abnormal pattern in the disc, often based on CT scan). Controversy exists about the relative importance of these two test results. Some authors(4,10) assign much greater importance to pain provocation; for example, Bogduk (1996)).(10) stated that "the morphology of the disc as revealed by discography is essentially irrelevant." By contrast, Buenaventura et al. (2007) cited disc morphology as the gold standard for discogenic pain, stating that "the imaging information is important since treating an anatomically normal disc, irrespective of its ability to cause pain, seems unethical."(11) Walsh (1990) proposed that a discography result should only be considered positive if the patient's typical pain was reproduced *and* the morphology was abnormal.(12)

The extent of spread of the contrast material from the nucleus pulposus determines disc morphology. The Dallas Discogram Description categorizes several levels of disruption of the disc annulus.(13,14) The first level (Grade 0) indicates normal morphology, when no contrast material leaked into the annulus. Grade 1 disruption indicates that the contrast material leaked into the annulus, but only within its inner third. Grade 2 disruption indicates that the contrast material leaked into two-thirds of the annulus, whereas Grade 3 denotes leakage into the outer one-third of the annulus. Grades 4 and 5 represent greater spread of the contrast material beyond the annulus, around the circumference of the disc, and/or into the epidural space. These six grades reflect the most recently proposed version of the Dallas Discogram Description.(14)

False Positive Rates: A Potential Concern

One major concern about discography is the rate of false positive results. Several authors have found that among people with no previous pain, the discography result can be positive.(15-22) Also, discography in lumbar discs has been reported to reproduce pain known to originate elsewhere in the body.(23) Various solutions have been proposed for these phenomena, including a more stringent definition of a positive test to require both typical pain provocation and abnormal morphology (Walsh definition),(12) the requirement that adjacent discs test negative,(24,25) and the avoidance of high pressure (≥22 pounds per square inch).(16) Carragee et al. (2006) found, however, that even when all of these conditions were met, the rate of false positives was still 25%.(26) Many have suggested that the origin of many false positives lies with the psychological status of some patients; a positive discography may be more likely in patients with psychological comorbidities who are predisposed to report pain.(17-20,22,23,27,28)

Diffusion

To estimate current usage of discography, Carrino et al. (2002) analyzed CPT-4 codes for spinal injection procedures (including lumbar discography) in the U.S. Medicare Population from 1993 to 1999.(62) They provided results separately for specialty categories (anesthesiology, surgery, physiatry, radiology, and other). Across specialties, the number of lumbar discography procedures increased from 4,520 in 1993, to 5,055 in 1996, to 8,605 in 1998, and finally to 11,323 in 1999. This represents an average annual increase of 16.6%, which is considerably larger than the 2.8% annual increase in the Medicare population over the same time period.

The annual increase was smallest for surgical specialty providers (5.2%) and largest for anesthesiologists (41.1%) (these percentage increases were calculated by ECRI Institute based on the data in Table 3 of the Carrino article).(62) By 1999, anesthesiology specialty providers conducted 35% of Medicare lumbar discography procedures, as compared to 33% for radiology and 21% for surgery.

Methods

Key Questions and Outcomes Assessed

In this report, we address the following six Key Questions:

Part I – Lumbar fusion surgery and nonsurgical treatments for chronic lumbar back pain

1) Does lumbar fusion surgery reduce pain and improve functional status/quality of life more effectively than nonsurgical treatments?
2) What are the rates of adverse events (perioperative, long-term events, and reoperations) for lumbar fusion surgery and nonsurgical treatments?
3) What patient characteristics (i.e., workers' compensation population, patients with chronic pain, psychological distress, and age-groups) are associated with differences in the benefits and adverse events of lumbar fusion surgery?

Part II – Role of discography prior to lumbar fusion surgery

1) In patients being considered for lumbar fusion surgery, what is the reliability of discography?
    a. Test-retest reliability
    b. Inter-reader reliability
2) In patients undergoing lumbar fusion surgery, do the results of pre-surgical discography predict the degree of pain reduction or improvement in functional status/quality of life after lumbar fusion surgery?
3) In patients being considered for lumbar fusion surgery, do patients who receive discography that influences the treatment choice have better treatment outcomes than patients who do not receive discography?

The primary outcomes of interest addressing Key Question 1 are functional status measured by the Oswestry Disability Index (ODI), back pain measured by a visual analog scale (VAS), and quality of life measured by a previously validated instrument; the only instrument used to measure quality of life in the available evidence base was the short-form (SF)-36 questionnaire. The ODI is comprised of 10 questions on pain and pain-related disability in activities of daily life and social participation. Each question has six response alternatives, and the overall score ranges from 0 (no disability) to 100 (totally disabled or bedridden). The VAS for back pain is also scored from 0 (no pain) to 100 (worst pain imaginable). A recent study calculated the minimal clinically important difference for the ODI and VAS of back pain using linear regression analysis of score change compared to pre-treatment scores. The authors determined that the minimal clinically important difference for the ODI was 10, and for the VAS of back pain it was 18-19.(29) Accordingly, we used a difference of 10 for the ODI and a difference of 20 for the VAS as the minimal clinically important difference in our assessment of these outcomes. The SF-36 is scored from 0 (worst health state) to 100 (best health state); we used a difference of 5 in the

SF-36 as the minimal clinically important difference based on data from an earlier study that investigated this issue.(30)

For Key Question 2, any reported adverse events or complications will be tabled. Key Question 3 does not involve an analysis of outcomes; instead, it requires an analysis of the relationship between patient characteristics and positive or negative outcomes of treatment.

This report is not intended to answer all clinical questions about discography, but rather focuses on three specific questions about discography as it relates to lumbar fusion surgery. Key Question 4 concerns reliability, comprising both test-retest reliability and inter-rater reliability. Key Question 5 concerns whether discography results can usefully predict the outcomes of spinal fusion surgery. This involves a comparison in surgical outcomes between those who had a positive discography before surgery and those who had a negative discography before surgery. Key Question 6 involves the clinical impact of discography: whether patients who underwent discography before surgery have better fusion outcomes than those who did not undergo discography. All other questions about discography, such as the optimal method for performing or interpreting discography, the correspondence between pain provocation and morphology, or the correspondence with other diagnostic procedures, are outside the scope of this report.

Literature Searches

The clinical studies included in this technology assessment were identified using a multi-staged study selection process, and were based on inclusion criteria that were determined *a priori*. Use of *a priori* inclusion criteria reduces the risk of bias because the decision to include or exclude each study is independent of the results of the study. In the first stage of the selection process, we performed a comprehensive literature search using broad criteria. In the second stage, we retrieved all articles that appeared to meet the *a priori* inclusion criteria, based on their published abstracts. In the final stage of the study selection, we reviewed the full text of each retrieved article, assessed its quality, and verified whether or not it met the *a priori* inclusion criteria.

One characteristic of a good technology assessment is a systematic and comprehensive search for information. Such searches distinguish systematic reviews from traditional literature reviews. Traditional literature reviews use a less rigorous approach to identifying and obtaining literature, making it possible for a reviewer to include primarily articles that agree with a particular perspective, and to ignore articles that do not. Our approach precludes this potential reviewer bias because we obtained and included articles according to explicitly determined *a priori* criteria.

Briefly, we searched 15 external and internal databases, including PubMed and Embase, for relevant studies. In addition, we searched more than 1,600 journals and supplements maintained in ECRI Institute's collections to determine if they contained relevant information. We also examined the bibliographies/reference lists from peer-reviewed and gray literature. (Gray literature includes reports and studies produced by local government agencies, private organizations, educational facilities, and corporations that do not appear in the peer-reviewed literature.) A complete list of the databases searched and the search strategy used to identify relevant studies are presented in Appendix A. The last search was conducted in August 2007.

Study Inclusion Criteria

Use of explicit inclusion criteria, decided upon before data have been extracted, is a vital tool in preventing reviewer biases. Some of these a priori criteria are based on study design, and other criteria ensure that the evidence is not derived from unusual patients or interventions and/or outmoded technologies. Finally, we also developed criteria to ensure that we focused our analysis on the outcomes that are of most interest to patients.

The inclusion criteria were:

1. Study must have reported on at least one of the outcomes that are the focus of this report.
   Other outcomes are beyond the scope of this report.
2. Study must be published in English.
   Moher et al. have demonstrated that exclusion of non-English language studies from meta-analyses has little impact on the conclusions drawn.(63) Juni et al. found that non-English studies typically were of lower methodological quality and that excluding them had little effect on effect size estimates in the majority of meta-analyses they examined.(64) Although we recognize that there may be situations in which exclusion of non-English studies could lead to bias, we believe that it is insufficiently likely that we cannot justify the time and cost of translations to identify studies of acceptable quality for inclusion in our reviews.
3. Study must be published as a peer-reviewed full article. Meeting abstracts will not be included.
   Published meeting abstracts have not been peer-reviewed and often do not include sufficient details about experimental methods to permit one to verify that the study was well designed.(65,66) *In addition, it is not uncommon for abstracts that are published as part of conference proceedings to describe studies that are never published as full articles.*(67-70)
4. Studies of efficacy must have enrolled 10 or more individuals per treatment arm. Cohort studies (including case series) examined for adverse effects must have included at least 100 patients.
   The results of case studies are typically more variable and less generalizable than those of larger studies.
5. When several sequential reports from the same study center are available, only outcome data from the largest and most recent report will be included. However, we will use relevant data from earlier and smaller reports if the report presents pertinent data not presented in the larger, more recent report.
6. At least 80% of treated patients must have contributed follow-up data to a given time point. Data from time points with <80% followup will be excluded from analysis.
   If >20% of patients are missing, the estimate of treatment effect may be inaccurate.
7. Patients had chronic (3+ months) of lumbar pain. At least 80% of the patients did not have any of the following medical conditions:
   - Radiculopathy
   - Functional neurologic deficits (motor weakness or EMG findings of radiculopathy)

- Spondylolisthesis (>Grade 1)
- Isthmic spondylolysis
- Primary neurogenic claudication associated with stenosis
- Fracture, tumor, infection, inflammatory disease
- Degenerative disease associated with significant deformity

8. For Key Questions 1, 2, 3, and 5, if the study enrolled some patients who did not receive lumbar fusion surgery, these patients must have comprised less than 20% of the enrolled patients.
9. For Key Questions 1, 2, 3, 5, and 6, the specific lumbar fusion procedure must not have been an outdated procedure.
10. Study design and publication date requirements:
    a. Key Questions 1 and 3. Only randomized trials comparing lumbar fusion to a nonsurgical approach, published in 1990 or later
    b. Key Question 2. Studies that either met criteria for KQ1 or 3, OR non-RCTs of lumbar fusion that enrolled at least 100 patients and were published in 1990 or later. Adverse event data on nonoperative approaches will be sought from comparative trials and from systematic reviews of nonoperative approaches.
    c. Key Question 4, test-retest reliability. Study must have reported data to determine the test-retest reliability of discography by using the same patients at different timepoints. Publication date 1990 or later.
    d. Key Question 4, inter-rater reliability. Study must have reported data to determine the inter-rater reliability of discography by employing different practitioners to apply discography to the same patients. Publication date 1990 or later.
    e. Key Question 5. Study must have reported data for two groups of patients (both of which received lumbar fusion surgery): one group of patients who had had a positive discography before lumbar fusion surgery, and another group of patients who had had a negative discography before lumbar fusion surgery. No publication date restriction.
    f. Key Question 6. Study must have reported data for two groups of patients: one group of patients who had received discography before treatment that influenced the choice of treatment (i.e., whether to perform lumbar fusion surgery; some may not have received fusion surgery, or some may have received different variants of fusion surgery), and another group of patients who had not received discography before treatment (again, some may not have received fusion surgery, or some may have received different variants of fusion surgery). Patient groups must have been well-matched at baseline. No publication date restriction.

11. For consideration of outcome data pertaining to pain, quality of life, functional status, the study must have used a previously validated instrument.
12. Specific criteria for systematic reviews of nonsurgical approaches in Key Question 2. We restricted our evaluation of the safety profile of nonsurgical alternatives to lumbar fusion to those reported in RCTs comparing surgical and nonsurgical therapies and systematic reviews of nonsurgical therapies. Systematic reviews capture data from many studies. The following inclusion criteria were applied to systematic reviews of nonsurgical therapies for Key Question 2:

    a) The review is published in 2000 to date of last search
    b) The review is on treatments for DDD, with a dedicated section on potential harms of nonsurgical approaches.
    c) A comprehensive literature search was performed using at least two electronic sources (e.g., Central, EMBASE, and MEDLINE).
    d) Inclusion and exclusion criteria for study selection were provided.
    The quality of included systematic reviews will be evaluated using a measurement tool for assessment of multiple systematic reviews (AMSTAR).(71)

Note that non-randomized uncontrolled studies were allowed for Key Question 2 (adverse events) but not Key Question 1 (efficacy of fusion). This is in part because the RCTs addressing Key Question 1 were designed to evaluate efficacy rather than safety. Adverse events were reported as one of several secondary outcomes, but the procedures for capturing these events are not generally reported. Also, two of the RCTs were relatively small and therefore unlikely to capture rare adverse events. Larger non-randomized and/or uncontrolled studies may capture events that did not occur in the RCTs. Furthermore, most adverse events reported in these studies could only occur in patients undergoing surgery, which also lessens the need for a nonsurgical control for adverse events.

In contrast, evaluation of efficacy requires RCTs due to the subjectivity of the outcomes (pain, functional status, quality of life). This subjectivity makes these outcomes vulnerable to measurement biases or regression to the mean (i.e., the patient's improvement may be due to a placebo effect or otherwise unrelated to the treatment). A parallel treatment group is necessary to control for measurement biases to which subjective outcomes are vulnerable.

Evaluation of the Stability and Strength of the Body of Evidence

To evaluate the stability and strength of a body of literature, we used a formal rating system.(72) This system employs decision points that collectively yield an overall category that describes the strength of the evidence for a *quantitative* estimate and *qualitative* conclusion as strong, moderate, weak, or unacceptably weak. The qualitative conclusion addresses the question, "Does it work?" The quantitative estimate addresses the question, "How well does it work?" This distinction allows flexibility in ratings of different aspects of the evidence. For example, an evidence base can be considered weak in terms of the precise *quantitative* estimate of effect (e.g., if estimates vary widely among studies), but strong or moderate with respect to the qualitative conclusion (e.g., if all studies nevertheless demonstrate the same direction of effect).

The system addresses five general aspects of the evidence: quality, quantity, consistency, robustness, and magnitude of effect. Quality refers to the degree of potential bias in the design or conduct of studies. Quantity refers to the number of studies and the number of enrolled patients. Consistency addresses the degree of agreement among the results of available studies. Robustness involves the constancy of conclusions in the face of minor hypothetical alterations in the data. Magnitude of effect concerns the quantitative amount of benefit that patients experience after treatment, and it is only considered in the qualitative section of the system. These concepts, and the rules we used to incorporate the concepts in this report, are described more fully in Appendix C.

Statistical Methods

Data calculation methods depended on type of data reported. We calculated individual study effect sizes from dichotomous data using the odds ratio. If there were no events in one or both of the study groups, the Peto odds ratio was used, as this method is appropriate for rare events in studies with no substantial imbalance in the number of patients in each comparison group.(73) Effect sizes for continuous data (e.g., change in ODI score) were calculated in the original metric (the weighted mean difference in change scores).

Whenever relevant data from three or more studies were available and could be combined (and assuming that the studies used similar enough clinical methods and patients that combining was considered appropriate), we summarized the results in terms of the statistic selected above using meta-analysis. In addition, in instances where the evidence base consisted of two studies and the median quality of the studies was moderate or high, we combined the studies in a meta-analysis in an attempt to reach a qualitative (but not quantitative) conclusion.

Meta-analysis involves pooling data from different studies to obtain an estimate of the average treatment effect. If a sufficient number of studies is available, it also provides a means for formally identifying and exploring important differences among the results of different studies (consistency). For a complete description of when studies can be combined in a meta-analysis, see Appendix C under Strength of Evidence Algorithm.

As the first step in meta-analysis, we tested the available data to determine whether the results of the studies included in the meta-analysis differed from one another by more than that expected by chance (heterogeneity testing) using the I^2 statistic ($I^2 \geq 50\%$ indicates notable unexplained inconsistency).(74) If study results did not substantially differ (i.e., the data were consistent), we pooled the data in a random effects model to obtain a summary estimate.(75) If the results were

heterogeneous (I^2 ≥50%), or if fewer than 75% of the studies reported information to permit effect size calculation, we did not attempt to obtain a quantitative estimate of effect. Instead, we consulted the 95% confidence interval to determine whether the summary statistic indicated the general direction of the effect.

If a meta-analysis included three or more studies, we then tested the quantitative and qualitative robustness of our findings using sensitivity analyses as recommended by Olkin.(76) This involved the removal and replacement of each study to determine whether any single study had a substantial influence on the summary statistic. We also performed cumulative meta-analysis, the systematic addition of each study, to determine the effect of adding studies sequentially on the summary statistic. Studies were added in order of publication date (earliest to latest). Finally, we re-calculated summary effects in a different metric (Hedges' g in place of the odds ratio or weighted mean difference) to see if this overturned the qualitative conclusions. We do not perform sensitivity analyses on a meta-analysis of two studies because by our definition such a meta-analysis is automatically not robust.

Results for Efficacy and Safety of Spinal Fusion

(Key Questions 1 through 3)

Evidence Base

Our searches identified 482 citations. Of those, 239 were excluded at the abstract level because they clearly did not meet our inclusion criteria for Key Questions 1 through 3. The remaining 243 citations that appeared to address Key Questions 1 through 3 were retrieved as full articles for further assessment. Of those 243 articles, we determined that 30 articles reporting on 27 studies satisfied the inclusion criteria. This selection process is presented in Figure 1. The included studies are listed in Table 1 and described below.

Included Studies

Four randomized controlled trials (RCTs) that enrolled a total of 767 patients met the inclusion criteria for Key Question 1, which required a comparison of lumbar fusion to non-operative treatment in patients with DDD. These same RCTs also reported treatment complications and therefore also met the inclusion criteria for Key Question 2.

In addition to the four RCTs described above, 23 studies with a total of 5,639 patients also met the inclusion criteria for Key Question 2. These studies were either case series of lumbar fusion or controlled studies (some randomized) that compared different lumbar fusion procedures.

Data from one separate publication of one RCT (also included in Key Question 1 and 2) that enrolled 294 patients met the inclusion criteria for Key Question 3.

Figure 1. Study Attrition Diagram

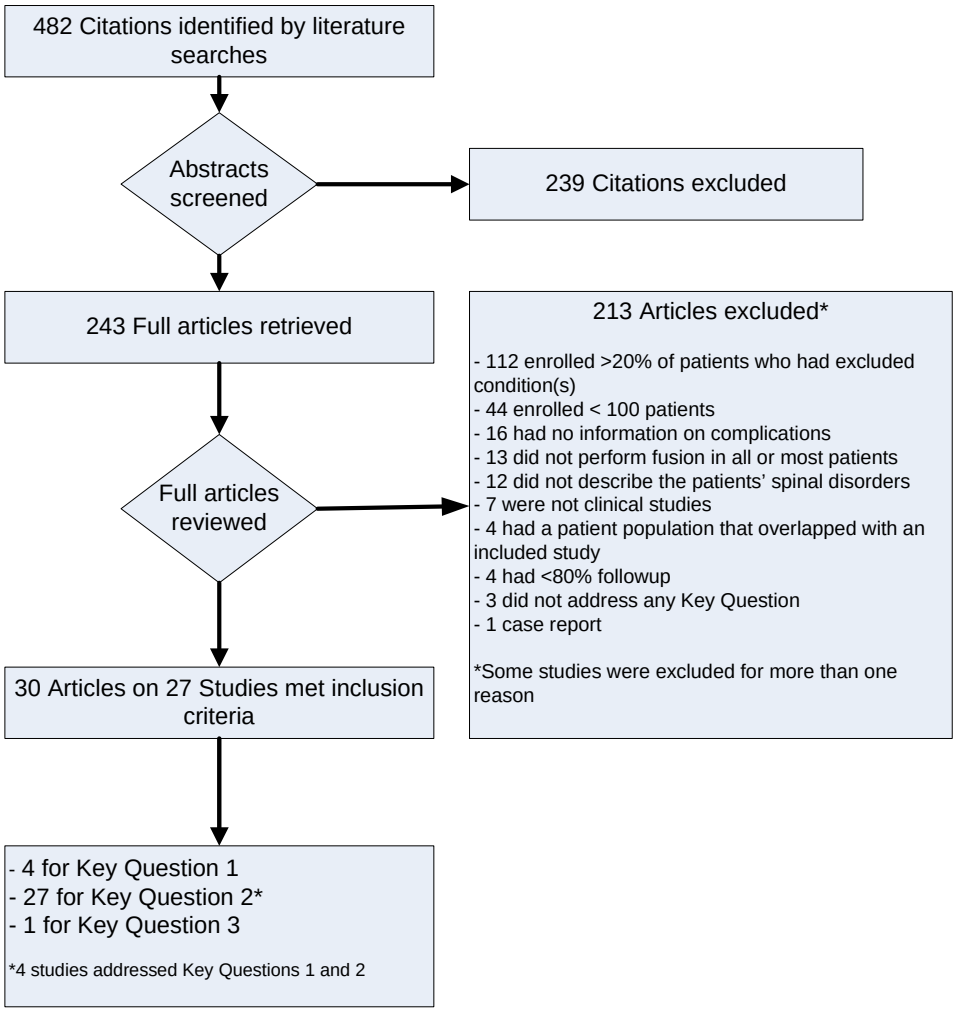


Table 1. Included Studies for Spinal Fusion (Key Questions 1-3)

Study	Study design	Interventions	Number of patients	Followup
Studies comparing fusion to non-operative therapy (Key Questions 1 and 2)				
Brox et al. 2006(31)	RCT	Posterolateral fusion (PLF) with pedicle screws + autologous bone graft Cognitive intervention + intensive exercise/rehabilitation (Note: all patients had undergone prior surgery for disc herniation)	29 (23 received surgery) 31 (29 received intervention)	1 year
Fairbank et al. 2005(77)	RCT	Spinal fusion (unspecified) Intensive cognitive behavioral-based rehabilitation	176 (139 received surgery) 173 (151 received intervention)	2 years
Brox et al. 2003(78)	RCT	PLF with pedicle screws + autologous bone graft + physical therapy Cognitive intervention + intensive exercise/rehabilitation	37 (33 received surgery) 27 (25 received intervention)	1 year
Fritzell et al. 2001(79); Fritzell et al. 2002(80)	RCT	PLF with or without pedicle screws, or circumferential (PLIF or ALIF) Routine (non-intensive) physical therapy + other non-operative therapies	222 (204 received surgery) 72 (65 received intervention)	2 years
Other studies addressing Key Question 2				
Martin et al. 2007(81)	Retrospective cohort study	Not specified	462 patients with herniated discs 515 patients with degenerative disc disease	11 years
Burkus et al. 2005(82); Burkus et al. 2006(83)	RCT	rhBMP-2and MD-II threaded cortical, bone dowel. Anterior lumbar interbody fusion (ALIF), open, with transperitoneal or retroperitoneal approach Autologous bone graft fusion, Anterior (ALIF), open	79 52	24 months
Sasso et al. 2005(84)	Prospective multicenter case series	Fusion with autograft, Anterior (ALIF), retroperitoneal or transperitoneal approach	208	2 years
Bezer et al. 2004(85)	RCT	Unspecified instrumentation with autologous bone graft – traditional harvest, Posterolatera approach (PLF) Unspecified instrumentation with autologous bone graft – interfascial harvest, PLF	59 58	2 years

Study	Study design	Interventions	Number of patients	Followup
Scaduto et al. 2003(86)	RCT	Various instruments, mostly cylindrical threaded titanium non- tapered implants, Anterior (ALIF) Various instruments, mostly cylindrical threaded titanium non- tapered implants, Posterior (lumbar interbody fusion (PLIF)	88 31	30 days
Burkus et al. 2002(87)	RCT	Fusion with iliac crest autograft, Anterior (ALIF), retroperitoneal or transperitoneal approach Fusion with rhBMP-2, Anterior (ALIF), retroperitoneal or transperitoneal approach	136 143	2 years
Christensen et al. 2002(88)	RCT	PLF with titanium Cotrel-Dubousset instrumentation	73	Mean: 14 months (1 day to 48 months)
McAfee et al. 2002(89)	Prospective controlled study	Anterior BAK instrumentation and fusion (ALIF), complete discectomy through open retroperitoneal approach Anterior BAK instrumentation and fusion (ALIF), partial discectomy through miniopen or laparascopic approach	50 50	2 years
Brantigan et al. 2000(90)	Prospective case series	Brantigan I/F Cage with Variable Pedicle Screw Placement System and Autologous bone graft, Posterior approach (PLIF)	221	2 years
Slosar et al. 2000(91)	Retrospective case series	Instrumented circumferential fusion	141	Mean: 37.2 months (Range: 24-53 months)
Thalgott et al. 2000(92)	Retrospective case series	Gasless endoscopic anterior lumbar interbody fusion (ALIF) utilizing the B.E.R.G approach	202	Perioperative (duration not reported)
Regan et al. 1999(93)	Prospective multicenter case series	BAK Cage, Anterior (ALIF), Open surgery, with Retroperitoneal Approach, Single-level BAK Cage, Laparascopic surgery, with transperitoneal Approach, Single-level	305 240	Postoperative
Greenough et al. 1998(94)	Retrospective case series	Pedicle screw fixation using variable screw plate, Posterolateral approach (PLF)	135	12-36 months
Kuslich et al. 1998(95); Kuslich et al. 2000(96)	Prospective multicenter case series	Bagby and Kuslich (BAK) interbody fusion using the anterior retroperitoneal approach (ALIF) or the posterior laminotomy (PLIF) approach	947	2 years
Malter et al. 1998(97)	Retrospective cohort study	Not specified	207 patients with herniated discs	Cumulative to 6 years

Study	Study design	Interventions	Number of patients	Followup
Penta and Fraser 1997(98)	Retrospective case series	Fusion with iliac crest autograft, Anterior approach (ALIF)	125	At least 10 years
Ray et al. 1997(99)	Prospective multicenter case series	Ray Titanium Cage, Posterior Approach (PLIF)	236	48 months
Thomsen et al. 1997(100); Christensen et al. 2002(101)†	RCT	PLF with Cotrel-Dubousset instrumentation and autologous bone implant	64	5 years
Christensen et al. 1996(102)	Retrospective case series	Anterior (ALIF)	132	5-13 years
Hall et al. 1996(103)	Prospective multicenter case series	Isola Spinal Implant System (staged anterior and posterior as well as posterior approach alone, but methods not well-described)	120	Operative
Greenough et al. 1994(104)	Retrospective case series	Anterior (ALIF)	151	Minimum 2 years, (Range: 24-82 months)
Gill and Blumenthal 1993(105)	Retrospective case series	Wiltse Pedicle Screw Fixation System with autograft, Posterior (PLIF) lateral implantation from unilateral approach with lateral/bilateral fusion	238	At least 2 years
Markwalder and Battaglia 1993(106)	Prospective case series	Fusion by various techniques, including Magerl translaminar screw fixation technique and Louis plate fixation method	171	Mean: 23.8 months
Study addressing Key Question 3				
Hagg et al. 2003(107)	RCT	PLF with or without pedicle screws, or circumferential (PLIF or ALIF) Physical therapy + other non-operative therapies	222 (204 received surgery) 72 (65 received intervention)	2 years

† A follow-up of Thomsen et al. 1997(100)

Quality (Internal Validity)

The results of our analysis of the quality of the RCTs comparing lumbar fusion to non-operative therapies are summarized in Table 2. We define quality as synonymous with internal validity (the likelihood that a study is free from bias that could influence the observed clinical outcomes). We based the quality ratings for these studies on the criteria and information presented in Table 13 of Appendix D. The average quality of these studies was moderate.

Limitations of these studies included lack of blinding of patients, providers, and outcome assessors (for the majority of outcomes) in all studies. This could lead to biased interpretation or reporting of outcomes, particularly of subjective outcomes.

The studies of Fairbank et al. (2005) and Brox et al. (2006) were further limited because >15% of the patients did not receive their assigned treatment. Some (but not all) of these patients crossed over to receive the alternative treatment (surgery to no surgery or vice versa). Under intent-to-treat principles, the data from patients who did not receive their assigned treatment is still counted as part of the original group to which they were assigned; this partly confounds the comparison of fusion to non-operative care, since some patients in both groups received the other groups' treatment. This was a particular problem in the Fairbank trial, where 28% of patients assigned to the exercise/rehabilitation group had undergone back surgery by the end of the two year follow-up period. Crossover to the non-assigned treatment tends to diminish a difference between treatments if one exists. An "on-treatment" analysis (based only on the treatment actually received) was not performed in these studies, and the data were not presented in a manner that would allow an independent on-treatment analysis. Although neither type of analysis is perfect, performing both types of analysis is advisable when crossovers occur.

The studies by Fritzell et al. (2001) and Brox et al. (2006) also showed substantial baseline differences in at least one important patient characteristic (proportion of patients with comorbidities or ratio of males/females) that could have influenced the observed effect of treatment. However, Brox et al. (2006) compensated for this by performing statistical adjustment of effect sizes based on differences in baseline patient characteristics (gender and treatment expectations).

Another potential limitation that might have affected these studies was a between-group difference in the proportion of patients receiving additional (ancillary) treatments. The only study that reported this information (Fairbank et al. 2005) revealed a substantial between-group difference in the number of patients receiving additional treatments (55% in surgical group versus 39% in rehabilitation group). The predominant ancillary treatment also differed between groups; it was physical therapy (34%) in the surgical group and surgery (25%) in the rehabilitation group. The influence of such ancillary treatments on the observed effect of fusion or non-operative therapy is uncertain. A small proportion (15%) of patients in this study underwent flexible stabilization rather than fusion, while no patients in the other three studies underwent flexible stabilization. Inclusion of these patients in the group undergoing fusion could have influenced the observed treatment effect for this patient group.

An internal validity rating was applied only to studies comparing lumbar fusion to non-operative therapy in Key Question 1. The remaining studies addressing Key Question 2 were not used to address comparative event rates of fusion and non-operative care; they were used only to provide additional data on adverse events and adverse event rates for lumbar fusion. Due to variability in

the way complications are reported among different studies, lists of complications do not lend themselves to evidence ratings.

Table 2. Internal Validity of Included Studies

Reference	Year	ECRI Quality Score (Rating)
RCTs addressing Key Question 1 – Efficacy Outcomes (ODI, VAS pain scores, QOL)		
Brox et al.(31)	2006	7.0 (Moderate)
Fairbank et al.(77)	2005	6.4 (Moderate)
Brox et al.(78)	2003	7.5 (Moderate)
Fritzell et al.(79); Fritzell et al.(80)	2001	6.6 (Moderate)
Mean quality score for RCTs		**6.9 (Moderate)**

Generalizability

Although four RCTs compared lumbar fusion to non-operative therapy, considerable between-study variation was evident in the intensity of non-operative therapy and the types of lumbar fusion used. In addition, there was variation among studies in the patient inclusion and exclusion criteria, and the characteristics of the patients who were actually enrolled in these studies.

The control treatment in one of the studies was substantially different from the other three studies. Fritzell et al. used non-intensive physical therapy as the control treatment, while the other three studies used more intensive rehabilitation as the control treatment. The control treatment in Fritzell et al. might be expected to provide less of a benefit than an intensive rehabilitation program, particularly since the patients in all of these studies had not responded to these same non-intensive conservative therapies prior to enrollment.(108) In contrast, the intensive rehabilitation programs had not been provided to any patients prior to enrollment, which might raise their expectations about improvement. However, intensive rehabilitation therapy is not readily available outside of clinical trials, so the comparison of fusion to intensive rehabilitation may be less generalizable to actual clinical practice.

Fusion strategies also differed, in terms of both surgical approach and instrumentation used. The two studies by Brox et al. exclusively used posterolateral fusion (PLF) with pedicle screws, but the two remaining studies used a greater variety of fusion procedures. Fritzell et al. used PLF with or without pedicle screws or circumferential fusion using two alternative approaches: posterior lumbar interbody fusion (PLIF) or anterior lumbar interbody fusion (ALIF). Fairbank et al. used an unspecified variety of fusion procedures (essentially whatever the individual surgeons decided to use). A small proportion (15%) of patients in this study underwent flexible stabilization rather than fusion, while no patients in the other three studies underwent flexible stabilization. As our report focuses on the efficacy and safety of lumbar fusion, patients undergoing flexible stabilization are not representative of the target patient population. However, the proportion of the population who were treated with flexible stabilization was small enough that the study met our inclusion criteria.

Differences in individual study patient inclusion/exclusion criteria (listed in Table 8 of Appendix D) resulted in differences in patient characteristics among these studies (shown in Table 3, below). The 2003 study by Brox et al. excluded patients with prior back surgery, while

their 2006 study only enrolled patients who had undergone prior discectomy. Since the patient population with degenerative disc disease undergoing lumbar fusion in clinical practice consists of a mixture of patients with prior back surgery and patients who have never had surgery, neither of the studies by Brox et al. is entirely generalizable to clinical practice. The other two RCTs included a majority of patients with no prior surgery but did not exclude patients with prior surgery: 18.7% of patients in the study by Fritzell et al. had prior discectomy, while 8% of patients in the study by Fairbank et al. had prior laminectomy. The latter study also included a small proportion (11%) of patients with spondylolisthesis of an unspecified grade. If all of these patients were Grade 1 or less, they would all be representative of the patient population of interest in this report. Any patients with a higher grade of spondylolisthesis are not generalizable to the target patient population. The average age of patients in all four RCTs was about 40-45 years, which is representative of the age at which a large proportion of patients with degenerative disc disease undergo surgery in clinical practice.

The generalizability of the remaining 23 studies used to address Key Question 2 is dependent mainly on the fusion procedures and the characteristics of the patient population within these studies, as they did not include a non-operative control group. These studies varied substantially in the types of fusion procedures employed; many studies used more than one type of fusion procedure. However, considered as a group, these studies covered the range of fusion procedures typically employed in clinical practice: anterior lumbar interbody fusion (11 studies), posterior lumbar interbody fusion (5 studies), posterolateral fusion (4 studies), and combined anterior/posterior (circumferential) fusion (4 studies). As noted, some studies used more than one of these methods. Two studies did not describe the methods clearly enough to determine which procedures were used. Fourteen studies reported the use of various types of instrumentation (pedicle screws, cages, etc.). The average age of the patients in these studies ranged from 39 to 54 years, which is representative of the age at which most patients with degenerative disease undergo surgery in clinical practice. The proportion of patients receiving workers' compensation varied considerably (ranging from 21% to 94%) in the 12 studies that reported this information.

Table 3. Characteristics of Patients in Included Studies for Key Question 1

Study	Number of Patients Referred	Number of Patients Randomized	Diagnoses	Duration of Symptoms	Proportion with Previous Spinal Fusion	Proportion with prior back surgery	Age	Proportion of Male patients	Proportion of Smokers	Proportion on Compensation/ Disability Pension	Proportion with Litigation Pending
Brox et al. 2006(31)	113	60	Chronic back pain after surgery for herniation, with degeneration at L4-L5 and/or L5-S1	8 years median (interquartile Range: 3-12.5)	0% (excluded)	100% (herniation surgery)	43 years median (interquartile Range: 35-50)	52%	65%	Not reported	Not reported
Fairbank et al. 2005(77)	Not reported	349	Chronic low back pain with or without referred pain. 11% had spondylolisthesis.	8 years mean (Range: 1-35)	0% (excluded)	8% (laminectomy)	12% <30 years, 37% 30-39 years, 35% 40-49 years, 15% ≥50 years	49%	43%	Not reported	13%
Brox et al. 2003(78)	121	64	Chronic back pain with degeneration at L4-L5 and/or L5-S1	10.8 years mean	0% (excluded)	0%	44.8 years mean	44%	44%	11%	Not reported
Fritzell et al. 2001(79), Fritzell et al. 2002(80)	310	294	Chronic low back pain with degeneration at L4-L5 and/or L5-S1 thought to be causing the back pain	8.0 years mean (Range: 2-40 years)	0% (excluded)	18.7% (discectomy)	43 years mean (Range: 25-64)	50%	32.7%	21%	61%

Key Question 1: Does lumbar fusion surgery reduce pain and improve functional status/quality of life more effectively than nonsurgical treatments?

ECRI Institute evidence assessments:

We did not find sufficient evidence that lumbar fusion surgery is more effective to a clinically meaningful degree than nonsurgical treatments for any of the following patient populations, comparisons and outcomes:

- **Meta-analysis of postoperative changes in Oswestry disability scores from two moderate quality RCTs (n = 413 patients) revealed no clinically meaningful difference between fusion and intensive exercise/rehabilitation plus cognitive behavioral therapy (CBT) in patients without prior back surgery (95% CI 0.2 to 7.5, *a priori* 10 point difference defined as clinically meaningful), although the difference slightly favored fusion. Strength of evidence: Weak.**
- **The evidence was insufficient to determine whether lumbar fusion provides a greater improvement in back pain (one moderate-quality RCT, n = 64 patients) or quality of life (no acceptable evidence) compared to intensive exercise/rehabilitation plus CBT in patients without prior back surgery.**
- **The evidence from one moderate quality RCT (n = 60 patients) was insufficient to determine the relative benefits of lumbar fusion compared to intensive exercise/rehabilitation in patients with prior back surgery.**
- **The evidence from one moderate quality RCT (n = 294 patients) was insufficient to determine the relative benefits of lumbar fusion compared to conventional physical therapy in patients with or without prior back surgery.**

As noted under *Generalizability*, the four trials that met our inclusion criteria for this question differ in potentially important ways. Based upon independent assessment by two methodologists, we assumed that one difference that was likely to create variation in the effect size among trials was the intensity of non-operative therapy in the control groups. Three trials (Brox et al. 2003; Brox et al. 2006; Fairbank et al. 2005) used more intensive exercise/rehabilitation with cognitive behavioral strategies, while the remaining trial (Fritzell et al. 2001) used non-intensive physical therapy as the main component of an unstructured nonsurgical treatment program. The more intensive therapy seems more likely to benefit patients than the less intensive treatment (which patients had undergone without improvement prior to enrollment). If the amount of patient benefit from surgery is assumed to be the same in all studies, then one would expect a greater difference in patient benefit between patients treated surgically and patients treated with conventional physical therapy compared with patients treated surgically and patients treated with multidisciplinary and intensive exercise/rehabilitation. This is important to our analysis because the mean difference measures the difference between treatment and control groups. Therefore, the mean difference would vary depending on the control selected, causing heterogeneity (differences) in study findings. For this reason, the data from Fritzell et al. was not combined with data from the other three trials.

Another factor that might create heterogeneity among effect sizes is whether the patients had back surgery before enrolling in the studies in question. Patients with prior back surgery may be less likely to benefit from further surgery than patients who have never had back surgery. One of the three trials that used intensive exercise/rehabilitation (Brox et al. 2006) included only patients who had undergone prior surgery for disc herniation (most likely discectomy or laminectomy, as none of the patients had undergone prior lumbar fusion). The authors mentioned that "the prognosis after a second operation is generally considered poor compared with the prognosis in patients without previous surgery for disc herniation".(31) Of the remaining two trials, Brox et al. (2003) included no patients with prior back surgery, while Fairbank et al. (2005) had a small proportion of patients (8%) who had undergone prior laminectomy. Based upon the differences in the patient populations, we determined that the data from Brox et al. (2006) should not be combined with data from the remaining two trials.

Although the control therapies and patient characteristics were similar in the trials by Brox et al. (2003) and Fairbank et al. (2005), the two trials differed in the types of fusion performed and the length of followup. Brox et al. (2003) exclusively used posterolateral fusion (PLF) with pedicle screws, while Fairbank et al. (2005) used an unspecified variety of fusion procedures. Also, Brox et al. reported treatment outcomes at one year of followup, while Fairbank et al. reported treatment outcomes at two years of followup. However, we considered differences in the fusion procedure and length of followup less likely to create heterogeneity in effect sizes than the other factors described above. Therefore, we determined that combining the data from these two trials was appropriate.

The four RCTs were therefore analyzed in three separate groups: fusion versus intensive exercise/rehabilitation plus CBT – divided into patients without prior back surgery (Brox et al. 2003, Fairbank et al. 2005) and patients with prior back surgery (Brox et al. 2006) – and fusion versus non-intensive physical therapy (Fritzell et al. 2006).

The Oswestry Disability Index (ODI) was the primary outcome of interest for Key Question 1 in this report; secondary outcomes were VAS back pain and quality of life.

Fusion versus Intensive Exercise/Rehabilitation Plus CBT in Patients without Prior Back Surgery

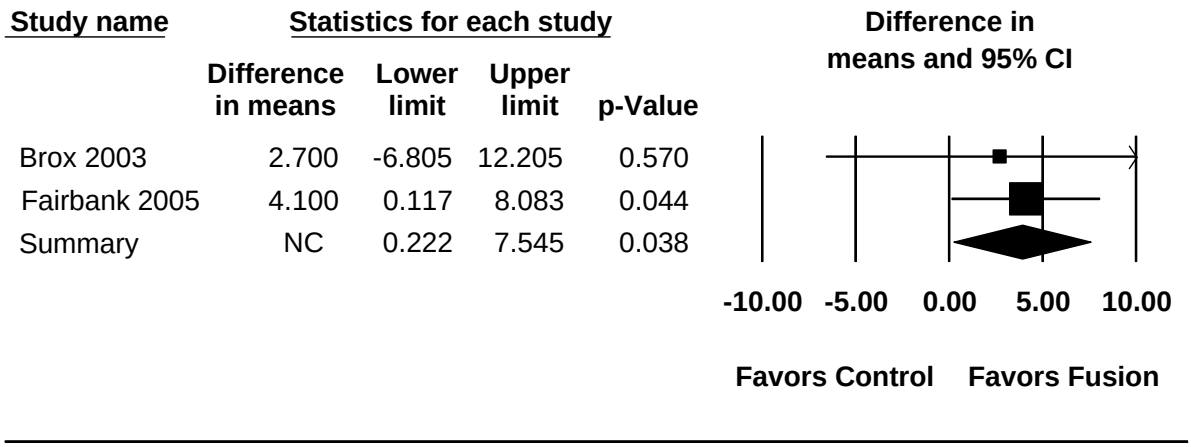
Oswestry Disability Index (ODI)

Two multicenter RCTs with a total of 413 patients compared intensive exercise/rehabilitation with cognitive behavioral therapy to fusion in patients who had not undergone back surgery before. Both studies reported the baseline (pre-treatment) and one- or two-year followup (post-treatment) ODI score; they also reported the between-group difference in the pre-post change in ODI score (see Brox et al. 2003 and Fairbank et al. 2005 in Table 13, Appendix D). Both studies also reported the change scores adjusted for baseline values by analysis of covariance (ANCOVA); this is the best method for adjusting for imbalances in patient characteristics.(32) Thus, our analysis is based on the adjusted change scores.

As described above, these studies were considered suitable for a combined data analysis (meta-analysis), so the change score data were combined in a random effects meta-analysis. As shown in Figure 3, fusion led to a small but statistically significant increase in ODI change scores compared to intensive exercise/rehabilitation plus CBT; however, the upper 95% confidence

limit (7.5) was below the minimum level that is considered clinically significant (ODI = 10). We therefore conclude that changes in ODI scores did not show a clinically meaningful difference between fusion and intensive exercise/rehabilitation plus CBT in patients without prior back surgery, although the difference slightly favored fusion (95% CI 0.2 to 7.5). Because the evidence base is of moderate quality and limited quantity, the strength of evidence supporting this conclusion is weak.

Figure 2. Meta-analysis of Difference in ODI Change Scores



NC – Not calculated.

Visual Analog Scale (VAS) for Back Pain

One RCT (Brox et al. 2003) with 64 patients addressed this comparison. This study reported no statistically significant difference in change in VAS scores between patients undergoing fusion and patients undergoing intensive exercise/rehabilitation plus CBT. Because the 95% CI overlapped with zero and the boundary of minimum clinical significance, the evidence is insufficient to allow a conclusion for this outcome.

Quality of Life

One trial (Fairbank et al.) with 349 patients measured quality of life; this study used the SF-36 instrument. Because fewer than 80% of patients completed the instrument, this study was excluded from analysis. Thus, no conclusion is possible regarding quality of life in patients undergoing lumbar fusion versus non-operative treatment.

Fusion versus Intensive Exercise/Rehabilitation Plus CBT in Patients with Prior Back Surgery

Oswestry Disability Index (ODI)

One RCT (Brox et al. 2006) with 60 patients studied the efficacy of exercise/rehabilitation plus cognitive behavioral therapy to fusion in patients who had previously undergone back surgery. This study reported the between-group difference in the pre-post change in ODI score, using ANCOVA to adjust for baseline between-group differences in gender and treatment expectations (see data in Table 13, Appendix D). The adjusted comparison showed a trend favoring a larger change in ODI in the control group. However, the results were inconclusive because the 95% CI overlapped with zero (not statistically significant) as well as the boundary of clinical significance (ODI = -10), meaning the true difference (if one exists) could favor either treatment.Thus, the evidence is insufficient for a conclusion regarding the relative benefit of fusion versus intensive exercise/rehabilitation plus CBT in patients with prior back surgery.

Visual Analog Scale (VAS) for Back Pain

One RCT (Brox et al. 2006) with 60 patients addressed this comparison. This study reported no statistically significant difference in change in VAS scores between patients undergoing fusion and patients undergoing intensive exercise/rehabilitation plus CBT. Although the 95% CI did not overlap with the boundary of minimum clinical significance, the results of a single moderate quality study are insufficient to allow a conclusion for this outcome.

Fusion versus Non-intensive Physical Therapy in Patients without Prior Back Surgery

Oswestry Disability Index (ODI)

One RCT (Fritzell et al. 2001) with 294 patients compared the efficacy of fusion to conventional physical therapy in patients who had not undergone previous back surgery; however, a minority of patients (18.7%) had prior discectomy. This study reported ODI pre-post change scores for each comparison group (see data in Table 13, Appendix D). A significantly larger improvement in ODI was observed in the fusion group compared to the physical therapy group (11.6 vs. 2.8, $p = 0.015$). The authors did not include group changers in their tabled data, but group changers were included in the analysis of difference. However, although the between-group difference in

change is statistically significant, the mean difference in change between groups (ODI = 8.8) is below the level of clinical significance (ODI = 10). Because this is a single trial of moderate quality, the evidence is insufficient to allow a conclusion for this comparison.

Visual Analog Scale (VAS) for Back Pain

One RCT (Fritzell et al. 2001) with 294 patients addressed this comparison; a minority of patients (18.7%) had prior discectomy. This study reported a statistically significant difference in the change in VAS score favoring fusion when compared to non-intensive physical therapy. However, the mean difference between groups (16.7) did not exceed the boundary of minimum clinical significance for VAS back pain (difference = 20). Because this study did not include group changers in their tabled data, we cannot be certain of the difference if group changers had been included. In any event, because this is a single study of moderate quality without a large effect, the evidence is inconclusive for this outcome.

Key Question 2: What are the rates of adverse events (perioperative, long-term events, and reoperations) for lumbar fusion surgery and nonsurgical treatments?

ECRI Institute evidence assessments:

- **Lumbar fusion leads to higher rates of both early and late adverse events compared to non-intensive physical therapy or intensive exercise/rehabilitation plus CBT.**
- **None of the four RCTs comparing fusion to non-intensive physical therapy or intensive exercise/rehabilitation plus CBT reported any adverse events occurring in patients who only received non-operative care. Most of the reported adverse events for patients in the surgical group could not have occurred in patients who did not undergo surgery (e.g., surgical complications).**
- **Categories of adverse events most frequently reported in fusion studies include reoperation (18/27 studies), infection (14/27 studies), various device-related complications (13/27 studies), neurologic complications (12/27 studies), thrombosis (11/27 studies), bleeding/vascular complications (10/27 studies), and dural injury (10/27 studies).**
- **The ranges of rates of the most frequently reported complications in fusion studies were: reoperation (0% to 46.1%), infection (0% to 9%), device-related complications (0% to 17.8%), neurologic complications (0.7% to 25.8%), thrombosis (0% to 4%), bleeding/vascular complications (0% to 12.8%), and dural injury (0.5% to 29%).**

Strength of evidence assessments were not performed for Key Question 2 because of variability in the reporting of adverse events across different studies. If enough studies reported adverse events in a consistent manner (as in prospective registries), one might be able to estimate rates of adverse events. However, the current evidence base does not have the necessary consistency of reporting. The most that can be stated is that certain adverse events occurred, most could only have occurred with surgery, and that the reported rates for different events covered a certain range.

Studies Comparing Lumbar Fusion to Nonsurgical Treatments

All four RCTs with 767 patients that met our inclusion criteria for Key Question 1 compared adverse event rates for lumbar fusion surgery and nonsurgical treatments. None of the trials reported the rate of total adverse events (from intraoperative to last followup). Instead, they generally divided complication rates by time of occurrence.

Two trials (Brox et al. 2003, Fritzell et al. 2001) separately reported "early" (usually meaning perioperative) and "late" complications (which either occur at a later time or are persistent or permanent). Fritzell et al. defined early as within the first two weeks post-treatment, while Brox et al. did not report the cutoff time for early complications (although it likely did not exceed one month). Another trial by Brox et al. (2006) appeared not to report all early complications; the authors stated that "early complications included two wound infections

among the 23 operated patients", but no other early complications are mentioned. Thus, we cannot be certain that these were the only early complications. However, the authors stated that no late complications occurred. The remaining trial (Fairbank et al. 2005) divided adverse events into intraoperative (during surgery) and post-operative (any time after surgery) categories, which is a somewhat different division than early and late. The only postoperative complications mentioned were need for reoperation; we cannot be certain that there were no late complications that did not require reoperation.

All trials calculated adverse event rates on a per protocol basis, meaning only patients who actually received surgery were included in calculations of surgical adverse events. This is the most conservative approach for analysis of adverse events; calculations on an intent-to-treat basis would underestimate the surgical complication rate, as some patients assigned to surgery never received it.

Overall Early Adverse Events

The results for overall early adverse events appear in Table 15, Appendix D. Despite variation in types of fusion and nonsurgical therapies used in these studies, the four trials had one factor in common; none of them identified any adverse event (early or late) resulting from nonsurgical treatment (intensive exercise/rehabilitation plus CBT in three trials, non-intensive physical therapy in one trial). The three trials that reported overall intraoperative or early adverse event rates found similar rates (range 12.7% to 18%) despite differences in the time period observed (intraoperative to one month). The differences between early adverse events in the surgical vs. nonsurgical groups was statistically significant in all three of these trials. The reported early adverse events in the surgical groups included bleeding, thrombosis, wound infection (deep and superficial), neurological (pain, sympathetic cord damage) complications, device-related (problems with screws or implants) complications, reoperations for various causes, and others (dural tears, peritoneal tears). A complete list of reported early complications and their occurrence rates in these trials appears in Table 17, Appendix D (note: some complications in this table may not be early; most studies did not report time cutoffs for the complications). Most of these complications could not have occurred in the absence of surgery.

Overall Late Adverse Events

Overall late adverse event rates showed more variation among studies, ranging from 0% to 7.4% (Table 15, Appendix D). A number of factors might account for this variation. It could have resulted from differences in the length of followup; the two trials with only one-year followup reported no late events, while the two trials with two-year followup reported that 6.2% and 7.4% of patients who underwent fusion had late events (in both trials, the difference in event rates between surgical and nonsurgical patients was statistically significant). The size of the trials may also have influenced these differences, as the two trials with one-year followup were also much smaller than the other trials, and therefore less likely to detect less common adverse events.
A third factor is that the authors of these trials may have had different definitions of what constitutes an adverse event. Reported late adverse events most frequently included reoperations for various problems (mostly infections and pseudoarthroses) and continuing pain at the donor site from bone graft harvesting. Specific causes of reoperations and other late complications and their rates are listed in Table 18, Appendix D. Again, these events could not have occurred in the absence of surgery.

Adverse Events in Additional Studies of Lumbar Fusion

We examined additional studies of lumbar fusion that lacked a non-operative control group to determine whether these studies report adverse events not reported in the four RCTs described above, and also to determine if the adverse event rates differed from those reported in the RCTs. We selected studies with at least 100 patients total that received any type of lumbar fusion procedure and met all of our other inclusion criteria.

Twenty-three studies with a total of 5,639 enrolled patients met our criteria for this question. Fourteen of these studies were prospective studies; of these 14, six were randomized trials comparing different fusion procedures (a comparison not addressed in this report). The remaining studies were retrospective. Some studies focused only on specific adverse events such as need for reoperation, while others reported all adverse events that occurred during the course of the study. Only eight studies reported any type of overall adverse event rates (operative, postoperative, total, etc.), and the studies varied considerably in the manner in which these events were summarized (Table 16, Appendix D). Because a patient may experience more than one adverse event, we could not calculate the percent of patients experiencing any adverse event when studies only reported rates for specific adverse events. These studies also showed considerable variation in the types of fusion procedures performed, which may contribute to variation in the types of adverse events that occurred in different studies.

A concise summary of reported ranges of specific adverse event rates appears in Table 4. These ranges combine data from the four RCTs described earlier with data from the 23 additional studies. In this table, we do not attempt to separate early from late events, as several studies did not report the specific time of occurrence for each event. Categories of adverse events most frequently reported in fusion studies include reoperation (18/27 studies), infection (14/27 studies), neurologic complications (12/27 studies), thrombosis (11/27 studies), bleeding/vascular complications (10/27 studies), and dural injury (10/27 studies). Death related to surgery was relatively rare, occurring only in 4/27 studies with a maximum reported rate of 2% (we assumed no deaths related to surgery occurred in the other 23 studies). Certain adverse events showed substantial variation in reported rates: these include reoperation (0% to 46.1%), dural injury (0.5% to 29%), neurologic complications (0.7% to 25.8%), and device-related complications (0% to 17.8%). Reported rates in the four RCTs comparing fusion to non-operative care were either at the low end (0% for death) or within the indicated ranges but below the maximum reported rate.

Complete information on the rates of all adverse events reported in these studies is summarized in Tables 19 and 20, Appendix D.

Table 4. Adverse Events Reported in Two or More Studies

Adverse event	No. of studies reporting event	Range of reported complication rates
Infection (deep or superficial)	14 (1 reported 0 events)	0% to 9%
Neurologic	12 (no study reported 0 events; other studies did not report neurologic events)	0.7% to 25.8%
Bleeding/vascular injury	10 (2 reported 0 events)	0% to 12.8%
Thrombosis	11 (1 reported 0 events)	0% to 4%
Dural injury	10 (no study reported 0 events; other studies did not report dural injuries)	0.5% to 29%
Hematoma	7 (no study reported 0 events; other studies did not report hematoma)	1% to 4%
Retrograde ejaculation	6 (no study reported 0 events; other studies did not report retrograde ejaculation)	0.7% to 6%
Device-related	13 (1 reported 0 events with a specific type of fusion)	0% to 17.8%
Reoperation	18 (1 reported 0 events)	0% to 46.1%
Death (surgically-related)	4 (the other 22 studies were assumed to have 0 surgically-related deaths)	0% to 2%

Safety Profile of Exercise/Physical Therapy and/or Cognitive Behavioral Approaches

We searched four electronic sources (MEDLINE, PsycINFO, EMBASE, and CINAHL) and also conducted a manual search for systematic reviews on nonoperative approaches to chronic low back pain. Our searches identified four systematic reviews that met our inclusion criteria.(109-112) However, only one of these reports included a section summarizing adverse events reported in their included studies.

Liddell et al. (2004) conducted a systematic review of RCTs that evaluated various types of exercise therapy for patients with chronic low back pain.(109) The review identified 54 relevant RCTs, of which 16 trials (with 1,730 patients) met the authors' criteria for inclusion in the review (based on their evaluation of individual study quality as moderate or better and the chronicity of symptoms in the study patient population). We assessed the quality of this report using the AMSTAR instrument, a validated tool for measuring the quality of systematic reviews.(71) We determined that the quality of this systematic review was moderate. The authors reported that adverse effects were described in six of the 16 RCTs. One trial reported a coronary occlusion and one reported a myocardial infarction, but both events were reported as being unrelated to the treatment programs. Another trial reported an increase in back pain after the start of treatment. The authors state that "it is difficult to establish from these results whether exercise programs cause adverse effects with chronic low back pain patients".(109) Based on these limited findings, the evidence is insufficient to determine the safety profile of nonsurgical interventions with an exercise/physical therapy component.

Key Question 3: What patient characteristics (i.e., workers' compensation population, patients with chronic pain, psychological distress, and age-groups) are associated with differences in the benefits and adverse events of lumbar fusion surgery?

ECRI Institute evidence assessments:

- **The evidence from one moderate-quality RCT (n = 294 patients) is insufficient to determine what patient characteristics are associated with differences in the benefits and adverse events of lumbar fusion surgery.**

One RCT (Hagg et al. 2003) with 294 patients met the inclusion criteria for this question. This was another publication derived from the Swedish Lumbar Spine Study originally described in Fritzell et al. (2001). The efficacy and safety findings of Fritzell et al. were discussed under Key Questions 1 and 2. In their subsequent publication, Hagg et al. presented data concerning prognostic factors that was not included in Fritzell et al. Hagg et al. conducted a multivariate analysis to identify factors that predicted various outcomes of treatment in the surgical and nonsurgical (non-intensive physical therapy) patient groups. The main outcome measures in their analysis included change of disability (measured as ≥50% reduction of the ODI score), patient global assessment of treatment effect (improvement/no improvement), and work status at followup. Stepwise, forward multiple logistic regression analyses were performed within each treatment group, with the outcomes as dependent variables.

As shown in Table 21 (Appendix D), only one patient characteristic (neurotic personality) showed a statistically significant association with change in disability in the surgical group; patients with neurotic personalities (assessed by the Karolinska Scales of Personality) were less likely to show improvement in the ODI score. No patient characteristic was significantly associated with improvement in ODI score in the nonsurgical group.

The study also identified patient characteristics significantly associated with the patient global assessment (improved or not improved). In the surgical group, neurotic personality was again associated with poor outcome (less likely to be improved), while disc height <50% was significantly associated with improvement. In the nonsurgical group, one patient characteristic (depressive symptoms) was significantly associated with poor outcome. No other factors were significantly associated with patient global assessment in either group.

Certain patient characteristics were significantly associated with work status at followup in both groups. Among surgical patients, older age and longer period of current sick leave were significant predictors of not working at followup. Among nonsurgical patients, only longer period of current sick leave was significantly associated with not working at followup. No positive predictors of working at followup were identified for either patient group.

The following variables did not show significant associations with any of the three outcomes at followup: pain (multiple measures), clinical findings (multiple measures), sociodemographics (disability pension, workers' compensation, unemployment, heavy job, comorbidity, smoking, prior surgery, gender, or marital status), other psychological measures (pain behavior, personality disorders), or radiographic indicators.

Although not specifically stated in the text of the study, it appears that patients who changed treatment groups after enrollment were not included in the analyses described above. The effect this might have on the observed associations is unknown.

Although multicenter, this was a single study of moderate quality; furthermore, none of the observed associations were large effects. Therefore, the evidence is insufficient to allow a conclusion regarding patient characteristics associated with differences in the benefits and adverse events of lumbar fusion surgery.

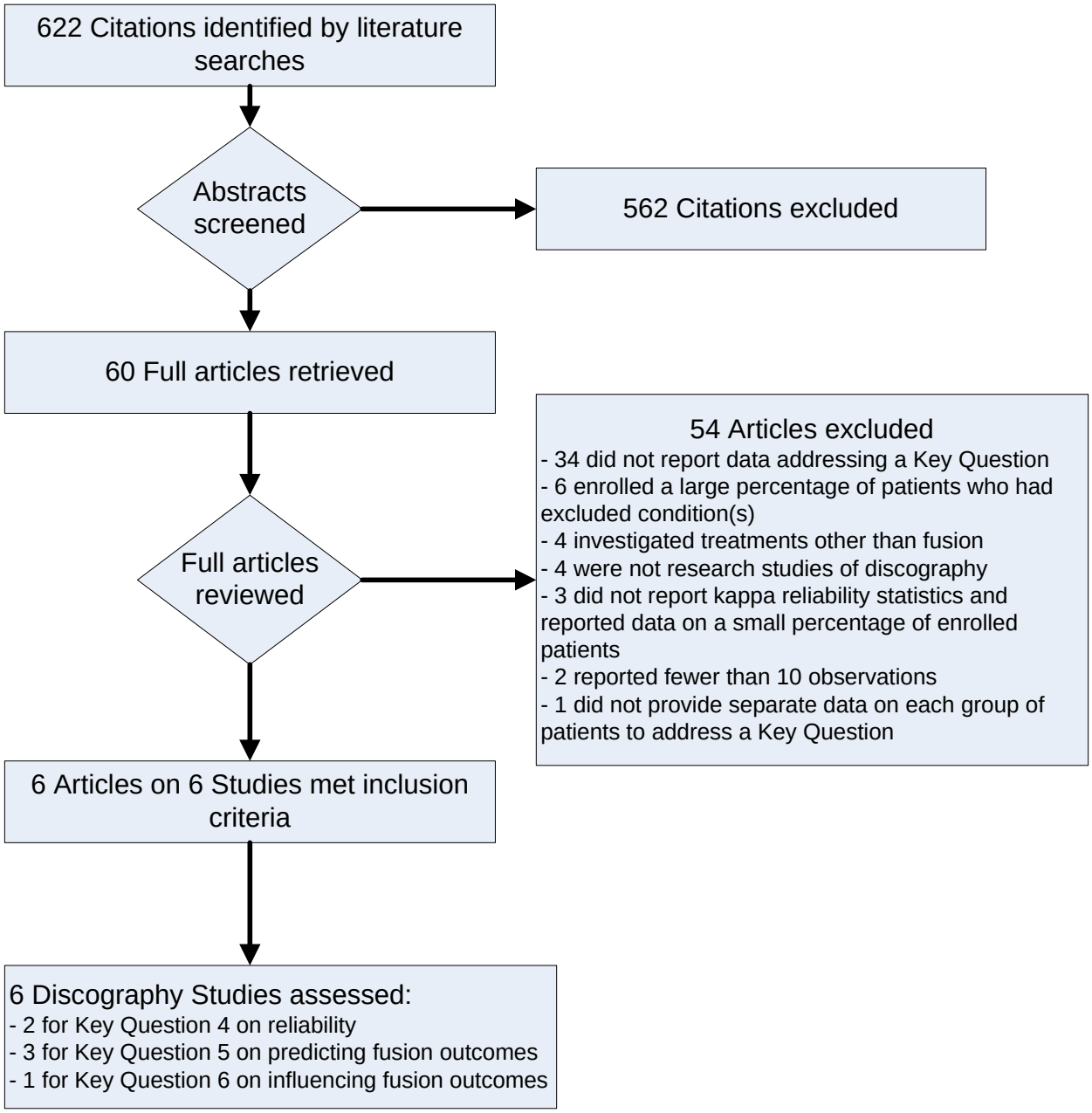
Results for Discography Prior to Lumbar Fusion

(Key Questions 4-6)

Evidence Base

Our searches identified 622 articles on the presurgical use of discography. We retrieved 60 of these that appeared to be relevant based upon their abstracts for possible inclusion. Only six of the articles met the inclusion criteria, and these are listed in Table 23 of Appendix E. The process of article identification is depicted in Figure 3. Reasons for exclusion of the other 54 articles are listed in Table 7 of Appendix B.

Figure 3. Study Attrition Diagram for Discography Key Questions 4-6



Key Question 4: In patients being considered for lumbar fusion surgery, what is the reliability of discography?

a. Test-retest reliability

b. Inter-reader reliability

ECRI Institute evidence assessments:

- **The evidence was insufficient to permit conclusions about the reliability of discography for patients with chronic uncomplicated lumbar degenerative disc disease.**

All evidence tables pertaining to this Key Question appear in Appendix E in Table 23 through Table 30. These tables provide general study characteristics, patient enrollment criteria, patient characteristics, discography details, fusion details, study quality assessments, outcomes for test-retest reliability, and outcomes for inter-rater reliability.

Two studies met the inclusion criteria for this Key Question[3].(33,34) Agorastides (2002)(33) reported data on both test-retest reliability and inter-rater reliability (133 discs in 72 patients), whereas Milette (1999)(34) only reported data on inter-rater reliability (132 discs in 45 patients).

Both studies investigated at least one specific type of reliability: whether a given discogram is judged to have the same morphology grade by the same reader at different times (i.e., test-retest) or by different readers (i.e., inter-rater). Notably, neither study performed two discography exams on the same disc to determine whether the results were consistent between discography injections. Also, neither study investigated the reliability of patients' reports of pain provocation or similarity to their typical pain. These types of reliability represent additional potential sources of variability in discography examinations that have not been assessed in patients with chronic uncomplicated lumbar degenerative disc disease.

We rated the quality of both studies as moderate (quality scores of 7.1 and 7.9). Both studies used consecutive enrollment, reported data on all or almost all enrolled patients, and the discograms were read without consultation of prior discograms or other clinical information about the patient. However, both were retrospective studies that did not report the funding source, and also the Agorastides study did not report whether patient inclusion/exclusion criteria were applied consistently to all patients.

For test-retest reliability, the Agorastides study observed good reliability (values for kappa ranging from 0.80 to 0.85 for the three raters),[4] but because it was a single moderate-quality study at a single center, we deemed this evidence insufficient quantity to permit conclusions. For inter-rater reliability, neither study observed large reliability (values for kappa ranging from 0.66 to 0.77), and neither study was multicenter. These factors, considered together with the moderate quality and limited quantity, mean that the evidence base was insufficient to permit conclusions.

---

[3] After finding only two studies, we removed the date requirement (that studies must have been published in 1990 or later), but when we examined earlier studies, none of them met the other inclusion criteria.

[4] Kappa measures chance-corrected agreement. 0 represents chance, and 1 represents perfect agreement. The standard interpretation of kappa values is that Below 0.0 is Poor agreement; 0.00-0.20 is Slight agreement; 0.21-0.40 is Fair agreement; 0.41-0.60 is Moderate agreement; 0.61-0.80 is Substantial agreement; 0.81-1.00 is Almost Perfect agreement.(35)

Key Question 5: In patients undergoing lumbar fusion surgery, do the results of pre-surgical discography predict the degree of pain reduction or improvement in functional status/quality of life after lumbar fusion surgery?

ECRI Institute evidence assessments:

- **Because of low quality and heterogeneous results from three studies (n = 330 patients), the evidence was insufficient to permit conclusions about the use of discography to predict fusion outcomes in patients with chronic uncomplicated lumbar degenerative disc disease.**

For this Key Question, tables of general study characteristics, patient enrollment criteria, patient characteristics, discography details, fusion details appear in Appendix E in Table 23 through Table 27. The study quality assessments and relevant outcomes appear in Table 31 through Table 33.

This question involves a comparison in surgical outcomes between those who had a positive discography before surgery and those who had a negative discography before surgery. Three studies met the inclusion criteria.(36-38) Willems (2007)(36) included 82 patients, Gill (1992)(37) included 53 patients, and Colhoun (1988)(38) included 195 patients.

Importantly, the three studies each used a different definition of a "positive" discography test:

- Willems (2007)(36) categorized two groups of patients based on *typical pain provocation* in *adjacent-disc(s)*: 1) patients whose adjacent lumbar disc(s) provoked typical pain on discography (N = 22); and 2) patients whose adjacent lumbar disc(s) did not provoke typical pain (or no pain) on discography (N = 60).
- Gill (1992)(37) categorized three groups of patients based on the *morphology* of the *suspected disc*: 1) annular tear beyond the periphery (N = 20); 2) annular tear and contrast extension to the periphery, but not beyond (N = 19); and 3) small annular tear that did not extend to the periphery (N = 14).
- Colhoun (1988)(38) categorized four groups of patients based on both *typical pain provocation and morphology* of the *suspected disc:*1) typical pain provocation and abnormal morphology (N = 137); no pain provocation and abnormal morphology (N = 25); 3) neither pain provocation nor abnormal morphology (N = 6); and 4) total disc resorption of contrast material thus morphology not assessable and pain provocation not reported (N = 27).

Also, the three studies assessed different surgical outcomes:

- Willems (2007)(36) reported mean VAS pain scores at followup as well as the percentage of patients who experienced at least 30% pain relief (at a mean followup 6.7 years)
- Gill (1992)(37) reported a composite outcome involving the percentage of patients showing "improvement on functional testing and pain report", which was based on three

items (Oswestry Pain Questionnaire, VAS, and pain drawing) (at a mean followup 3 years)

- Colhoun (1988)(38) reported a composite outcome involving the percentage of patients who were considered a "success", which was defined as meeting all three conditions: 1) complete relief or significant subjective improvement in symptoms; 2) resumption of work and/or normal duties; 3) no intake of analgesics (at a mean followup 3.6 years).

Furthermore, the three studies reported qualitatively different results (the data appear in Table 33 of Appendix E):

- Willems (2007)(36) found evidence of no statistical difference in VAS pain scores at followup between the two groups, suggesting that discography results do not predict surgical outcomes.
- Gill (1992)(37) did not enroll enough patients to determine whether their data demonstrated a difference or no difference, leaving open the question of whether discography results predict surgical outcomes.
- Colhoun (1988)(38) found evidence of a difference in success rates, suggesting that discography results do predict surgical outcomes. Specifically, "success" was found to be more likely among patients with positive pain provocation and abnormal morphology (88%) than for other groups (52% to 85%).

We rated the quality of all three studies as low (with scores ranging from 4.1 to 4.3). All three were retrospective, non-randomized, unblinded studies. Only one of the three studies (Willems) reported baseline data to assess comparability of patient groups at baseline or attempted to enhance comparability using statistical methods.

Given the low quality, the different definitions of a positive discography, the different outcomes examined, and the qualitatively different results reported, we drew no conclusions about whether discography results predict surgical outcomes.

Key Question 6: In patients being considered for lumbar fusion surgery, do patients who receive discography that influences the treatment choice have better treatment outcomes than patients who do not receive discography?

ECRI Institute evidence assessments:

- **No evidence of acceptable quality was available to address this question; thus, the evidence was insufficient to permit conclusions about the influence of discography on fusion outcomes in patients with chronic uncomplicated lumbar degenerative disc disease.**

For this Key Question, tables of general study characteristics, patient enrollment criteria, patient characteristics, discography details, fusion details appear in Appendix E in Table 23 through Table 27. The study quality assessments and relevant outcomes tables appear in Table 34 through Table 36.

This question involves comparison of treatment outcomes between patients who had received discography before treatment and patients who had not received discography before treatment. Only one study met the inclusion criteria. Madan (2002)(39) retrospectively compared the surgical outcomes of two groups of patients at a single center: 32 patients who were seen between January 1998 and January 1999 and had a positive discography result; and 2) 41 patients who were seen prior to 1998 and had not received discography. All patients underwent the same surgical procedure (instrumented PLIF with posterolateral fusion).

Our quality assessment indicated that the study was very low quality (score 3.4), therefore we excluded the study from further consideration. The primary factors influencing this quality rating were a retrospective, non-concurrent, non-randomized, unblinded design in which the groups were not well-matched at baseline and authors had not attempted statistical methods that may have enhanced group comparability. Due to the lack of evidence of sufficient quality, we drew no conclusions about whether performing discography influences surgical outcomes.

Discussion

General Considerations

In general, positive discography was not reported to be a required indication for surgery in studies of fusion for chronic uncomplicated lumbar degenerative disc disease. Specifically, none of the four RCTs that we included for Key Questions 1-3 mentioned that they required a positive discography as an inclusion criterion. Of the 23 additional studies included in Key Question 2, only three studies even mentioned discography as an inclusion criterion: one only used it when necessary, and two others required positive discography.

Previous Systematic Reviews

Our searches identified three published systematic reviews that evaluated outcomes of lumbar fusion and one systematic review of discography for the diagnosis of uncomplicated lumbar degenerative disc disease. These reviews are summarized below.

One recently-published systematic review evaluated the same four RCTs that compared lumbar fusion to non-operative therapy (non-intensive physical therapy or CBT plus intensive exercise/rehabilitation) and addressed Key Question 1 in the current report. Mirza and Deyo (2007) presented a qualitative summary (no meta-analysis was performed) of the four trials. They noted the same limitations and differences among the trials that we summarized under Key Question 1. They concluded that "surgery may be more efficacious than unstructured nonsurgical care for chronic back pain but may not be more efficacious than structured cognitive behavior therapy. Methodological limitations of the randomized trials prevent firm conclusions."(113)This agrees with the conclusion of the present report in that no firm conclusions were presented.

An earlier Cochrane review on the more general topic of surgery for degenerative lumbar spondylosis included the two earlier RCTs (Fritzell et al. 2001; Brox et al. 2003) that compared fusion to non-operative care. These were only a small part of a much larger review that also addressed numerous comparisons of different surgical treatments. For fusion versus non-operative care, the review stated that the two trials found conflicting results, one showing that "fusion gave better clinical outcomes than conventional physical therapy, while the other showed that fusion was no better than a modern exercise and rehabilitation program." The authors' statement that the findings of Fritzell et al. supported better clinical outcomes for fusion than physical therapy was based solely on consideration of statistical significance in the between-group difference in effect for certain outcomes. They did not assess whether the observed differences in outcomes were clinically significant. Their only statement about complications was for a comparison of instrumented to non-instrumented fusion, where they mentioned that "there is other evidence that [instrumented fusion] may be associated with higher complication rates." The main conclusion of the review was that "there is still insufficient evidence on the effectiveness of surgery on clinical outcomes to draw any firm conclusions. Further studies are needed."(108) Again, we are in agreement that firm conclusions are not possible with the current evidence base.

A recent systematic review on spinal fusion for degenerative disease was sponsored by the Agency for Healthcare Research and Quality (AHRQ) Technology Assessment Program and prepared by the Duke University Evidence-Based Practice Center (McCrory et al. 2006).(114) This report (currently available only in draft form) was primarily focused on outcomes of lumbar fusion in patients age 65 or older with DDD compared to nonsurgical management or other surgical strategies. The report evaluated several of the same studies evaluated in our report, including the four RCTs that directly compared lumbar fusion to nonsurgical therapies. The tentative conclusions of the draft report were that the evidence does not conclusively demonstrate short-term or long-term benefits compared with nonsurgical treatment for degenerative disc disease, which is in agreement with the conclusions of our report.

Another systematic review (Fenton et al. 2007) focused specifically on assessing the adverse outcomes of lumbar interbody fusion using stand-alone cage devices. The authors identified 30 eligible studies with 3,228 patients, and used meta-analytic methods to quantify heterogeneity among seven pre-specified adverse events that they felt would be commonly reported (nonunion, reoperation, major vessel injury, retrograde ejaculation, neurologic injury, dural injury, and infection). They found substantial or marked heterogeneity among rates of nonunion, reoperation, neurologic injury, and dural injury. Random effects meta-regression determined that potential author conflict of interest was associated with significantly lower rates of nonunion. Heterogeneity among other outcomes was not significantly associated with conflicts of interest or other study characteristics.(115) This review focused on a specific type of fusion and the studies evaluated in the review included patients with disorders that were beyond the scope of our report. Performing a meta-analysis of adverse events is risky (even though the authors pre-specified which events they would analyze) because of the likelihood of inconsistent reporting among different studies. This is one possible explanation for the substantial heterogeneity in reported adverse event rates that was not specifically associated with study characteristics.

We located one systematic review of discography in the diagnosis of uncomplicated lumbar degenerative disc disease. Buenaventura et al. (2006)(11) reviewed 81 studies of discography: 69 of lumbar discography, 10 of cervical discography, and 2 of thoracic discography). The review updated an earlier review by Shah et al. (2005).(116)The authors did not present specific key questions, but concluded overall that "Evidence is strong for the diagnostic accuracy of discography as an imaging tool."(11) The conclusion of the review addresses a question (diagnostic accuracy) that was not addressed in our report, because there is no gold standard against which to test discography.

Other Population-Based Studies

Six large population-based cohort studies that did not meet the inclusion criteria but otherwise addressed one or more of Key Questions 2 and 3 may be worth noting. A retrospective study of the Washington State Workers' Compensation system administrative database evaluated data from 1,950 patients who had undergone lumbar fusion between 1994 and 2001 (Maghout-Juratli et al. 2006).(117) This study evaluated the risk of persistent disability, reoperation, and other complications as well as predictors of disability and reoperation in patients with diagnoses of radiculopathy, disc degeneration, disc herniation, spondylolisthesis, and spinal stenosis. The study did not meet our inclusion criteria because >20% of patients had an excluded disorder. The authors reported that 63.9% of patients were receiving work disability payments two years after surgery, 11.8% of patients had some type of postoperative complication within three months

after surgery, and 22% of patients underwent reoperation within two years after surgery. The rates for postoperative complications and reoperations are similar (or within the range) to those reported in studies that met our inclusion criteria. Multivariate analyses identified several factors significantly associated with worse postoperative disability (older age at fusion, presence of psychologic comorbidity, attorney involvement, fusion of more than two vertebral levels, and longer preoperative disability). Use of cages or instrumentation was significantly associated with increased complication risk. Factors significantly associated with reoperation within two years included undergoing discography and undergoing fusion at two or more levels. Two factors significantly associated with lower reoperation risk included DDD and concurrent decompression procedures.

An earlier study of the same Washington State database analyzed data from 388 patients who underwent lumbar fusion during a one-year period from 1986-1987 (Franklin et al. 1994).(118) Overall, 23% of lumbar fusion patients had a reoperation within two years of the index procedure, a number that was in the range of rates reported in the studies that met our inclusion criteria. Multiple logistic regression indicated that fusion with instrumentation was a statistically significant predictor of reoperation. Similarly, multiple logistic regression identified several baseline markers of severity that were significantly associated with work disability status at two years following surgery, including older age at injury, greater time interval between injury to index fusion, greater time on work disability prior to fusion, and greater number of levels fused during index fusion.

A retrospective cohort study of the Workers' Compensation Fund of Utah database evaluated the relationship between biopsychosocial variables and medical/compensation costs for 203 patients who had undergone lumbar fusion between 1990 and 1995 (DeBerard et al. 2003).(119) Multiple regression identified several biological and social variables significantly related to higher total costs, including male gender, older age, greater number of vertebral levels fused, lower levels of education, and completed presurgical psychological evaluation.

A retrospective cohort study based on records of patients covered by the largest Belgian sickness fund reported on factors related to a bad outcome (incapacity for work >1 year) among all 5,808 patients who underwent surgery for lumbar disc herniation between 1992 to 1994 (Donceel and Du Bois, 1998).(120) The percentage of patients who underwent fusion was not reported, but the authors did separately analyze factors that were associated with a bad outcome in this subgroup. They identified unemployment, preoperative work incapacity >6 months, hospital stay >10 days, and age >30 years as significantly associated with work incapacity >1 year in patients who had undergone lumbar fusion.

A retrospective cohort study from the Swedish National Register for Lumbar Spine Surgery reported outcomes on 2,553 patients who underwent lumbar fusion in 1999 (Stromqvist et al. 2001).(121) However, this study mixed other types of spine surgery (discectomy, decompression) with lumbar fusion. The authors did separately report early complications for different types of surgery; early complication rates for different fusion procedures ranged from 8.1% to 13%, but the number of patients who underwent these procedures were not reported. The range of rates reported here is somewhat lower than the range among studies that met our inclusion criteria, although there is some overlap. The study did not meet our inclusion criteria because >20% of patients had an excluded diagnosis.

A retrospective cohort study based on data from the Canadian Institute for Health Information evaluated the rate of early complications and reoperation among 4,772 patients who underwent back surgery in Ontario during a one-year interval from 1990 to 1991 (Hu et al. 1997).(122) The authors reported the total rate as well as separate rates for different types of surgery; for patients undergoing fusion (382 patients) or fusion plus decompression (639 patients), the early complication rates were 22.8% and 15.3%, while the reoperation rates during four years of followup were 9.2% and 10.2%, respectively. These rates were similar (or fell within the ranges) to those reported in studies that met our inclusion criteria. The study did not meet our inclusion criteria because >20% of patients had an excluded diagnosis.

Recent Randomized Trials Comparing Different Lumbar Fusion Procedures

A recent systematic review on spinal fusion for degenerative disease was sponsored by the Agency for Healthcare Research and Quality (AHRQ) Technology Assessment Program and prepared by the Duke University Evidence-Based Practice Center (McCrory et al. 2006).(114) This report (currently available only in draft form) addressed some questions not addressed in the present report, including comparisons of outcomes for different lumbar fusion procedures. A few RCTs comparing different fusion procedures have been published since the search cutoff date of the AHRQ report. We summarize the relevant trials below.

Anterior Lumbar Interbody Fusion (ALIF)

The AHRQ draft report noted that according to a recent review, ALIF procedures accounted for the lowest proportion of fusions performed in the 1980s and 1990s.(114) ALIF fusion rates during that period, based on uncontrolled studies with 583 patients, were 86%. Fusion rates for posterior approaches during the same period were 85% to 91%. The AHRQ draft report summarized data from ten studies (six controlled) of ALIF for patients with DDD. None of the studies included a nonsurgical or conservative management control group. The six controlled trials compared either different variations of ALIF (four studies) or compared ALIF to posterior approaches (two studies). Although the mean patient improvements in ODI after surgery exceeded 15 points in every study that reported pre- and post-treatment ODI scores, one cannot determine the extent of patient improvement that might have occurred in a nonsurgical control group. Our searches identified no new studies comparing ALIF to the posterior approaches.

Posterior Surgical Approaches for Lumbar Fusion

The AHRQ draft report noted that posterior fusions have been recommended for axial back pain despite the potential for dorsal muscle damage during dissection. The report also cited the conclusions of a Cochrane Review comparing Posterior Lumbar Fusion (PLF) to circumferential (ALIF and PLF combined, or A/P) fusion. The review found no difference in fusion failure, complications, or patient-judged improvements. Twenty-three reports of PLF, Posterior Lumbar Interbody Fusion (PLIF), Transforaminal Lumbar Interbody Fusion (TLIF), and circumferential fusion were examined in the AHRQ draft report.(114) Comparisons between surgical approaches were reported in only two of these studies (one RCT and one retrospective study). The RCT showed no differences in ODI response among PLF, PLF plus pedicle screw fixation, and PLIF, but the retrospective study showed better improvement in the PLF and ALIF groups compared to

PLIF, TLIF, and circumferential fusion. However, baseline differences in patient characteristics may have confounded these results.

Our searches identified two new RCTs of posterior lumbar fusion, one comparing circumferential fusion to PLF with instrumentation, and one comparing three fusion methods (PLF, PLIF, and PLF combined with PLIF). Each study measured pain and ODI.

Videbaek et al.(123) randomized patients with severe chronic low back pain due to localized lumbar or lumbosacral segmental instability to PLF with instrumentation (n = 73) or circumferential fusion (n = 73). More than 90% of the patients were available for long-term followup at five to nine years. All outcome measures significantly favored circumferential fusion at long-term followup (median ODI: PLF = 40, circumferential fusion = 28; median low back pain, 0-10 scale: PLF = 6, circumferential fusion = 3). The authors suggest that the significant difference in back pain score may be related to the anterior support provided by a circumferential fusion, especially for patients with disc degeneration.

Kim et al.(124) randomized patients with disabling back pain due to spinal stenosis or spondylolisthesis to three posterior surgical approaches (PLF = 62, PLIF = 57, combined PLF and PLIF = 48). Pedicle screw instrumentations were performed in all patients. After three years, all three groups showed significant reductions in low back pain (from approximately 7.4 at baseline to 2.2 after three years on a 0-10 scale). Each group also showed significant improvements in ODI (from approximately 60 at baseline to 25 after three years). Fusion rates were 92% for PLF, 95% for PLIF, and 96% for PLF+PLIF; the rates were not significantly different. The authors concluded that PLIF alone (without PLF) had advantages of the elimination of donor site pain (only local bone was used for fusion), shorter operating time, and less blood loss.

Both of these studies support the use of bone grafts in the disc space to support the anterior spine rather than bone grafting of the transfer processes in the posterior spine.

Augmentation

The AHRQ draft report refers to several studies that support the use of recombinant human bone morphogenetic protein-2 (rhBMP-2) as an alternative to autograft bone to augment spinal fusion and to one study that did not support the use of autologous growth factor (AGF) gel for augmentation.(114) Our searches identified two new studies, one on each type of augmentation for spinal fusion.

Dimar et al.(125) examined the use of rhBMP-2 in PLF with instrumentation. Patients with symptomatic, single level lumbosacral degenerative disease were randomized to PLF and instrumentation using autogenous iliac crest bone for grafting material (ICBG, n = 45 at two years) or PLF and instrumentation using rhBMP-2 in bovine collagen and tricalcium/hydroxyapatite for grafting material (rhBMP-2, n = 53 at two years). The ODI scores were similar in both groups at each of the time points measured over two years and showed statistically significant improvement compared to preoperative scores. Solid fusions were reported in 73% of the ICBG group and 91% of the rhBMP-2 group ($p = 0.051$). rhBMP-2 may therefore provide a better alternative to autogenous ICBG for spinal fusion because it eliminates iliac bone harvest and its associated morbidities while providing equal or better fusion rates.

Jenis et al.(126) examined the use of AGF gel in circumferential (A/P) fusion). Patients with chronic low back pain due to degenerative disc disease or spondylolisthesis were assigned to

A/P fusion using ICBG (autograft group, n = 22) or allograft combined with AGF gel (AGF group, n = 15). Assignment to treatment was based on the availability of the equipment necessary to prepare the AGF gel. The gel is prepared from each patient's own blood by concentrating the platelet portion using a proprietary system. At six months the fusion rate was 56% in both groups. By 24 months the fusion rates were 85% for the autograft group and 89% for the AGF group. Improvements in pain and ODI were also similar between groups. This study suggests that AGF combined with allograft bone may be a suitable substitute for ICBG, however the authors believe more research is needed to determine the optimal carriers and platelet concentrations for this technology.

Ongoing Clinical Trials

We identified 14 relevant ongoing clinical trials of the United States National Institute of Health clinical trials registry (www.clinicaltrials.gov). Most are randomized controlled trials (RCTs), but a few observational studies are also registered. Ten are on a specific spinal implant device, or devices, all financially sponsored by the manufacturer. The others are assessing fusion with screws or rods that are not identified as being a particular brand. All 14 ongoing trials are summarized in Table 5 below.

Table 5. Ongoing Clinical Trials

Study	Sponsor	Design	Purpose	Start Date	Expected	Indication
Five-year Follow-up of the CHARITE Artificial Disc Compared to Anterior Lumbar Interbody Fusion with the BAK Cage NCT00215332	DePuy Spine	Observational	"To assess the clinical and radiographic outcomes through 5-years following treatment with CHARITE Artificial Disc vs. the BAK Cage for treatment of degenerative disc disease at one level (L4-S1)."	March 2005	367	Participated in either training or randomized arm of the CHARITE Artificial Disc IDE (Investigational device exemption) study, and still have the original implant from that study
Clinical Outcome Study of the Triad Allograft for Posterior Lumbar Fusion NCT 00205101	University of Wisconsin, NuVasive	Observational	"To prospectively measure pain, function, and patient satisfaction in 70 consecutive patients treated by lumbar fusion using the Triad allograft. Results of the Triad allograft will be compared to those of other anterior lumbar interbody fusion (ALIF), transforaminal lumbar interbody fusion (TLIF), and posterior lumbar interbody fusion (PLIF)."	September 2004	70	Isthmic spondylolisthesis, degenerative spondylolisthesis, pseudoarthritis, severe foraminal stenosis, or prior failed discectomy with need for TLIF or PLIF fusion at 1, 2, or 3 levels as determined by surgeon
Dynamic Stabilization for Lumbar Spinal Stenosis with Stabilimax NZ Dynamic Spine Stabilization System NCT 00479544	Applied Spine Technologies	Randomized controlled trial (RCT)	"To assess whether the Stabilimax NZ is at least as safe and effective as the control therapy of fusion in patients receiving decompression surgery for the treatment of clinically symptomatic spinal stenosis at one or two contiguous vertebral levels from L1s1."	February 2007	480	Degenerative spinal stenosis of lumbar spine with evidence of thecal sac and/or cauda equina compression, erve root impingement, hypertrophic facets with canal encroachments, with or without spondylolisthesis up to grade 1 on radiographic image.

Study	Sponsor	Design	Purpose	Start Date	Expected	Indication
Evaluation of Radiographic and Patient Outcomes Following Lumbar Spine Fusion Using Demineralized Bone Matrix (DBM) Mixed with Autograft NCT00254852	Exactech	RCT	"To compare Optecure as an autograft extender (treatment) to autograft alone (control_ in patients undergoing 1 or 2 level fusion of the lumbar spine (one level is defined as two adjacent vertebrae), L2 and below."	October 2005	150	Lumbar stenosis or spondylolisthesis with indication for lumbar fusion of 1 or 2 segments L2-S1
Greenwich Lumbar Stenosis SLIP Study NCT00109213	Greenwich Hospital	RCT	"To determine the proper use of lower back screws and rods (instrumentation) and bony fusion in subjects with one level of degenerative spinal narrowing (stenosis) compressing nerves to the legs with one spinal bone slipping forward on another (spondylolisthesis)."	May 2002	75	Spinal stenosis with a grade I spondylolisthesis
Lumbar Interbody Fusion Using the Telamon Peek Versus the Telamon Hydrosorb Fusion Device NCT 00095095	Medtronic Bakken Research Center	RCT	"To compare two fusion devices, which are used in spinal surgery in order to promote the fusion of two lumbar vertebrae."	October 2004	210	Chronic low back pain with evidence of degenerative changes at L4-L5 or L5-S1 (spondylosis) on plain radiographs and/or CT scan, and/or MRI with PLIF needed at a single level as determined by surgeon

Study	Sponsor	Design	Purpose	Start Date	Expected	Indication
Safety and Effectiveness Study of the TOPS System, a Total Posterior Arthroplasty Implant Designed to Alleviate Pain Resulting from Moderate to Severe Lumbar Stenosis NCT00405691	Impliant, Ltd.	RCT	"To determine whether the TOPS device will effectively treat moderate to severe lumbar stenosis."	September 2006	450	Moderate to severe lumbar spinal stenosis at single level between L3-L5, with radiographic confirmation of thecal sac and/or cauda equina compression, nerve root impingement, hypertrophic facets with canal encroachment, at least 25% reduction in A/P dimension of the central and/or lateral foramen
Safety and Efficacy Study of Healos as a Bone Replacement to Treat Degenerative Disc Disease NCT 00316121	Regenerative, DePuy Spine	RCT	"To determine the safety and effectiveness of HEALOS compared with autograft using the transforaminal lumbar interbody fusion (TLIF) method)."	April 2006	400	Pain with objective evidence of significant disc degeneration at one or two adjacent lumbar level(s) from L2/L3 to L5/S1
Spine Fusion Instrumented with BMP-2 vs. Uninstrumented with Infuse BMP-2 Alone NCT 00405600	Capital District Health Authority, Canada	RCT	"To relieve pain and/or increase stability in painful or unstable spine joints. A patient may or may not receive rods and screws with the use of bone graft materials to facilitation bone growth and a fusion thus preventing movement of the bones of the spine."	November 2006	50	One or two levels contiguous involvement from L1-S1 requiring fusion
Spine Research with Roentgen Stereophotogrammetric Analysis (Spine RSA) NCT00493558	Capital District Health Authority, Canada	Case series	"To gather information on the effectiveness of a new spine implant for patients who require spinal fusion surgery."	July 2007	25	Clinical and radiologic history of spondylolisthesis, no greater than grade 1, spinal stenosis, degenerative disc disease

Study	Sponsor	Design	Purpose	Start Date	Expected	Indication
Study of a Facet Replacement to Treat Spinal Stenosis NCT 00401518	Facet Solutions, Inc.	RCT	"To determine if the Anatomic Facet Replacement System is effective in the treatment of spinal stenosis. The primary objective of the study is to evaluate the overall success rate of the Anatomic Facet Replacement System in patients with spinal stenosis when compared to a posterior spinal fusion control."	November 2006	300	Lateral, lateral recess and/or central canal stenosis Disc height at least half of either adjacent level Spondylolisthesis must be grade 1 or less
Surgical Treatment Comparison for Recurrent Lumbar Disc Herniation NCT 00444405	St. John's Health System, Zimmer Spine	Observational	"To compare patients who underwent decompression/discectomy with pedicle screw fusion to patients who received decompression/discectomy without fusion."	March 2007	50	Recurrent lumbar disc herniation evident on radiography with symptoms including radicular leg pain the improved after the first surgery Sponylolisthesis or spondylolisthesis with less than 3 mm of movement excluded
Total Facet Arthroplasty System (TFAS) Clinical Trial NCT 00418197	Archus Orthopedics Inc.	RCT	"The clinical trial is intended to demonstrate restoration of stability and sagittal balance to the spine."	August 2005	450	Moderate to severe lumbar spinal stenosis L3-L4 or L4-L5 Spondylolisthesis at stenotic level no greater than grade 1, and no more than 3 lumbar levels of degeneration total

Study	Sponsor	Design	Purpose	Start Date	Expected	Indication
Study of Disc Anesthesia for the Preoperative Diagnosis of Chronic Lower Back Pain (SODA) NCT 00443781	Kyphon	Observational	To document and compare diagnostic test results and procedure safety in subjects undergoing both Functional Anaesthetic Discography™ (F.A.D.) and provocative discography (PD).	August 2007	100	Chronic axial low back pain without radicular pain for >six months, not responding to at least 3 months of nonsurgical treatment One or two discs at L5/S1, L4/L5 or L3/L4 with abnormal findings by MRI, including any of the following: loss of disc hydration, disc height, high intensity zone (HIZ), Modic changes at adjacent vertebral endplates, or herniation without nerve root compression Pre-treatment ODI >40 Pre-treatment low back pain by numerical rating scale (NRS) score >4

Clinical Practice Guidelines and Position Statements

We searched the National Guideline Clearinghouse™ (NGC) and conducted hand searches to identify clinical practice guidelines and position statements on spinal fusion and discography for patients with chronic discogenic back pain. Some guidelines addressed either fusion or discography, and some addressed both.

On Fusion

Our searches of the National Guideline Clearinghouse identified sixteen potentially relevant clinical practice guidelines on lumbar fusion. Upon closer evaluation, five were subsequently excluded from this section on fusion. One was excluded for not addressing lumbar fusion, another was excluded for addressing acute back pain only, two were excluded for not addressing low back pain due to degenerative disc disease, and one was not covered in this section because it is more relevant to discography, and is included in that section.

In this section we discuss the eleven relevant guidelines. These include one clinical practice guideline from the Work Loss Data Institute, published in 2006(127), one from the Washington State Department of Labor and Industries published in 2002(54), one from the American College of Occupational and Environmental Medicine published in 2004(128), and eight from the American Association of Neurological Surgeons, all published in 2005.(129-136):

Imaging

Presurgical discography is discussed in the section on clinical practice guidelines on discography, which follows this section on fusion.

The American Association of Neurological Surgeons issued a guideline regarding radiographic assessment of fusion, stating(129):

- "Lateral flexion and extension radiography is recommended as an adjunct to determine the presence of lumbar fusion postoperatively. The lack of motion between vertebrae, in the absence of rigid instrumentation, is highly suggestive of successful fusion."
- "Technetium-99 bone scanning is not recommended as a means to assess lumbar fusion."

But in another guideline they reported, "the correlation between fusion status [and radiological outcome] is not strong."(130)

Recommending a Patient for Surgical Consultation

We identified one guideline recommending patients for surgical consultation. The American College of Occupational and Environmental Medicine published a guideline(128) on low back complaints with the recommendation that "surgical consultation is indicated for patients who have:

- Severe and disabling lower leg symptoms in a distribution consistent with abnormalities on imaging studies (radiculopathy), preferably with accompanying objective signs of neural compromise
- Activity limitations due to radiating leg pain for more than one month or extreme progression of lower leg symptoms
- Clear clinical, imaging, and electrophysiological evidence of a lesion that has been shown to benefit in both the short and long term from surgical repair
- Failure of conservative treatment to resolve disabling radicular symptoms."

Indications for Lumbar Fusion

We identified two guidelines citing indications for lumbar fusion, one from the Work Loss Data Institute, and one from the Washington State Department of Labor and Industries.

The Work Loss Data Institute listed the following indication for spinal fusions for patients with chronic low back pain who had problems for at least six months, except in the presence of fracture or dislocation:(127)

- "Neural arch defect – spondyloloytic spondylolisthesis, congenital unilateral neural arch hyperplaseia
- Segmental instability – excessive motion, as in degenerative spondylolisthesis, surgically induced segmental instability
- Primary mechanical back pain/functional spinal unit failure (in cases other than workers' comp)
- Revision surgery for failed previous operation(s) if significant functional gains are anticipated. Revision surgery for purposes of pain relief must be approached with extreme caution due to the less than 50% success rate reported in medical literature
- Infection, tumor, or deformity of the lumbosacral spine that cause intractable pain, neurological deficit, and/or functional disability"

The Washington State Department of Labor and Industries list the following indications for lumbar fusion in their guideline:(54)

- "If conservative care has failed to relieve symptoms and the patient has had no prior surgery, lumbar fusion should be considered only if the patient has one or more of the following:
    - Mechanical (non-radicular) low back pain with instability
        - Instability of the lumbar segment is defined as at least 4mm of anterior/posterior translation at L3-4 and L4-5, or 5mm of translation at L5-S1 or 11 degrees greater end plate angular change at a single level, compared to an adjacent level. Adequate

flexion/extension views should be taken utilizing techniques that minimize the potential contribution of hip motion to perceived lumbar flexion or extension"

- "Spondylolisthesis with one or more of the following:
  - Objective signs/symptoms of neurogenic claudication OR
  - Objective signs/symptoms of unilateral or bilateral radiculopathy, which are corroborated by neurological examination and by MRI or CT (with or without myelography) OR
  - Instability of the lumbar segment" (Defined above)

- "If conservative care has failed to relieve symptoms and the patient has had a prior laminectomy, discectomy, or other decompressive procedure at the same level, lumbar fusion should be considered only if the patient has one or more of the following:
  - Mechanical (non-radicular) low back pain with instability [as defined above] at the same or adjacent levels OR
  - Mechanical (non-radicular) low back pain with pseudospondylolisthesis, rotational deformity, or other condition leading to a progressive (measurable) deformity, OR
  - Objective signs/symptoms compatible with neurogenic claudication or lumbar radiculopathy that is supported by MRI or CT (with or without myelography) and by a detailed clinical neurological examination OR
  - Evidence from a post-laminectomy structural study of either
    1. 100% loss of a facet surface area unilaterally, OR
    2. 50% combined loss of facet surface area bilaterally"

- "If conservative care has failed to relieve symptoms and the patient has had a prior fusion at the same level, lumbar fusion should be considered only if the patient has one or more of the following:
  - Pseudarthrosis with or without hardware failure, confirmed by objective evidence of pseudarthrosis (e.g., abnormal thin slice CT scan)
  - Neurogenic claudication supported by either MRI, CT, or myelography
  - Lumbar radiculopathy supported by either MRI, CT, or myelography, or supported by a detailed clinical neurological or neurosurgical examination"

- "If conservative care has failed to relieve symptoms and the patient has had a prior fusion at a level adjacent to the new one being considered, lumbar fusion

should be considered only if the patient meets the same criteria as described for patients with no history of spine surgery" (above).

Contraindications to Lumbar Fusion

We identified one guideline citing absolute contraindications to surgery, and three reporting relative contraindications.

Absolute Contraindications:

The Washington State Department of Labor and Industries (2004) cited the following as an absolute contraindication for lumbar fusion(54):

- Initial laminectomy/discectomy related to unilateral compression of a lumbar nerve root

Relative Contraindications:

The Washington State Department of Labor and Industries (2004) cited the following as relative contraindications for lumbar fusion(54):

- Severe physical deconditioning
- Current smoking
- Multiple level degenerative disease of the lumbar spine
- Greater than 12 months of disability (time-loss compensation benefits) prior to consideration of fusion
- No evidence of functional recovery (return to work) for at least six months following the most recent spine surgery
- Psychosocial factors that are correlated with poor outcome, such as
    - History of drug or alcohol abuse
    - High degrees of somatization on clinical or psychological evaluation
    - Presence of a personality disorder or major psychiatric illness
    - Current evidence of a factitious disorder

The American Association of Neurological Surgeons reviewed evidence on lumbar fusion for the treatment of disc herniation and radiculopathy, and concluded: "There is insufficient evidence to recommend a treatment guideline."(135) However, they did comment that lumbar spinal fusion is not recommended as a routine treatment following primary disc excision in patients with a herniated lumbar disc causing radiculopathy, though it may be of use for patients with herniated discs and evidence of preoperative lumbar spinal disability or deformity, for patients with significant chronic axial low back pain and radiculopathy due to disc herniation, or for patients with recurrent lumbar disc herniation.

For patients with low back complaints in general, the American College of Occupational and Environmental Medicine (2005)(128) wrote that patients with co-morbidities including cardiac or respiratory disease, diabetes, or mental illness, as poor candidates for back surgery in general.

Fusion Materials and Methods

We identified three guidelines from the American Association of Neurological Surgeons in 2005 on the recommended types of fusion to be performed or the materials to use.

Regarding type of fusion, they issued the following guideline: "In the context of single-level stand-alone ALIF or ALIF with posterior instrumentation, the addition of a PLF is not recommended as it increases operating room time and blood loss without increasing the likelihood of fusion or the functional outcome."(131)

On the use of pedicle screw fixation as an adjunct to posterolateral fusion for low-back pain, they formed the conclusion, "There is insufficient evidence to recommend a treatment guideline."(132)

They also evaluated the use of bone graft extenders and substitutes and concluded that there were insufficient evidence to recommend a treatment guideline, but note that "Recombinant human BMP-2 [bone morphogenic protein] in combination with HA [hydroxyapatite] and tricalcium phosphate may be used in substitute for autograft bone in some cases of PLF."(133)

Intraoperative Assessment of Fusion

We identified one guideline on intraoperative assessment of fusion. The American Association of Neurological Surgeons reviewed the medical literature on the use of electrophysiological monitoring intraoperatively with lumbar surgery and issued the guideline: "Use of intraoperative SSEP [somatosensory evoked potential] or DSEP [dermatomal sensory evoked potential] monitoring is recommended as an adjunct in those circumstances during instrumented lumbar spinal fusion procedures in which the surgeon desires immediate intraoperative information regarding the potential of a neurological injury," although they acknowledge that there is a high false-positive rate.

Postoperative Assessment of Fusion

We identified one guideline on the postoperative assessment of outcomes. The American Association of Neurological Surgeons reviewed the medical literature but did not form a guideline based upon their findings, stating: "There is insufficient evidence to recommend a guideline for assessment of functional outcome following fusion for lumbar degenerative disc disease."(134) However, they did recommend the use of validated scales when assessing functional outcomes.

Alternatives to Fusion

We identified one guideline on alternatives to fusion. The American Association of Neurological Surgeons reviewed the medical literature on the use of brace therapy as an adjunct to or substitute for lumbar fusion and issued the guideline, "The short-term use of a rigid lumbar support (1-3 weeks) is recommended as a treatment for low-back pain of

relatively short duration (<6 months). The use of a lumbar brace for patients with chronic low-back pain is not recommended because there is no pertinent medical evidence of any long-term benefit or evidence that brace therapy is effective in the treatment of patients with chronic (>6 months) low-back pain."(136)

On Discography

Our searches of the National Guideline Clearinghouse™ (NGC) yielded a total of nine guidelines on the use of discography, and hand searches produced one additional guideline. Upon closer examination, four guidelines were found not to be relevant to either discography or discogenic back pain, and a fifth was superseded by a later guideline.

In this section we discuss the five remaining guidelines, one each from the American Society of Interventional Pain Physicians(58) published in 2007, the Work Loss Data Institute, 2006(127); the American Association of Neurological Surgeons, 2005(24), the Washington State Department of Labor and Industries, 2002(54), and individuals Richard D. Guyer and Donna D. Ohnmeiss from the Texas Back Institute, 2003.(57)

Imaging for Discogenic Back Pain

None of the guidelines recommend the use of discography as a stand-alone preoperative diagnostic test. The Work Loss Data Institute does not recommend the use of discography at all. The American Association of Neurological Surgeons, the American Society of Interventional Pain Physicians, and Guyer and Ohnmeiss recommend the use of discography for patients with equivocal or inconclusive MRI findings. The American Society of Interventional Pain Physicians likewise states that discography could be used if diagnostic tests "such as MRI" provide insufficient diagnostic information. Both Guyer and Ohnmeiss and the American Association of Neurological Surgeons recommend magnetic resonance imaging (MRI) as the diagnostic test of choice for patients with low back pain.

When Discography and MRI Disagree

If MRI is normal, the American Association of Neurological Surgeons recommends against treatment, including surgical treatment, of the disc spaces. Even if discography is positive, in the absence of "MR imaging evidence of disc degeneration", they recommend that surgery not be considered. Similarly, the Washington State Department of Labor and Industries does not consider positive discography as a definitive indication for fusion.

Guyer and Ohnmeiss and the American Academy of Interventional Pain Physicians consider discography appropriate when diagnostic tests (such as MRI) have "failed to reveal clear confirmation of a suspected disc as the source of pain." Similarly, the American Academy of Neurological Surgeons recommended that "discography be reserved for use in patients with equivocal MR imaging findings".

Indications for Discography

Guyer and Ohnmeiss reviewed medical literature on the use of discography and provided specific criteria for usage; they reported no search strategies or search dates in their publication. The American Society of Interventional Pain Physicians released nearly

identical recommendations regarding indications for discography. They recommend the use of discography in the following situations:

1. "Further evaluation of demonstrably abnormal discs to help assess the extent of abnormality or correlation of the abnormality with the clinical symptoms. Such symptoms may include recurrent pain from a previously operated disc and lateral disc herniation.
2. Patients with persistent, severe symptoms in whom other diagnostic tests have failed to reveal clear confirmation of a suspected disc as the source of pain
3. Assessment of patients who have failed to respond to surgical intervention to determine if there is painful pseudarthrosis or a symptomatic disc in a posteriorly fused segment and to help evaluate possible recurrent disc herniation.
4. Assessment of discs before fusion to determine if the discs within the proposed fusion segment are symptomatic, and to determine if discs adjacent to this segment are normal
5. Assessment of candidates for minimally invasive surgical intervention to confirm a contained disc herniation or to investigate dye distribution pattern before chemonucleolysis or percutaneous procedures."

Conclusions

In this section, we summarize the six clinical questions and the conclusions we drew based on the evidence (for more detailed descriptions of the evidence, please consult the *Results* section).

1. Does lumbar fusion surgery reduce pain and improve functional status/quality of life more effectively than nonsurgical treatments?

   ECRI Institute evidence assessments:

   We did not find sufficient evidence that lumbar fusion surgery is more effective to a clinically meaningful degree than nonsurgical treatments for any of the following patient populations, comparisons and outcomes:

   - Meta-analysis of postoperative changes in Oswestry disability scores from two moderate quality RCTs (n = 413 patients) revealed no clinically meaningful difference between fusion and intensive exercise/rehabilitation plus cognitive behavioral therapy (CBT) in patients without prior back surgery (95% CI 0.2 to 7.5, *a priori* 10 point difference defined as clinically meaningful), although the difference slightly favored fusion. Strength of evidence: Weak.
   - The evidence was insufficient to determine whether lumbar fusion provides a greater improvement in back pain (one moderate-quality RCT, n = 64 patients) or quality of life (no acceptable evidence) compared to intensive exercise/rehabilitation plus CBT in patients without prior back surgery.
   - The evidence from one moderate quality RCT (n = 60 patients) was insufficient to determine the relative benefits of lumbar fusion compared to intensive exercise/rehabilitation in patients with prior back surgery.
   - The evidence from one moderate quality RCT (n = 294 patients) was insufficient to determine the relative benefits of lumbar fusion compared to conventional physical therapy in patients with or without prior back surgery.

2. What are the rates of adverse events (perioperative, long-term events, and reoperations) for lumbar fusion surgery and nonsurgical treatments?

   ECRI Institute evidence assessments:

   - Lumbar fusion leads to significantly higher rates of early adverse events compared to non-intensive physical therapy or intensive exercise/rehabilitation plus CBT.
   - Lumbar fusion leads to significantly higher rates of late adverse events at two-year followup compared to non-intensive physical therapy or intensive exercise/rehabilitation plus CBT.

- None of the four RCTs comparing fusion to non-intensive physical therapy or intensive exercise/rehabilitation plus CBT reported any adverse events occurring in patients who only received non-operative care. Most of the reported adverse events could not have occurred in patients who did not undergo surgery.

- Categories of adverse events most frequently reported in fusion studies include reoperation (18/27 studies), infection (14/27 studies), various device-related complications (13/27 studies), neurologic complications (12/27 studies), thrombosis (11/27 studies), bleeding/vascular complications (10/27 studies), and dural injury (10/27 studies).
- The ranges of rates of the most frequently reported complications in fusion studies were: reoperation (0% to 46.1%), infection (0% to 9%), device-related complications (0% to 17.8%), neurologic complications (0.7% to 25.8%), thrombosis (0% to 4%), bleeding/vascular complications (0% to 12.8%), and dural injury (0.5% to 29%).

3. What patient characteristics (i.e., workers' compensation population, patients with chronic pain, psychological distress, and age-groups) are associated with differences in the benefits and adverse events of lumbar fusion surgery?

   ECRI Institute evidence assessments:

   - The evidence from one moderate-quality RCT (n = 294 patients) is insufficient to determine what patient characteristics are associated with differences in the benefits and adverse events of lumbar fusion surgery.

4. In patients being considered for lumbar fusion surgery, what is the reliability of discography?
   a. Test-retest reliability
   b. Inter-reader reliability

   ECRI Institute evidence assessments:

   - The evidence was insufficient to permit conclusions about the reliability of discography for patients with chronic uncomplicated lumbar degenerative disc disease.

5. In patients undergoing lumbar fusion surgery, do the results of pre-surgical discography predict the degree of pain reduction or improvement in functional status/quality of life after lumbar fusion surgery?

   - Because of low quality and heterogeneous results from three studies (n = 330 patients), the evidence was insufficient to permit conclusions about the use of discography to predict fusion outcomes in patients with chronic uncomplicated lumbar degenerative disc disease.

6. In patients being considered for lumbar fusion surgery, do patients who receive discography that influences the treatment choice have better treatment outcomes than patients who do not receive discography?

    - No evidence of reasonable quality was available to address this question; thus, the evidence was insufficient to permit conclusions about the influence of discography on fusion outcomes in patients with chronic uncomplicated lumbar degenerative disc disease.

References

1. Urban JP, Roberts S. Degeneration of the intervertebral disc. Arthritis Res Ther 2003;5(3):120-30

2. Resnick DK. Spinal fusion for discogenic back pain: Patient selection, operative techniques, and outcomes. Tech Neurosurg 2003;8(3):176-190

3. Errico TJ. Lumbar disc arthroplasty. Clin Orthop Relat Res 2005 Jun;(435):106-17

4. Patel RK, Slipman CW. Lumbar degenerative disk disease. [internet]. Omaha (NE): eMedicine; 2007 Jan 18 [accessed 2007 Jul 3]. [20 p]. Available: http://www.emedicine.com/pmr/topic67.htm

5. Manchikanti L. Epidemiology of low back pain. Pain Phys 2000;3(2):167-92

6. Quebec Task Force. Quebec Task Force Report on Spinal Disorders. Spine 1981;12(75 S):1-59

7. Houben MH, Hensen EF, Rauws EA, Hulst RW, Hoff BW, Ende AV, Kate FJ, Tytgat GN. Randomized trial of omeprazole and clarithromycin combined with either metronidazole or amoxycillin in patients with metronidazole-resistant or -susceptible Helicobacter pylori strains. Aliment Pharmacol Ther 1999 Jul;13(7):883-9

8. Kishner S, Babigumira E, Laborde JM. Degenerative disk disease. In: eMedicine [database online]. Omaha (NE): eMedicine.com, Inc.; 1996- [updated 2006 Aug 18]. [accessed 2007 Jul 31]. [26 p]. Available: http://www.emedicine.com/orthoped/topic480.htm

9. Mayo Clinic. Back surgery: coming back from pain. [internet]. Rochester (MN): Mayo Clinic; 2000 Oct 25 [accessed 2003 Jan 28]. [3 p]. Available: http://www.mayoclinic.com

10. Bogduk N, Modic MT. Lumbar discography. Spine 1996 Feb 1;21(3):402-4

11. Buenaventura RM, Shah RV, Patel V, Benyamin R, Singh V. Systematic review of discography as a diagnostic test for spinal pain: an update. Pain Phys 2007 Jan;10(1):147-64

12. Walsh TR, Weinstein JN, Spratt KF, Lehmann TR, Aprill C, Sayre H. Lumbar discography in normal subjects. A controlled, prospective study. J Bone Joint Surg Am 1990 Aug;72(7):1081-8

13. Sachs BL, Vanharanta H, Spivey MA, Guyer RD, Videman T, Rashbaum RF, Johnson RG, Hochschuler SH, Mooney V. Dallas discogram description. A new classification of CT/discography in low-back disorders. Spine 1987 Apr;12(3):287-94

14. Schellhas KP, Pollei SR, Gundry CR, Heithoff KB. Lumbar disc high-intensity zone. Correlation of magnetic resonance imaging and discography. Spine 1996 Jan 1;21(1):79-86

15. Holt EP Jr. The question of lumbar discography. J Bone Joint Surg Am 1968 Jun;50(4):720-6

16. Derby R, Lee SH, Kim BJ, Chen Y, Aprill C, Bogduk N. Pressure-controlled lumbar discography in volunteers without low back symptoms. Pain Med 2005 May-Jun;6(3):213-21; discussion 222-4

17. Carragee EJ, Tanner CM, Khurana S, Hayward C, Welsh J, Date E, Truong T, Rossi M, Hagle C. The rates of false-positive lumbar discography in select patients without low back symptoms. Spine 2000 Jun 1;25(11):1373-80; discussion 1381

18. Carragee EJ, Alamin TF, Miller J, Grafe M. Provocative discography in volunteer subjects with mild persistent low back pain. Spine J 2002 Jan-Feb;2(1):25-34

19. Carragee EJ, Paragioudakis SJ, Khurana S. 2000 Volvo Award winner in clinical studies: Lumbar high-intensity zone and discography in subjects without low back problems. Spine 2000 Dec 1;25(23):2987-92

20. Carragee EJ, Chen Y, Tanner CM, Truong T, Lau E, Brito JL. Provocative discography in patients after limited lumbar discectomy: A controlled, randomized study of pain response in symptomatic and asymptomatic subjects. Spine 2000 Dec 1;25(23):3065-71

21. Massie W, Stevens D. A critical evaluation of discography. J Bone Joint Surg Am 1967;49:1243-4

22. Ohnmeiss DD, Vanharanta H, Guyer RD. The association between pain drawings and computed tomographic/discographic pain responses. Spine 1995 Mar 15;20(6):729-33

23. Carragee EJ, Tanner CM, Yang B, Brito JL, Truong T. False-positive findings on lumbar discography. Reliability of subjective concordance assessment during provocative disc injection. Spine 1999 Dec 1;24(23):2542-7

24. Resnick DK, Choudhri TF, Dailey AT, Groff MW, Khoo L, Matz PG, Mummaneni P, Watters WC 3rd, Wang J, Walters BC, Hadley MN, American Association of Neurological Surgeons/Congress of Neurological Surgeons. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 6: magnetic resonance imaging and discography for patient selection for lumbar fusion. J Neurosurg Spine 2005 Jun;2(6):662-9

25. Carragee EJ, Lincoln T, Parmar VS, Alamin T. A gold standard evaluation of the 'discogenic pain' diagnosis as determined by provocative discography. Spine 2006 Aug 15;31(18):2115-23

26. Carragee EJ, Alamin TF, Carragee JM. Low-pressure positive Discography in subjects asymptomatic of significant low back pain illness. Spine 2006 Mar 1;31(5):505-9

27. Block AR, Vanharanta H, Ohnmeiss DD, Guyer RD. Discographic pain report. Influence of psychological factors. Spine 1996 Feb 1;21(3):334-8

28. Carragee EJ. Psychological and functional profiles in select subjects with low back pain. Spine J 2001 May-Jun;1(3):198-204

29. Hagg O, Fritzell P, Nordwall A, Swedish Lumbar Spine Study Group. The clinical importance of changes in outcome scores after treatment for chronic low back pain. Eur Spine J 2003 Feb;12(1):12-20

30. Samsa G, Edelman D, Rothman ML, Williams GR, Lipscomb J, Matchar D. Determining clinically important differences in health status measures: a general approach with illustration to the Health Utilities Index Mark II. Pharmacoeconomics 1999 Feb;15(2):141-55

31. Brox JI, Reikeras O, Nygaard O, Sorensen R, Indahl A, Holm I, Keller A, Ingebrigtsen T, Grundnes O, Lange JE, Friis A. 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 May;122(1-2):145-55

32. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practiceand problems. Stat Med 2002 Oct 15;21(19):2917-30

33. Agorastides ID, Lam KS, Freeman BJ, Mulholland RC. The Adams classification for cadaveric discograms: inter- and intra-observer error in the clinical setting. Eur Spine J 2002 Feb;11(1):76-9

34. Milette PC, Fontaine S, Lepanto L, Cardinal E, Breton G. Differentiating lumbar disc protrusions, disc bulges, and discs with normal contour but abnormal signal intensity. Magnetic resonance imaging with discographic correlations. Spine 1999 Jan 1;24(1):44-53

35. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics 1977 Mar;33(1):159-74

36. Willems PC, Elmans L, Anderson PG, van der Schaaf DB, de Kleuver M. Provocative discography and lumbar fusion: is preoperative assessment of adjacent discs useful? Spine 2007 May 1;32(10):1094-9; discussion 1100

37. Gill K, Blumenthal SL. Functional results after anterior lumbar fusion at L5-S1 in patients with normal and abnormal MRI scans. Spine 1992 Aug;17(8):940-2

38. Colhoun E, McCall IW, Williams L, Cassar Pullicino VN. Provocation discography as a guide to planning operations on the spine. J Bone Joint Surg Br 1988 Mar;70(2):267-71

39. Madan S, Gundanna M, Harley JM, Boeree NR, Sampson M. Does provocative discography screening of discogenic back pain improve surgical outcome. J Spinal Disord Tech 2002 Jun;15(3):245-51

40. ECRI. Interbody cage with bone morphogenetic protein (InFUSE/LT-CAGE) for degenerative disc disease. Plymouth Meeting (PA): ECRI Health Technology Assessment Information Service; 2004 Dec. 78 p. (Windows on medical technology; no. 120)

41. Trent CG. The surgical treatment of lumbar degenerative disc disease. J S C Med Assoc 2000 Nov;96(11):454-8

42. Bajnoczy S. Artificial disc replacement--evolutionary treatment for degenerative disc disease. AORN J 2005 Aug;82(2):192, 195-202, 205-6; quiz

43. Lee CK, Langrana NA. A review of spinal fusion for degenerative disc disease: need for alternative treatment approach of disc arthroplasty. Spine J 2004 Nov-Dec;4(6 Suppl):173S-176S

44. Swezey RL. Pathophysiology and treatment of intervertebral disk disease. Rheum Dis Clin North Am 1993 Aug;19(3):741-58

45. Yong-Hing K, Kirkaldy-Willis WH. The pathophysiology of degenerative disease of the lumbar spine. Orthop Clin North Am 1983 Jul;14(3):491-504

46. Smith SE, Darden BV, Rhyne AL, Wood KE. Outcome of unoperated discogram-positive low back pain. Spine 1995 Sep 15;20(18):1997-2000; discussion 2000-1

47. Rhyne AL III, Smith SE, Wood KE, Darden BV II, Deyo RA. Outcome of unoperated discogram-positive low back pain. Spine 1995;20(18):1997-2001

48. Sasso RC. Screws, cages or both? [internet]. Wheaton (IL): SpineUniverse; 2005 Jun 27 [accessed 2007 Jul 20]. [13 p]. Available: http://www.spineuniverse.com/displayarticle.php/article1363.html

49. Lau S, Lam KS. Lumbar stabilisation techniques. Curr Orthop 2007;21(1):25-39

50. Stulik J, Vyskocil T, Sebesta P, Kryl J. Atlantoaxial fixation using the polyaxial screw-rod system. Eur Spine J 2007;16(4):479-484

51. Lee KJ, Roper JG, Wang JC. Demineralized bone matrix and spinal arthrodesis. Spine J 2005;5(6 Suppl):217S-223S

52. Fraser RD. Interbody, posterior, and combined lumbar fusions. Spine 1995 Dec 15;20(24 Suppl):167S-177S

53. Auerbach JD, Wills BP, McIntosh TC, Balderston RA. Lumbar disc arthroplasty versus fusion for single-level degenerative disc disease: Two-year results from a randomized prospective study. Semin Spine Surg 2005;17(4):310-318

54. Washington State Department of Labor and Industries. Guidelines for lumbar fusion (arthrodesis). Olympia (WA): Washington State Department of Labor and Industries; 2002 Aug. 5 p

55. Blue Cross and Blue Shield of North Carolina. Health care cost estimator: inpatient hospital admissions. [internet]. Durham (NC): Blue Cross and Blue Shield of North Carolina; 2007 [accessed 2007 Jul 31]. [3 p]. Available: http://www.bcbsnc.com/apps/cost-estimator/report.do?type=inpatient&sub=8

56. Polly DW Jr, Glassman SD, Schwender JD, Shaffrey CI, Branch C, Burkus JK, Gornet MF, Lumbar Spine Study Group. SF-36 PCS benefit-cost ratio of lumbar fusion comparison to other surgical interventions: a thought experiment. Spine 2007 May 15;32(11 Suppl):S20-6

57. Guyer RD, Ohnmeiss DD, NASS. Lumbar discography. Spine J 2003 May-Jun;3(3 Suppl):11S-27S

58. Boswell MV, Trescot AM, Datta S, Schultz DM, Hansen HC, Abdi S, Sehgal N, Shah RV, Singh V, Benyamin RM, Patel VB, Buenaventura RM, Colson JD, Cordner HJ, Epter RS, Jasper JF, Dunbar EE, Atluri SL, Bowman RC, Deer TR, Swicegood JR, Staats PS, Smith HS, Burton AW, Kloth DS, Giordano J, Manchikanti L,. Interventional techniques: evidence-based practice guidelines in the management of chronic spinal pain. Pain Phys 2007 Jan;10(1):7-111

59. Sachs BL, Spivey MA, Vanharanta H, Guyer RD, Rashbaum RF, Hochschuler SH, Smith GV, Scala AD. Technique for lumbar discography and ct/discography in clinical practice. Appl Radiol 1989;18(11):28-30

60. Reitman CA, Hipp JA, Kirking BC, Haas S, Esses SI. Posterior annular strains during discography. J Spinal Disord 2001 Aug;14(4):347-52

61. Murtagh FR, Arrington JA. Computer tomographically guided discography as a determinant of normal disc level before fusion. Spine 1992 Jul;17(7):826-30

62. Carrino JA, Morrison WB, Parker L, Schweitzer ME, Levin DC, Sunshine JH. Spinal injection procedures: volume, provider distribution, and reimbursement in the U.S. medicare population from 1993 to 1999. Radiology 2002 Dec;225(3):723-9

63. Moher D, Pham B, Klassen TP, Schulz KF, Berlin JA, Jadad AR, Liberati A. What contributions do languages other than english make on the results of meta-analyses? J Clin Epidemiol 2000 Sep;53(9):964-72

64. Juni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. Int J Epidemiol 2002 Feb;31(1):115-23

65. Chalmers I, Adams M, Dickersin K, Hetherington J, Tarnow-Mordi W, Meinert C, Tonascia S, Chalmers TC. A cohort study of summary reports of controlled trials. JAMA 1990 Mar 9;263(10):1401-5

66. Neinstein LS. A review of Society for Adolescent Medicine abstracts and Journal of Adolescent Health Care articles. J Adolesc Health Care 1987 Mar;8(2):198-203

67. De Bellefeuille C, Morrison CA, Tannock IF. The fate of abstracts submitted to a cancer meeting: factors which influence presentation and subsequent publication. Ann Oncol 1992 Mar;3(3):187-91

68. Scherer RW, Langenberg P. Full publication of results initially presented in abstracts. In: Cochrane Library [Cochrane methodology review]. Issue 2. Oxford: Update Software; 2001 [accessed 2001 Apr 23]. [35 p]. Available: http://www.cochrane.org/index.htm

69. Marx WF, Cloft HJ, Do HM, Kallmes DF. The fate of neuroradiologic abstracts presented at national meetings in 1993: rate of subsequent publication in peer-reviewed, indexed journals. Am J Neuroradiol 1999 Jun-Jul;20(6):1173-7

70. Yentis SM, Campbell FA, Lerman J. Publication of abstracts presented at anaesthesia meetings. Can J Anaesth 1993 Jul;40(7):632-4

71. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, Porter AC, Tugwell P, Moher D, Bouter LM. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10

72. Treadwell JT, Tregear SJ, Reston JT, Turkelson CM. A system for rating the stability and strength of medical evidence. BMC Med Res Methodol 2006 Oct 19;6:52. Also available: http://www.biomedcentral.com/1471-2288/6/52

73. Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. Stat Med 2007 Jan 15;26(1):53-77

74. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002 Jun 15;21(11):1539-58

75. Hedges LV. Fixed effects models. In: Cooper H, Hedges LV, editors. The handbook of research synthesis. New York (NY): Russell Sage Foundation; 1994. p. 285-99

76. Olkin I. Diagnostic statistical procedures in medical meta-analysis. Stat Med 1999 Sep 15;18(17-18):2331-41

77. Fairbank J, Frost H, Wilson-MacDonald J, Yu LM, Barker K, Collins R, Spine Stabilisation Trial Group. Randomised controlled trial to compare surgical stabilisation of the lumbar spine with an intensive rehabilitation programme for patients with chronic low back pain: the MRC spine stabilisation trial. BMJ 2005 May 28;330(7502):1233

78. Ivar Brox J, Sorensen R, Friis A, Nygaard O, Indahl A, Keller A, Ingebrigtsen T, Eriksen HR, Holm I, Koller AK, Riise R, Reikeras O. Randomized clinical trial of lumbar instrumented fusion and cognitive intervention and exercises in patients with chronic low back pain and disc degeneration. Spine 2003 Sep 1;28(17):1913-21

79. Fritzell P, Hagg O, Wessberg P, Nordwall A. 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 Dec 1;26(23):2521-34

80. Fritzell P, Hagg O, Wessberg P, Nordwall A, Swedish Lumbar Spine Study Group. Chronic low back pain and fusion: a comparison of three surgical techniques: a prospective multicenter randomized study from the Swedish lumbar spine study group. Spine 2002 Jun 1;27(11):1131-41

81. Martin BI, Mirza SK, Comstock BA, Gray DT, Kreuter W, Deyo RA. Reoperation rates following lumbar spine surgery and the influence of spinal fusion procedures. Spine 2007 Feb 1;32(3):382-7

82. Burkus JK, Sandhu HS, Gornet MF, Longley MC. Use of rhBMP-2 in combination with structural cortical allografts: clinical and radiographic outcomes in anterior lumbar spinal surgery. J Bone Joint Surg Am 2005 Jun;87(6):1205-12

83. Burkus JK, Sandhu HS, Gornet MF. Influence of rhBMP-2 on the healing patterns associated with allograft interbody constructs in comparison with autograft. Spine 2006 Apr 1;31(7):775-81

84. Sasso RC, LeHuec JC, Shaffrey C, Spine Interbody Research Group. Iliac crest bone graft donor site pain after anterior lumbar interbody fusion: a prospective patient satisfaction outcome assessment. J Spinal Disord Tech 2005 Feb;18 Suppl:S77-81

85. Bezer M, Kocaoglu B, Aydin N, Guven O. Comparison of traditional and intrafascial iliac crest bone-graft harvesting in lumbar spinal surgery. Int Orthop 2004 Dec;28(6):325-8

86. Scaduto AA, Gamradt SC, Yu WD, Huang J, Delamarter RB, Wang JC. Perioperative complications of threaded cylindrical lumbar interbody fusion devices: anterior versus posterior approach. J Spinal Disord Tech 2003 Dec;16(6):502-7

87. Burkus JK, Gornet MF, Dickman CA, Zdeblick TA. Anterior lumbar interbody fusion using rhBMP-2 with tapered interbody cages. J Spinal Disord Tech 2002 Oct;15(5):337-49

88. Christensen FB, Hansen ES, Eiskjaer SP, Hoy K, Helmig P, Neumann P, Niedermann B, Bunger CE. Circumferential lumbar spinal fusion with Brantigan cage versus posterolateral fusion with titanium Cotrel-Dubousset instrumentation: a prospective, randomized clinical study of 146 patients. Spine 2002 Dec 1;27(23):2674-83

89. McAfee PC, Lee GA, Fedder IL, Cunningham BW. Anterior BAK instrumentation and fusion: complete versus partial discectomy. Clin Orthop Relat Res 2002 Jan;(394):55-63

90. Brantigan JW, Steffee AD, Lewis ML, Quinn LM, Persenaire JM. Lumbar interbody fusion using the Brantigan I/F cage for posterior lumbar interbody fusion and the variable pedicle screw placement system. Two-year results from a Food and Drug Administration investigational device exemption clinical trial. Spine 2000 Jun 1;25(11):1437-46

91. Slosar PJ, Reynolds JB, Schofferman J, Goldthwaite N, White AH, Keaney D. Patient satisfaction after circumferential lumbar fusion . Spine 2000 Mar 15;25(6):722-6

92. Thalgott JS, Chin AK, Ameriks JA, Jordan FT, Daubs MD, Giuffre JM, Fritts K, Timlin M. Gasless endoscopic anterior lumbar interbody fusion utilizing the B. E.R.G. approach. Surg Endosc 2000 Jun;14(6):546-52

93. Regan JJ, Yuan H, McAfee PC. Laparoscopic fusion of the lumbar spine: minimally invasive spine surgery. A prospective multicenter study evaluating open and laparoscopic lumbar fusion. Spine 1999 Feb 15;24(4):402-11

94. Greenough CG, Peterson MD, Hadlow S, Fraser RD. Instrumented posterolateral lumbar fusion. Results and comparison with anterior interbody fusion. Spine 1998 Feb 15;23(4):479-86

95. Kuslich SD, Ulstrom CL, Griffith SL, Ahern JW, Dowdle JD. The Bagby and Kuslich method of lumbar interbody fusion. History, techniques, and 2-year follow-up results of a United States prospective, multicenter trial. Spine 1998 Jun 1;23(11):1267-78; discussion 1279

96. Kuslich SD, Danielson G, Dowdle JD, Sherman J, Fredrickson B, Yuan H, Griffith SL. Four-year follow-up results of lumbar spine arthrodesis using the Bagby and Kuslich lumbar fusion cage. Spine 2000 Oct 15;25(20):2656-62

97. Malter AD, McNeney B, Loeser JD, Deyo RA. 5-year reoperation rates after different types of lumbar spine surgery. Spine 1998 Apr 1;23(7):814-20

98. Penta M, Fraser RD. Anterior lumbar interbody fusion. A minimum 10-year follow-up. Spine 1997 Oct 15;22(20):2429-34

99. Ray CD. Threaded titanium cages for lumbar interbody fusions. Spine 1997 Mar 15;22(6):667-79; discussion 679-80

100. Thomsen K, Christensen FB, Eiskjaer SP, Hansen ES, Fruensgaard S, Bunger CE. 1997 Volvo Award winner in clinical studies. The effect of pedicle screw instrumentation on functional outcome and fusion rates in posterolateral lumbar spinal fusion: a prospective, randomized clinical study. Spine 1997 Dec 15;22(24):2813-22

101. Bjarke Christensen F, Stender Hansen E, Laursen M, Thomsen K, Bunger CE. Long-term functional outcome of pedicle screw instrumentation as a support for posterolateral spinal fusion: randomized clinical study with a 5-year follow-up. Spine 2002 Jun 15;27(12):1269-77

102. Christensen FB, Karlsmose B, Hansen ES, Bunger CE. Radiological and functional outcome after anterior lumbar interbody spinal fusion. Eur Spine J 1996;5(5):293-8

103. Hall BB, Asher MA, Zang RH, Quinn LM. The safety and efficacy of the Isola Spinal Implant System for the surgical treatment of degenerative disc disease. A prospective study. Spine 1996 Apr 15;21(8):982-94

104. Greenough CG, Taylor LJ, Fraser RD. Anterior lumbar fusion. A comparison of noncompensation patients with compensation patients. Clin Orthop Relat Res 1994 Mar;(300):30-7

105. Gill K, Blumenthal SL. Posterior lumbar interbody fusion. A 2-year follow-up of 238 patients. Acta Orthop Scand Suppl 1993;251:108-10

106. Markwalder TM, Battaglia M. Failed back surgery syndrome. Part II: Surgical techniques, implant choice, and operative results in 171 patients with instability of the lumbar spine. Acta Neurochir 1993;123(3-4):129-34

107. Hagg O, Fritzell P, Ekselius L, Nordwall A, Swedish Lumbar Spine Study. Predictors of outcome in fusion surgery for chronic low back pain. A report from the Swedish Lumbar Spine Study. Eur Spine J 2003 Feb;12(1):22-33

108. Gibson JN, Waddell G. Surgery for degenerative lumbar spondylosis (review). In: The Cochrane Database of Systematic Reviews [internet]. Issue 4. Hoboken (NJ): John Wiley & Sons, Ltd.; 2005 [accessed 2005 Oct 19]. [Art. No. CD001352.pub3]. Available: DOI: 10.1002/14651858.CD001352.pub3

109. Liddle SD, Baxter GD, Gracey JH. Exercise and chronic low back pain: what works. Pain 2004 Jan;107(1-2):176-90

110. Tveito TH, Hysing M, Eriksen HR. Low back pain interventions at the workplace: a systematic literature review. Occup Med 2004;54(1):3-13

111. Cleland J, Schulte C, Durall C. The role of therapeutic exercise in treating instability-related lumbar spine pain: a systematic review. J Back Musculoskeletal Rehabil 2002;16(2-3):105-15

112. Carneiro Machado LA, von Sperling de Souza M, Ferreira PH, Ferreira ML. The McKenzie method for low back pain: a systematic review of the literature with a meta-analysis approach. Spine 2006;31(9):E254-E262

113. Mirza SK, Deyo RA. Systematic review of randomized trials comparing lumbar fusion surgery to nonoperative care for treatment of chronic back pain. Spine 2007 Apr 1;32(7):816-23

114. McCrory DC, Turner DA, Patwardhan MB, Richardson WJ. Spinal fusion for treatment of degenerative disease affecting the lumbar spine [draft]. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2006 Nov 1. 93 p

115. Fenton JJ, Mirza SK, Lahad A, Stern BD, Deyo RA. Variation in reported safety of lumbar interbody fusion: influence of industrial sponsorship and other study characteristics. Spine 2007 Feb 15;32(4):471-80

116. Shah RV, Everett CR, McKenzie-Brown AM, Sehgal N. Discography as a diagnostic test for spinal pain: a systematic and narrative review. Pain Phys 2005 Apr;8(2):187-209

117. Maghout Juratli S, Franklin GM, Mirza SK, Wickizer TM, Fulton-Kehoe D. Lumbar fusion outcomes in Washington State workers' compensation. Spine 2006 Nov 1;31(23):2715-23

118. Franklin GM, Haug J, Heyer NJ, McKeefrey SP, Picciano JF. Outcome of lumbar fusion in Washington State workers' compensation. Spine 1994 Sep 1;19(17):1897-903; discussion 1904

119. DeBerard MS, Masters KS, Colledge AL, Holmes EB. Presurgical biopsychosocial variables predict medical and compensation costs of lumbar fusion in Utah workers' compensation patients. Spine J 2003 Nov-Dec;3(6):420-9

120. Donceel P, Du Bois M. Fitness for work after surgery for lumbar disc herniation: a retrospective study. Eur Spine J 1998;7(1):29-35

121. Stromqvist B, Jonsson B, Fritzell P, Hagg O, Larsson B, Lind B. The Swedish national register for lumbar spine surgery: Swedish Society for Spinal Surgery. Acta Orthop Scand 2001 Apr;72(2):99-106

122. Hu RW, Jaglal S, Axcell T, Anderson G. A population-based study of reoperations after back surgery. Spine 1997 Oct 1;22(19):2265-70; discussion 2271

123. Videbaek TS, Christensen FB, Soegaard R, Hansen ES, Hoy K, Helmig P, Niedermann B, Eiskjoer SP, Bunger CE. Circumferential fusion improves outcome in comparison with instrumented posterolateral fusion: long-term results of a randomized clinical trial. Spine 2006 Dec 1;31(25):2875-80

124. Kim KT, Lee SH, Lee YH, Bae SC, Suk KS. Clinical outcomes of 3 fusion methods through the posterior approach in the lumbar spine. Spine 2006 May 20;31(12):1351-7; discussion 1358

125. Dimar JR, Glassman SD, Burkus KJ, Carreon LY. Clinical outcomes and fusion success at 2 years of single-level instrumented posterolateral fusions with recombinant human bone morphogenetic protein-2/compression resistant matrix versus iliac crest bone graft. Spine 2006 Oct 15;31(22):2534-9; discussion 2540

126. Jenis LG, Banco RJ, Kwon B. A prospective study of Autologous Growth Factors (AGF) in lumbar interbody fusion. Spine J 2006 Jan-Feb;6(1):14-20

127. Work Loss Data Institute. Low back - lumbar & thoracic (acute & chronic). Corpus Christi (TX): Work Loss Data Institute; 2006. 431 p

128. Low back complaints. Elk Grove Village (IL): American College of Occupational and Environmental Medicine (ACOEM); 2004. 41 p

129. Resnick DK, Choudhri TF, Dailey AT, Groff MW, Khoo L, Matz PG, Mummaneni P, Watters WC 3rd, Wang J, Walters BC, Hadley MN, American Association of Neurological Surgeons/Congress of Neurological Surgeons. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 4: radiographic assessment of fusion. J Neurosurg Spine 2005 Jun;2(6):653-7

130. Resnick DK, Choudhri TF, Dailey AT, Groff MW, Khoo L, Matz PG, Mummaneni P, Watters WC 3rd, Wang J, Walters BC, Hadley MN, American Association of Neurological Surgeons/Congress of Neurological Surgeons. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 5: correlation between radiographic and functional outcome. J Neurosurg Spine 2005 Jun;2(6):658-61

131. Resnick DK, Choudhri TF, Dailey AT, Groff MW, Khoo L, Matz PG, Mummaneni P, Watters WC 3rd, Wang J, Walters BC, Hadley MN, American Association of Neurological Surgeons/Congress of Neurological Surgeons. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 11: interbody techniques for lumbar fusion. J Neurosurg Spine 2005 Jun;2(6):692-9

132. Resnick DK, Choudhri TF, Dailey AT, Groff MW, Khoo L, Matz PG, Mummaneni P, Watters WC 3rd, Wang J, Walters BC, Hadley MN, American Association of Neurological Surgeons/Congress of Neurological Surgeons. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 12: pedicle screw fixation as an adjunct to posterolateral fusion for low-back pain. J Neurosurg Spine 2005 Jun;2(6):700-6

133. Resnick DK, Choudhri TF, Dailey AT, Groff MW, Khoo L, Matz PG, Mummaneni P, Watters WC 3rd, Wang J, Walters BC, Hadley MN, American Association of Neurological Surgeons/Congress of Neurological Surgeons. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 16: bone graft extenders and substitutes. J Neurosurg Spine 2005 Jun;2(6):733-6

134. Resnick DK, Choudhri TF, Dailey AT, Groff MW, Khoo L, Matz PG, Mummaneni P, Watters WC 3rd, Wang J, Walters BC, Hadley MN, American Association of Neurological Surgeons/Congress of Neurological Surgeons. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 2: assessment of functional outcome. J Neurosurg Spine 2005 Jun;2(6):639-46

135. Resnick DK, Choudhri TF, Dailey AT, Groff MW, Khoo L, Matz PG, Mummaneni P, Watters WC 3rd, Wang J, Walters BC, Hadley MN, American Association of Neurological Surgeons/Congress of Neurological Surgeons. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 8: lumbar fusion for disc herniation and radiculopathy. J Neurosurg Spine 2005 Jun;2(6):673-8

136. Resnick DK, Choudhri TF, Dailey AT, Groff MW, Khoo L, Matz PG, Mummaneni P, Watters WC 3rd, Wang J, Walters BC, Hadley MN, American Association of Neurological Surgeons/Congress of Neurological Surgeons. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 14: brace therapy as an adjunct to or substitute for lumbar fusion. J Neurosurg Spine 2005 Jun;2(6):716-24

137. Peul WC, van Houwelingen HC, van den Hout WB, Brand R, Eekhof JA, Tans JT, Thomeer RT, Koes BW, Leiden-The Hague Spine Intervention Prognostic Study Group. Surgery versus prolonged conservative treatment for sciatica. N Engl J Med 2007 May 31;356(22):2245-56

138. Weinstein JN, Lurie JD, Tosteson TD, Hanscom B, Tosteson AN, Blood EA, Birkmeyer NJ, Hilibrand AS, Herkowitz H, Cammisa FP, Albert TJ, Emery SE, Lenke LG, Abdu WA, Longley M, Errico TJ, Hu SS. Surgical versus nonsurgical treatment for lumbar degenerative spondylolisthesis. N Engl J Med 2007 May 31;356(22):2257-70

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

140. Bentsen SB, Wahl AK, Hanestad BR, Strand LI. Outcomes for patients with chronic low back pain treated using instrumented fusion. Scand J Caring Sci 2007 Mar;21(1):71-8

141. Glassman SD, Carreon L, Djurasovic M, Campbell MJ, Puno RM, Johnson JR, Dimar JR. Posterolateral lumbar spine fusion with INFUSE bone graft. Spine J 2007 Jan-Feb;7(1):44-9

142. Mirovsky Y, Floman Y, Smorgick Y, Ashkenazi E, Anekstein Y, Millgram MA, Giladi M. Management of deep wound infection after posterior lumbar interbody fusion with cages. J Spinal Disord Tech 2007 Apr;20(2):127-31

143. Ronnberg K, Lind B, Zoega B, Halldin K, Gellerstedt M, Brisby H. Patients' satisfaction with provided care/information and expectations on clinical outcome after lumbar disc herniation surgery. Spine 2007;32(2):256-61

144. Sasso RC, Garrido BJ. Computer-assisted spinal navigation versus serial radiography and operative time for posterior spinal fusion at L5-S1. J Spinal Disord Tech 2007 Apr;20(2):118-22

145. Sinikallio S, Aalto T, Airaksinen O, Herno A, Kroger H, Savolainen S, Turunen V, Viinamaki H. Lumbar spinal stenosis patients are satisfied with short-term results of surgery - Younger age, symptom severity, disability and depression decrease satisfaction. Disabil Rehabil 2007;29(7):537-44

146. Andersen T, Christensen FB, Bunger C. Evaluation of a Dallas Pain Questionnaire classification in relation to outcome in lumbar spinal fusion. Eur Spine J 2006 Nov;15(11):1671-85

147. Anderson PA, Schwaegler PE, Cizek D, Leverson G. Work status as a predictor of surgical outcome of discogenic low back pain. Spine 2006 Oct 1;31(21):2510-5

148. Anand N, Hamilton JF, Perri B, Miraliakbar H, Goldstein T. Cantilever TLIF with structural allograft and RhBMP2 for correction and maintenance of segmental sagittal lordosis: long-term clinical, radiographic, and functional outcome. Spine 2006 Sep 15;31(20):E748-53

149. Best NM, Sasso RC. Efficacy of translaminar facet screw fixation in circumferential interbody fusions as compared to pedicle screw fixation. J Spinal Disord Tech 2006;19(2):98-103

150. Cakir B, Richter M, Huch K, Puhl W, Schmidt R. Dynamic stabilization of the lumbar spine. Orthopedics 2006;29(8):716-722

151. Epstein NE. Efficacy of pneumatic compression stocking prophylaxis in the prevention of deep venous thrombosis and pulmonary embolism following 139 lumbar laminectomies with instrumented fusions. J Spinal Disord Tech 2006 Feb;19(1):28-31

152. Fogel GR, Toohey JS, Neidre A, Brantigan JW. Outcomes of L1-L2 posterior lumbar interbody fusion with the Lumbar I/F cage and the variable screw placement system: reporting unexpected poor fusion results at L1-L2. Spine J 2006 Jul-Aug;6(4):421-7

153. Glassman S, Gornet MF, Branch C, Polly D Jr, Peloza J, Schwender JD, Carreon L. MOS short form 36 and Oswestry Disability Index outcomes in lumbar fusion: a multicenter experience. Spine J 2006 Jan-Feb;6(1):21-6

154. Hsu CJ, Chou WY, Chang WN, Wong CY. Clinical follow up after instrumentation-augmented lumbar spinal surgery in patients with unsatisfactory outcomes. J Neurosurg Spine 2006 Oct;5(4):281-6

155. Kim YJ, Bridwell KH, Lenke LG, Rhim S, Cheh G. Pseudarthrosis in long adult spinal deformity instrumentation and fusion to the sacrum: prevalence and risk factor analysis of 144 cases. Spine 2006 Sep 15;31(20):2329-36

156. Kim YJ, Bridwell KH, Lenke LG, Rhim S, Cheh G. Sagittal thoracic decompensation following long adult lumbar spinal instrumentation and fusion to L5 or S1: causes, prevalence, and risk factor analysis. Spine 2006 Sep 15;31(20):2359-66

157. Motosuneya T, Asazuma T, Tsuji T, Watanabe H, Nakayama Y, Nemoto K. Postoperative change of the cross-sectional area of back musculature after 5 surgical procedures as assessed by magnetic resonance imaging. J Spinal Disord Tech 2006 Jul;19(5):318-22

158. Neen D, Noyes D, Shaw M, Gwilym S, Fairlie N, Birch N. Healos and bone marrow aspirate used for lumbar spine fusion: a case controlled study comparing healos with autograft. Spine 2006 Aug 15;31(18):E636-40

159. Okuda S, Miyauchi A, Oda T, Haku T, Yamamoto T, Iwasaki M. Surgical complications of posterior lumbar interbody fusion with total facetectomy in 251 patients. J Neurosurg Spine 2006 Apr;4(4):304-9

160. Okuda S, Takenori O, Miyauchi A, Haku T, Yamamoto T, Iwasaki M. Surgical outcomes of posterior lumbar interbody fusion in elderly patients. J Bone Joint Surg Am 2006;88(12):2714-20

161. Okuda S, Oda T, Miyauchi A, Haku T, Yamamoto T, Iwasaki M. Surgical outcomes of posterior lumbar interbody fusion in elderly patients. J Bone Joint Surg Am 2006 Dec;88(12):2714-20

162. Pappou IP, Papadopoulos EC, Sama AA, Girardi FP, Cammisa FP. Postoperative infections in interbody fusion for degenerative spinal disease. Clin Orthop Relat Res 2006 Mar;444:120-8

163. Raffo CS, Lauerman WC. Predicting morbidity and mortality of lumbar spine arthrodesis in patients in their ninth decade. Spine 2006 Jan 1;31(1):99-103

164. Ringel F, Stoffel M, Stuer C, Meyer B. Minimally invasive transmuscular pedicle screw fixation of the thoracic and lumbar spine. Neurosurgery 2006 Oct;59(4 Suppl 2):ONS361-6

165. Satoh I, Yonenobu K, Hosono N, Ohwada T, Fuji T, Yoshikawa H. Indication of posterior lumbar interbody fusion for lumbar disc herniation. J Spinal Disord Tech 2006 Apr;19(2):104-8

166. Sengupta DK, Truumees E, Patel CK, Kazmierczak C, Hughes B, Elders G, Herkowitz HN. Outcome of local bone versus autogenous iliac crest bone graft in the instrumented posterolateral fusion of the lumbar spine. Spine 2006 Apr 20;31(9):985-91

167. Soegaard R, Christensen FB, Lauerberg I, Bunger CE. Lumbar spinal fusion patients' demands to the primary health sector: evaluation of three rehabilitation protocols. A prospective randomized study.[erratum appears in Eur Spine J. 2006;15(5):657 Note: Lauersen, Ida [corrected to Lauerberg, Ida]]. Eur Spine J 2006 May;15(5):648-56

168. Suda K, Ito M, Abumi K, Haba H, Taneichi H, Kaneda K. Radiological risk factors of pseudoarthrosis and/or instrument breakage after PLF with the pedicle screw system in isthmic spondylolisthesis. J Spinal Disord Tech 2006 Dec;19(8):541-6

169. Swan J, Hurwitz E, Malek F, van den Haak E, Cheng I, Alamin T, Carragee E. Surgical treatment for unstable low-grade isthmic spondylolisthesis in adults: a prospective controlled study of posterior instrumented fusion compared with combined anterior-posterior fusion. Spine J 2006 Nov-Dec;6(6):606-14

170. Villavicencio AT, Burneikiene S, Bulsara KR, Thramann JJ. Perioperative complications in transforaminal lumbar interbody fusion versus anterior-posterior reconstruction for lumbar disc degeneration and instability. J Spinal Disord Tech 2006 Apr;19(2):92-7

171. Weinstein JN, Lurie JD, Tosteson TD, Skinner JS, Hanscom B, Tosteson AN, Herkowitz H, Fischgrund J, Cammisa FP, Albert T, Deyo RA. Surgical vs nonoperative treatment for lumbar disk herniation: The Spine Patient Outcomes Research Trial (SPORT) observational cohort. JAMA 2006 Nov 22;296(20):2451-9

172. Yamashita K, Ohzono K, Hiroshima K. Patient satisfaction as an outcome measure after surgical treatment for lumbar spinal stenosis: Testing the validity and discriminative ability in terms of symptoms and functional status. Spine 2006;31(22):2602-8

173. Yi S, Yoon do H, Kim KN, Kim SH, Shin HC. Postoperative spinal epidural hematoma: risk factor and clinical outcome. Yonsei Med J 2006 Jun 30;47(3):326-32

174. Aiki H, Ohwada O, Kobayashi H, Hayakawa M, Kawaguchi S, Takebayashi T, Yamashita T. Adjacent segment stenosis after lumbar fusion requiring second operation. J Orthop Sci 2005 Sep;10(5):490-5

175. Atlas SJ, Keller RB, Wu YA, Deyo RA, Singer DE. Long-term outcomes of surgical and nonsurgical management of lumbar spinal stenosis: 8 to 10 year results from the maine lumbar spine study. Spine 2005 Apr 15;30(8):936-43

176. Bednar DA, Al-Tunaib W. Failure of reconstitution of open-section, posterior iliac-wing bone graft donor sites after lumbar spinal fusion. Observations with implications for the etiology of donor site pain. Eur Spine J 2005 Feb;14(1):95-8

177. Blumenthal S, McAfee PC, Guyer RD, Hochschuler SH, Geisler FH, Holt RT, Garcia R Jr, Regan JJ, Ohnmeiss DD. A prospective, randomized, multicenter Food and Drug Administration investigational device exemptions study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part I: evaluation of clinical outcomes. Spine 2005 Jul 15;30(14):1565-75; discussion E387

178. McAfee PC, Cunningham B, Holsapple G, Adams K, Blumenthal S, Guyer RD, Dmietriev A, Maxwell JH, Regan JJ, Isaza J. A prospective, randomized, multicenter Food and Drug Administration investigational device exemption study of lumbar total disc replacement with the CHARITE artificial disc versus lumbar fusion: part II: evaluation of radiographic outcomes and correlation of surgical technique accuracy with clinical outcomes. Spine 2005 Jul 15;30(14):1576-83; discussion E388

179. Geisler FH, Blumenthal SL, Guyer RD, McAfee PC, Regan JJ, Johnson JP, Mullin B. Neurological complications of lumbar artificial disc replacement and comparison of clinical results with those related to lumbar arthrodesis in the literature: results of a multicenter, prospective, randomized investigational device exemption study of Charité intervertebral disc. Invited submission from the Joint Section Meeting on Disorders of the Spine and Peripheral Nerves, March 2004. J Neurosurg Spine 2004 Sep;1(2):143-54

180. Bostelmann R, Benini A. Computer-navigated pedicle screw insertion in the lumbar spine. Oper Orthop Traumatol 2005 Jun;17(2):178-94

181. Burton CV. Spinal instrumentation and attendant problems. Surg Neurol 2005;63(4):349

182. Chang Y, Singer DE, Wu YA, Keller RB, Atlas SJ. The effect of surgical and nonsurgical treatment on longitudinal outcomes of lumbar spinal stenosis over 10 years. J Am Geriatr Soc 2005 May;53(5):785-92

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

184. Epstein D. Intensive rehabilitation may be more cost effective than surgical stabilisation for chronic low back pain: Commentary. Aust J Physiother 2005;51(4):269

185. Hsu CJ, Chou WY, Teng HP, Chang WN, Chou YJ. Coralline hydroxyapatite and laminectomy-derived bone as adjuvant graft material for lumbar posterolateral fusion. J Neurosurg Spine 2005 Oct;3(4):271-5

186. Inamasu J, Guiot BH. Laparoscopic anterior lumbar interbody fusion: A review of outcome studies. Minim Invasive Neurosurg 2005;48(6):340-347

187. Jang JS, Lee SH. Clinical analysis of percutaneous facet screw fixation after anterior lumbar interbody fusion. J Neurosurg Spine 2005 Jul;3(1):40-6

188. Kilincer C, Steinmetz MP, Sohn MJ, Benzel EC, Bingaman W. Effects of age on the perioperative characteristics and short-term outcome of posterior lumbar fusion surgery. J Neurosurg Spine 2005 Jul;3(1):34-9

189. Lamberg TS, Remes VM, Helenius IJ, Schlenzka DK, Yrjonen TA, Osterman KE, Tervahartiala PO, Seitsalo SK, Poussa MS. Long-term clinical, functional and radiological outcome 21 years after posterior or posterolateral fusion in childhood and adolescence isthmic spondylolisthesis. Eur Spine J 2005;14(7):639-44

190. Lettice JJ, Kula TA, Derby R, Kim BJ, Lee SH, Seo KS. Does the number of levels affect lumbar fusion outcome? Spine 2005 Mar 15;30(6):675-81

191. Lidar Z, Beaumont A, Lifshutz J, Maiman DJ. Clinical and radiological relationship between posterior lumbar interbody fusion and posterolateral lumbar fusion. Surg Neurol 2005 Oct;64(4):303-8; discussion 308

192. Matsudaira K, Yamazaki T, Seichi A, Takeshita K, Hoshi K, Kishimoto J, Nakamura K. Spinal stenosis in grade I degenerative lumbar spondylolisthesis: a comparative study of outcomes following laminoplasty and laminectomy with instrumented spinal fusion. J Orthop Sci 2005 May;10(3):270-6

193. Moffett JK. Spinal fusion slightly more effective than intensive rehabilitation for chronic low back pain: Commentary. Aust J Physiother 2005;51(4):268

194. Potter BK, Freedman BA, Verwiebe EG, Hall JM, Polly DW Jr, Kuklo TR. Transforaminal lumbar interbody fusion: clinical and radiographic results and complications in 100 consecutive patients. J Spinal Disord Tech 2005 Aug;18(4):337-46

195. Rampersaud YR, Pik JH, Salonen D, Farooq S. Clinical accuracy of fluoroscopic computer-assisted pedicle screw fixation: a CT analysis. Spine 2005 Apr 1;30(7):E183-90

196. Sasso RC, Best NM, Mummaneni PV, Reilly TM, Hussain SM. Analysis of operative complications in a series of 471 anterior lumbar interbody fusion procedures. Spine 2005 Mar 15;30(6):670-4

197. Schuler TC, Burkus JK, Gornet MF, Subach BR, Zdeblick TA. The correlation between preoperative disc space height and clinical outcomes after anterior lumbar interbody fusion. J Spinal Disord Tech 2005 Oct;18(5):396-401

198. Shabat S, Folman Y, Arinzon Z, Adunsky A, Catz A, Gepstein R. Gender differences as an influence on patients' satisfaction rates in spinal surgery of elderly patients. Eur Spine J 2005;14(10):1027-1032

199. Tuli SK, Eichler ME, Woodard EJ. Comparison of perioperative morbidity in translaminar facet versus pedicle screw fixation. Orthopedics 2005 Aug;28(8):773-8

200. Wenger M, Sapio N, Markwalder TM. Long-term outcome in 132 consecutive patients after posterior internal fixation and fusion for Grade I and II isthmic spondylolisthesis. J Neurosurg Spine 2005 Mar;2(3):289-97

201. Brau SA, Delamarter RB, Schiffman ML, Williams LA, Watkins RG. Vascular injury during anterior lumbar surgery. Spine J 2004 Jul-Aug;4(4):409-12

202. Burkus JK, Schuler TC, Gornet MF, Zdeblick TA. Anterior lumbar interbody fusion for the management of chronic lower back pain: current strategies and concepts. Orthop Clin North Am 2004 Jan;35(1):25-32

203. Cammisa FP Jr, Lowery G, Garfin SR, Geisler FH, Klara PM, McGuire RA, Sassard WR, Stubbs H, Block JE. Two-year fusion rate equivalency between Grafton DBM gel and autograft in posterolateral spine fusion: a prospective controlled trial employing a side-by-side comparison in the same patient. Spine 2004 Mar 15;29(6):660-6

204. Christensen FB. Lumbar spinal fusion. Outcome in relation to surgical methods, choice of implant and postoperative rehabilitation. Acta Orthop Scand Suppl 2004 Oct;75(313):2-43

205. Gaetani P, Aimar E, Panella L, Debernardi A, Tancioni F, Rodriguez y Baena R. Surgery for herniated lumbar disc disease: factors influencing outcome measures. An analysis of 403 cases. Funct Neurol 2004 Jan-Mar;19(1):43-9

206. Ghiselli G, Wang JC, Bhatia NN, Hsu WK, Dawson EG. Adjacent segment degeneration in the lumbar spine. J Bone Joint Surg Am 2004 Jul;86-A(7):1497-503

207. Greiner-Perth R, Boehm H, Allam Y, Elsaghir H, Franke J. Reoperation rate after instrumented posterior lumbar interbody fusion: a report on 1680 cases. Spine 2004 Nov 15;29(22):2516-20

208. Kornblum MB, Fischgrund JS, Herkowitz HN, Abraham DA, Berkower DL, Ditkoff JS. Degenerative lumbar spondylolisthesis with spinal stenosis: a prospective long-term study comparing fusion and pseudarthrosis. Spine 2004 Apr 16;29(7):726-33; discussion 733-4

209. Korovessis P, Papazisis Z, Koureas G, Lambiris E. Rigid, semirigid versus dynamic instrumentation for degenerative lumbar spinal stenosis: a correlative radiological and clinical analysis of short-term results. Spine 2004 Apr 16;29(7):735-42

210. Lai PL, Chen LH, Niu CC, Fu TS, Chen WJ. Relation between laminectomy and development of adjacent segment instability after lumbar fusion with pedicle fixation. Spine 2004 Nov 15;29(22):2527-32; discussion 2532

211. Lai PL, Chen LH, Niu CC, Chen WJ. Effect of postoperative lumbar sagittal alignment on the development of adjacent instability. J Spinal Disord Tech 2004 Oct;17(5):353-7

212. Seal C, Gelb D, Ludwig S. Complications of surgical treatment for lumbar stenosis in the elderly population. Curr Opin Orthop 2004;15(3):172-4

213. Vaccaro AR, Patel T, Fischgrund J, Anderson DG, Truumees E, Herkowitz HN, Phillips F, Hilibrand A, Albert TJ, Wetzel T, McCulloch JA. A pilot study evaluating the safety and efficacy of OP-1 Putty (rhBMP-7) as a replacement for iliac crest autograft in posterolateral lumbar arthrodesis for degenerative spondylolisthesis. Spine 2004 Sep 1;29(17):1885-92

214. Andersen T, Christensen FB, Hansen ES, Bunger C. Pain 5 years after instrumented and non-instrumented posterolateral lumbar spinal fusion. Eur Spine J 2003 Aug;12(4):393-9

215. Beutler WJ, Peppelman WC Jr. Anterior lumbar fusion with paired BAK standard and paired BAK Proximity cages: subsidence incidence, subsidence factors, and clinical outcome. Spine J 2003 Jul-Aug;3(4):289-93

216. Carreon LY, Puno RM, Dimar JR 2nd, Glassman SD, Johnson JR. Perioperative complications of posterior lumbar decompression and arthrodesis in older adults. J Bone Joint Surg Am 2003 Nov;85-A(11):2089-92

217. Christensen FB, Laurberg I, Bunger CE. Importance of the back-cafe concept to rehabilitation after lumbar spinal fusion: a randomized clinical study with a 2-year follow-up. Spine 2003 Dec 1;28(23):2561-9

218. Gillet P. The fate of the adjacent motion segments after lumbar fusion. J Spinal Disord Tech 2003 Aug;16(4):338-45

219. Glaser J, Stanley M, Sayre H, Woody J, Found E, Spratt K. A 10-year follow-up evaluation of lumbar spine fusion with pedicle screw fixation. Spine 2003 Jul 1;28(13):1390-5

220. Glassman SD, Alegre G, Carreon L, Dimar JR, Johnson JR. Perioperative complications of lumbar instrumentation and fusion in patients with diabetes mellitus. Spine J 2003 Nov-Dec;3(6):496-501

221. Fritzell P, Hagg O, Nordwall A, Swedish Lumbar Spine Study Group. Complications in lumbar fusion surgery for chronic low back pain: comparison of three surgical techniques used in a prospective randomized study. A report from the Swedish Lumbar Spine Study Group. Eur Spine J 2003 Apr;12(2):178-89

222. Hagg O, Fritzell P, Hedlund R, Moller H, Ekselius L, Nordwall A, Swedish Lumbar Spine Study. Pain-drawing does not predict the outcome of fusion surgery for chronic low-back pain: a report from the Swedish Lumbar Spine Study. Eur Spine J 2003 Feb;12(1):2-11

223. Hakkinen A, Ylinen J, Kautiainen H, Airaksinen O, Herno A, Kiviranta I. Does the outcome 2 months after lumbar disc surgery predict the outcome 12 months later? Disabil Rehabil 2003 Sep 2;25(17):968-72

224. Holt RT, Majd ME, Vadhva M, Castro FP. The efficacy of anterior spine exposure by an orthopedic surgeon. J Spinal Disord Tech 2003 Oct;16(5):477-86

225. Ivanic GM, Pink TP, Achatz W, Ward JC, Homann NC, May M. Direct stabilization of lumbar spondylolysis with a hook screw: mean 11-year follow-up period for 113 patients. Spine 2003 Feb 1;28(3):255-9

226. Klara PM, Freidank SA, Rezaiamiri S. Comparison of lumbar interbody fusion techniques using ray threaded fusion cages and pedicle screw fixation systems. Neurosurg Q 2003;13(1):20-29

227. North American Spine Society Board of Directors. Spine Patient Outcome Research Trial (SPORT): multi-center randomized clinical trial of surgical and non-surgical approaches to the treatment of low back pain. Spine J 2003 Nov-Dec;3(6):417-9

228. Rainville J, Sobel J, Hartigan C. Does failed spine surgery affect the outcomes from rehabilitation of chronic low back pain? Eura Medicophys 2003;39(4):171-9

229. Sasso RC, Kenneth Burkus J, LeHuec JC. Retrograde ejaculation after anterior lumbar interbody fusion: transperitoneal versus retroperitoneal exposure. Spine 2003 May 15;28(10):1023-6

230. Shah RR, Mohammed S, Saifuddin A, Taylor BA. Radiologic evaluation of adjacent superior segment facet joint violation following transpedicular instrumentation of the lumbar spine. Spine 2003 Feb 1;28(3):272-5

231. Wang MY, Green BA, Shah S, Vanni S, Levi AD. Complications associated with lumbar stenosis surgery in patients older than 75 years of age. Neurosurg Focus 2003 Feb 15;14(2):e7

232. Arai Y, Takahashi M, Kurosawa H. Comparative study of iliac bone graft and carbon cage with local bone graft in posterior lumbar interbody fusion. J Orthop Surg 2002;10(1):1-7

233. Balderston RA, Brummett RS. The use of the bacfix instrumentation system for fixation of the lumbar spine. Semin Spine Surg 2002;14(4):278-289

234. Burkus JK, Transfeldt EE, Kitchel SH, Watkins RG, Balderston RA. Clinical and radiographic outcomes of anterior lumbar interbody fusion using recombinant human bone morphogenetic protein-2. Spine 2002 Nov 1;27(21):2396-408

235. DeBerard MS, Colledge AL, Masters KS, Schleusener RL, Schlegel JD. Outcomes of posterolateral versus BAK titanium cage interbody lumbar fusion in injured workers: a retrospective cohort study. J South Orthop Assoc 2002 Fall;11(3):157-66

236. Deberard MS, Masters KS, Colledge AL, Schleusener RL, Schlegel JD. Pre-surgical psychological screenings for lumbar fusion: a look at real world practice. Psychol Health Med 2002;7(4):411-424

237. du Toit G, Vlok G. The radiological outcome of lumbar spinal fusion using a South African-developed dynamic spinal fixation system. S Afr Med J 2002 Oct;92(10):821-5

238. Gehrchen PM, Dahl B, Katonis P, Blyme P, Tondevold E, Kiaer T. No difference in clinical outcome after posterolateral lumbar fusion between patients with isthmic spondylolisthesis and those with degenerative disc disease using pedicle screw instrumentation: a comparative study of 112 patients with 4 years follow. Eur Spine J 2002 Oct;11(5):423-7

239. Hirunyachote P, Adulkasem W. Posterolateral fusion with autogenous laminospinous process bone graft. J Med Assoc Thai 2002 Oct;85(10):1105-12

240. Korsgaard M, Christensen FB, Thomsen K, Hansen ES, Bunger C. The influence of lumbar lordosis on spinal fusion and functional outcome after posterolateral spinal fusion with and without pedicle screw instrumentation. J Spinal Disord Tech 2002 Jun;15(3):187-92

241. Linovitz RJ, Pathria M, Bernhardt M, Green D, Law MD, McGuire RA, Montesano PX, Rechtine G, Salib RM, Ryaby JT, Faden JS, Ponder R, Muenz LR, Magee FP, Garfin SA. Combined magnetic fields accelerate and increase spine fusion: a double-blind, randomized, placebo controlled study. Spine 2002 Jul 1;27(13):1383-9; discussion 1389

242. Mayer HM, Wiechert K. Microsurgical anterior approaches to the lumbar spine for interbody fusion and total disc replacement. Neurosurgery 2002 Nov;51(5 Suppl):159-65

243. Pradhan BB, Nassar JA, Delamarter RB, Wang JC. Single-level lumbar spine fusion: a comparison of anterior and posterior approaches. J Spinal Disord Tech 2002 Oct;15(5):355-61

244. Vaccaro AR. Spinal applications of bioabsorbable implants. Orthopedics 2002 Oct 1;25(10):s1115-s1120

245. Andersen T, Christensen FB, Laursen M, Hoy K, Hansen ES, Bunger C. Smoking as a predictor of negative outcome in lumbar spinal fusion. Spine 2001 Dec 1;26(23):2623-8

246. Atlas SJ, Keller RB, Chang Y, Deyo RA, Singer DE. Surgical and nonsurgical management of sciatica secondary to a lumbar disc herniation: five-year outcomes from the Maine Lumbar Spine Study. Spine 2001 May 15;26(10):1179-87

247. Bednar DA. Failure of external spinal skeletal fixation to improve predictability of lumbar arthrodesis. J Bone Joint Surg Am 2001 Nov;83-A(11):1656-9

248. Block AR, Ohnmeiss DD, Guyer RD, Rashbaum RF, Hochschuler SH. The use of presurgical psychological screening to predict the outcome of spine surgery. Spine J 2001 Jul-Aug;1(4):274-82

249. Cook SD, Barbera J, Rubi M, Salkeld SL, Whitecloud TS 3rd. Lumbosacral fixation using expandable pedicle screws. an alternative in reoperation and osteoporosis. Spine J 2001 Mar-Apr;1(2):109-14

250. DeBerard MS, Masters KS, Colledge AL, Schleusener RL, Schlegel JD. Outcomes of posterolateral lumbar fusion in Utah patients receiving workers' compensation: a retrospective cohort study. Spine 2001 Apr 1;26(7):738-46; discussion 747

251. Finkenberg J, Banta C, Cross GL 3rd, Dawson E, Gutzman D, Highland T, Kucharzyk D, Lenderman L, Murphy J, Neely W, Rogozinski A, Rogozinski C. Evaluation and analysis of patient outcomes with an intrasegmental fixation system in lumbar spinal fusion. Spine J 2001 Mar-Apr;1(2):102-8

252. Goldstein JA, Macenski MJ, Griffith SL, McAfee PC. Lumbar sagittal alignment after fusion with a threaded interbody cage. Spine 2001 May 15;26(10):1137-42

253. Izumi Y, Kumano K. Analysis of sagittal lumbar alignment before and after posterior instrumentation: Risk factor for adjacent unfused segment. Eur J Orthop Surg Traumatol 2001;11(1):9-13

254. Janssen ME, Lam C, Beckham R. Outcomes of allogenic cages in anterior and posterior lumbar interbody fusion. Eur Spine J 2001 Oct;10(Suppl 2):S158-68

255. Jolles BM, Porchet F, Theumann N. Surgical treatment of lumbar spinal stenosis. J Bone Joint Surg Br 2001;83(7):949-953

256. Parisini P, Di Silvestre M, Greggi T, Paderni S, Lippo C, Macchiagodena M. Circumferential fusion by posterior approach in lumbosacral instability with or without pedicle screw fixation: a comparison of methods. Chir Organi Mov 2001 Apr-Jun;86(2):127-42

257. Robertson PA, Wray AC. Natural history of posterior iliac crest bone graft donation for spinal surgery: a prospective analysis of morbidity. Spine 2001 Jul 1;26(13):1473-6

258. Wong CB, Chen WJ, Chen LH, Niu CC, Lai PL. Clinical outcomes of revision lumbar spinal surgery: 124 patients with a minimum of two years of follow-up. Chang Gung Med J 2002 Mar;25(3):175-81

259. Barbera J. The Omega 21 spinal fixator. Analysis of the results in pedicle instrumented lumbar fusion after a two year postoperative follow up. Neurocirugia 2000;11(6):409-18

260. Barrick WT, Schofferman JA, Reynolds JB, Goldthwaite ND, McKeehen M, Keaney D, White AH. Anterior lumbar fusion improves discogenic pain at levels of prior posterolateral fusion. Spine 2000 Apr 1;25(7):853-7

261. Brara HS, Fessler RG. The role of anterior lumbar interbody allograft bone dowel fusion as an adjunct to posterior segmental lumbar fixation. Clin Neurosurg 2000;47:528-33

262. Cornefjord M, Byrod G, Brisby H, Rydevik B. A long-term (4- to 12-year) follow-up study of surgical treatment of lumbar spinal stenosis. Eur Spine J 2000 Dec;9(6):563-70

263. Hodges SD, Humphreys SC, Eck JC, Murphy RB. Intraoperative loosening of Bagby and Kuslich cages during anterior lumbar interbody fusion. J Spinal Disord 2000;13(6):535-537

264. Keskimaki I, Seitsalo S, Osterman H, Rissanen P. Reoperations after lumbar disc surgery. Spine 2000 Jun 15;25(12):1500-1508

265. Laine T, Lund T, Ylikoski M, Lohikoski J, Schlenzka D. Accuracy of pedicle screw insertion with and without computer assistance: A randomised controlled clinical study in 100 consecutive patients. Eur Spine J 2000;9(3):235-41

266. Matge G, Leclercq TA. Rationale for interbody fusion with threaded titanium cages at cervical and lumbar levels. Results on 357 cases. Acta Neurochir 2000;142(4):425-33

267. Moller H, Hedlund R. Instrumented and noninstrumented posterolateral fusion in adult spondylolisthesis--a prospective randomized study: part 2. Spine 2000 Jul 1;25(13):1716-21

268. Mulholland RC. Cages: outcome and complications. Eur Spine J 2000 Feb;9 Suppl 1:S110-3

269. Rodts GE Jr, McLaughlin MR, Zhang J, Subach BR, Haid RW Jr. Laparoscopic anterior lumbar interbody fusion. Clin Neurosurg 2000;:47:541-56

270. Taylor VM, Deyo RA, Ciol M, Farrar EL, Lawrence MS, Shonnard NH, Leek KM, McNeney B, Goldberg HI. Patient-oriented outcomes from low back surgery: a community-based study. Spine 2000 Oct 1;25(19):2445-52

271. Weinstein MA, McCabe JP, Cammisa FP Jr. Postoperative spinal wound infection: a review of 2,391 consecutive index procedures. J Spinal Disord 2000 Oct;13(5):422-6

272. Cavagna R, Daculsi G, Bouler JM. Macroporous calcium phosphate ceramic: a prospective study of 106 cases in lumbar spinal fusion. J Long Term Eff Med Implants 1999;9(4):403-12

273. Etebar S, Cahill DW. Risk factors for adjacent-segment failure following lumbar fixation with rigid instrumentation for degenerative instability. J Neurosurg 1999 Apr;90(4 Suppl):163-9

274. Goodwin CB, Brighton CT, Guyer RD, Johnson JR, Light KI, Yuan HA. A double-blind study of capacitively coupled electrical stimulation as an adjunct to lumbar spinal fusions. Spine 1999 Jul 1;24(13):1349-56; discussion 1357

275. Hanakita J, Suwa H, Mizuno M. Surgical treatment of lumbar canal stenosis in the elderly. Neurol Med Chir (Tokyo) 1999 Jul;39(7):519-22; discussion 522-3

276. Kim NH. Anterior interbody fusion in the treatment of the lumbar herniated nucleus pulposus. Yonsei Med J 1999 Jun;40(3):256-64

277. Kucharzyk DW, Kahanovitz N. A controlled prospective outcome study of implantable electrical stimulation with spinal instrumentation in a high-risk spinal fusion population. Spine 1999 Mar 1;24(5):465-9

278. Lonstein JE, Denis F, Perra JH, Pinto MR, Smith MD, Winter RB. Complications associated with pedicle screws. J Bone Joint Surg Am 1999 Nov;81(11):1519-28

279. Okuyama K, Abe E, Suzuki T, Tamura Y, Chiba M, Sato K. Posterior lumbar interbody fusion: a retrospective study of complications after facet joint excision and pedicle screw fixation in 148 cases. Acta Orthop Scand 1999 Aug;70(4):329-34

280. Rompe JD, Eysel P, Zollner J, Nafe B, Heine J. Degenerative lumbar spinal stenosis. Long-term results after undercutting decompression compared with decompressive laminectomy alone or with instrumented fusion. Neurosurg Rev 1999 Oct;22(2-3):102-6

281. Snider RK, Krumwiede NK, Snider LJ, Jurist JM, Lew RA, Katz JN. Factors affecting lumbar spinal fusion. J Spinal Disord 1999 Apr;12(2):107-14

282. Stambough JL. Lumbosacral instrumented fusion: Analysis of 124 consecutive cases. J Spinal Disord 1999 Feb;12(1):1-9

283. Brown CA, Lenke LG, Bridwell KH, Geideman WM, Hasan SA, Blanke K. Complications of pediatric thoracolumbar and lumbar pedicle screws. Spine 1998 Jul 18;23(14):1566-71

284. Buttermann GR, Garvey TA, Hunt AF, Transfeldt EE, Bradford DS, Boachie-Adjei O, Ogilvie JW. Lumbar fusion results related to diagnosis. Spine 1998 Jan 1;23(1):116-27

285. Christensen FB, Thomsen K, Eiskjaer SP, Gelinick J, Bunger CE. Functional outcome after posterolateral spinal fusion using pedicle screws: comparison between primary and salvage procedure. Eur Spine J 1998;7(4):321-7

286. Gertzbein SD, Hollopeter M, Hall SD. Analysis of circumferential lumbar fusion outcome in the treatment of degenerative disc disease of the lumbar spine. J Spinal Disord 1998 Dec;11(6):472-8

287. Glassman SD, Minkow RE, Dimar JR, Puno RM, Raque GH, Johnson JR. Effect of prior lumbar discectomy on outcome of lumbar fusion: a prospective analysis using the SF-36 measure. J Spinal Disord 1998 Oct;11(5):383-8

288. Grob D, Humke T. Translaminar screw fixation in the lumbar spine: technique, indications, results. Eur Spine J 1998;7(3):178-86

289. Humke T, Grob D, Dvorak J, Messikommer A. Translaminar screw fixation of the lumbar and lumbosacral spine. A 5-year follow-up. Spine 1998 May 15;23(10):1180-4

290. Ray CD. Spinal interbody fusion with Ray threaded titanium cages. Tech Neurosurg 1998;4(3):235-45

291. Andreshak TG, An HS, Hall J, Stein B. Lumbar spine surgery in the obese patient. J Spinal Disord 1997 Oct;10(5):376-9

292. Faraj AA, Webb JK. Early complications of spinal pedicle screw. Eur Spine J 1997;6(5):324-6

293. Frazier DD, Lipson SJ, Fossel AH, Katz JN. Associations between spinal deformity and outcomes after decompression for spinal stenosis. Spine 1997 Sep 1;22(17):2025-9

294. Katz JN, Lipson SJ, Lew RA, Grobler LJ, Weinstein JN, Brick GW, Fossel AH, Liang MH. Lumbar laminectomy alone or with instrumented or noninstrumented arthrodesis in degenerative lumbar spinal stenosis. Patient selection, costs, and surgical outcomes. Spine 1997 May 15;22(10):1123-31

295. Pihlajamaki H, Myllynen P, Bostman O. Complications of transpedicular lumbosacral fixation for non-traumatic disorders. J Bone Joint Surg Br 1997 Mar;79(2):183-9

296. Shapiro SA, Snyder W. Spinal instrumentation with a low complication rate. Surg Neurol 1997 Dec;48(6):566-74

297. Schwarzenbach O, Berlemann U, Jost B, Visarius H, Arm E, Langlotz F, Nolte LP, Ozdoba C. Accuracy of computer-assisted pedicle screw placement. An in vivo computed tomography analysis. Spine 1997 Feb 15;22(4):452-8

298. Bailey SI, Bartolozzi P, Bertagnoli R, Boriani S, Van Beurden AFA, Cross AT, Friedl HP, Gurr KR, Halm H, Kruls HJA, Metz- Stavenhagen P, Schulze K-J, Esses SI. The BWM spinal fixator system: A preliminary report of a 2-year prospective, international multicenter study in a range of indications requiring surgical intervention for bone grafting and pedicle screw fixation. Spine 1996;21(17):2006-15

299. Calderone RR, Garland DE, Capen DA, Oster H. Cost of medical care for postoperative spinal infections. Orthop Clin North Am 1996 Jan;27(1):171-82

300. Calderone RR, Thomas JC Jr, Haye W, Abeles D. Outcome assessment in spinal infections. Orthop Clin North Am 1996 Jan;27(1):201-5

301. Fritsch EW, Heisel J, Rupp S. The failed back surgery syndrome: reasons, intraoperative findings, and long-term results: a report of 182 operative treatments. Spine 1996 Mar 1;21(5):626-33

302. Gertzbein SD, Betz R, Clements D, Errico T, Hammerberg K, Robbins S, Shepherd E, Weber A, Kerina M, Albin J, Wolk D, Ensor K. Semirigid instrumentation in the management of lumbar spinal conditions combined with circumferential fusion. A multicenter study. Spine 1996 Aug 15;21(16):1918-25; discussion 1925-

303. Glassman SD, Dimar JR, Puno RM, Johnson JR. Salvage of instrumental lumbar fusions complicated by surgical wound infection. Spine 1996 Sep 15;21(18):2163-9

304. Hadjipavlou A, Enker P, Dupuis P, Katzman S, Silver J. The causes of failure of lumbar transpedicular spinal instrumentation and fusion: a prospective study. Int Orthop 1996;20(1):35-42

305. Franklin GM, Haug J, Heyer NJ, McKeefrey SP, Picciano JF, Hardy RW Jr. Outcome of lumbar fusion in Washington State Workers' compensation. Neurosurg Q 1996;6(1):73-5

306. Junge A, Frohlich M, Ahrens S, Hasenbring M, Sandler AJ, Grob D, Dvorak J, Main CJ. Predictors of bad and good outcome of lumbar spine surgery: A prospective clinical study with 2 years' follow-up. Spine 1996;21(9):1056-65

307. Nachemson A, Zdeblick TA, O'Brien JP, Boden SD, McLain RF. Lumbar disc disease with discogenic pain: what surgical treatment is most effective? Spine 1996;21(15):1835-1838

308. O'Brien JP. Should backache be treated with spinal fusion? Spinal fusion is the only treatment for discogenic pain. Br Med J 1996;312(7022):38-40

309. Pfeiffer M, Griss P, Haake M, Kienapfel H, Billion M. Standardized evaluation of long-term results after anterior lumbar interbody fusion. Eur Spine J 1996;5(5):299-307

310. Tiusanen H, Hurri H, Seitsalo S, Osterman K, Harju R. Functional and clinical results after anterior interbody lumbar fusion. Eur Spine J 1996;5(5):288-92

311. Wood KB, Geissele AE, Ogilvie JW. Pelvic fractures after long lumbosacral spine fusions. Spine 1996;21(11):1357-1362

312. Bosacco SJ, Berman AT, Bosacco DN, Levenberg RJ. Results of lumbar disk surgery in a city compensation population. Orthopedics 1995 Apr;18(4):351-5

313. Penta M, Sandhu A, Fraser RD. Magnetic resonance imaging assessment of disc degeneration 10 years after anterior lumbar interbody fusion. Spine 1995 Mar 15;20(6):743-7

314. Schwab FJ, Nazarian DG, Mahmud F, Michelsen CB. Effects of spinal instrumentation on fusion of the lumbosacral spine. Spine 1995 Sep 15;20(18):2023-8

315. Beguiristain JL, Martinez Peric R, Barrios RH, Villas C. Lumbosacral arthrodesis with louis technique. Review of 186 cases. Eur Spine J 1994;3(3):169-71

316. Ciol MA, Deyo RA, Kreuter W, Bigos SJ. Characteristics in Medicare beneficiaries associated with reoperation after lumbar spine surgery. Spine 1994 Jun 15;19(12):1329-34

317. Laus M, Tigani D, Pignatti G, Alfonso C, Malaguti C, Monti C, Giunti A. Posterolateral spinal fusion: a study of 123 cases with a long-term follow-up. Chir Organi Mov 1994 Jan-Mar;79(1):69-79

318. Little DG, MacDonald D, Katz JN. The use of the percentage change in Oswestry disability index score as an outcome measure in lumbar spinal surgery. Spine 1994;19(19):2139-43

319. Oldridge NB, Yuan Z, Stoll JE, Rimm AR. Lumbar spine surgery and mortality among Medicare beneficiaries, 1986. Am J Public Health 1994 Aug;84(8):1292-8

320. Ransom N, La Rocca SH, Thalgott J. The case for pedicle fixation of the lumbar spine. Spine 1994 Dec 1;19(23):2702-6

321. Yuan HA, Garfin SR, Dickman CA, Mardjetko SM. A Historical Cohort Study of Pedicle Screw Fixation in Thoracic, Lumbar, and Sacral Spinal Fusions. Spine 1994 Oct 15;19(20 Suppl):2279S-2296S

322. Bernard TN Jr. Repeat lumbar spine surgery. Factors influencing outcome. Spine 1993 Nov;18(15):2196-200

323. Blumenthal S, Gill K. Complications of the Wiltse Pedicle Screw Fixation System. Spine 1993;18(13):186771

324. Deyo RA, Ciol MA, Cherkin DC, Loeser JD, Bigos SJ. Lumbar spinal fusion. A cohort study of complications, reoperations, and resource use in the Medicare population. Spine 1993 Sep 1;18(11):1463-70

325. Esses SI, Sachs BL, Dreyzin V, Zdeblick T. Complications associated with the technique of pedicle screw fixation: A selected survey of ABS members. Spine 1993;18(15):2231-9

326. Nachemson AL. Evaluation of results in lumbar spine surgery. Acta Orthop Scand Suppl 1993;251:130-3

327. Roy-Camille R, Benazet JP, Desauge JP, Kuntz F. Lumbosacral fusion with pedicular screw plating instrumentation. A 10-year follow-up. Acta Orthop Scand Suppl 1993;251(Suppl):100-4

328. van Akkerveeken PF. Anterior lumbar interbody fusion. Acta Orthop Scand Suppl 1993;251:105-7

329. Willner S. Lumbar spine fusion--conclusions. Acta Orthop Scand Suppl 1993;251:123-4

330. Dhar S, Porter RW. Failed lumbar spinal surgery. Int Orthop 1992;16(2):152-6

331. Dickman CA, Fessler RG, MacMillan M, Haid RW. Transpedicular screw-rod fixation of the lumbar spine: operative technique and outcome in 104 cases. J Neurosurg 1992 Dec;77(6):860-70

332. Grubb SA, Lipscomb HJ. Results of lumbosacral fusion for degenerative disc disease with and without instrumentation. Two- to five-year follow-up. Spine 1992 Mar;17(3):349-55

333. Zucherman J, Hsu K, Picetti G 3rd, White A, Wynne G, Taylor L. Clinical efficacy of spinal instrumentation in lumbar degenerative disc disease. Spine 1992 Jul;17(7):834-7

334. Frennered AK, Danielson BI, Nachemson AL, Nordwall AB. Midterm follow-up of young patients fused in situ for spondylolisthesis. Spine 1991 Apr;16(4):409-16

335. North RB, Campbell JN, James CS, Conover-Walker MK, Wang H, Piantadosi S, Rybock JD, Long DM. Failed back surgery syndrome: 5-year follow-up in 102 patients undergoing repeated operation. Neurosurgery 1991 May;28(5):685-90; 690-1

336. Antti-Poika I, Soini J, Tallroth K, Yrjonen T, Konttinen YT. Clinical relevance of discography combined with CT scanning. A study of 100 patients. J Bone Joint Surg Br 1990 May;72(3):480-5

337. Brodsky AE, Binder WF. Lumbar discography. Its value in diagnosis and treatment of lumbar disc lesions. Spine 1979 Mar-Apr;4(2):110-20

338. Buirski G, Watt I. Dynamic CT discography: an evaluation of a new technique. Australas Radiol 1988 May;32(2):197-202

339. Carragee EJ, Chen Y, Tanner CM, Hayward C, Rossi M, Hagle C. Can discography cause long-term back symptoms in previously asymptomatic subjects. Spine 2000 Jul 15;25(14):1803-8

340. Carragee EJ, Barcohana B, Alamin T, van den Haak E. Prospective controlled study of the development of lower back pain in previously asymptomatic subjects undergoing experimental discography. Spine 2004 May 15;29(10):1112-7

341. Collins CD, Stack JP, O'Connell DJ, Walsh M, McManus FP, Redmond OM, Ennis JT. The role of discography in lumbar disc disease: a comparative study of magnetic resonance imaging and discography. Clin Radiol 1990 Oct;42(4):252-7

342. Collis JS Jr, Gardner WJ. Lumbar discography. An analysis of one thousand cases. J Neurosurg 1962 Jun;19:452-61

343. Derby R, Howard MW, Grant JM, Lettice JJ, Van Peteghem PK, Ryan DP. The ability of pressure-controlled discography to predict surgical and nonsurgical outcomes. Spine 1999 Feb 15;24(4):364-71; discussion 371-2

344. Esses SI, Botsford DJ, Kostuik JP. The role of external spinal skeletal fixation in the assessment of low- back disorders. Spine 1989 Jun;14(6):594-601

345. Feinberg SB. The place of diskography in radiology as based on 2,320 cases. Am J Roentgenol Radium Ther Nucl Med 1964 Dec;92:1275-81

346. Friedman J, Goldner MZ. Discography in evaluation of lumbar disk lesions. Radiology 1955 Nov;65(5):653-62

347. Gresham JL, Miller R. Evaluation of the lumbar spine by diskography and its use in selection of proper treatment of the herniated disk syndrome. Clin Orthop Relat Res 1969 Nov-Dec;67:29-41

348. Hartman JT, Kendrick JI, Lorman P. Discography as an aid in evaluation for lumbar and lumbosacral fusion. Clin Orthop Relat Res 1971 Nov-Dec;81:77-81

349. Jacobs GB, Pillone PR, Mazzola R. Lumbar discography in the diagnosis of herniated disks. Int Surg 1975 Jan;60(1):6-9

350. Knox BD, Chapman TM. Anterior lumbar interbody fusion for discogram concordant pain. J Spinal Disord 1993 Jun;6(3):242-4

351. Kostuik JP. Decision making in adult scoliosis. Spine 1979 Nov-Dec;4(6):521-5

352. Lee CK, Vessa P, Lee JK. Chronic disabling low back pain syndrome caused by internal disc derangements. The results of disc excision and posterior lumbar interbody fusion. Spine 1995 Feb 1;20(3):356-61

353. Min K, Leu HJ, Perrenoud A. Discography with manometry and discographic CT: their value in patient selection for percutaneous lumbar nucleotomy. Bull Hosp Jt Dis 1996;54(3):153-7

354. Moneta GB, Videman T, Kaivanto K, Aprill C, Spivey M, Vanharanta H, Sachs BL, Guyer RD, Hochschuler SH, Raschbaum RF, et al. Reported pain during lumbar discography as a function of anular ruptures and disc degeneration. A re-analysis of 833 discograms. Spine 1994 Sep 1;19(17):1968-74

355. Newman MH, Grinstead GL. Anterior lumbar interbody fusion for internal disc disruption. Spine 1992 Jul;17(7):831-3

356. Parker LM, Murrell SE, Boden SD, Horton WC. The outcome of posterolateral fusion in highly selected patients with discogenic low back pain. Spine 1996 Aug 15;21(16):1909-16; disc 1916-7

357. Patrick BS. Lumbar discography: a five year study. Surg Neurol 1973 Sep;1(5):267-73

358. Roberts A, Loupe J, Goldsmith J, Comeaux L, Wickstrom J. Lumbar diskography using a posterolateral approach with a guide. South Med J 1972 Mar;65(3):358-60

359. Sachs BL, Spivey MA, Vanharanta H, Guyer RD, Rashbaum RF, Hochschuler SH, Scala AD. Techniques for lumbar discography and computed tomography/discography in clinical practice. Orthop Rev 1990 Sep;19(9):775-8

360. Schechter NA, France MP, Lee CK. Painful internal disc derangements of the lumbosacral spine: discographic diagnosis and treatment by posterior lumbar interbody fusion. Orthopedics 1991 Apr;14(4):447-51

361. Simmons EH, Segil CM. An evaluation of discography in the localization of symptomatic levels in discogenic disease of the spine. Clin Orthop Relat Res 1975 May;(108):57-69

362. Simmons JW, Aprill CN, Dwyer AP, Brodsky AE. A reassessment of Holt's data on: "The question of lumbar discography". Clin Orthop Relat Res 1988 Dec;(237):120-4

363. Smith BM, Hurwitz EL, Solsberg D, Rubinstein D, Corenman DS, Dwyer AP, Kleiner J. Interobserver reliability of detecting lumbar intervertebral disc high-intensity zone on magnetic resonance imaging and association of high-intensity zone with pain and anular disruption. Spine 1998 Oct 1;23(19):2074-80

364. Taveras J. Is discography a useful diagnostic procedure? J Can Assoc Radiol 1967 Jun;18(2):294-5

365. Urasaki T, Muro T, Ito S, Hattori Y, Ozaki S. Consistency of lumbar discograms of the same disc obtained twice at a 2-week interval: influence of needle tip position. J Orthop Sci 1998;3(5):243-51

366. Vanharanta H, Sachs BL, Spivey MA, Guyer RD, Hochschuler SH, Rashbaum RF, Johnson RG, Ohnmeiss D, Mooney V. The relationship of pain provocation to lumbar disc deterioration as seen by CT/discography. Spine 1987 Apr;12(3):295-8

367. Vanharanta H, Guyer RD, Ohnmeiss DD, Stith WJ, Sachs BL, Aprill C, Spivey M, Rashbaum RF, Hochschuler SH, Videman T, et al. Disc deterioration in low-back syndromes. A prospective, multi-center CT/discography study. Spine 1988 Dec;13(12):1349-51

368. Vanharanta H, Sachs BL, Ohnmeiss DD, Aprill C, Spivey M, Guyer RD, Rashbaum RF, Hochschuler SH, Terry A, Selby D, et al. Pain provocation and disc deterioration by age. A CT/discography study in a low-back pain population. Spine 1989 Apr;14(4):420-3

369. Weatherley CR, Prickett CF, O'Brien JP. Discogenic pain persisting despite solid posterior fusion. J Bone Joint Surg Br 1986 Jan;68(1):142-3

370. Wetzel FT, LaRocca SH, Lowery GL, Aprill CN. The treatment of lumbar spinal pain syndromes diagnosed by discography. Lumbar arthrodesis. Spine 1994 Apr 1;19(7):792-800

371. Whitecloud TS 3rd, Seago RA. Cervical discogenic syndrome. Results of operative intervention in patients with positive discography. Spine 1987 May;12(4):313-6

372. Wilson DH, MacCarty WC. Discography: its role in the diagnosis of lumbar disc protrusion. J Neurosurg 1969 Nov;31(5):520-3

373. West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, Lux L. Systems to rate the strength of scientific evidence. (Prepared by Research Triangle Institute - University of North Carolina Evidence-based Practice Center under Contract no. 290-97-0011). AHRQ Publication no. 02-E016. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2002 Apr. 199 p. (Evidence report/technology assessment; no. 47). Also available: http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hstat1.chapter.70996

374. Hagg O, Fritzell P, Nordwall A, Swedish Lumbar Spine Study Group. Sexual function in men and women after anterior surgery for chronic low back pain. Eur Spine J 2006 May;15(5):677-82

375. Sasso RC, Kitchel SH, Dawson EG. A prospective, randomized controlled clinical trial of anterior lumbar interbody fusion using a titanium cylindrical threaded fusion device. Spine 2004 Jan 15;29(2):113-22; discussion 121-2

376. Greenough CG, Taylor LJ, Fraser RD. Anterior lumbar fusion: results, assessment techniques and prognostic factors. Eur Spine J 1994;3(4):225-30

377. Office of Device Evaluation, Center for Devices and Radiological Health. BAK Interbody Fusion System with instrumentation premarket approval (PMA) application. Summary of safety and effectiveness. P950002. Rockville (MD): Food and Drug Administration (FDA); 1996 Sep 20

378. Adams MA, Dolan P, Hutton WC. The stages of disc degeneration as revealed by discograms. J Bone Joint Surg Br 1986 Jan;68(1):36-41

Appendix A. Literature Search Methods

Electronic Database Searches

The following databases have been searched for relevant information:

Database	Date limits	Platform/provider
CINAHL (Cumulative Index to Nursing and Allied Health Literature)	1982 through August 13, 2007	OVID
The Cochrane Central Register of Controlled Trials (CENTRAL)	through 2007, Issue 3	http://www.thecochranelibrary.com
The Cochrane Database of Methodology Reviews (Methodology Reviews)	through 2007, Issue 3	http://www.thecochranelibrary.com
The Cochrane Database of Systematic Reviews (Cochrane Reviews)	through 2007, Issue 3	http://www.thecochranelibrary.com
Database of Abstracts of Reviews of Effects (DARE)	through 2007, Issue 3	http://www.thecochranelibrary.com
ECRI Institute Library Catalog	August 23, 2007	ECRI Institute
EMBASE (Excerpta Medica)	1980 through February 2, 2007	OVID
Health Technology Assessment Database (HTA)	through 2007, Issue 3	http://www.thecochranelibrary.com
Healthcare Standards	1975 through August 2007	ECRI Institute
International Health Technology Assessment (IHTA)	Through August 2007	ECRI Institute
MEDLINE	1950 through February 2, 2007	OVID
PsycINFO	1967 through August 16, 2007	OVID
PubMed (PreMEDLINE, Publisher)	Searched July 31, 2007	http://pubmed.gov
U.K. National Health Service Economic Evaluation Database (NHS EED)	through 2007, Issue 1	http://www.thecochranelibrary.com
U.S. National Guideline Clearinghouse™ (NGC™)	through August 2007	www.ngc.gov

Hand Searches of Journal and Nonjournal Literature

Journals and supplements maintained in ECRI's collections were routinely reviewed. Nonjournal publications and conference proceedings from professional organizations, private agencies, and government agencies were also screened. Other mechanisms used to retrieve additional relevant information included review of bibliographies/reference lists from peer-reviewed and gray literature. (Gray literature consists of reports, studies, articles, and monographs produced by federal and local government agencies, private organizations, educational facilities, consulting firms, and corporations. These documents do not appear in the peer-reviewed journal literature.)

Detailed Search Strategies

The search strategies employed combinations of freetext keywords as well as controlled vocabulary terms including (but not limited to) the following concepts. The strategy below is presented in OVID syntax; the search was simultaneously conducted across EMBASE, MEDLINE, and PsycINFO. A parallel strategy was used to search the databases comprising the Cochrane Library.

Medical Subject Headings (MeSH), Emtree, PsycINFO and Keywords

Conventions:

OVID

$	=	truncation character (wildcard)
exp	=	"explodes" controlled vocabulary term (e.g., expands search to all more specific related terms in the vocabulary's hierarchy)
.de.	=	limit controlled vocabulary heading
.fs.	=	floating subheading
.hw.	=	limit to heading word
.md.	=	type of methodology (PsycINFO)
.mp.	=	combined search fields (default if no fields are specified)
.pt.	=	publication Type
.ti.	=	limit to title
.tw.	=	limit to title and abstract fields

PubMed

[mh]	=	MeSH heading
[majr]	=	MeSH heading designated as major topic
[pt]	=	Publication Type
[sb]	=	Subset of PubMed database (PreMedline, Systematic, OldMedline)
[sh]	=	MeSH subheading (qualifiers used in conjunction with MeSH headings)
[tiab]	=	keyword in title or abstract
[tw]	=	Text word

Spinal Fusion - Topic-specific Search Terms

Bone Morphogenic Protein

Controlled Vocabulary	Text Words
exp bone morphogenic proteins/	bone morphogenic protein2
exp bone morphogenic protein/	"INFUSE"
exp recombinant proteins	morpohogen$
	"Ne-OSTEO"
	"OP-1"
	rhBMP

Lumbar

Controlled Vocabulary	Text Words
exp low back/	low back pain
exp lumbar vertebrae/	lumbar
exp lumbosacral region/	lumbar spine
	lumbosacral

Spinal Fusion

Controlled Vocabulary	Text Words
exp spinal fusion/	ALIF
exp spine fusion/	arthrodesis
	cage$
	fusion$
	instrumentation
	interbody
	lumbar
	PLIF
	pedlicle screw
	pedicl$ adj screw$
	spinal
	tapered

CINAHL/EMBASE/MEDLINE/PsycINFO
English language, human

Set Number	Concept	Search statement
1	Spinal fusion	Exp spinal fusion or exp spine fusion or ((spinal or lumbar or interbody or tapered) adj3 (cage$ or fusion$ or instrumentation or arthrodesis) or pedicl$ adj screw$)) or (ALIF or PLIF)
2	Bone morphogenic protein	Exp bone morphogenic proteins/ or bone morphogenetic protien/ or bone morphogenetic protein 2.de. or exp recombinant proteins/ or (rhBMP adj 2 or hr adj bmp adj 2 or morphogen$ or "InFUSE" or OP-1" or "Ne-Osteo"
3	Combine sets	1 or 2
4	Limit to lumbar	3 and (lumbar or lumbosacral or exp lumbar vertebrae/ or lumbar vertebrae or lumbar spine or lumbosacral spine or exp lumbosacral region/or low back pain/ or low back)
5	Limit by study type	4 and ((Randomized controlled trials or random allocation or double-blind method or single-blind method or placebos or cross-over studies or crossover procedure or double blind procedure or single blind procedure or placebo or latin square design or crossover design or double-blind studies or single-blind studies or triple-blind studies or random assignment or exp controlled study/ or exp clinical trial/ or exp comparative study/ or cohort analysis or follow-up studies.de. or intermethod comparison or parallel design or control group or prospective study or retrospective study or case control study or major clinical study or evaluation studies or follow-up studies).de. or random$.hw. or random$.ti. or placebo$ or ((singl$ or doubl$ or tripl$ or trebl$) and (dummy or blind or sham)) or latin square or ISRTCN)
6	Limit by publication type	5 not ((letter or editorial or news or comment or case reports or note or conference paper).de. or (letter or editorial or news or comment or case reports).pt.)
7	Limit by language	6, English, English language
8	Limit by population	7, human, humans
9	Eliminate overlap	8, remove duplicates
10	Adverse events	9 and ((adverse effects or complications or side effect or contraindication).fs. or (harm$ or iatrogen$ or nosocom$ or hazard$ or safety or nnh) ti.ab. or (morbid$ or mortal$.).fs.mp. or (treatment outcome or patient satisfaction or reoperation).de. or exp *pain/ or exp postoperative complications/)

Discography - Topic-specific Search Terms

Discography

Controlled Vocabulary	Text Words
Discography.de	discography dis?ograph$ dis?ogram$

Injection

Controlled Vocabulary	Text Words
injection,spinal.de.	

Intervertebral Disc

Controlled Vocabulary	Text Words
intervertebral disk.de.	

Lumbar

Controlled Vocabulary	Text Words
exp low back/ exp lumbar vertebrae/ exp lumbosacral region/	low back pain lumbar lumbar spine lumbosacral

Reliability

Controlled Vocabulary	Text Words
	reliab$ or repeatab$ replicat$).

Validity

Controlled Vocabulary	Text Words
observer variation.de	intraobserver intra-observer o interobserver inter-observer interpret$ o kappa observer bias observer variability reader$ reader concordance.tw.

CINAHL/EMBASE/MEDLINE/PsycINFO
English language, human

Set Number	Concept	Search statement
1	Lumbar	lumbar or lumbosacral or exp lumbar vertebrae/ or lumbar vertebra/r lumbar spine or lumbosacral spine or exp lumbosacral region/or low back pain/ or low back
2	Discography	Discography.de. or dis?ography$.mp. or dis?ogram$.mp.
3	Intervertebral disk injection	injections, spinal.de. and intervertebral disk.de
4	Combine sets	1 and 2 and 3
5	Limit by study type	4 and ((Randomized controlled trials or random allocation or double-blind method or single-blind method or placebos or cross-over studies or crossover procedure or double blind procedure or single blind procedure or placebo or latin square design or crossover design or double-blind studies or single-blind studies or triple-blind studies or random assignment or exp controlled study/ or exp clinical trial/ or exp comparative study/ or cohort analysis or follow-up studies.de. or intermethod comparison or parallel design or control group or prospective study or retrospective study or case control study or major clinical study or evaluation studies or follow-up studies).de. or random$.hw. or random$.ti. or placebo$ or ((singl$ or doubl$ or tripl$ or trebl$) and (dummy or blind or sham)) or latin square or ISRTCN)
6	Limit by publication type	5 not ((letter or editorial or news or comment or case reports or note or conference paper).de. or (letter or editorial or news or comment or case reports).pt.)
7	Limit by language	6, English, English language
8	Limit by population	7, human, humans
9	Eliminate overlap	8, remove duplicates
10	Validity	9 and ((intraobserver or intra-observer or interobserver or inter-observer or interpret$ or kappa or observer bias or observer variability or reader$ or reader concordance).tw. or observer variation.de.)
11	Reliability	9 and (reliab$ or repeatab$ or replicat$).mp.

Hand Searches of Journal and Nonjournal Literature

Journals and supplements maintained in ECRI's collections were routinely reviewed. Nonjournal publications and conference proceedings from professional organizations, private agencies, and government agencies were also screened. Other mechanisms used to retrieve additional relevant information included review of bibliographies/reference lists from peer-reviewed and gray literature. (Gray literature consists of reports, studies, articles, and monographs produced by federal and local government agencies, private organizations, educational facilities, consulting firms, and corporations. These documents do not appear in the peer-reviewed journal literature.)

Appendix B. Excluded Studies

Table 6. Excluded Studies for Efficacy and Safety of Spinal Fusion (Key Questions 1-3)

Study	Reason for exclusion
KQ1 and KQ2 (RCTs comparing spine surgery to nonsurgical therapy)	
Peul et al. 2007(137)	Patients did not undergo fusion, patients had an excluded disorder
Weinstein et al. 2007(138)	Patients had spinal stenosis with neurogenic claudication
Moller and Hedlund 2001(139)	>20% patients had an excluded disorder
KQ2	
Bentsen et al. 2007(140)	No information on complications
Glassman et al. 2007(141)	>20% of patients had excluded disorders, <100 patients in study
Mirovsky et al. 2007(142)	Spinal disorders not reported for most patients
Ronnberg et al. 2007(143)	No patients underwent fusion (other disc surgery was performed)
Sasso and Garrido 2007(144)	No information on complications
Sinikallio et al. 2007(145)	All patients had an excluded disorder
Andersen et al. 2006(146)	>20% patients had an excluded disorder
Anderson et al. 2006(147)	No information on complications
Anand et al. 2006(148)	Cannot determine if >20% of patients had an excluded disorder
Best and Sasso 2006(149)	<100 patients
Cakir et al. 2006(150)	<100 patients
Dimar et al. 2006(125)	<80% followup
Epstein 2006(151)	Excluded disorder
Fogel et al. 2006(152)	<100 patients
Glassman et al. 2006(153)	No information on complications or spinal disorders
Hsu et al. 2006(154)	>20% patients had an excluded disorder
Kim et al. 2006(155): Kim et al. 2006(124) (same study)	>20% patients had an excluded disorder
Kim et al. 2006(156)	<100 patients, >20% patients had an excluded disorder
Kim et al. 2006(124)	>20% patients had an excluded disorder
Maghout-Juratli et al. 2006(117)	>20% patients had an excluded disorder
Motosuneya et al. 2006(157)	<100 patients
Neen et al. 2006(158)	Spine disorders not described
Okuda et al. 2006(159)	>20% patients had an excluded disorder
Okuda et al. 2006(160); Okuda et al. 2006(161) (same study)	>20% patients had an excluded disorder
Pappou et al. 2006(162)	>20% patients had an excluded disorder
Raffo and Lauerman 2006(163)	>20% patients had an excluded disorder
Ringel et al. 2006(164)	>20% patients had an excluded disorder
Satoh et al. 2006(165)	Cannot determine if >20% patients had an excluded disorder

Study	Reason for exclusion
Sengupta et al. 2006(166)	Excluded disorder, <100 patients
Soegaard et al. 2006(167)	Compares rehabilitation programs after fusion. Does not address any Key Question.
Suda et al. 2006(168)	>20% patients had an excluded disorder
Swan et al. 2006(169)	>20% patients had an excluded disorder
Videbaek et al. 2006(123)	No information on complications
Villavicencio et al. 2006(170)	Cannot determine if >20% patients had an excluded disorder
Weinstein et al. 2006(171)	Excluded disorder, most patients did not undergo fusion
Yamashita et al. 2006(172)	Excluded disorder, most patients did not undergo fusion
Yi et al. 2006(173)	<100 patients
Aiki et al. 2005(174)	>20% patients had an excluded disorder
Atlas et al. 2005(175)	>20% patients had an excluded disorder, most patients did not undergo fusion
Bednar and Al-Tunaib 2005(176)	Spinal disorders not described
Blumenthal et al. 2005(177); McAfee et al. 2005(178); Geisler et al. 2004(179) (same study)	<100 patients underwent fusion
Bostelmann and Benini 2005(180)	>20% of patients had excluded disorders, <100 patients in study
Burkus et al. 2005(82)	No information on complications
Burton 2005(181)	Case report
Chang et al. 2005(182)	Excluded disorder, most patients did not undergo fusion
Ekman et al. 2005(183)	>20% patients had an excluded disorder
Epstein 2005(184)	Commentary on primary study, no results presented(184)
Hsu et al. 2005(185)	<100 patients, excluded disorder
Inamasu and Gulot(186)	Review article
Jang and Lee 2005(187)	<100 patients
Kilincer et al. 2005(188)	>20% patients had an excluded disorder
Lamberg et al. 2005(189)	>20% patients had an excluded disorder
Lettice et al. 2005(190)	Spinal disorders not described
Lidar et al. 2005(191)	>20% patients had an excluded disorder
Matsudaira et al. 2005(192)	<100 patients
Moffett 2005(193)	Commentary on included study(77)
Potter et al. 2005(194)	>20% patients had an excluded disorder
Rampersaud et al. 2005(195)	<100 patients
Sasso et al. 2005(196)	Spinal disorders not described
Schuler et al. 2005(197)	No information on complications
Shabat et al. 2005(198)	Patients did not undergo fusion (laminectomy without fusion was performed)
Tulli et al. 2005(199)	<100 patients
Wenger et al. 2005(200)	>20% patients had an excluded disorder
Brau et al. 2004(201)	Spinal disorders not described, proportion of patients who underwent fusion not reported
Burkus et al. 2004(202)	No information on complications, cannot determine if >20% patients had an excluded disorder
Cammisa et al. 2004(203)	<80% followup
Christensen et al. 2004(204)	Cannot determine if >20% patients had an excluded disorder

Study	Reason for exclusion
Gaetani et al. 2004(205)	No patients underwent fusion (other disc surgery was performed)
Ghiselli et al. 2004(206)	>20% patients had an excluded disorder
Greiner-Perth et al. 2004(207)	>20% patients had an excluded disorder
Kornblum et al. 2004(208)	Excluded disorder, <100 patients
Korovessis et al. 2004(209)	Excluded disorder, <100 patients
Lai et al. 2004(210)	Excluded disorder
Lai et al. 2004(211)	<100 patients, excluded disorder
Seal et al. 2004(212)	>20% had excluded disorder
Vaccaro et al. 2004(213)	<100 patients
Andersen et al. 2003(214)	Cannot determine if >20% patients had an excluded disorder
Beutler and Peppelman 2003(215)	No information on complications
Carreon et al. 2003(216)	>20% patients had an excluded disorder, <100 patients
Christensen et al. 2003(217)	Compares rehabilitation programs after fusion. Does not address any Key Question
DeBerard et al. 2003(119)	No information on complications
Gillet 2003(218)	Cannot determine if >20% patients had an excluded disorder
Glaser et al. 2003(219)	>20% patients had an excluded disorder
Glassman et al. 2003(220)	<100 patients
Hagg et al. 2003(29)	Relevant data reported in another publication(221)
Hagg et al. 2003(222)	No relevant outcomes
Hakkinen et al. 2003(223)	Cannot determine how many patients (if any) underwent fusion
Holt et al. 2003(224)	>20% patients had an excluded disorder
Ivanic et al. 2003(225)	Patients did not undergo fusion
Klara et al. 2003(226)	<100 patients
North American Spine Society Board of Directors 2003(227)	Position statement, not a clinical study
Rainville et al. 2003(228)	No surgical outcomes
Sasso et al. 2003(229)	Repeats information from an included study(87)
Shah et al. 2003(230)	No information on complications, >20% patients had an excluded disorder
Wang et al. 2003(231)	<100 patients
Arai et al. 2002(232)	<100 patients
Balderston and Brummett 2002(233)	>20% had excluded disorders
Burkus et al. 2002(234)	<100 patients
DeBerard et al. 2002(235)	<80% followup for reported harms (reoperation)
DeBerard et al. 2002(236)	>20% patients had an excluded disorder
Du Toit and Vlok 2002(237)	<80% followup, cannot determine if >20% of patients had an excluded disorder
Gehrchen et al. 2002(238); Gehrchen et al. 2002(238) (same study)	>20% patients had an excluded disorder
Hirunyachote and Adulkasem 2002(239)	<100 patients
Korsgaard et al. 2002(240)	No information on complications
Linovitz et al. 2002(241)	>20% patients had an excluded disorder

Study	Reason for exclusion
Mayer and Weichert 2002(242)	>20% patients had an excluded disorder
Pradhan et al. 2002(243)	>20% patients had an excluded disorder
Vaccaro and Madigan 2002(244)	Not a clinical study, no results presented
Andersen et al. 2001(245)	Cannot determine if >20% patients had an excluded disorder
Atlas et al. 2001(246)	>20% patients had an excluded disorder
Bednar 2001(247)	>20% patients had an excluded disorder
Block et al. 2001(248)	No information on complications, <100 patients underwent fusion
Cook et al. 2001(249)	>20% patients had an excluded disorder
DeBerard et al. 2001(250)	>20% patients had an excluded disorder
Finkenberg et al. 2001(251)	>20% patients had an excluded disorder
Goldstein et al. 2001(252)	Patient population overlaps with patients in included study(252)
Izumi and Kumano 2001(253)	No information on complications, disorders not described
Janssen et al. 2001(254)	Cannot determine if >20% patients had an excluded disorder
Jolles et al. 2001(255)	>20% patients had an excluded disorder
Parisini et al. 2001(256)	<100 patients
Robertson and Wray 2001(257)	>20% patients had an excluded disorder
Stromqvist et al. 2001(121)	>20% patients had an excluded disorder, complications reported by procedure but not by diagnosis
Wong et al. 2001(258)	>20% patients had an excluded disorder
Barbera 2000(259)	>20% patients had an excluded disorder
Barrick et al. 2000(260)	<100 patients
Brara and Fessler 2000(261)	<100 patients
Cornefjord et al. 2000(262)	All patients had an excluded disorder
Hodges et al. 2000(263)	<100 patients
Keskimaki et al. 2000(264)	Spinal disorders not described, most patients did not undergo fusion
Laine et al. 2000(265)	>20% patients had an excluded disorder
Matge and Leclercq 2000(266)	>20% patients had an excluded disorder
Moller and Hedlund 2000(267)	<100 patients, >20% patients had an excluded disorder
Mulholland 2000(268)	Spinal disorders not described, patient characteristics not described
Rodts et al. 2000(269)	<100 patients
Taylor et al. 2000(270)	>20% patients had an excluded disorder
Weinstein et al. 2000(271)	>20% patients had an excluded disorder
Cavagna et al. 1999(272)	>20% patients had an excluded disorder
Etebar and Cahill 1999(273)	Spinal disorders not described
Goodwin et al. 1999(274)	>20% patients had an excluded disorder
Hanakita et al. 1999(275)	Excluded disorder, most patients did not undergo fusion
Kim 1999(276)	>20% patients had an excluded disorder
Kucharzyk 1999(277)	>20% patients had an excluded disorder
Lonstein et al. 1999(278)	>20% patients had an excluded disorder
Okuyama et al. 1999(279)	>20% patients had an excluded disorder
Rompe et al. 1999(280)	Excluded disorder, most patients did not undergo fusion
Snider et al. 1999(281)	>20% patients had an excluded disorder

Study	Reason for exclusion
Stambough 1999(282)	>20% patients had an excluded disorder
Brown et al. 1998(283)	>20% patients had an excluded disorder
Buttermann et al. 1998(284)	>20% patients had an excluded disorder
Christensen et al. 1998(285)	>20% patients had an excluded disorder
Donceel and Du Bois 1998(120)	No information on complications
Gertzbein et al. 1998(286)	<100 patients
Glassman et al. 1998(287)	<100 patients
Grub and Humke 1998(288)	>20% patients had an excluded disorder
Humke et al. 1998(289)	>20% patients had an excluded disorder
Ray 1998(290)	Duplicate publication of findings from an included study(290)
Andreshak et al. 1997(291)	>20% patients had an excluded disorder
Faraj and Webb 1997(292)	<100 patients, >20% patients had an excluded disorder
Frazier et al. 1997(293)	>20% patients had an excluded disorder
Hu et al. 1997(122)	>20% patients had an excluded disorder
Katz et al. 1997(294)	Excluded disorder
Pihlajamaki et al. 1997(295)	>20% patients had an excluded disorder
Shapiro and Snyder 1997(296)	Spinal disorders not described for most patients
Schwarzenbach et al. 1997(297)	<100 patients
Bailey et al. 1996(298)	>20% patients had an excluded disorder
Calderone et al. 1996(299)	Spinal disorders not described
Calderone et al. 1996(300)	<100 patients, spinal disorders not described
Fritsch et al. 1996(301)	Patients did not undergo fusion (other disc surgery was performed)
Gertzbein et al. 1996(302)	<100 patients, cannot determine if >20% patients had an excluded disorder
Glassman et al. 1996(303)	<100 patients
Glassman et al. 1996(303)	>20% patients had an excluded disorder
Hadjipavlou et al. 1996(304)	>20% patients had an excluded disorder
Hardy 1996(305)	Commentary on an excluded study(118)
Junge et al. 1996(306)	>20% patients had an excluded disorder, surgery not adequately described
Nachemson et al. 1996(307)	Not a clinical study, no results presented
O'Brien 1996(308)	Not a clinical study, no results presented
Pfeiffer et al. 1996(309)	>20% patients had an excluded disorder
Tiusanen et al. 1996(310)	>20% patients had an excluded disorder
Wood et al. 1996(311)	<100 patients
Bosacco et al. 1995(312)	<100 patients
Penta et al. 1995(313)	<100 patients, no relevant outcomes
Schwab et al. 1995(314)	>20% patients had an excluded disorder
Beguiristain et al. 1994(315)	>20% patients had an excluded disorder
Ciol et al. 1994(316)	>20% patients had an excluded disorder, not all patients underwent fusion
Franklin et al. 1994(118)	>20% patients had an excluded disorder
Laus et al. 1994(317)	>20% patients had an excluded disorder
Little and MacDonald 1994(318)	>20% patients had an excluded disorder
Oldridge et al. 1994(319)	>20% patients had an excluded disorder

Study	Reason for exclusion
Ransom et al. 1994(320)	>20% patients had an excluded disorder
Yuan et al. 1994(321)	>20% patients had an excluded disorder
Bernard 1993(322)	<100 patients
Blumenthal and Gill 1993(323)	Cannot determine if >20% patients had an excluded disorder
Deyo et al. 1993(324)	>20% patients had an excluded disorder
Esses et al. 1993(325)	>20% patients had an excluded disorder
Nachemson 1993(326)	Not a clinical study, no results presented
Roy-Camille et al. 1993(327)	<100 patients
Van Akkerveeken 1993(328)	<100 patients
Willner 1993(329)	Not a clinical study, no results presented
Dhar and Porter 1992(330)	Most patients did not undergo fusion
Dickman et al. 1992(331)	>20% patients had an excluded disorder
Grubb and Lipscomb 1992(332)	Complications collected differently for each group
Zucherman et al. 1992(333)	>20% patients had an excluded disorder
Frennered et al. 1991(334)	>20% patients had an excluded disorder
North et al. 1991(335)	Most patients did not undergo fusion

Table 7. Excluded Studies for Discography Prior to Lumbar Fusion (Key Questions 4-6)

Study	Reason for Exclusion
Antti-Poika (1990)(336)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Block (1996)(27)	All patients had radiculopathy (an excluded condition)
Brodsky (1979)(337)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Buirski (1988)(338)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Carragee (1999)(23)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Carragee (2000)(339)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Carragee (2000)(20)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Carragee (2000)(19)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Carragee (2000)(17)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Carragee (2001)(28)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Carragee (2002)(18)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Carragee (2006)(25)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Carragee (2006)(26)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Carragee (2007)(340)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Collins 1990 p252 (1990)(341)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Collis (1962)(342)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Derby (1999)(343)	Study did not report the results separately for patients who had undergone fusion but did not have chemically sensitive discs on discography.
Derby (2005)(16)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Esses (1989)(344)	Treatment was external fixation, not fusion
Feinberg (1964)(345)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Friedman (1955)(346)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.

Study	Reason for Exclusion
Gresham (1969)(347)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Hartman (1971)(348)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Holt (1968)(15)	Patients did not have back pain
Jacobs (1975)(349)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Knox (1993)(350)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Kostuik (1979)(351)	Scoliosis
Lee (1995)(352)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Maghout-Juratli (2006)(117)	54% of patients had excluded medical conditions (e.g., radiculopathy, spondylolisthesis, spinal stenosis), and the discography analysis combined their data with data from patients with included conditions
Min (1996)(353)	Patients underwent discography prior to hemilaminectomy, not prior to fusion.
Moneta (1994)(354)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Murtagh (1992)(61)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Newman (1992)(355)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Ohnmeiss (1995)(22)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Parker (1996)(356)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Patrick (1973)(357)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Roberts (1972)(358)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Sachs (1987)(13)	Many patients had excluded medical conditions (e.g., radiculopathy), and their data were combined with data from patients with included conditions
Sachs (1990)(359)	Not a research study
Schechter (1991)(360)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Schellhas 1996 p79 (1996)(14)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Simmons (1975)(361)	Treatment was discotomy, not fusion
Simmons (1988)(362)	Not a research study
Smith (1998)(363)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.
Taveras (1967)(364)	Not a research study

Study	Reason for Exclusion
Urasaki (1998)(365)	Patients were receiving steroid intradiscal therapy, and were not being considered for fusion surgery.
Vanharanta (1987)(366)	Kappa statistic was not reported, and only 46/225 discs were included in pertinent analyses
Vanharanta (1988)(367)	Kappa statistic was not reported, and only 11/300 patients were included in pertinent analyses
Vanharanta (1989)(368)	Secondary publication of Vanharanta(367)
Walsh 1990 p1081 (1990)(12)	Study enrolled fewer than 10 patients with back pain
Weatherly (1986)(369)	Not a study of discography
Wetzel (1994)(370)	The group with noncontained contrast on morphology examination comprised fewer than 10 observations
Whitecloud (1987)(371)	Cervical not lumbar
Wilson (1969)(372)	Study did not report the reliability of discography, nor data on whether discography results predict fusion outcomes, nor data on the impact of discography on fusion outcomes.

Appendix C. Quality of Literature and Evidence Strength

Study Quality Assessment

For all Key Questions except Key Question 4, we applied a 22-item quality assessment instrument used to assess the quality of controlled studies (below). Each question is answered with "Yes", "No" or "NR" (not reported).

1. Were patients randomly assigned to the study's groups?
2. Did the study use appropriate randomization methods?
3. Was there concealment of group allocation?
4. For non-randomized trials, did the study employ any other methods to enhance group comparability?
5. Was the process of assigning patients to groups made independently from physician and patient preference?
6. Did patients in different study groups have similar levels of performance on the outcome of interest at the time they were assigned to groups?
7. Were the study groups comparable for all other important factors at the time they were assigned to groups?
8. Did the study enroll all suitable patients or consecutive suitable patients within a time period?
9. Was the comparison of interest prospectively planned?
10. If patients received ancillary treatments, was there a ≤5% difference between groups in the proportion of patients receiving each specific ancillary treatment?
11. Were the two groups treated concurrently?
12. Was compliance with treatment ≥85% in both of the study's groups?
13. Were patients blinded to the treatment they received?
14. Was the healthcare provider blinded to the groups to which the patients were assigned?
15. Were those who assessed the patient's outcomes blinded to the group to which the patients were assigned?
16. Was the integrity of blinding of patients, physicians or outcome raters tested and found to be preserved?
17. Was the outcome measure of interest objective and was it objectively measured?
18. Was a standard instrument used to measure the outcome?
19. Was there ≤15% difference in the length of follow-up for the two groups?

20. Did ≥85% of the patients complete the study?
21. Was there a ≤15% difference in completion rates in the study's groups?
22. Was the funding for this study derived from a source that would not benefit financially from results in a particular direction?

For Key Question 4, we assessed studies of the reliability of discography using the following 10 items:

1. Was the study prospective?
2. Were the patients enrolled consecutively?
3. Were the patient inclusion/exclusion criteria applied consistently to all patients?
4. Were data reported for at least 85% of enrolled patients?
5. Was the funding for this study derived from a source that does not have a financial interest in its results?
6. If two injections were performed on each patient, did patients receive the same instructions for pain reporting during the two discography examinations being compared?
7. If two injections were performed on for each patient, did the same injector perform the two discography examinations being compared?
8. Was discography interpreted without knowledge of other discography results in this patient?
9. Was discography interpreted without knowledge of other clinical information about this patient?
10. TEST-RETEST RELIABILITY ONLY. Did the same person interpret the two discography results?

We scored the quality for each outcome/timepoint by coding +1 for each Yes, -1 for each No, and 0 for each NR. The numbers were added, and then we transformed the total so that the best possible study would score 10 (i.e., all Yes's), and the worst possible study would score 0 (i.e., all No's). If the resulting combined score was <4, we considered the study Very Low quality and excluded that data from further consideration (but the study may have been included for other outcomes or other timepoints). If the score was ≥4 and ≤6, we categorized the quality as Low; if the score was >6 or ≤8, we categorized quality as Moderate; if it was >8, we categorized the quality as High. We then used these quality categories to proceed through the Strength of Evidence system, described next.

Strength of Evidence System

In evaluating the stability and strength of a body of literature, we used the ECRI Institute strength-of-evidence system.(72) This system employs decision points that collectively yield an overall category that describes the strength of the evidence for a quantitative estimate and qualitative conclusion as strong, moderate, weak, or unacceptably weak. The qualitative conclusion addresses the question, "Does it work?" The quantitative estimate addresses the question, "How well does it work?" This distinction allows an evidence base to be considered weak in terms of the quantitative estimate of effect (e.g., if estimates vary widely among studies) but strong or moderate with respect to the qualitative conclusion (e.g., if all studies nevertheless demonstrate the same direction of effect).

The system addresses five general aspects of the evidence: quality, quantity, consistency, robustness, and magnitude of effect. Quality refers to the degree of potential bias in the design or

conduct of studies. Quantity refers to the number of studies and the number of enrolled patients. Consistency addresses the degree of agreement among the results of available studies. Robustness involves the constancy of conclusions in the face of minor hypothetical alterations in the data. Magnitude of effect concerns the quantitative amount of benefit (or harm) that patients experience after treatment, and it is only considered in the qualitative section of the system.

The system outputs two ratings: a stability rating (which pertains to a quantitative conclusion) and a strength rating (which pertains to a qualitative conclusion). Interpretations of the two types of ratings appear in the table below.

Table 8. Interpretation of Different Categories of Strength of Evidence Supporting Conclusion

Strength of Evidence	Interpretation
Qualitative Conclusion (Direction of Effect)	
Strong Evidence	Evidence supporting the qualitative conclusion is convincing. It is highly unlikely that new evidence will lead to a change in this conclusion.
Moderate Evidence	Evidence supporting the qualitative conclusion is somewhat convincing. There is a small chance that new evidence will overturn or strengthen our conclusion. ECRI Institute recommends regular monitoring of the relevant literature at this time.
Weak Evidence	Although some evidence exists to support the qualitative conclusion, this evidence is tentative and perishable. There is a reasonable chance that new evidence will overturn or strengthen our conclusions. ECRI Institute recommends frequent monitoring of the relevant literature at this time.
Insufficient	Although some evidence exists, this evidence is not of sufficient strength to warrant drawing an evidence-based conclusion from it. ECRI Institute recommends frequent monitoring of the relevant literature at this time.
Quantitative Conclusion (Magnitude of Effect)	
High Stability	The estimate of treatment effect included in the conclusion is stable. It is highly unlikely that the magnitude of this estimate will change substantially as a result of the publication of new evidence.
Moderate Stability	The estimate of treatment effect included in the conclusion is somewhat stable. There is a small chance that the magnitude of this estimate will change substantially as a result of the publication of new evidence. ECRI Institute recommends regular monitoring of the relevant literature at this time.
Low Stability	The estimate of treatment effect included in the conclusion is likely to be unstable. There is a reasonable chance that the magnitude of this estimate will change substantially as a result of the publication of new evidence. ECRI Institute recommends frequent monitoring of the relevant literature at this time.
Unstable	Estimates of the treatment effect are too unstable to allow a quantitative conclusion to be drawn at this time. ECRI Institute recommends frequent monitoring of the relevant literature.

To determine strength and stability ratings for Key Questions, we applied the ECRI Strength and Stability of Evidence System.(72) The methods we used to resolve these decision points appear below.

Decision Point 1: Determine Quality of Individual Studies

For this decision point, we excluded any study that was determined to be of Very Low quality (score <4, see previous section). The remaining studies constituted the evidence base for the rest of the system.

Decision Point 2: Determine Quality of Evidence Base

We classified the overall quality of the evidence base by taking the median quality category of the individual studies. We used the median because it is the appropriate measure of central tendency to represent the "typical" quality category, and is less sensitive to outliers than the mean. Depending on the overall quality categories for each outcome, we then followed the high, moderate, or low quality branch of the system (as illustrated in Figure 4). If the median score fell between two categories, we proceeded with the lower quality category. Because the quality was determined separately for each outcome, a study that scored as moderate quality for one outcome might score as low quality for another outcome.

Decision Point 3: Is Quantitative Estimate Possible?

The answer to Decision Point 3 depends upon the adequacy of reporting in available studies, as well as the number of available studies in an evidence base.

We conducted quantitative analysis of a given outcome using meta-analysis when the data for that outcome was reported in at least three studies in a statistically compatible manner. When this was the case, we proceeded to Decision Point 4.

There are two possible scenarios for which we do not attempt to draw a quantitative conclusion, but rather proceed directly to Decision Point 8 to attempt to form a qualitative conclusion. If fewer than three studies are available, no quantitative estimate is warranted, regardless of reporting. This is because heterogeneity cannot be adequately assessed if there are only one or two studies. If three or more studies are available, but fewer than 75% of them permit determination of the effect size and its dispersion (either by direct reporting from the trial or calculations based on reported information) we do not attempt a quantitative estimate.

Decision Point 4: Are Data Quantitatively Consistent (Homogeneous)?

This decision point was used only if a quantitative estimate is possible in Decision Point 3.

Consistency refers to the extent to which the results of studies in an evidence base agree with each other.(373) The more consistent the evidence, the more precise the summary estimate of treatment effect derived from the evidence base. We measured quantitative consistency in an evidence base using a meta-analytic test, Higgins and Thompson's I^2 statistic.(74) We considered the evidence base to be quantitatively consistent (not substantially heterogenous) when I^2 <50%.(74)

When I^2 <50%, we combined the results from the studies to yield a meta-analytic summary statistic. We then tested the robustness of this summary estimate in Decision Point 5. If it was substantially heterogenous ($I^2 \geq 50\%$), then we proceeded to Decision Point 6.

Decision Point 5: Are Findings Stable (Quantitatively Robust)?

To be considered stable, or quantitatively robust, the summary estimate must have met all three of the following conditions:

1) Sufficiently narrow confidence interval around the summary effect size. This is defined as a total interval (from the lower bound of the 95% confidence interval to the upper bound) that is not bigger than twice the level of clinical significance for that outcome (clinical significance is defined in the text under Key Questions and Outcomes Assessed).
2) After removal and replacement of one study at a time, the summary effect size never strays further than 1 unit of clinical significance away from the all-study effect size.
3) Cumulative robustness test by year, using the same criterion as for removal of one study at a time.

If the summary estimate met all three of the above criteria, it was considered quantitatively robust. If the summary estimate did not meet all three of these conditions, it was deemed not quantitatively robust.

Decision Points 6 and 7: Exploration of Heterogeneity

Decision Points 6 and 7 are relevant only when one has, during a quantitative analysis, found that the findings of the studies that comprise an evidence base are determined to be substantially heterogeneous (see Decision Point 4).

Decision Point 6: Does Meta-regression Explain Heterogeneity?

In the case of substantial heterogeneity in an evidence base comprised of fewer than 5 studies, we did not attempt to determine a quantitative estimate. Whenever this was the case, we proceeded to Decision Point 8.

If we observed substantial heterogeneity in an evidence base comprised of at least 5 studies, we attempted to explain the heterogeneity using meta-regression. We planned *a priori* to use the following factors as predictor variables in meta-regression:

- The type of fusion procedure performed
- Whether the trial had a crossover rate (to non-assigned treatment) ≥15%.
- The actual percentage of patients with reported data to the timepoint of interest
- The overall quality category (high, moderate, low)

Meta-regression was considered to have explained the heterogeneity if the covariate was statistically significant ($p < 0.05$), and if the resulting I^2 was less than 50%.

Decision Point 7: Is Meta-regression Model Stable?

The purpose of Decision Point 7 is to test the stability of any quantitative findings that may emanate from meta-regression analysis. We used the same robustness tests as in Decision Point 5.

Decision Point 8: Are Qualitative Findings Robust?

To be considered qualitatively robust, the conclusion must have met both of the following conditions:

1) After removal of one study at a time, the qualitative conclusion was unchanged.
2) During cumulative robustness test by year, the qualitative conclusion was unchanged

If the analysis did not meet both of these conditions, it was deemed not qualitatively robust.

Decision Point 9: Are Data Qualitatively Consistent?

This Decision Point is used only when two studies comprise the evidence base for an outcome and meta-analysis is either not possible or inappropriate. We considered two-study evidence bases qualitatively consistent if they met *either* of the following two criteria:

1) Both studies showed a statistically significant effect in the same direction; or

2) Neither study showed a statistically significant effect.

If the evidence base is of moderate or high quality, meta-analysis of two studies may be performed in an attempt to reach a qualitative conclusion about the direction of effect (i.e., does the treatment work?).

Decision Point 10: Is Magnitude of Treatment Effect Extremely Large?

When considering the strength of evidence supporting a qualitative conclusion based on only one or two studies, magnitude of effect becomes very important. The larger the size of the summary effect, the more confident one can be that new evidence will not overturn a general conclusion that the treatment is beneficial.

In this Decision Point, the system divides the magnitude of effect into two categories: large and not large. An effect size was judged to be large if it was ≥3 times greater than the minimum effect size considered to be clinically significant.

Other parts of the system

Some parts of the system are not formally called "Decision Points", and yet some decisions must be made in order to apply them. These are described next.

Informative?

When there are only a small number of patients in an evidence base, statistical tests generally do not perform well. Under such circumstances, statistics cannot determine whether a true difference exists between treatments. This means that no clear conclusion can be drawn. For this decision point, we determined whether the precision of an evidence base was sufficient to permit a conclusion. Statistically significant results are potentially conclusive because they mean that a treatment effect may exist. Statistically non-significant results are also potentially conclusive, but only if they exclude the possibility that a clinically significant treatment effect exists.

When considering the summary effect size from a meta-analysis (or the effect size from a single study), there are three ways in which the effect can be "informative":

1) The summary effect size is statistically significantly different from 0. This would be indicated whenever the confidence interval does not overlap 0.

2) The summary effect size is not statistically significantly different from 0, but the confidence intervals are narrow enough to exclude the possibility that a *clinically significant difference* exists (see below for definitions of clinical significance).

3) The summary effect size is not statistically significantly different from 0, but the confidence intervals are narrow enough to exclude the possibility that a *substantial difference* exists. This possibility is included to address situations when even a very small

effect can be considered "clinically significant" (e.g., a difference in mortality rates), but the effect may not be "substantial".

Multicenter Trial?

We defined a multicenter trial as one that meets the following two conditions: 1) ≥3 centers and 2) either ≥100 patients or at least 3 centers enrolled ≥20 patients percenter.

Figure 4. Entry into System

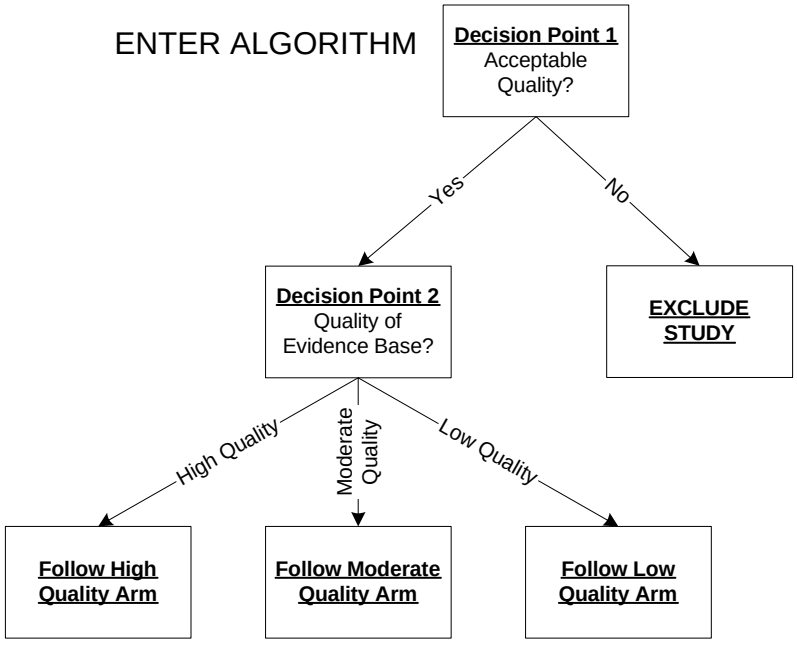


Figure 5. High-Quality Arm

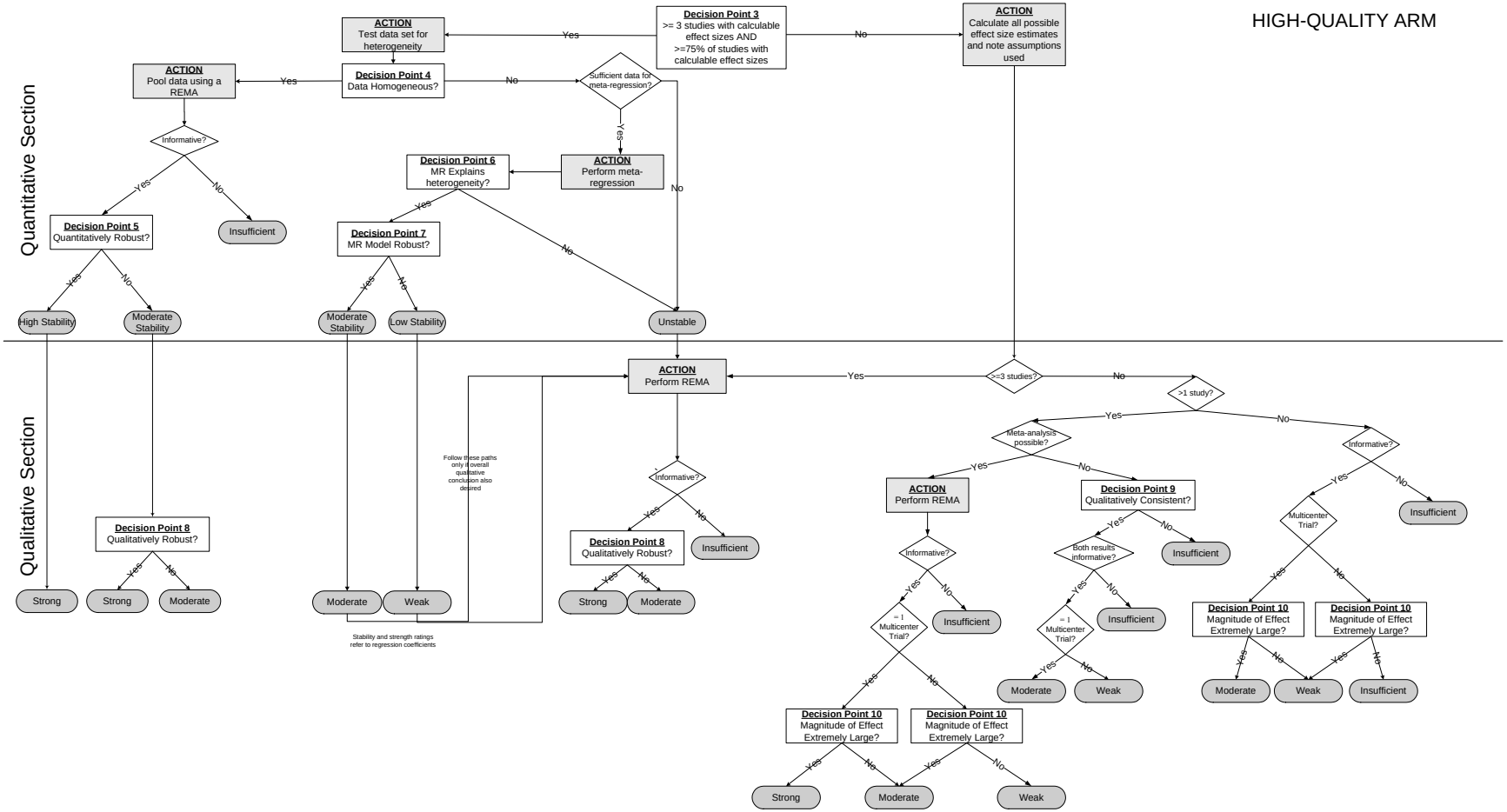


Figure 6. Moderate-Quality Arm

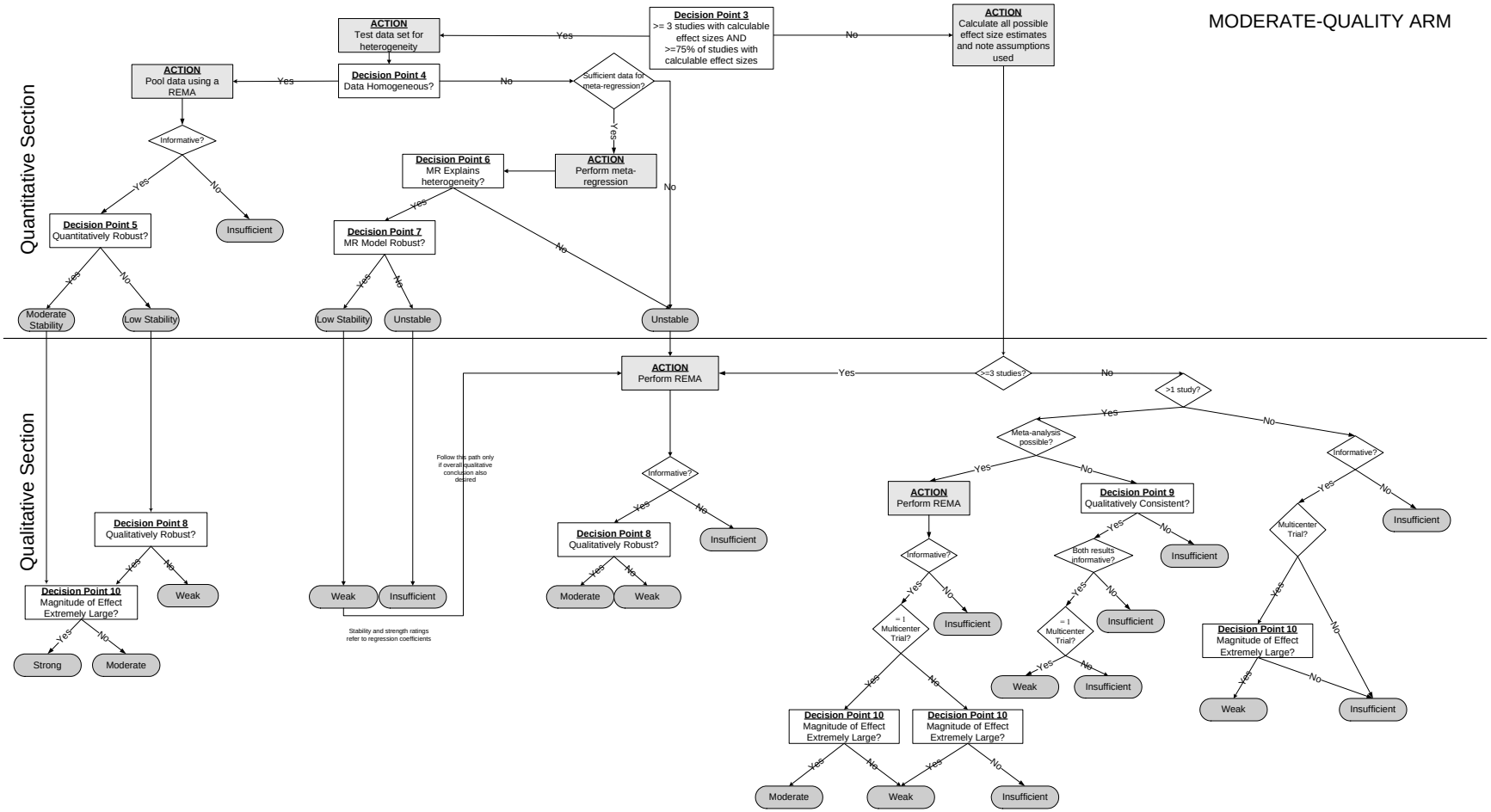
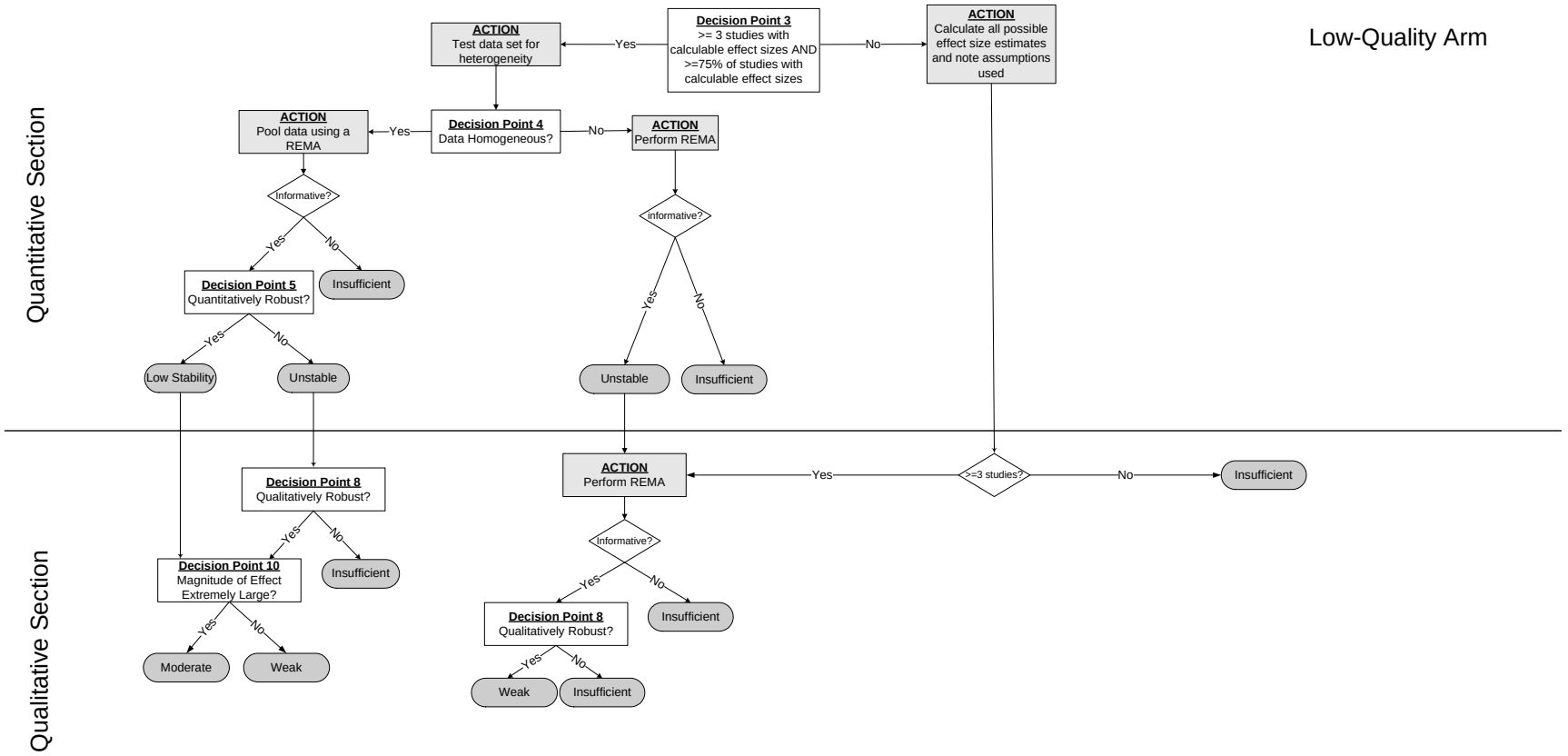


Figure 7. Low-Quality Arm



Appendix D. Evidence Tables for Spinal Fusion (Key Questions 1-3)

Table 9. Patient Selection Criteria of Included Studies (Randomized Trials Addressing Key Questions 1 and 2)

Study	Inclusion Criteria					Exclusion Criteria	
	Primary Diagnosis or Condition	Severity	Diagnostic Criteria	Minimum Duration of Condition	Age	Exclusion Criteria, Back-Related	Exclusion Criteria: Other Comorbidity
Brox et al. 2006(31)	Chronic low back pain and previous surgery for disc herniation	≥30 on Oswestry disability index	Degeneration of L4-L5 and/or L5-S1 on plain radiograph	1 year	25-60 years	Spinal stenosis with reduced walking distance and neurological signs Recurrent disc herniation Lateral recess stenosis with clinical signs of radiculopathy Previous spinal fracture Previous lumbar fusion Generalized disc degeneration on plain X-ray examination	Widespread myofascial pain Inflammatory disease Ongoing somatic or psychiatric disease that excludes on or both treatments Registered medical abuse Reluctance to accept one or both of the treatment regiments
Fairbank et al. 2005(77)	Chronic low back pain, with or without referred pain, and with or without previous root decompression or discectomy	Not reported	Not reported	1 year	18-55 years	Previous surgical stabilization of the spine	Co morbidities that could complicate surgery, including infection, inflammatory disease, tumours, fractures Psychiatric disease Pregnancy

Study	Inclusion Criteria					Exclusion Criteria	
	Primary Diagnosis or Condition	Severity	Diagnostic Criteria	Minimum Duration of Condition	Age	Exclusion Criteria, Back-Related	Exclusion Criteria: Other Comorbidity
Brox et al. 2003(78)	Chronic low back pain	≥30 on Oswestry disability index	Degeneration of L4-L5 and/or L5-S1 on plain radiograph	1 year	25-60 years	Spinal stenosis with reduced walking distance and neurological signs Recurrent disc herniation Lateral recess stenosis with clinical signs of radiculopathy Previous spinal fracture Previous spine surgery Generalized disc degeneration on plain X-ray examination	Widespread myofascial pain Inflammatory disease Ongoing somatic or psychiatric disease that excludes on or both treatments Registered medical abuse Reluctance to accept one or both of the treatment regiments
Fritzell et al. 2001(79); Fritzell et al. 2002(80); Hagg et al. 2006(374)	Chronic low back pain with back pain worse than leg pain and no signs of nerve root compression and unsuccessful nonsurgical treatment	Severe; ≥7/10 on authors' disability scale	Degenerative changes at L4-L5 and/or L5-S1 on plain radiographs and/or CT and/or MRI	2 years; on sick leave at least 1 year	25-65	Precious 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 radiologic findings, such as spondylolisthesis, new or old fractures, infection, inflammatory process, neoplasm, or stenosis Anamnestic signs of spinal stenosis	Obvious ongoing psychiatric illness Obvious painful and disabling arthritic hip joints

CT Computed tomographic imaging.
MRI Magnetic resonance imaging.

Table 10. Patient Selection Criteria of Included Studies (Other Studies Addressing Key Question 2)

Study	Inclusion Criteria					Exclusion Criteria	
	Primary Diagnosis or Condition	Severity	Diagnostic Criteria	Minimum Duration of Condition	Age	Exclusion Criteria, Back-Related	Exclusion Criteria: Other Comorbidity
Martin et al. 2007(81)	Lumbar surgery for degenerative disease, ICD-9-CM codes for disc degeneration, disc herniation, spinal stenosis, and spondylolisthesis	Not reported	ICD-9-CM classification in database	None reported	≥20 years	Any record of lumbar surgery or procedure code indicating repeat spinal surgery within preceding 3 years	Diagnosis of fracture, dislocation, infection, pregnancy Individuals with cervical and thoracic spinal condition
Burkus et al. 2006(83); Burkus et al. 2005(82)	Single-level lumbar degenerative disc disease with up to Grade I spondylolisthesis, unresponsive to treatment	Disabling	Not reported	6 months	Not reported	Spinal conditions other than degenerative disc disease Degenerative disc disease at disc space levels other than L4-L5 or L5-S1 Previous anterior arthrodesis at the involved level	Obesity (>40% above ideal weight) Active bacterial infection Medical condition requiring medication that could interfere with fusion (e.g., steroids or nonsteroidal antiinflammatory medication)
Sasso et al. 2005(84)	Symptomatic degenerative disc disease	Not reported	Symptoms of intractable back and/or leg pain with positive diagnostic imaging findings	6 months	19 to 81 years	None reported	None reported

Bezer et al. 2004(85)	Degenerative disease of lumbar spine treated with spinal decompression with posterior fusion using iliac bone graft and instrumentation	Not reported	Not reported	Not reported	Not reported	None reported	None reported
Sasso et al. 2004(375)	Not reported	Not reported	Not reported	Not reported	Not reported	None reported	None reported
Scaduto et al. 2003(86)	Not reported (Retrospective study on all instrumented lumbar interbody fusions during a time period)	Not reported	Not reported	Not reported	Not reported	None reported	None reported
Burkus et al. 2002(87)	Single-level degenerative lumbar disc disease with disabling low back and/or leg pain	Disabling	Not reported	Not reported	Not reported	None reported	None reported

Christensen et al. 2002(88)	Chronic low back pain and leg pain, static or dynamic, from lumbar or lumbosacral instability caused by isthmic spondylolisthesis (grades 1 and 2) primary degeneration (with no previous surgery), secondary or accelerating degeneration after decompressive surgery	Severe	Not reported	Not reported	20-65 years	Previous fusion	Metabolic bone disease, including previously diagnosed osteoporosis or osteoporosis diagnosed at surgical clinic by radiography and bone mineral density testing One or more comorbidities Psychosocial instability Previous retroperitoneal surgery
McAfee et al. 2002(89)	Internal disc disruption, postlaminectomy syndrome, herniated nucleus pulposus, discogenic instability	Severe	Severe, disabling, intractable back pain; no previous interbody arthrodesis at the target levels; an absence of degeneration at adjacent neighboring disc spaces; and no greater than Meyerding Grade I spondylolisthesis, disabling back pain for at least 12 months, refractory to nonoperative care; substantial loss of disc height and mobility	At least 12 months	22-75 years	Disc space height of more than 12 mm, prior fusion at target level, >Grade 1 spondylolisthesis	Not reported
Slosar et al. 2000(91)	Low back pain, with or without radicular pain	Severe	Not reported	6-12 months	Not reported	None reported	None reported

Thalgott et al. 2000(92)	Not reported (Retrospective study of patients who had ALIF 1 or 2 levels)	Not reported	Not reported	Not reported	Not reported	None reported	None reported
Kuslich et al. 2000(96); Kuslich et al. 1998(95); FDA PMA P950002	Degenerative disc disease at 1 or 2 contiguous levels of the lumbar spine with chronic low back pain not responsive to conservative treatment	Disabling	Patient history, physical exam, radiography, magnetic resonance, and when necessary, discography	6 months	21-65	Spondylolisthesis greater than Grade 1	Active systemic or local infection Ostoepenia Symptomatic vascular disease Malignancy Gross obesity Pregnancy
Brantigan et al. 2000(90); FDA PMA P960025	Degenerative disc disease or herniation at one to four spinal levels from L2-S1 For Brantigan et. al; subgroup of patients with at least one failed discectomy operation	Not reported	Discogenic back pain with degeneration of the disc confirmed by history and radiographic studies	6 weeks	18-89	Spondyloptosis Clinically significant abnormalities at more than three levels	Significant osteoporosis or metabolic bone disease Past or present infection of disc or spine Past or present illicit drug use Current alcohol use Prisoners

Regan et al. 1999(93)	Painful degenerative disc disease with or without herniation	Not reported	Patient history of mechanical back and leg pain with magnetic resonance findings with disc space narrowing at 1 or 2 contiguous levels L4-L5 and L5-S1	6 months	Not reported	Grade II spondylolisthesis or greater	Active infection Metabolic bone disease Poor general health (e.g., cancer, significant cardiac disease) Obesity ≥40% of ideal body weight Pregnancy Multiple complicating psychosocial factors Patients with history of abdominal surgery that potentially left significant intraperitoneal scarring were not included in laparoscopic arm
Greenough et al. 1998(94)	Not reported (Retrospective study of consecutive patients who underwent lumbar posterolateral fusion with pedicle screw fixation using the variable screw plate)	Not reported	Not reported	Not reported	Not reported	None reported	None reported

Malter et al. 1998(97)	Lumbar surgery for degenerative disease, ICD-9-CM codes for disc degeneration, disc herniation, spinal stenosis, and possible instability	Not reported	ICD-9-CM classification in database	None reported	Not reported	Patients with diagnoses or surgical procedures related to the cervical or thoracic spine	Patients with diagnoses or procedures indicative of malignancy, spinal infection, spinal trauma, fracture, or inflammatory spondylitis
Penta and Fraser 1997(98)	Not reported (Retrospective study of consecutive patients undergoing fusion)	Not reported	For discogenic back pain, abnormal discography and concordant pain reproduction at the affected level(s)	Not reported	Not reported	None reported	None reported
Ray 1997(99); Ray 1998(290); FDA PMA P950019	Symptomatic degenerative disc disease at 1 or 2 levels from L2-S1	Severely disabling	Low back pain with or without sciatical, pain reproduction on discography, annular degeneration, herniation, loss of disc height and/or osteophytes	1 year	≥18 years	Prior interbody fusion at target level(s) Degeneration of adjacent levels, whether painful or not Spondylolisthesis above Grade 1 Need for fusion at 3 or more levels Anatomic abnormalities of bone to be fused Significant endplate sclerosis at diseased level Cervical or thoracic degenerative disc disease	Concomitant conditions requiring steroids Systemic or terminal illness Active drug abuse Active infection Pregnancy

Thomsen et al. 1997(100); Christensen et al. 2002(101)	Chronic low back pain from spondylolisthesis Grades 1 and 2 or primary or secondary degenerative instability	Severe	Not reported	Not reported	20-70 years	Previous fusion Metabolic bone disease (e.g., osteoporosis)	Comorbidity, not specified Psychosocial instability
Hall et al. 1996(103)	Degenerative disc disease	Not reported	Increased surface area of the vertebral endplates by ridging or osteophyte formation on radiograph, increased surface area of facet area with change in geometric configuration, scarring or thickening of outer layer of anulus fibrosis, ligamentum flavum, or facet joint capsule, instability of motion segment unit	Not reported	Not reported	Absence of evidence of change of disc space on radiograph Stable motion segments Asymptomatic clinical status No stenosis of the canal caused by degenerative process of disc	None reported
Christensen et al. 1996(102)	Not reported (Retrospective study of patients who underwent anterior lumbar spondylodesis)	Not reported	Not reported	Not reported	Not reported	None reported	None reported
Greenough et al. 1994(376); Greenough et al. 1994(104)	Not reported (Retrospective study of patients who underwent anterior interbody fusion)	Not reported	Not reported	Not reported	Not reported	None reported	None reported

Markwalder and Battaglia 1993(106)	Not reported (Retrospective study of patients who underwent instability-related failed back surgery syndrome)	Not reported	Not reported	Not reported	Not reported	None reported	None reported
Gill and Blumenthal 1993(105)	Not reported (Retrospective study of patients who underwent spinal fusion)	Not reported	Not reported	Not reported	Not reported	None reported	None reported

Table 11. Characteristics of Patients in Included Studies (Randomized Trials Addressing Key Questions 1 and 2)

Study	Number of Patients Screened	Number of Patients Referred	Number of Patients Randomized	Diagnoses	Duration of Symptoms	Proportion with Previous Spinal Fusion	Proportion with prior back surgery	Age	Proportion of Men	Proportion of Smokers	Proportion on Compensation/ Disability Pension	Proportion with Litigation Pending
Brox et al. 2006(31)	Not reported	113	60	Chronic back pain after surgery for herniation, with degeneration at L4-L5 and/or L5-S1	96 months median (interquartile Range: 36-150)	0% (excluded)	100% (herniation surgery)	43 years median (interquartile range 35-50)	52%	65%	Not reported	Not reported
Fairbank et al. 2005(77)	Not reported	Not reported	349	Chronic low back pain with or without referred pain. 11% had spondylolisthesis.	8 years mean (Range: 1-35)	0% (excluded)	8% (laminectomy)	12% <30 years, 37% 30-39 years, 35% 40-49 years, 15% ≥50 years	49%	43%	Not reported	13%
Brox et al. 2003(78)	Not reported	121	64	Chronic back pain with degeneration at L4-L5 and/or L5-S1	10.8 years mean	0% (excluded)	0%	44.8 years mean	44%	44%	11%	Not reported
Fritzell et al. 2001(79), Fritzell et al. 2002(80)	Not reported	310	294	Chronic low back pain with degeneration at L4-L5 and/or L5-S1 thought to be causing the back pain	8.0 years mean (Range: 2-40 years)	0% (excluded)	18.7% (discectomy)	43 years mean (Range: 25-64)	50%	32.7%	21%	61%

Table 12. Characteristics of Patients in Included Studies (Other Studies Addressing Key Question 2)

Study	Number of Patients Referred	Number of Patients Randomized	Number of patients undergoing surgery	Diagnoses	Duration of Symptoms	Proportion with Previous Spinal Fusion	Proportion with prior back surgery	Age	Proportion of Men	Proportion of Smokers	Proportion on Workers' Compensation	Proportion with Litigation Pending
Martin et al. 2007(81)	NR	NA	977	Herniated disc (462 patients), degenerative disc (515 patients) 1,368 fusion patients with stenosis or spondylolisthesis did not meet our inclusion criteria	NR	0	0	~ 51 years (not reported separately for subgroups that met our inclusion criteria)	~48%	NR	~21%	NR
Burkus et al. 2005(82); Burkus et al. 2006(83)	NR	131	131	Single-level symptomatic degenerative disc disease, spondylolisthesis (grade ≤1)	>6 months	0	35.1%	42	39%	32.8%	30.5%	9.9%
Sasso et al. 2005(84)	NR	NR	208	Degenerative disc disease	At least 6 months	NR	NR	42.3 (19 to 81)	52.9%	NR	NR	NR
Bezer et al. 2004(85)	NR	117	117	Degenerative disease of the lumbar spine	NR	NR	NR	49.5	43.6%	NR	NR	NR
Sasso et al. 2004(375)	NR	140	140	Degenerative disc disease, spondylolisthesis (≤Grade 1)	At least 6 months	0%	42.1%	41 (18 to 64)	45%	30.7%	38.6%	NR
Scaduto et al. 2003(86)	NR	NA	119	Painful degenerative disc disease	NR	NR	37.8%	45 (20 to 70)	43.7%	NR	NR	NR

Study	Number of Patients Referred	Number of Patients Randomized	Number of patients undergoing surgery	Diagnoses	Duration of Symptoms	Proportion with Previous Spinal Fusion	Proportion with prior back surgery	Age	Proportion of Men	Proportion of Smokers	Proportion on Workers' Compensation	Proportion with Litigation Pending
Burkus et al. 2002(87)	NR	279	279	Degenerative lumbar disc disease	≥ 6 months	NR	0	43	52.3%	34.4%	33.7%	14.3%
Christensen et al. 2002(88)	NR	148	148	Primary and secondary degeneration, isthmic spondylolisthesis (Grades 1 and 2)	<2 years (18.2%), >2 years (81.8%)	0%	40.5%	45.5 (20 to 63)	39.2%	NR	NR	NR
McAfee et al. 2002(89)	NR	NA	100	Internal disc disruption, herniated nucleus pulposis, discogenic instability, postlaminectomy syndrome	At least 12 months	NR	27% had prior laminectomy	47 (22 to 75)	50%	22%	NR	NR
Brantigan et al. 2000(90); FDA PMA P960025	NR	NA	221	Recurrent disc disease (49.8%), spondylolisthesis (23.1%), failed fusion (27.1%)	Mean: 8 years	27.1%	76.9%	44.3 (24 to 77)	57.0%	62%	40%	22%
Slosar et al. 2000(91)	NR	NA	141	Painful degenerative discs (63.1%), spondylolisthesis (10.6%), pseudarthrosis (17%), internal disc disruption (5.7%), foraminal stenosis (2.8%)	At least 6 months	NR, but 17% had pseudarthrosis from prior attempted fusion	69%	38.8 (21 to 58)	53%	NR	55%	NR

Study	Number of Patients Referred	Number of Patients Randomized	Number of patients undergoing surgery	Diagnoses	Duration of Symptoms	Proportion with Previous Spinal Fusion	Proportion with prior back surgery	Age	Proportion of Men	Proportion of Smokers	Proportion on Workers' Compensation	Proportion with Litigation Pending
Thalgott et al. 2000(92)	NR	NA	202	Internal disc disruption, failed laminectomy, spondylolisthesis, degenerative disc disease, spinal stenosis, severe disc herniation, pseudarthrosis, aseptic discitis	NR	NR	50%	45 (23 to 74)	50%	NR	NR	NR
Regan et al. 1999(93)	NR	NA	540	Degenerative disc disease (some also had disc herniation or spondylolisthesis)	Mean: 64 months	0	37.9%	40.6	50.8%	NR	51.5%	NR
Greenough et al. 1998(94)	NR	NA	135	Discogenic/ mechanical back pain, failed previous surgery, spondylolysis/ spondylolisthesis	NR	NR	11.1%	43 (22 to 79)	43.7%	NR	48.1%	NR
Kuslich et al. 1998(95); Kuslich et al. 2000(96); FDA PMA P950002	NR	NA	947	Degenerative disc disease, some with additional disc herniation or spondylolisthesis (≤ Grade 1)	At least 6 months	5%	36%	41.5 (19 to 73)	54%	26%	57%	NR

Study	Number of Patients Referred	Number of Patients Randomized	Number of patients undergoing surgery	Diagnoses	Duration of Symptoms	Proportion with Previous Spinal Fusion	Proportion with prior back surgery	Age	Proportion of Men	Proportion of Smokers	Proportion on Workers' Compensation	Proportion with Litigation Pending
Malter et al. 1998(97)	NR	NA	1041	Herniated disc (207 patients), degenerative disc (265 patients) 569 fusion patients with stenosis, possible instability or miscellaneous did not meet our inclusion criteria	NR	NR	~12%	~47	~54	NR	NR	NR
Ray et al. 1997(99); Ray 1998(290); FDA PMA P950019	NR	NR	236	Degenerated painful disc space, spondylolisthesis ≤Grade 1	≥1 year	NR	45%	41.2 (18 to 79)	62%	NR	NR	NR
Penta and Fraser 1997(98)	NR	NA	125	Discogenic back pain (78%), failed previous fusion (10.4%), spondylolisthesis (6%)	NR	10.4%	NR	48 (28 to 73)	41.7%	NR	58.1%	NR
Thomsen et al. 1997(100); Christensen et al. 2002(101)†	NR	130	130	Primary degenerative instability (40.8%), Secondary degenerative instability (31.5%), isthmic spondylolisthesis grades 1 and 2 (26.9%)	>2 years (78.5%), <2 years (20.8%)	0	40.8%	44 (20 to 67)	46.2%	51.5%	NR	NR

Study	Number of Patients Referred	Number of Patients Randomized	Number of patients undergoing surgery	Diagnoses	Duration of Symptoms	Proportion with Previous Spinal Fusion	Proportion with prior back surgery	Age	Proportion of Men	Proportion of Smokers	Proportion on Workers' Compensation	Proportion with Litigation Pending
Christensen et al. 1996(102)	NR	NA	132	Disc Degeneration (52.5%), spondylolisthesis (47.5%, mostly Grade 1)	NR	0	45.8%	Range: 15 to 55	41.7%	NR	NR	NR
Hall et al. 1996(103)	NR	NA	120	Degenerative disc disease	NR	0	63%	54 (25 to 83)	40.8%	32%	NR	NR
Greenough et al. 1994(376); Greenough et al. 1994(104)	NR	NA	151	Discogenic or mechanical back pain (65%), failed prior surgery (23%), motion segment instability (3%), spondylolysis or spondylolisthesis (9%)	NR	NR	23%	41 (17 to 62)	51%	NR	70.2%	NR
Gill and Blumenthal 1993(105)	NR	NA	238	Degenerative disc disease (33%), herniated degenerative disc disease (40%), postlaminectomy syndrome (11%), spondylolisthesis (10%), recurrent herniated disc (5%), pseudarthrosis of prior fusion (1%)	Mean: 1 year	1%	NR	39 (21 to 57)	73.5%	57.1%	94.5%	NR
Markwalder and Battaglia 1993(106)	NR	NA	171	Failed back surgery syndrome (100%)	NR	NR	100%	NR	NR	NR	NR	NR

* A secondary publication on the male patients in both arms of Burkus et al. 2002(87)

† A follow-up of Thomsen et al. 1997(100)

Table 13. Study Quality Assessments – RCTs Addressing Key Question 1

ECRI study quality scale - questions	Efficacy outcomes (ODI and VAS back pain)			
	Brox 2006	Fairbank 2005	Brox 2003	Fritzell 2001
1. Were patients randomly assigned to groups?	Yes	Yes	Yes	Yes
2. Did the study employ appropriate randomization methods?	Yes	Yes	Yes	Yes
3. Was there concealment of allocation?	Yes	Yes	Yes	Yes
4. Were methods other than randomization used to make the groups comparable?	Yes	Yes	Yes	Yes
5. Were patients assigned to groups based on factors other than patient or physician preference?	Yes	Yes	Yes	Yes
6. Were the characteristics of the patients in different groups comparable?	Yes	Yes	Yes	No
7. Did the patients in the different study groups have similar levels of performance on outcomes at baseline?	Yes	Yes	Yes	Yes
8. Did study enroll all suitable or consecutive patients within a time period?	Yes	Yes	Yes	Yes
9. Was the study prospectively planned?	Yes	Yes	Yes	Yes
10. ≤5% difference between groups in the proportion of patients receiving ancillary treatments?	NR	No	NR	NR
11. Were all of the study's groups concurrently treated?	Yes	Yes	Yes	Yes
12. Was compliance with treatment greater than or equal to 85% in both of the groups?	No	No	Yes	Yes
13. Were subjects blinded to treatment?	No	No	No	No
14. Was the treating physician blinded to group assignment?	No	No	No	No
15. Were the outcome assessors blinded to group assignment?	No	No	No	No
16. Was blinding of patients, physicians, or outcome raters tested and found to be preserved?	No	No	No	No
17. Was the outcome of interest objective?	No	No	No	No
18. Was the instrument used to measure the outcome standard?	Yes	Yes	Yes	Yes
19. Were the follow-up times in all of the study's relevant groups approximately equal (difference ≤15%)?	Yes	Yes	Yes	Yes
20. Did 85% or more of the patients complete the study?	Yes	No	Yes	Yes
21. Was there a less than 16% difference in completion rates in the study's groups?	Yes	Yes	Yes	Yes
22. Was the study free from potential financial conflict of interest?	Yes	Yes	Yes	No
Quality score	**7.0**	**6.4**	**7.5**	**6.6**
Quality rating	**Moderate**	**Moderate**	**Moderate**	**Moderate**

Table 14. Data for Key Question 1 – Change in Oswestry Disability Index (ODI) Scores

Study	Surgical procedure / nonsurgical therapy	Number of Patients Enrolled	Number of patients who changed treatment groups (crossovers)	Duration of Followup	Number of Patients at Follow up	Oswestry Disability Index (ODI)[a] at Baseline	Oswestry Disability Index (ODI)[a] at Followup	Difference in Change Score[b]
Brox et al. 2006(31)	Posterolateral fusion (PLF) with pedicle screws + autologous bone graft	29 (23 received surgery)	0	1 year	28	47 (9.4)	38.1 (20.1)	-9.7 (-21.7 to 1.7) p = 0.09 (adjusted for gender and treatment expectations)
	Cognitive intervention + exercises	31 (29 received intervention)	2		29	45.1 (9.1)	32.3 (19.1)	
Fairbank et al. 2005(77)	Spinal fusion (unspecified)	176 (139 received surgery)	7	2 years	138	46.5 (14.6)	34.0 (21.1)	4.1 (0.1 to 8.1) p = 0.045 (adjusted for baseline measures)
	Intensive cognitive behavioral-based rehabilitation	173 (151 received intervention)	10 (but an additional 38 had surgery after rehabilitation before the 2-year followup)		146	44.8 (14.8)	36.1 (20.6)	
Brox et al. 2003(78)	PLF with pedicle screws + autologous bone graft + physical therapy	37 (33 received surgery)	0	1 year	35	42.0 (11.0)	26.4 (16.4)	2.7 (-6.8 to 12.2) p = 0.33 (adjusted for gender and treatment expectations)
	Cognitive intervention + exercises	27 (25 received intervention)	1		26	43.0 (13.0)	29.7 (19.6)	

Study	Surgical procedure / nonsurgical therapy	Number of Patients Enrolled	Number of patients who changed treatment groups (crossovers)	Duration of Followup	Number of Patients at Follow up	Oswestry Disability Index (ODI)[a] at Baseline	Oswestry Disability Index (ODI)[a] at Followup	Difference in Change Score[b]
Fritzell et al. 2001(79); Fritzell et al. 2002(80)	PLF with or without pedicle screws, or circumferential (PLIF or ALIF)	222 (204 received surgery)	18	2 years	201	47.3 (11.4)	35.7 (18.0)	8.8 p = 0.015
	Physical therapy + other non-operative therapies	72 (65 received intervention)	7		63	48.4 (11.9)	45.6 (16.1)	

[a] ODI score ranges from 0-100; higher scores mean greater disability

[b] Positive differences favor fusion; negative differences favor nonsurgical therapy

Table 15. Data for Key Question 1 – Change in VAS for Back Pain

Study	Surgical procedure / nonsurgical therapy	Number of Patients Enrolled	Number of patients who changed treatment groups (crossovers)	Duration of Followup	Number of Patients at Follow up	VAS Back Pain[a] at Baseline	VAS Back Pain[a] at Followup	Difference in Change[b]
Brox et al. 2006(31)	Posterolateral fusion (PLF) with pedicle screws + autologous bone graft	29 (23 received surgery)	0	1 year	28	64.6 (15.4)	50.7 (27.3)	-5.2 (-18.0 to 7.6)
	Cognitive intervention + exercises	31 (29 received intervention)	2		29	64.7 (11.1)	49.5 (20.0)	p = 0.42 (Adjusted for gender)
Brox et al. 2003(78)	PLF with transpedicular screws + autologous bone graft + physical therapy	37 (33 received surgery)	0	1 year	35	62.1 (14.5)	39.4 (25.5)	8.6 (-3.0 to 20.1)
	Cognitive intervention + exercises	27 (25 received intervention)	1		26	64.1 (13.7)	48.7 (24.0)	p = 0.14
Fritzell et al. 2001(79);	PLF with or without pedicle screws, or circumferential (PLIF or ALIF)	222 (204 received surgery)	18	2 years	201	64.2 (14.3)	43.2 (25.2)	16.7
Fritzell et al. 2002(80)	Physical therapy + other non-operative therapies	72 (65 received intervention)	7		63	62.6 (14.3)	58.3 (18.8)	p = 0.0002

[a] VAS ranges from 0-100; higher scores mean greater disability

[b] Positive differences favor fusion; negative differences favor nonsurgical therapy

Table 16. Overall Early and Late Adverse Event Rates Associated with Fusion (RCTs Addressing Key Questions 1 and 2)

Study	Surgical procedure (type of fusion) and control intervention	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Overall early adverse event rate	Overall late adverse event rate	Overall total adverse event rate
Brox et al. 2006(31)	Posterolateral fusion (PLF) with pedicle screws + autologous bone graft	29 (23 received surgery)	Early: 1 month?	23	NR	0 (0%)	NR
	Cognitive intervention + exercises	31 (29 received intervention)	Late: >1 month? to 1 year	29	NR	0 (0%	NR
Fairbank et al. 2005(77)	Spinal fusion (unspecified)	176 (139 received surgery)	Intraoperative	149 (includes 10 crossover patients)	Intraoperative: 19 (12.8%)	Postoperative: 11 (7.9%)	NR
	Intensive cognitive behavioral-based rehabilitation	173 (151 received intervention)	Postoperative (1 day to 2 years)	158 (includes 7 crossover patients)	0 (0%) p = 0.000004	0 (0%) p = 0.0005	NR
Brox et al. 2003(78)	PLF with pedicle screws + autologous bone graft + physical therapy	37 (33 received surgery)	Early: 1 month?	33	6 (18%)	0 (0%)	NR
	Cognitive intervention + exercises	27 (25 received intervention)	Late: >1 month? to 1 year	25	0 (0%) p = 0.03	0 (0%)	NR
Fritzell et al. 2001(79), Fritzell et al. 2002(80)	PLF with or without pedicle screws, or circumferential (PLIF or ALIF)	222 (204 received surgery)	Early: within first 2 weeks	211 (includes 7 crossover patients)	37 (17.5%)	13 (6.2%)	NR
	Physical therapy + other non-operative therapy (varied)	72 (65 received intervention)	Late: >2 weeks to 2 years	65	0 (0%) p = 0.0003	0 (0%) p = 0.04	NR

NR Not reported.

Table 17. Overall Early and Late Adverse Event Rates Associated with Fusion (Other Studies Addressing Key Question 2)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Overall early adverse event rate	Overall late adverse event rate	Overall total adverse event rate	Other
Martin et al. 2007(81)	Not specified	462 patients with herniated discs	11 years	462	NR	NR	NR	NR
		515 patients with degenerative disc disease		515	NR	NR	NR	NR
Burkus et al. 2005(82); Burkus et al. 2006(83)	rhBMP-2and MD-II threaded cortical, bone dowel. Anterior lumbar interbody fusion (ALIF), open, with transperitoneal or retroperitoneal approach	79	24 months	79	NR	NR	NR	NR
	Autologous bone graft fusion, Anterior (ALIF), open	52		52	NR	NR	NR	NR
Sasso et al. 2005(84)	Fusion with autograft, Anterior lumbar interbody fusion (ALIF), retroperitoneal or transperitoneal approach	208	2 years	208	NR	NR	NR	NR

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Overall early adverse event rate	Overall late adverse event rate	Overall total adverse event rate	Other
Bezer et al. 2004(85)	Unspecified instrumentation with autologous bone graft – traditional harvest, Posterolateral approach (PLF)	59	2 years	59	NR	NR	12 (20%)	Major: 4 (7%) Minor: 8 (14%)
	Unspecified instrumentation with autologous bone graft – interfascial harvest, PLF	58		58	NR	NR	5 (8.6%)	Major: 1 (2%) Minor: 4 (7%)
Scaduto et al. 2003(86)	Various instruments, mostly cylindrical threaded titanium non- tapered implants, Anterior (ALIF)	88	30 days	88	Intraoperative: 0 (0%) Major postoperative: 3 (3%) Minor postoperative: 10 (11%)	NR	Total (both procedures combined) 26 (22%)	NR
	Various instruments, mostly cylindrical threaded titanium non- tapered implants, Posterior lumbar interbody fusion (PLIF)	31		31	Intraoperative: 9 (29%) Major postoperative: 8 events in 6 patients (19%) Minor postoperative: 1 (3%)	NR		NR

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Overall early adverse event rate	Overall late adverse event rate	Overall total adverse event rate	Other
Burkus et al. 2002(87)	Fusion with iliac crest autograft, Anterior (ALIF), retroperitoneal or transperitoneal approach	136	2 years	136	NR	NR	NR	NR
	Fusion with rhBMP-2, Anterior (ALIF), retroperitoneal or transperitoneal approach	143		132	NR	NR	NR	NR
Christensen et al. 2002(88)	Posterolateral fusion (PLF) with titanium Cotrel-Dubousset instrumentation	73	Mean: 14 months (Range: 1 day to 48 months)	73	NR	NR	NR	NR
Brantigan et al. 2000(90)	Brantigan I/F Cage with Variable Pedicle Screw Placement System and Autologous bone graft, Posterior approach (PLIF)	221	2 years	221	NR	NR	Device-related: 0 major, 30 (14%) minor Not device-related: 23 (10%) major, 29 (13%) minor	NR
Slosar et al. 2000(91)	Instrumented circumferential fusion	141	Mean: 37.2 months (Range: 24-53 months)	133	NR	NR	27 (20%)	NR

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Overall early adverse event rate	Overall late adverse event rate	Overall total adverse event rate	Other
Thalgott et al. 2000(92)	Gasless endoscopic anterior lumbar interbody fusion utilizing the B.E.R.G approach	202	Perioperative (duration not reported)	202	NR	NR	NR	NR
Regan et al. 1999(93)	BAK Cage, Anterior (ALIF), Open surgery, with Retroperitoneal Approach, Single-level	305	Postoperative	305	NR	NR	NR	Major: 6 (2%) Minor: 43 (14%)
	BAK Cage, Laparascopic surgery, with transperitoneal Approach, Single-level	240		215	NR	NR	NR	Major: 0 Minor: 41 (19.1%)
Greenough et al. 1998(94)	Pedicle screw fixation using variable screw plate, Posterolateral approach (PLF)	135	12-36 months	135	NR	NR	NR	NR
Kuslich et al. 1998(95); Kuslich et al. 2000(96)	Bagby and Kuslich (BAK) interbody fusion using the anterior retroperitoneal approach (ALIF) or the posterior laminotomy (PLIF) approach	947	2 years	947	Intraoperative: 78 (8.2%)	Postoperative: 90 (9.5%) (Cannot separate early from late)	NR	Total major complications (2.0%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Overall early adverse event rate	Overall late adverse event rate	Overall total adverse event rate	Other
Malter et al. 1998(97)	Not specified	207 patients with herniated discs	Cumulative to 6 years	207	NR	NR	NR	NR
Penta and Fraser 1997(98)	Fusion with iliac crest autograft, Anterior approach (ALIF)	125	At least 10 years	103	NR	NR	10% of patients	NR
Ray et al. 1997(99)	Ray Titanium Cage, Posterior Approach (PLIF)	236	48 months	211	NR	NR	NR	NR
Thomsen et al. 1997(100); Christensen et al. 2002(101)†	Posterolateral fusion (PLF) with Cotrel-Dubousset instrumentation and autologous bone implant	64	5 years	64	NR	NR	NR	NR
Christensen et al. 1996(102)	Anterior (ALIF)	132	5-13 years	120	NR	NR	NR	NR
Hall et al. 1996(103)	Isola Spinal Implant System (staged anterior and posterior as well as posterior approach alone, but methods not well-described)	120	Operative	120	Operative: 14 events in 12 (10%) patients	Postoperative: 37 events in 28 (23%)patients	NR	NR
Greenough et al. 1994(104)	Anterior (ALIF)	151	Minimum 2 years, Range: 24-82 months	136	NR	NR	30 events in 26 patients	NR

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Overall early adverse event rate	Overall late adverse event rate	Overall total adverse event rate	Other
Gill and Blumenthal 1993(105)	Wiltse Pedicle Screw Fixation System with autograft, Posterior (PLIF), lateral implantation from unilateral approach with lateral/bilateral fusion	238	At least 2 years	238	NR	NR	NR	NR
Markwalder and Battaglia 1993(106)	Fusion by various techniques, including Magerl translaminar screw fixation technique and Louis plate fixation method	171	Mean: 23.8 months	163	NR	NR	NR	NR

† A follow-up of Thomsen et al. 1997(100)

NR Not reported.

Table 18. Individual Operative and Postoperative Adverse Events (RCTs Addressing Key Questions 1 and 2)

Study	Surgical procedure (type of fusion) and control intervention	Number of Patients Enrolled	Duration of Follow- up for early events	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Brox et al. 2006(31)	Posterolateral fusion (PLF) with pedicle screws + autologous bone graft	29 (23 received surgery)	1 month?	23	NR	Wound infection 2 (9%)	NR	NR	NR	NR
Fairbank et al. 2005(77)	Spinal fusion (unspecified)	176 (139 received surgery)	Intraoperative	139	Excessive bleeding: 3 (2%) Vascular injury: 1 (0.7%) Hemorrhage 1 (0.7%)	NR	NR	Vascular injury: 1 (0.7%) Loss of swab 1 (0.7%) Peritoneal tear 2 (1%)	Broken drain 1 (0.7%) Implant problems 5 (4%)	Dural tear 5 (4%) Bone fracture 1 (0.7%) Loss of purchase of fixation: 3 (2%)
Brox et al. 2003(78)	PLF with pedicle screws + autologous bone graft + physical therapy	37 (33 received surgery)	Perioperative	33	Bleeding: 2 (6%) Venous thrombosis: 1 (3%)	Wound infection: 2 (6%)	NR	NR	NR	Dural tear 1 (3%)

Study	Surgical procedure (type of fusion) and control intervention	Number of Patients Enrolled	Duration of Follow- up for early events	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Fritzell et al. 2001(79); Fritzell et al. 2002(80)	PLF with or without pedicle screws, or circumferential (PLIF or ALIF)	222 (204 received surgery)	Within first 2 weeks	211	Major complications:					
			Within first 2 weeks	211	Major bleedings during surgery 2 (1%) Thrombosis + pulmonary embolus 1 (0.5%) Thrombosis 1 (0.5%) Heart failure + GI bleeding 1 (0.5%)	Deep wound infections 3 (1.4%) (2 implant related)	New sensation of nerve root pain, no re-operation 6 (2.8%)	NR	Nerve root hit by pedicle screw, re-operated 3 (1.4%)	Aspiration-sepsis-ARDS 1 (0.5%) Pulmonary edema 1 (0.5%)
Fritzell et al. 2001(79); Fritzell et al. 2002(80)	PLF with or without pedicle screws, or circumferential (PLIF or ALIF)	222 (204 received surgery)	Within first 2 weeks	211	Minor complications:					
			Within first 2 weeks	211	Gastro-intestinal bleeding 3 (1.4%) Reoperation: hematoma at donor site 2 (1.0%)	Reopera-tion: Superficial wound infection 2 (1%)	Sympathetic cord damage with symptoms 2 (1%) Pain in arm after surgery 1 (0.5%)	Donor site pain 9 (4%) (1 day to 2 years)	Laterally placed screw 2 (1.0%)	Skin problem during surgery 2 (1%) Dural tear 1 (0.5%) Wing scapula after surgery 1 (0.5%)

NR Not reported.

Table 19. Reoperation and Late Events (RCTs Addressing Key Questions 1 and 2)

Study	Surgical procedure (type of fusion) and control intervention	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (total)	Reoperation (specific causes)	Other late complication
Fairbank et al. 2005(77)	Spinal fusion (unspecified)	176 (139 received surgery)	1 day to 24 months	139	11 (7.9%)	NR	NR
Fritzell et al. 2001(79), Fritzell et al. 2002(80)	PLF with or without pedicle screws, or circumferential (PLIF or ALIF)	222 (204 received surgery)	1 day to 24 months	211	16 (7.6%)	Major: Nerve root hit by pedicle screw (early) 3 (1.4%) Deep wound infections (early) 3 (2.4%) Deep wound infections (late) 2 (1.0%) Patient operated on wrong level (reoperated late) 1 (0.5%) Pseudoarthroses (late) 2 (1.0%)	Donor site pain 9 (4%)
						Minor: Reoperation: hematoma at donor site (early) 2 (1%) Reoperation: Superficial wound infection (early) 2 (1%) Reoperation: late deep wound infection 2 (0.9%) Reoperation: pseudarthrosis 2 (0.9%)	NR

NR Not reported.

Table 20. Individual Operative and Postoperative Adverse Events (Other Studies Addressing Key Question 2)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Sasso et al. 2005(84)	Fusion with autograft, Anterior lumbar interbody fusion (ALIF), retroperitoneal or transperitoneal approach	208	Length of hospital stay	202	NR	NR	NR	Iliac crest bone graft donor site pain 200 (99%)	NR	NR
Bezer et al. 2004(85)	Unspecified instrumentation with autologous bone graft – traditional harvest, Posterolateral approach (PLF)	59	< 6 months	59	Major complications					
					0	0	0	0	0	0
					Minor complications					
					Serous hematoma 2 (4%)	NR	Temporary sensory loss 4 (7%)	Pain over the donor site lasting more than one month 2 (4%)	NR	NR

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Bezer et al. 2004(85)	Unspecified instrumentation with autologous bone graft – interfascial harvest, Posterolateral approach (PLF)	58	< 6 months	58	Major complications					
					0	0	0	Sacroiliac penetration 1 (2%)	0	0
					Minor complications					
					Serous hematoma 1 (2%)	NR	NR	Temporary sensory loss 2 (3%) Pain over the donor site lasting more than one month 1 (2%)	NR	NR

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Sasso et al. 2004(375)	INTER FIX device with autogenous iliac crest-derived bone, anterior approach (ALIF), single-level interbody fusion	77	2 years (early and late events not noted separately)	77	Vascular intraoperative 9 (11.5%) Vascular postoperative 1 (1.3%)	0	14 (17.9%)	Sacroiliac pain 7 (8.9%) Retrograde ejaculation 1 (1.3%)	0	Incisional 5 (6.4%) Spinal event 11 (14.1%) Urological 2 (2.6%) Gastro-intestinal 5 (6.4%) Other 7 (9.0%) Back pain 4 (5.1%) Respiratory 6 (7.7%) Leg pain 2 (2.6%) Trauma 1 (1.3%) Peritoneal 3 (3.8%) Bone fracture 2 (2.6%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
	Femoral ring allograft with autogenous iliac crest-derived bone, anterior approach (ALIF), single-level interbody fusion		2 years	62	Vascular intraoperative 2 (3.2%) Vascular postoperative 2 (3.2%)	0	Neurological 16 (25.8%)	Sacroiliac pain 3 (4.8%) Graft site pain 1 (1.6%) Retrograde ejaculation 0 (0%)	Implant breakage 5 (8.1%) Implant displacement/ loosening 6 (9.7%)	Death 1 (1.6%) Back pain 14 (22.5%) Incisional 8 (12.9%) Spinal event 4 (6.5%) Urological 2 (3.2%) Gastro-intestinal 5 (8.1%) Respiratory 1 (1.6%) Meningitis 1 (1.6%) Other pain 6 (9.7%) Other 5 (8.1%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Scaduto et al. 2003(86)	Various instruments, mostly cylindrical threaded titanium non- tapered implants, Anterior approach (ALIF)	88	30 days	88	Deep vein thrombosis 2 (3%)	NR	NR	NR	Radiculopathy from bone spike 1 (1%)	Ileus 5 (6%) Readmission for pain, negative workup 2 (3%) Urinary retention 1 (1%) Atelectasis 1 (1%) Transient brachial plexus palsy 1 (1%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Scaduto et al. 2003(86)	Various instruments, mostly cylindrical threaded titanium non- tapered implants, Posterior lumbar interbody fusion (PLIF)	31	30 days	31	Epidural hematoma 1 (3%)	NR	NR	NR	Reoperation: radiculopathy from cage placement 1 (3%)	Durotomy 9 (29%) New weakness, negative workup 1 (3%) Persistent cerebrospinal fluid leak or pseudomeningocele 5 (16%) Meningitis 1 (3%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Burkus et al. 2002(87)	Fusion with iliac crest autograft, Anterior (ALIF), retroperitoneal or transperitoneal approach	136	Intraoperative or perioperative	136	Intraoperative vascular events 5 (4%) Deep vein thrombosis 2 (2%)	NR	NR	Autograft: injuries to lateral femoral cutaneous nerve 3 (2%) Autograft related: Avulsion fracture of anterior superior iliac crest 2 (2%) Autograft-related: infection 1 (0.8%) Autograft-related: hematoma 1 (0.8%)	NR	NR
Burkus et al. 2002(87)	Fusion with rhBMP-2, Anterior (ALIF), retroperitoneal or transperitoneal approach	143	Intraoperative	132	Intraoperative vascular events 6 (5%)	NR	NR	0	NR	NR

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Burkus et al. 2002(87)	Fusion with rhBMP-2 or autograft, Anterior (ALIF), retroperitoneal or transperitoneal approach	146 males	2 years	146	NR	NR	NR	Retrograde ejaculation 6 (4%), 4 (3%) permanent	NR	NR
Christensen et al. 2002(88)	Posterolateral fusion (PLF) with titanium Cotrel-Dubousset instrumentation	73	Perioperative	73	Hematoma 2 (2.8%)	Superficial infection 1 (1.4%)	NR	NR	Nerve root injury due to screw misplacement 1 (1.4%)	Dura lesion 1 (1.4%) Urinary tract infection 3 (4.1%)
	Circumferential fusion (anterior lumbar interbody fusion (ALIF) with Brantigan cage, using a retroperitoneal approach to the lumbar discs + posterolateral fusion)	75	Perioperative	75	Vascular injury 4 (5.3%) Hematoma 1 (1.3%)	Deep infection 1 (1.3%)	NR	NR	Nerve root injury due to screw misplacement 3 (4%)	Urinary tract infection 4 (5.3%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Brantigan et al. 2000(90)	Brantigan I/F Cage with Variable Pedicle Screw Placement System and Autologous bone graft, Posterior approach (PLIF)	221	Intraoperative to several days after surgery	221	Major complications					
					Deep venous thrombosis 2 (0.9%)	NR	Increasing motor deficit after surgery 3 (1%), 1 permanent (0.5%) Reflex sympathetic dystrophy 3 (1.4%)	NR	NR	Death due to intraoperative complications (massive bleeding, myocardial infarction) 2 (0.9%) Myocardial infarction 1 (0.5%)
					Minor complications					
					NR	NR	NR	NR	NR	NR

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Slosar et al. 2000(91)	Instrumented circumferential fusion	141	Operative or perioperative	141	Death due to massive pulmonary embolism 1 (0.7%) Iliac artery thrombosis, treated in recovery without permanent effect 1 (0.7%) Deep vein thrombosis 2 (2%)	Superficial posterior infection 4 (3%) Deep posterior infection 5 (4%)	Transient motor weakness 9 (7%)	Retrograde ejaculation 1 (0.7%) Anterior graft extrusion 3 (2%)	NR	NR
Thalgott et al. 2000(92)	Gasless endoscopic anterior lumbar interbody fusion utilizing the B.E.R.G approach	202	Operative or perioperative (duration not reported)	202	Deep vein thrombosis 1 (0.7%)	NR	Foot drop 2 (1%)	Conversion to open approach 34 (17%) Retrograde ejaculation 1 (0.7%)	NR	NR

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Regan et al. 1999(93)	BAK Cage, Anterior (ALIF), Open surgery, with Retroperitoneal Approach, Single-level	305	Perioperative or postoperative (duration not reported)	305	Great vessel (aorta or vena cava) damage 2 (0.7%) Pulmonary embolism 1 (0.3%) Hematoma/ seroma 3 (1%) Great vessel damage (aorta or vena cava) 2 (0.7%) Pulmonary embolism 1 (0.3%)	Infection 6 (2%)	Leg pain 2 (0.7%)	Retrograde ejaculation 7 (2%)	Implant Migration 4 (1%)	Ileus 10 (3%) Atelectasis/ pneumonia 2 (0.7%) Urologic 3 (1%) Wound dehiscence/ incisional hernia 3 (1%) Other, minor 3 (1%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
	BAK Cage, Laparascopic surgery, with transperitoneal Approach, Single-level	240	Perioperative or postoperative (duration not reported)	215	Thrombosis/ thrombo-phlebitis 1 (0.5%)	Infection 3 (1.4%)	Leg pain 1 (0.5%)	Retrograde ejaculation 11 (5.1%)	Implant migration 1 (0.5%)	Conversion to open procedure: 25 (10.4%) Atelectasis/ pneumonia 2 (0.9%) Disc herniation 6 (2.8%) Spondylosis (fractures) 3 (1.4%) Other 3 (1.4%)
Greenough et al. 1998(94)	Pedicle screw fixation using variable screw plate, Posterolateral approach (PLF)	135	Perioperative	135	Deep vein thrombosis 1 (1%) Pulmonary embolism without evident deep venous thrombosis 1 (1%)	Superficial wound infection 1 (1%)	Signs of nerve root compression 3 (2%)	NR	Pedicle fracture 1 (1%)	Urinary tract infection secondary to catheteriza-tion 13 (10%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Kuslich et al. 1998(95); Kuslich et al. 2000(96)	Bagby and Kuslich (BAK) interbody fusion using the anterior retroperitoneal approach (ALIF) or the posterior laminotomy (PLIF) approach	947	Intraoperative or postoperative to 2 months	947	Vessel damage, bleeding 11 (1.2%) Hematoma, seroma 12 (1.3%) Phlebitis, pulmonary embolism 5 (0.5%) Other (GI bleed, anemia) 11 (1.2%)	Superficial infection 23 (2.4%)	Neurologic 26 (2.7%)	Retrograde ejaculation 11 (1.2%) Fatigue fracture S1 1 (0.1%)	Implant migration with reoperation 11 (1.2%)	Dural complications 36 (3.8%) Atelectasis, pneumonia 11 (1.2%) Urologic 9 (1.0%) Ileus 14 (1.5%)
Penta and Fraser 1997(98)	Fusion with iliac crest autograft, Anterior approach (ALIF)	125	Duration not reported	103	Pulmonary embolus 4 (3.9%), (including 1 deep vein thrombosis)	Superficial wound infection 2 (2%) Chest infection 1 (1%)		Prolonged donor site pain 2 (2%)		Urinary retention 1 (1%) Superficial wound dehiscence 1 (1%) Urinary tract infection 2 (2%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Ray et al. 1997(99)	Ray Titanium Cage, Posterior Approach (PLIF)	236	Up to 6 weeks	211	NR	Superficial wound infection 5 (2%) Deep wound infection 2 (1%)	Foot weakness 24 (11%) (2 did not resolve within 6 weeks)	NR	NR	Dural tear 13 (6%)
Thomsen et al. 1997(100) and Christensen et al. 2002(101)†	Posterolateral fusion (PLF) with or without Cotrel-Dubousset instrumentation and autologous bone implant	110	Perioperative to discharge	110	Minor complications (no major complications)					
Thomsen et al. 1997(100)	Posterolateral fusion (PLF) with Cotrel-Dubousset instrumentation and autologous bone implant	64	Perioperative (no perioperative complications in non-instrumented group)	64	0	0	0	0	0	Dural tear 1 (2%)
	With or without Cotrel Dubousset instrumentation	110	Postoperative to discharge	110	NR	Superficial wound infection 2 (1.8%)	NR	NR	NR	Urinary tract infection 2 (1.8%) Stress ulcus 2 (1.8%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Christensen et al. 1996(102)	Anterior (ALIF)	132	Postoperative (patients followed for 5-13 years)	120	Died from pulmonary embolism 2 (2%)	NR	NR	Retrograde ejaculation, permanent "at least" 8 men (6%)	NR	NR
Hall et al. 1996(103)	Isola Spinal Implant System (staged anterior and posterior as well as posterior approach alone, but methods not well-described)	120	Operative	120	Excessive bleeding 3 (3%) Aortic thrombosis 1 (1%)	NR	Paraplegia 1 (1%)	NR	Split or fractured pedicle during screw insertion 2 (2%)	Dural leak 7 (6%)
			Postoperative	120	NR	Superficial 3 (2.5%) Deep: 4 (3.3%) (3 of 4 occurred late)	Neurologic 6 (5%) Reflex sympathetic dystrophies 2 (1.7%)	NR	Nerve root irritation by malpositioned pedicle screws 4 (3.3%)	Adynamic ilius 4 (3.3%) Cardiac 4 (3.3%) Pulmonary 3 (2.5%) Other 7 (5.8%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Greenough et al. 1994(104)	Anterior interbody fusion (ALIF), retroperitoneal approach	151	Minimum 2 years, Range: 24-82 months	136	Deep vein thrombosis without pulmonary embolism 2 (1%) Deep vein thrombosis with pulmonary embolism 4 (3%)	NR	NR	Persistent symptoms from donor site 1 (0.7%)	Knodt rod fracture 2 (1%)	Urinary tract infection secondary to catheteriza-tion 13 (9.6%) Diskitis 1 (0.7%)
Gill and Blumenthal 1993(105)	Wiltse Pedicle Screw Fixation System with autograft, Posterior (PLIF) lateral implantation from unilateral approach with lateral/bilateral fusion	238	Time of latest complication not reported. Most appear early.	238	Thrombo-phlebitis 1 (0.4%)	Superficial wound infection 13 (5%) Deep wound infection 4 (2%)	Drop foot 1 (0.4%) New onset leg pain 6 (3%) Severe postoperative sciatica requiring surgical exploration 1 (0.4%)	Retropulsion of graft 2 (1%)	Fractured screws, evident on radiograph by asymptomatic 3 (1%) Loosening of screws 26 (10%)	Dural leak 1 (0.4%) Acute urinary retention 1 (0.4%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Bleeding or vascular complication	Infection	Neurological complication	Associated with surgical approach or graft harvest	Device-related	Other
Markwalder and Battaglia 1993(106)	Fusion by various techniques, including Magerl translaminar screw fixation technique and Louis plate fixation method	171	Early	171	NR	Infection after reoperation 1 (0.6%)	Foot drop 3 (1.8%)	NR	Correction of malpositioned pedicle screw 2 (1.2%)	NR

† A follow-up of Thomsen et al. 1997(100)

NR Not reported.

Table 21. Reoperation and Late Events (Other Studies Addressing Key Question 2)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (any cause)	Reoperation (specific causes)	Other late or permanent events
Martin et al. 2007(81)	Not specified	462 patients with herniated discs	Cumulative to 2 years	462	58 (13%)	NR	NR
			Cumulative to 4 years	462	77 (17%)	NR	NR
			Cumulative to 6 years	462	89 (19%)	NR	NR
			Cumulative to 8 years	462	101 (22%)	NR	NR
			Cumulative to 10 years	462	107 (23%)	NR	NR
			Cumulative to 11 years	462	115 (25%)	NR	NR
Martin et al. 2007(81)	Not specified	515 patients with degenerative disc disease	Cumulative to 2 years	515	51 (10%)	NR	NR
			Cumulative to 4 years	515	66 (13%)	NR	NR
			Cumulative to 6 years	515	76 (15%)	NR	NR
			Cumulative to 8 years	515	83 (16%)	NR	NR
			Cumulative to 10 years	515	92 (18%)	NR	NR
			Cumulative to 11 years	515	105 (23%)	NR	NR

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (any cause)	Reoperation (specific causes)	Other late or permanent events
Burkus et al. 2005(82); Burkus et al. 2006(83)	rhBMP-2and MD-II threaded cortical, bone dowel. Anterior lumbar interbody fusion (ALIF), open, with transperitoneal or retroperitoneal approach	79	"More than 24 months"	79	2 (2.5%)	Supplemental posterior fixation 2 (2.5%)	NR
Burkus et al. 2005(82); Burkus et al. 2006(83)	Autologous bone graft fusion, Anterior (ALIF), open	52	24 months	52	Reoperation:	To remove residual disc material 1 (2%) Supplemental posterior fixation 8 (15.4%)	NR
Sasso et al. 2005(84)	Fusion with autograft, Anterior (ALIF), retroperitoneal or transperitoneal approach	208	6 weeks postoperative	199	NR	NR	Iliac crest bone graft donor site pain 165 (83%)
			3 months postoperative	199	NR	NR	114 (57%)
			6 months postoperative	192	NR	NR	79 (41%)
			1 year postoperative	168	NR	NR	55 (33%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (any cause)	Reoperation (specific causes)	Other late or permanent events
Bezer et al. 2004(85)	Unspecified instrumentation with autologous bone graft – traditional harvest, Posterolateral approach (PLF)	59	2 years	59	0	0	Pain over the donor site lasting more than 1 year 3 (5%) Residual numbness over donor site 1 (2%)
	Unspecified instrumentation with autologous bone graft – interfascial harvest, Posterolateral approach (PLF)	58	2 years	58	0	0	NR
Burkus et al. 2002(87)	Fusion with iliac crest autograft, Anterior (ALIF), retroperitoneal or transperitoneal approach	136	2 years	124	NR	NR	Autograft-related: Pain at donor site 24 months after surgery 44 (32%)
	Fusion with rhBMP-2, Anterior (ALIF), retroperitoneal or transperitoneal approach	143	2 years	132	NR	NR	NR
Burkus et al. 2002(87)	Fusion with rhBMP-2 or autograft, Anterior (ALIF), retroperitoneal or transperitoneal approach	146 males	2 years	146	NR	NR	Permanent retrograde ejaculation 4 (3%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (any cause)	Reoperation (specific causes)	Other late or permanent events
Christensen et al. 2002(88)	Posterolateral fusion (PLF) with titanium Cotrel-Dubousset instrumentation	73	Mean: 14 months (1 day to 48 months)	73	16 (22%)	NR	NR
	Circumferential fusion (anterior lumbar interbody fusion (ALIF) with Brantigan cage, using a retroperitoneal approach to the lumbar discs + posterolateral fusion)	75	Mean: 14 months (1 day to 48 months)	75	5 (7%)	NR	NR
McAfee et al. 2002(89)	Anterior BAK instrumentation and fusion (ALIF), complete discectomy through open retroperitoneal approach	50	2 years	50	0%	0%	NR
	Anterior BAK instrumentation and fusion (ALIF), partial discectomy through miniopen retroperitoneal approach, miniopen laporotomy, or laparascopically	50	2 years	50	8 (16%)	Pseudarthroses 7 (14%) Early postop cage displacement 1 (2%)	NR

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (any cause)	Reoperation (specific causes)	Other late or permanent events
Brantigan et al. 2000(90)	Brantigan I/F Cage with Variable Pedicle Screw Placement System and Autologous bone graft, Posterior approach (PLIF)	221			102 (46.1%)	To treat deep infections 8 (3.6%) Broken screw requiring hardware removal 6 (3%) Elective removal of pedicle screws and VSP plates 78 (35%) Repair of dural tears 6 (3%) Removal of broken drains 3 (1.4%) Subsequent surgical treatment of new disc level 7 (3%)	Death due to apparently unrelated cause 2 (0.9%) Death due to suicide 2 (0.9) Permanent motor deficit 1 (0.5%) Broken screws 13 (5.9%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (any cause)	Reoperation (specific causes)	Other late or permanent events
Regan et al. 1999(93)	BAK Cage, Anterior (ALIF), Open surgery, with Retroperitoneal Approach, Single-level	305	6 months	305	7 (2%)	Implant migration requiring reoperation 3 (1%) Revision of implant 3 (1%) Removal of implant 2 (0.7%) Bone graft augmentation 1 (0.3%) Additional stabilization 1 (0.3%)	

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (any cause)	Reoperation (specific causes)	Other late or permanent events
Greenough et al. 1998(94)	Pedicle screw fixation using variable screw plate, posterolateral approach (PLF)	135	To 36 months	127	23 (18%)	Nonunion 4 (3.1%) Fusion at different level 2 (1.6%) Surgical exploration 1 (0.8%) Lateral decompression 3 (2.4%) Plates removed for persistent symptoms 11 (8.7%) Further surgery at other centers 2 (1.6%) Rhizolysis of adjacent facet joint 1 (0.8%)	Permanent weakness of knee extension due to screw misplacement 1 (0.8%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (any cause)	Reoperation (specific causes)	Other late or permanent events
Kuslich et al. 1998(95); FDA PMA P950002(377)	Bagby and Kuslich (BAK) interbody fusion using the anterior retroperitoneal approach (ALIF) or the posterior laminotomy (PLIF) approach	947	To 24 months	947	Total: 69 (7.3%) Total device-related: 42 (4.4%) Device-related within first 100 days: 14 (1.5%)	Implant migration with reoperation 11 (1.2%) Additional stabilization 26 (2.7%) Additional level fusion 9 (1.0%) Leg pain 9 (1.0%) Dura-related 6 (0.6%) Implant reposition 4 (0.4%) Other decompression 2 (0.2%) Anterior ligament penetration 1 (0.1%) Fractured sacrum 1 (0.1%)	NR

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (any cause)	Reoperation (specific causes)	Other late or permanent events
Malter et al. 1998(97)	Not specified	207 patients with herniated discs	Cumulative to 2 years	207	19 (9.2%)	NR	NR
			Cumulative to 4 years	207	28 (13.5%)	NR	NR
			Cumulative to 6 years	207	34 (16.4%)	NR	NR
		265 patients with degenerative disc disease	Cumulative to 2 years	265	35 (13.2%)	NR	NR
			Cumulative to 4 years	265	48 (18.1%)	NR	NR
			Cumulative to 6 years	265	66 (24.9%)	NR	NR
Ray et al. 1997(99)	Ray Titanium Cage, Posterior Approach (PLIF)	236	48 months	211	3 (1%) (within first 2 weeks)	Adjustment of anterior-posterior depth of penetration 3 (1%)	Long-term foot weakness 2 (1.0%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (any cause)	Reoperation (specific causes)	Other late or permanent events
Thomsen et al. 1997(100); Christensen et al. 2002(101)†	Posterolateral fusion (PLF) with Cotrel-Dubousset instrumentation and autologous bone implant	64	5 years	64	Cumulative reoperation rate to 5 years 16 (3%)	Deep wound infections requiring reoperation 2 (3%) Removal of implant, with or without additional surgical procedure at time of surgery 14 (2%) Waiting list 3 (0.5%) Fusion 1 (2%)	NR
Christensen et al. 2002(101)†	Posterolateral fusion (PLF) with no instrumentation and autologous bone implant	66	5 years	62	Cumulative reoperation rate to 5 years 9 (15%)	Fusion 7 (11%) Decompression 1 (2%) Waiting list 1 (2%)	NR
Christensen et al. 1996(102)	Anterior (ALIF)	132	5-13 years	120	NR	NR	Retrograde ejaculation, permanent “at least” 8 men (6%)

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (any cause)	Reoperation (specific causes)	Other late or permanent events
Hall et al. 1996(103)	Isola Spinal Implant System (staged anterior and posterior as well as posterior approach alone, but methods not well-described)	120	Operative	120	35 events in 31 (26%)patients	Device-related for nerve root impingement by screws 4 (4%) Loose or broken implants, asymptomatic 4 (4%) Local pain; relieved by implant removal 2 (2%) Elective removal of implants for possible implant-related pain 10 (10%) Infection 4 (4%) Pseudarthrosis 5 (5%) Neurological problem deemed not related to implant 2 (2%) Wound infection 2 (2%) Wound hematoma 1 (1%) Embolectomy 1 (1%)	NR

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (any cause)	Reoperation (specific causes)	Other late or permanent events
Greenough et al. 1994(104)	Anterior (ALIF)	151	Minimum 2 years, Range: 24-82 months	136	24 (17.6%)	Nonunion 12 (8.8%) Removal of distraction rods due to breakage or persistent symptoms 5 (4%) Further operation at different level 6 (4.4%) Further surgery for nonunion at other center 1 (0.7%)	NR
Gill and Blumenthal 1993(105)	Wiltse Pedicle Screw Fixation System with autograft, Posterior (PLIF) lateral implantation from unilateral approach with lateral/bilateral fusion	238	At least 2 years	238	72 (30%)	Severe postoperative sciatica requiring surgical exploration 1 (0.4%) Hardware removal 72 (30%)	Infection after reoperation 1

Study	Surgical procedure (type of fusion)	Number of Patients Enrolled	Duration of Follow- up	Number of Patients at Follow up	Reoperation (any cause)	Reoperation (specific causes)	Other late or permanent events
Markwalder and Battaglia 1993(106)	Fusion by various techniques, including Magerl translaminar screw fixation technique and Louis plate fixation method	171	Mean: 23.8 months	163	29	Removal of osteosynthetic material with loosening implants causing pain 18 Correction of malpositioned pedicle screw 2 Fusion of adjacent motion segments because of reactive overload 9 Pseudoarthrosis formation 3 (overlaps with other surgeries)	

† A follow-up of Thomsen et al. 1997(100)

NR Not reported.

Table 22. Data for Key Question 3 – Patient Characteristics Associated with Positive or Negative Outcomes of Lumbar Fusion

Study	Surgical procedure/ nonsurgical therapy	Number of Patients Enrolled	Duration of Followup	Number of Patients at Follow up	Factors correlating with outcomes (multiple logistic regression)	Regression coefficient (β)	Odds ratio[a] (95% CI)
					Outcome: Change of disability (ODI)		
Hagg et al. 2003(107) (part of an earlier study by Fritzell et al. 2001)	PLF with or without pedicle screws, or circumferential (PLIF or ALIF)	222 (204 received surgery)	2 years	201	Neurotic personality (KSP)	-0.096	0.91 (0.87 to 0.95)
	Physical therapy + other non-operative therapies	72 (65 received therapy)	2 years	63	None identified		
					Outcome: Patient global assessment (improved/not improved)		
	PLF with or without pedicle screws, or circumferential (PLIF or ALIF)	222 (204 received surgery)	2 years	201	Neurotic personality (KSP) Disc height <50%	-0.052 0.787	0.95 (0.92 to 0.98) 2.20 (1.14 to 4.24)
	Physical therapy + other non-operative therapies	72 (65 received therapy)	2 years	63	Depressive symptoms (ZDS)	0.074	1.08 (1.02 to 1.14)
					Outcome: Work status at followup (working/not working)		
	PLF with or without pedicle screws, or circumferential (PLIF or ALIF)	222 (204 received surgery)	2 years	201	Age (years) Current sick leave (months)	-0.05 -0.036	0.95 (0.91 to 0.99) 0.96 (0.94 to 0.99)
	Physical therapy + other non-operative therapies	72 (65 received therapy)	2 years	63	Current sick leave (months)	-0.505	0.96 (0.92 to 0.99)

[a] An odds ratio <1 means that a factor (e.g., neurotic personality) was associated with a poor outcome; an odds ratio >1 means that a factor was associated with improvement in an outcome.

KSP Karolinska Scales of Personality.
ODI Oswestry Disability Index.
ZDS Zung Depression Scale.

Appendix E. Evidence Tables for Discography (Key Questions 4-6)

Table 23. General Characteristics of Studies Addressing KQ4, KQ5, and KQ6

Study	Dates of enrollment	Location	Number of patients	Number of disc levels studies with discography
Studies addressing Key Question 4 (reliability)				
Agorastides (2002)(33)	5/1995 to 10/1997	Queen's Medical Centre, Nottingham, United Kingdom	72	133
Milette (1999)(34)	Not reported	Hospital Saint-Luc, Montreal, Quebec, Canada	45	132
Studies addressing Key Question 5 (prediction of fusion results)				
Willems (2007)(36)	4/1990 to 10/1999	Nijmegen, The Netherlands	82	164[a]
Gill (1992)(37)	Not reported	Southwest Orthopaedics Institute, Dallas, Texas	53	53
Colhoun (1988)(38)	Not reported	The Robert Jones and Agnes Hunt Orthopaedic Hospital, Oswestry, England	195	585[b]
Studies addressing Key Question 6 impact on fusion results)				
Madan (2002)(39)	1997-1999	Southampton University Hospital, New York, NY	73	96[c]

[a] Willems (2007) reported that each patient "generally" received discography on the two discs adjacent to the suspect disc, and the number 164 corresponds to exactly two discs per patient with reported fusion data.

[b] Colhoun (1988) reported that all patients had "at least" two levels studied, "the majority" had three levels, and "occasionally" at four levels. The number 585 corresponds to three levels in all 195 patients.

[c] Madan only performed discography in 32 of 73 patients; these 32 patients all received discography at three lumbar levels, which corresponds to a total of 96 discs.

Table 24. Inclusion and Exclusion Criteria for Studies Addressing KQ4, KQ5, and KQ6

Study	Inclusion Criteria	Exclusion Criteria
Studies addressing Key Question 4 (reliability)		
Agorastides (2002)(33)	Chronic back pain being considered for spinal fusion and underwent discography	None reported
Milette (1999)(34)	Underwent both lumbar discography and lumbar magnetic resonance imaging	Previous lumbar disc surgery, spondylolysis, spondylolisthesis, delays of more than 12 months between discography and magnetic resonance imaging, incomplete imaging files.
Studies addressing Key Question 5 (prediction of fusion results)		
Willems (2007)(36)	Suffered from incapacitating low back pain for at least one year, and the indication for lumbar fusion remained uncertain after routine diagnostic tests,	Objective neurologic deficit
Gill (1992)(37)	History of trauma, as described by Crock	None reported
Colhoun (1988)(38)	Persistent low back pain, failed to respond to conservative treatment, two years of follow-up	Previous back surgery
Studies addressing Key Question 6 (impact on fusion results)		
Madan (2002)(39)	Received lumbar fusion surgery between 1997 and 1999. Between 1/1998 and 1/1999, all patients had to undergo screening with provocative discography, and these were indicated for surgery only if pain was partially or wholly reproduced during discography. The other group of patients had not undergone discography and received the same form of surgery.	None reported

Table 25. Patient Characteristics in Studies Addressing KQ4, KQ5, and KQ6

Study	N	Age (range)	% male	Conditions	Duration of condition	Previous spinal surgery	Other previous treatment
Studies addressing Key Question 4 (reliability)							
Agorastides (2002)(33)	72	41 (23-69)	69% (49/71) [a]	Low back pain being considered for spinal fusion	"Chronic"; mean duration not reported	0%	Not reported
Milette (1999)(34)	45	38 (22-64)	69% (31/45)	Low back pain without spondylolisthesis or spondlyosis	Average 5 years	0%	Not reported
Studies addressing Key Question 5 (prediction of fusion results)							
Willems (2007)(36)	82	40b	32% (63/197) [b]	"Incapacitating" low back pain, with no objective neurologic motor deficit, and being considered for lumbar fusion surgery	More than one year	67%	Not reported
Gill (1992)(37)	53	34 (21-50)	68% (36/53)	Low back pain with internal disc disruption	Average 0.9 years	Not reported	All patients had previous unsuccessful conservative treatment
Colhoun (1988)(38)	195	39 (17-70)	51% (99/195)	Low back pain	"Persistent"; mean duration not reported	0%	All patients had previous unsuccessful conservative treatment
Studies addressing Key Question 6 (impact on fusion results)							
Madan (2002)(39)	73	41 (15-68)	59% (43/73)	Low back pain	Not reported	Not reported	All patients had previous unsuccessful conservative treatment

[a] In the Agorastides study, the male-female distribution was reported based on 71 patients
[b] In the Willems study, the age and male-female distribution were reported on a large set of patients (N = 197) that included some patients who did not receive surgery.

Table 26. Discography Details in Studies Addressing Key Question KQ4, KQ5, and KQ6

Study	Injectors	Readers	CT?	Patient positioning	Control disc(s) examined by discography	Specific lumbar levels examined	How results were interpreted
Studies addressing Key Question 4 (reliability)							
Agorastides (2002)(33)	1	3 [a]	No	Prone	Where possible, one level below or above the degenerative disc was injected	L2-3: 6 discs; L3-4: 30 discs; L4-5: 58 discs; L5-S1: 39 discs	Adams classification(378)
Milette (1999)(34)	1	2	No	Prone	NR	L2-3: 7 discs; L3-4: 42 discs; L4-5: 42 discs; L5-S1: 41 discs	Original Dallas Discogram Description(13)
Studies addressing Key Question 5 (prediction of fusion results)							
Willems (2007)(36)	>1	NR	NR	Left lateral decubitus	All patients had at least one disc examined by discography that was adjacent to the suspected disc	NR	Adams classification(378)
Gill (1992)(37)	NR	NR	NR	NR	NR	L5-S1 in all patients	Discography results were categorized as type I, II or III. Type I was "indicating a small annular tear that did not extend to the periphery. Type II constituted "annular tear and contrast extension to the periphery". Type III constituted "annular tear to the periphery and beyond to the epidural space".

Study	Injectors	Readers	CT?	Patient positioning	Control disc(s) examined by discography	Specific lumbar levels examined	How results were interpreted
Colhoun (1988)(38)	NR	NR	NR	NR	Adjacent discs did receive discography in some patients, but this was not referred to as a control discography.	"Lumbar discography was attempted in all patients at L4-5 and L5-S1 levels, at L3-4 in the majority of patients and occasionally at L2-3"	Discography was interpreted according to two factors: whether it reproduced typical pain, and whether morphology was abnormal. The study did not report additional details.
Studies addressing Key Question 6 (impact on fusion results)							
Madan (2002)(39)	>1	NR	NR	NR	NR	The three lower lumbar levels in all patients, and additional levels in "some" patients (actual number not reported)	Patients who received discography received fusion only if "their discogenic pain was partially or wholly reproduced during discography".

[a] Agorastides was the only study that reported test-rest reliability; the time between discographies was three weeks. The study by Millete only reported inter-rater reliability.

NR Not reported.

Table 27. Lumbar Fusion Details in Studies Addressing Key Question KQ4, KQ5, and KQ6

Study	Procedure(s)	Additional Details	Length of follow-up after surgery
Studies addressing Key Question 4 (reliability)			
Agorastides (2002)(33)	NA	NA	NA
Milette (1999)(34)	NA	NA	NA
Studies addressing Key Question 5 (prediction of fusion results)			
Willems (2007)(36)	58 patients received instrumented posterolateral intertransverse process fusion, and 24 patients received ALIF	No other details provided	Mean follow-up 6.7 years (Range: 1.3 to 12)
Gill (1992)(37)	Modified Crock ALIF	The approach was anterior retroperitoneal in all patients. Allograft in 48 patients, autogenous iliac crest in 5 patients.	Mean follow-up 3 years (Range: 2.0 to 5.3 years)
Colhoun (1988)(38)	Some ALIF, some PLIF (numbers not reported)	No other details provided	Mean follow-up 3.6 years (Range: 2 to 10)
Studies addressing Key Question 6 (impact on fusion results)			
Madan (2002)(39)	Instrumented PLIF with posterolateral fusion	Midline subperiosteal approach. Autologous iliac creast cancellous bone was used. Instrumentation included pedicle screws and Isola rods.	Mean follow-up 2.6 years (Range: 2 to 4.2)

ALI Anterior lumbar interbody fusion.
NA Not applicable because the study did not report fusion outcomes; the study was included only for reliability analysis.
PLIF Posterior lumbar interbody fusion.

Table 28. Quality Assessment of Studies of the Reliability of Discography (Key Question 4)

Study	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Score (quality category)
Studies addressing Key Question 4 for test-retest reliability											
Agorastides (2002)(33)	No	Yes	NR	Yes	NR	NA	NA	Yes	Yes	Yes	7.5 (Moderate)
Studies addressing Key Question 4 for inter-rater reliability											
Agorastides (2002)(33)	No	Yes	NR	Yes	NR	NA	NA	Yes	Yes	NA	7.1 (Moderate)
Milette (1999)(34)	No	Yes	Yes	Yes	NR	NA	NA	Yes	Yes	NA	7.9 (Moderate)

NA Not applicable.
NR Not reported.

The quality items for Key Question 4 were:

1) Was the study prospective?
2) Were the patients enrolled consecutively?
3) Were the patient inclusion/exclusion criteria applied consistently to all patients?
4) Were data reported for at least 85% of enrolled patients?
5) Was the funding for this study derived from a source that does not have a financial interest in its results?
6) If two injections were performed on each patient, did patients receive the same instructions for pain reporting during the two discography examinations being compared?
7) If two injections were performed on each patient, did the same injector perform the two discography examinations being compared?
8) Was discography interpreted without knowledge of other discography results in this patient?
9) Was discography interpreted without knowledge of other clinical information about this patient?
10) TEST-RETEST RELIABILITY ONLY. Did the same person interpret the two discography results?

See Appendix C for details on how we scored the quality assessments.

Table 29. Outcome Data for Test-Retest Reliability of Discography (Key Question 4)

Study	Number of discs	Test-retest kappa (95% CI)[a]			Notes
		Rater 1	Rater 2	Rater 3	
Agorastides (2002)(33)	133	0.80 (95%CI: 0.71 to 0.89)	0.85 (95%CI: 0.77 to 0.93)	0.80 (95%CI: 0.7 to 0.9)	Of the 41 total disagreements (10% of 399 ratings), 39 were one-grade disagreements, 1 was a two-grade disagreement, and 1 was a three-grade disagreement.[b]

NOTE: These data apply to the test-retest reliability of interpreting a single discogram at different times. Each rater viewed a single discogram at two different times (three weeks apart), and asked to judge the Adams morphology grade (1, 2, 3, 4 or 5). This was repeated for each of 133 discograms.

[a] Kappa measures chance-corrected agreement. 0 represents chance, and 1 represents perfect agreement. The standard interpretation of kappa values is that Below 0.0 is Poor agreement; 0.00-0.20 is Slight agreement; 0.21-0.40 is Fair agreement; 0.41-0.60 is Moderate agreement; 0.61-0.80 is Substantial agreement; 0.81-1.00 is Almost Perfect agreement.(35) The 95% CI around each kappa was calculated by ECRI based on the reported standard errors.

[b] The Adams classificaton of discograms contains five levels:(378) Level 1 is “No signs of degeneration. Soft white amorphous nucleus”; Level 2 is “Mature disc with nucleus starting to coalescse into fibrous lumps”; Level 3 is “Degenerated disc with fissures and clefts in the nucleus and inner annulus”; Level 4 is “Degenerated disc with radial fissure leading to the outer edge of the annulus”; Level 5 is “Disc has a complete radial fissure hat allows injected fluid to escape. Can be in any state of degeneration”.

Table 30. Outcome Data for Inter-Rater Reliability of Discography (Key Question 4)

Study	Number of discs	System	Inter-rater reliability kappa (95% CI)[a]				Notes
			Overall	Rater 1 & Rater 2	Rater 1 & Rater 3	Rater 2 & Rater 3	
Agorastides (2002)(33)	133	Adams classification	0.77 (0.66 to 0.87)	0.70 (0.59 to 0.81)	0.79 (0.69 to 0.89)	0.82 (0.73 to 0.91)	Of the 25 total disagreements (19% of 133 discograms), 23 were one-grade disagreements, 1 was a two-grade disagreement, and 1 was a three-grade disagreement.
Milette (1999)(34)	132	DDD classification of annular *degeneration*	0.67 (0.55 to 0.78)	NA	NA	NA	
Milette (1999)(34)	132	DDD classification of annular *disruption*	0.66 (0.56 to 0.76)	NA	NA	NA	"Most interobserver disagreement occurred in the differentiation between Stages 0 and 1 disruption: one reader interpreted as normal 15 of 20 discs (75%) showing Stage 1 disruption according to the other reader."

NOTE: These data apply to the inter-rater reliability of viewing a discogram and gradings its morphology.

[a] Kappa measures chance-corrected agreement. 0 represents chance, and 1 represents perfect agreement. The standard interpretation of kappa values is that Below 0.0 is Poor agreement; 0.00-0.20 is Slight agreement; 0.21-0.40 is Fair agreement; 0.41-0.60 is Moderate agreement; 0.61-0.80 is Substantial agreement; 0.81-1.00 is Almost Perfect agreement.(35) The 95% CI around each kappa was calculated by ECRI based on the reported standard errors.

NA Not applicable because the Milette study only had two raters.

DDD Original Dallas Discogram Description. This classification method considers separately two aspects of the discogram: *degeneration* and *disruption*. For degeneration, 0 indicates no change, 1 indicates Local (<10% degeneration), 2 indicates Partial (10-50% degeneration), and 3 indicates total (>50% degeneration). For disruption, 0 indicates none, 1 indicates disruption into inner annulus, 2 indicates disruption into outer annulus, and 3 indicates disruption beyond outer annulus.(13)

Table 31. Quality Assessment of Studies on Discography to Predict Fusion Outcomes (Key Question 5)

Quality item	Willems (2007)(36)	Gill (1992)(37)	Colhoun (1988)(38)
Item 1	No	No	No
Item 2	No	No	No
Item 3	No	No	No
Item 4	Yes	No	No
Item 5	Yes	Yes	Yes
Item 6	Yes	Yes	Yes
Item 7	No	NR	NR
Item 8	NR	NR	NR
Item 9	No	No	Yes
Item 10	NR	Yes	NR
Item 11	NR	NR	NR
Item 12	Yes	Yes	Yes
Item 13	No	No	No
Item 14	No	No	No
Item 15	No	No	No
Item 16	No	No	No
Item 17	No	No	No
Item 18	NR	Yes	No
Item 19	Yes	NR	NR
Item 20	Yes	Yes	Yes
Item 21	Yes	Yes	Yes
Item 22	NR	NR	NR
Quality score and category	**4.3 (Low)**	**4.3 (Low)**	**4.1 (Low)**

The 22 quality items for Key Question 5 and 6 are listed in Appendix C , along with details on how we scored the quality assessments.

Table 32. Baseline Group Comparability in Studies Addressing Key Question 5

*Willems (2007)(36)	POSITIVE: Discography provoked pain on adjacent disc(s) (N = 22)	NEGATIVE: Discography did not provoke pain on adjacent disc(s) (N = 60)	Well-matched at baseline?
Mean age (SD)	39 (SD: 8.5)	39 (SD: 7.8)	Yes
Percentage male	38% (6/16)	43% (18/42)	Yes
Degenerative disc disease	18% (4/22)	33% (20/60)	No
Spondylolysis/Spondylolisthesis[a]	9% (2/22)	5% (3/60)	Yes
Prior spine surgery	73% (16/22)	62% (37/60)	Yes
Mean baseline VAS pain score (SD)[b]	72 (SD: 19)	75 (SD: 15)	Yes
Mean baseline Oswestry disability score	NR	NR	NR

*NOTE: Only the Willems study is listed because that was the only one of the three studies for this question that reported comparative patient characteristics before surgery.

[a] The percentage of patients with either spondylolysis or spondylolisthesis was low enough (6%) that the study still met inclusion criteria

[b] The VAS scores as reported in Table 3 of the article were erroneous because their ranges went above 100, which is nonsensical. Therefore, ECRI Institute estimated these scores based on Figures 1 and 2 by assuming that the pain scores of patients in a given category (e.g., VAS 21-40) were at the midpoint of the category (e.g., 30.5).

ALIF Anterior lumbar interbody fusion.
DDD Degenerative Disc Disease.
NR Not reported.
SD Standard deviation.

Table 33. Outcome Data for Discography to Predict Fusion Outcomes (Key Question 5)

Discography interpretation based ONLY on pain provocation						
Willems (2007)(36)	**Mean length of follow-up in years (range)**	**POSITIVE: Discography provoked pain on adjacent disc(s) (N = 22)**	**NEGATIVE: Discography did not provoke pain on adjacent disc(s) (N = 60)**	**Between group difference (95% CI)**	**Statistically different?**	**Statistically equivalent?[b]**
VAS pain score at follow-up (95% CI)[a]	6.7 (1.3 to 12)	51 (39 to 64)	52 (45 to 60)	1.1 (-13.6 to +15.8)	No	Yes
Percentage of patients with at least 30% pain relief	6.7 (1.3 to 12)	45.5% (10/22)	45.0% (27/60)	0.5% (-24% to +25%)	No	No
ECRI Institute re-analysis of Colhoun (1988)(38)	**Mean length of follow-up in years (range)**	**POSITIVE PAIN PROVOCATION (N = 137)**	**NEGATIVE PAIN PROVOCATION (N = 58)**	**Between group difference (95% CI)**	**Statistically different?**	**Statistically equivalent?[b]**
Percentage of patients with "Success"[c]	3.6 (2 to 10)	88% (121/137)	67% (39/58)	21% (+9% to +33%)	Yes	No
Discography interpretation based ONLY on morphology						
Gill (1992)(37)	**TYPE III RESULT: Annular tear beyond the periphery (N = 20)**	**TYPE II RESULT: Annular tear and contrast extension to the periphery, but not beyond (N = 19)**	**TYPE I RESULT: Small annular tear that did not extend to the periphery (N = 14)**	**Chi square test result[d]**	**Statistically different?**	**Statistically equivalent?[b]**
Percentage of patients showing "improvement on functional testing and pain report"	75% (15/20)	74% (14/19)	50% (7/14)	$X^2(2) = 2.81$; p = 0.24	**No**	**No**
ECRI Institute re-analysis of Colhoun (1988)(38)	**POSITIVE MORPHOLOGY (N = 162)**	**NEGATIVE MORPHOLOGY (N = 6)**	**Between group difference (95% CI)**	**Chi square test result[d]**	**Statistically different?**	**Statistically equivalent?[b]**
Percentage of patients with "Success"[c]	83% (134/162)	50% (3/6)	33% (+1% to +64%)	$X^2(1) = 4.12$; p = 0.042	Yes	**No**

Discography interpretation based on both pain provocation AND morphology							
Colhoun (1988)(38)	**Pain provocation and positive morphology (N = 137)**	**No pain provocation and positive morphology (N = 25)**	**Total disc resorption thus morphology not assessable, and pain provocation not reported (N = 27)**	**No pain provocation and negative morphology (N = 6)**	**Chi square test result[d]**	**Statistically different?**	**Statistically equivalent?[b]**
Percentage of patients with "Success"[c]	88% (121/137)	52% (13/25)	85% (23/27)	50% (3/6)	$X^2(3) = 23.35$; p = 0.000034	Yes	No

[a] The VAS scores as reported in Table 3 of the Willems article were erroneous because their ranges went above 100, which is nonsensical because the maximum VAS score is 100. Therefore, ECRI Institute estimated VAS scores based on Figures 1A and qB by assuming that the pain scores of patients in a given category (e.g., VAS 21-40) were at the midpoint of the category (e.g., 30.5).

[b] Equivalence was defined as statistically significantly less than the minimum clinically significant difference. For VAS pain scores, this was defined as 20 points. For the difference in the percentage of patients with 30% pain relief, this was defined as 15%.(36)

[c] "Success" was defined as meeting all three of the following conditions: 1) Complete relief or significant subjective improvement in symptoms; 2) Resumption of work and/or normal duties; 3) No intake of analgesics.

[d] Chi square test performed by ECRI Institute

CI Confidence interval.

Table 34. Quality Assessment of Studies on Discography to Impact Fusion Outcomes (Key Question 6)

Quality item	Madan (2002)(39)
Item 1	No
Item 2	No
Item 3	No
Item 4	No
Item 5	Yes
Item 6	NR
Item 7	No
Item 8	Yes
Item 9	No
Item 10	NR
Item 11	No
Item 12	Yes
Item 13	No
Item 14	No
Item 15	No
Item 16	No
Item 17	No
Item 18	Yes
Item 19	No
Item 20	Yes
Item 21	Yes
Item 22	NR
Quality score and category	**3.4 (Very Low)**

The 22 quality items for Key Question 5 and 6 are listed in Appendix C, along with details on how we scored the quality assessments.

Table 35. Baseline Group Comparability in Studies Addressing Key Question 6

Madan et al. (2002)(39)	Received discography prior to fusion (N = 32)	Did not receive discography prior to fusion (N = 41)	Well-matched at baseline?
Mean age (range)	42.1 (30-47)	40.8 (15-68)	Yes
Percentage male	50% (16/32)	66% (27/41)	No
Duration of condition	NR	NR	NR
Percentage with degenerative disc disease	NR	NR	NR
MRI grade 2 disc changes	13% (4/32)	7% (3/41)	Yes
MRI grade 3 disc changes	44% (14/32)	27% (11/41)	No
MRI grade 4 disc changes	44% (14/32)	66% (27/41)	No
Prior spine surgery	NR	NR	NR
Mean baseline VAS pain score (SD)[b]	NR	NR	NR
Mean baseline Oswestry disability score	NR	NR	NR

NR Not reported
SD Standard deviation

Table 36. Outcome Data for the Impact of Discography on Fusion Outcomes (Key Question 6)

Madan et al. (2002)(39)	Received discography prior to fusion (N = 32)	Did not receive discography prior to fusion (N = 41)	Statistical test result, or between group difference (95% CI)	Statistically different?	Statistically equivalent?
Mean Oswestry disability score (range)	34.17 (4-94)	34.15 (0-86)	Not computable, because authors only reported that p >0.05.	No	NR
Percentage with Oswestry score <20	63% (20/32)	59% (24/41)	X^2 (3) = 1.642; p = 0.65	No	NR
Percentage with Oswestry score 20 to 40	19% (6/32)	17% (7/41)			
Percentage with Oswestry score 40 to 60	19% (6/32)	20% (8/41)			
Percentage with Oswestry score >60	0% (0/32)	5% (2/41)			
Percentage with Oswestry score <40	81% (26/32)	76% (31/41)	6% (-13% to +25%)	No	No
Psychologic Score (range)[a]	21.5 (7-36)	15 (7-22)	Not computable, because authors only reported p >0.05.	No	NR
VAS pain rating (range)	4.25 (1-9)	4.4 (0-10)	Not computable, because authors only reported p >0.05.	No	NR
Core set (range)[b]	23.75 (10-48)	25.2 (10-48)	Not computable, because authors only reported p >0.05.	No	NR

NOTE: The mean length of follow-up in this study was 2.6 years (range 2 to 4.2)

[a] The Psychologic Score was the sum of the Zung Depression Scale and the Modified Somatic Perception Questionnaire.

[b] The Core set involved seven questions about surgical outcome. The score ranges from 7 (indicating the best possible outcome) to 50 (indicating the worse possible outcome).