Draft Systematic Review

Number xx

The Performance of Fusion Procedures for Degenerative Disease of the Lumbar Spine

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

and

Patient-Centered Outcomes Research Institute
1333 New Hampshire Avenue, NW, Ste. 1200
Washington, DC 20036
www.pcori.org

Contract No. xxx-xx-xxxx

Prepared by: To be added for final version.

Investigators: To be added for final version.

AHRQ Publication No. xx-EHCxxx

PCORI Publication No. XXXX-SR-XX

This report is based on research conducted by the [name to be added for final] Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. XXX-20XX-XXXXX). The Patient-Centered Outcomes Research Institute (PCORI®) funded the report (PCORI® Publication No. XXXX-SR-XX). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ or PCORI®, its Board of Governors or Methodology Committee. Therefore, no statement in this report should be construed as an official position of AHRQ, PCORI® or of the U.S. Department of Health and Human Services.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

The information in this report is intended to help healthcare decision makers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of healthcare services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

A representative from AHRQ served as a Contracting Officer's Representative and reviewed the contract deliverables for adherence to contract requirements and quality. AHRQ did not directly participate in the literature search, determination of study eligibility criteria, data analysis, interpretation of data, or preparation or drafting of this report.

AHRQ appreciates appropriate acknowledgment and citation of its work. Suggested language for acknowledgment: This work was based on an evidence report, The Performance of Fusion Procedures for Degenerative Disease of the Lumbar Spine, by the Evidence-based Practice Center Program at the Agency for Healthcare Research and Quality (AHRQ).

Suggested citation: <Authors>. <Topic in Title Caps>. <Report Series Name in Title Caps No.> <#>. (Prepared by the <EPC Name> Evidence-based Practice Center under Contract No. <##>.) AHRQ Publication No. XX-EHCXXX-EF. Rockville, MD: Agency for Healthcare Research and Quality. <Month Year>. Posted final reports are located on the Effective Health Care Program Webpage.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of systematic reviews to assist public- and private-sector organizations in their efforts to improve the quality of healthcare in the United States. The Patient-Centered Outcomes Research Institute (PCORI®) requested this report from the EPC Program at AHRQ. AHRQ assigned this report to the XX EPC (Contract Number <##>).

AHRQ EPC reviews provide comprehensive, science-based information on common, costly medical conditions, and new healthcare technologies and strategies.

The Patient-Centered Outcomes Research Institute (PCORI®) was established to fund research that helps patients and caregivers make better informed health care choices. To fulfill its authorizing mandate, PCORI® partners with AHRQ to generate evidence synthesis products and make comparative effectiveness research more available to patients and providers.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews can help clarify whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about AHRQ EPC systematic reviews, see https://effectivehealthcare.ahrq.gov/about/epc/evidence-synthesis

AHRQ expects that these systematic reviews will be helpful to health plans, providers, purchasers, government programs, and the healthcare system as a whole. Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the website (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an e-mail list to learn about new program products and opportunities for input.

If you have comments on this systematic review, they may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857, or by email to epc@ahrq.hhs.gov.

Mamatha Pancholi, M.S.
Acting Director
Agency for Healthcare Research and Quality

Therese Miller Ph.D.
Director
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Christine Chang, M.D., M.P.H.
Director
Evidence-based Practice Center Program
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Kim Wittenberg, M.A.
Task Order Officer
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Nakela Cook, M.D., M.P.H.
Executive Director
Patient-Centered Outcomes Research Institute

Greg Martin, B.A.
Chief of Engagement Dissemination, and Implementation
Patient-Centered Outcomes Research Institute

Michelle Althuis, Ph.D.
Associate Director
Patient-Centered Outcomes Research Institute

Rachael Robinson, M.P.H.
Program Officer
Research Synthesis
Patient-Centered Outcomes Research Institute

Acknowledgments

The authors gratefully acknowledge the following individuals for their contributions to this project: To be added for the final version.

Key Informants

In designing the study questions, the EPC consulted several Key Informants who represent the end-users of research. The EPC sought the Key Informant input on the priority areas for research and synthesis. Key Informants are not involved in the analysis of the evidence or the writing of the report. Therefore, in the end, study questions, design, methodological approaches, and/or conclusions do not necessarily represent the views of individual Key Informants.

Key Informants must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any conflicts of interest.

The list of Key Informants who provided input to the report will be added for the final version.

Technical Expert Panel

In designing the study questions and methodology at the outset of this report, the EPC consulted several technical and content experts. Broad expertise and perspectives were sought. Divergent and conflicted opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, in the end, study questions, design, methodologic approaches, and/or conclusions do not necessarily represent the views of individual technical and content experts.

Technical Experts must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

The list of Technical Experts who provided input to the report will be added for the final version.

Peer Reviewers

Prior to publication of the final evidence report, EPCs sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report do not necessarily represent the views of individual reviewers. AHRQ may also seek comments from other Federal agencies when appropriate.

Peer Reviewers must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential non-financial conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential non-financial conflicts of interest identified.

The list of Peer Reviewers who reviewed the report will be added for the final version.

The Performance of Fusion Procedures for Degenerative Disease of the Lumbar Spine

Abstract

Background. Optimal management of symptomatic degenerative lumbar disease (DLD), including symptomatic degenerative lumbar spondylolisthesis (DLS), is unclear.

Objective: To evaluate the effectiveness and harms of lumbar fusion in patients with DLS and of specific nonsurgical procedures for patients with chronic low back pain (CLBP) due to degenerative spine disease.

Data sources. Three electronic databases from 2000 to October 23, 2024 for surgical questions (Key Questions 1 to 4) and to October 23, 2024 from inception for nonsurgical questions (Key Questions 5 and 6); reference lists; and submissions in response to a Federal Register.

Review methods. Using predefined criteria and dual review, we selected randomized controlled trials (RCTs) and nonrandomized studies of interventions (NRSIs) that evaluated the benefits and harms of fusion when added to decompression in patients with stable DLS, as well as the use of bone extenders, biologic substitutes, and intraoperative neuromonitoring (IONM) during lumbar fusion. We also evaluated specific interventions in patients with degenerative CLBP. Study risk of bias was assessed using predefined criteria. Strength of evidence (SOE) was assessed for the primary outcomes of pain, function, quality of life, fusion and reoperation rates, and harms. Random effects meta-analysis was conducted when feasible, and effects were classified as small, moderate, or large using previously defined criteria.

Results. We included 35 RCTs and 13 comparative NRSIs; most were assessed as moderate risk of bias. In patients with grade I DLS and spinal stenosis, pain relief and functional improvement were similar after fusion plus decompression versus decompression alone (5 RCTs). The likelihood of reoperation at 1 year was slightly lower when fusion was done but estimates were imprecise (2 RCTs, SOE: low). In a small trial in patients undergoing spinal fusion, addition of an interbody cage to pedicle screw instrumentation was associated with a higher likelihood of fusion (arthrodesis) but no difference in measures of function (SOE: low). Another small RCT found no difference in fusion rates for demineralized bone matrix compared to autograft in patients with DLS (SOE: low). One study found similar odds of developing neurological complications in general with and without IONM during fusion for degenerative lumbar disease, (SOE: low) but specific intraoperative and post-surgical clinical outcomes were not detailed. Compared with placebo, epidural steroid injections (ESI) in patients with nondiscogenic CLBP (5 RCTs, SOE: moderate) and facet joint injections (4 RCTs, SOE: low) and medial branch blocks (2 RCTs, SOE: low) in patients with facet-related CLBP provided similar improvements in pain, function and quality of life. Improvement in pain and function with continuous radiofrequency ablation (RFA) versus sham in patients with facet-related CLBP was not consistently seen across six trials. At 3 months, pooled estimates across trials of RFA found no difference between groups, but two trials report improvements at 6 months versus sham. There

was insufficient evidence on the use of interventional procedures such as ESI, facet join injection, RFA or medial branch block for predicting fusion outcomes. Data were not sufficient to evaluate outcomes in subgroups of patients with different condition severity or concurrent degenerative spine disease or to assess publication bias.

Conclusions. Our findings may facilitate shared decision-making and the development of evidence-based recommendations. In patients with DLS, adding fusion to decompression was not associated with improved pain or function, although some improvement in QOL and slightly lower likelihood of reoperation at 1 year was seen. Evidence comparing the use of interbody cages with PLF is sparse, but fusion may be more common with the use of cages. ESI, facet joint injection, and medial bundle branch block were not associated with improved pain or function versus placebo in patients with CLBP. Continuous RFA may be an option in patients facet joint pain who would like to avoid surgery or who may not be good candidates for surgery, however evidence of benefit was inconsistent across studies and time frames. Research is needed to clarify optimal approaches to fusion surgery for stable DLS and to determine the benefits of IONM during fusion surgery for degenerative lumbar disease and the predictive value of interventional procedures regarding fusion outcomes.

Contents

Tables

Figures

Appendixes

Executive Summary

Main Points

- Pain and function improvement and risk of serious adverse events may be similar for the addition of fusion to decompression versus decompression alone in patients with Grade I degenerative lumbar spondylolisthesis and concomitant stenosis.
- Interbody fusion and posterolateral fusion may have similar functional improvement, but interbody fusion may have higher rates of fusion in patients with degenerative lumbar spondylolisthesis (DLS); evidence on harms was insufficient.
- Evidence on clinical benefits of intraoperative neurologic monitoring in patients undergoing fusion for degenerative lumbar disease is lacking, but this is routinely used in some settings and may be of value for detecting intraoperative neurological events.
- In patients with chronic low back pain (CLBP) due to degenerative lumbar disease without herniated disc, improvement in pain and function is probably similar for epidural steroid injections versus placebo. Improvements in these outcomes may be similar to placebo for medial bundle branch block and facet joint injections.
- There was heterogeneity in results regarding improvement in pain and function with continuous radiofrequency ablation (RFA) versus sham in patients with facet-related CLBP.
- RCTs using clear definitions of DLS stability that compare fusion plus decompression versus decompression alone and RCTs comparing interbody fusion with PLF in patients with DLS are needed. Studies that help identify which patients with DLS may benefit most from surgery are needed.
- High quality studies of utility of IONM during fusion procedures for degenerative lumbar disease are needed.

Background and Purpose

Degenerative lumbar spine disease is common and is associated with age. In some patients, degenerative lumbar disease causes clinically relevant pain and loss of function from nerve compression associated with spondylolisthesis. A decompressive procedure can alleviate the symptoms associated with neurological compression; however, decompression alone can result in progression of the vertebral misalignment. The Spine Patient Outcomes Research Trial (SPORT) established that surgery may have a role in patients with stenosis associated with degenerative spondylolisthesis. However, while the SPORT established a role for surgery in the setting of refractory symptoms of nerve compression, primarily in patients with spinal stenosis, it did not lead to clarity regarding the ideal candidate for decompression alone, decompression with fusion, and the role of instrumentation in fusion procedures.[1] In 2014, reviews conducted for guidelines from the Congress of Neurologic Surgeons and the National Association of Spine Surgeons confirmed that these gaps were still unresolved.[2-4] For example, while there is agreement that fusion has a role when degenerative lumbar spondylolisthesis is advanced or there is evidence of instability, it is not often indicated when there is no spondylolisthesis or evidence of instability. There is disagreement about the need for fusion in patients with spinal stenosis who have grade I DLS, as well as evidence that surgeons cannot always identify which patients will benefit from fusion with decompression vs decompression alone.

We assessed whether various procedures associated with fusion improve the effectiveness and harms of surgery for patients with degenerative spondylolisthesis. We also assessed nonsurgical procedures for treatment in patients with CLBP due to degenerative spine disease. The intended audiences for this review are guideline developers, clinicians, policymakers, patients, their caregivers, and researchers. The Patient Centered Outcomes Research Institute (PCORI) provided funding for this review, and the Congress of Neurological Surgeons (CNS) plans to use the review to inform an updated guideline

Methods

We employed methods consistent with those outlined in the Agency for Healthcare Research and Quality Evidence-based Practice Center Program methods guidance (https://effectivehealthcare.ahrq.gov/topics/cer-methods-guide/overview). We described these in detail in the full report. Our searches covered publication dates from 2000 up to October 23, 2024 for surgical questions, and no start date limit for nonsurgical questions. We sought studies in patients with stable degenerative lumbar spondylolisthesis as the focus for Key Question 1, defined as <3 mm slip on extension/flexion radiographs. We also included studies that did not specify the proportions of stable and unstable DLS based on this definition. For Key Questions 2 and 3, we included studies of patients with either stable or unstable DLS. We sought studies in patients with degenerative lumbar disease undergoing fusion for Key Question 4. Because we anticipated that the evidence specifically for DLS would be sparse, we sought studies of patients with CLBP without symptoms attributable to herniated disc undergoing select nonsurgical procedures for Key Questions 5 and 6. Study risk of bias was assessed using predefined criteria. We analyzed effects and assessed strength of evidence (SOE) for the primary outcomes of pain, function, quality of life, reoperation rates, fusion rates, persistent neurological damage and harms.

Results

We included 48 studies (in 58 publications; 60% at moderate risk of bias): 35 randomized controlled trials (RCTs) (in 43 publications) and 13 comparative nonrandomized studies of interventions (NRSIs) (15 publications). For the surgical questions, the number of included studies is as follows: Key Question 1 comparing fusion plus decompression to decompression alone (5 RCTs in 9 publications; 7 NRSIs in 8 publications); Key Question 2 comparing interbody fusion with posterolateral fusion (2 RCTs, 4 NRSIs); Key Question 3 comparing bone graft extenders and biologic substitutes versus autografts (1 RCT); and Key Question 4 evaluating the impact of intraoperative neuromonitoring (IONM) (1 NRSI). For Key Questions, 5 and 6, nonsurgical procedures, we identified 17 RCTs for radiofrequency ablation, 5 RCTs on epidural steroid injections, 4 RCTs of facet joint injections, and 2 RCTs (3 publications) on medial branch block. One NRSI met inclusion criteria for Key Question 6. Key findings with at least low SOE are summarized below. Effect sizes reported as small, moderate or large in the tables favor the intervention unless otherwise noted.

Key Questions on Lumbar Fusion (Key Questions 1-4)

For Key Question 1, we identified five trials.[5,6 7-9] One of these, the Spinal Laminectomy versus Instrumented Pedicle Screw (SLIP) trial (n=60),[5] focused specifically on patients with stable DLS as defined above. Approximately 80 percent of patients in the Nordsten trial (N=267)

met this definition, however patients with >3 mm slip were not excluded.[6] The other three RCTs[7-9] did not specify the proportions of patients with stable and unstable DLS; patients in all trials had concomitant stenosis. Improvement in pain and function was similar as was the risk of serious adverse events with the addition of fusion to decompression versus decompression alone (SOE: low) (**Table ES-1**).

Table ES-1. Summary of evidence of decompression plus fusion versus decompression alone (Key Question 1) in populations with degenerative spondylolisthesis of stable, mixed, or unknown stability

Outcome	Time Point	Effect Size/SOE[a]
Mean improvement in back and leg pain	3 months, 1, 2, 5 years	Similar +
Successful function outcome[b]	5 years	Similar +
Mean improvement in function[c]	3, 6 months, 1, 2, 3, 4, 5 years	Similar +
Mean improvement in QOL	3 months, 1, 5 years	Similar +
	6 months, 3, 4 years	Small +
	2 years	EQ-5D: Similar + SF-36 PCS: Small +
Reoperation	1, 3 months	Similar +
	1 year	Small decrease +
Serious AE, dural tear, deep infection, PE, DVT, heart attack, or stroke	Any time	Similar +
Neurological deterioration	5 years	Small increase +

Abbreviations: AE = adverse event; DVT = deep vein thrombosis; EQ-5D = EuroQol 5 dimensions; IONM = intraoperative neuromonitoring; PE = pulmonary embolism; QOL = quality of life; SF-36 PCS= Short-form 36 questionnaire Physical Component Score; SOE = strength of evidence.

[a] Effect size: Similar (no effect), small, moderate, or large difference favoring the intervention unless otherwise noted; SOE: + = low, ++ = moderate, +++ = high.

[b] Defined as ≥30% improvement (i.e., reduction) in Oswestry Disability Index scores.

[c] Based on Oswestry Disability Index scores; evidence for Japanese Orthopaedic Association scores was insufficient.

For Key Question 2, two RCTs provided insufficient evidence to draw reliable conclusions for interbody fusion versus posterolateral fusion (PLF) in patients with DLS for most outcomes. One of the RCTs (N=60)[10] found no difference between transforaminal interbody fusion (TLIF) plus PLF and PLF in the proportion of patients achieving Oswestry Disability Index (ODI) scores between 0 and 20 percent (ODI is scored 0%-100% where 0%-20% represents mild disability; 21%-40% moderate disability; ≥41% severe disability) at 2 years but found TLIF plus PLF was associated with substantially higher likelihood of fusion versus PLF by 2 years. For Key Question 3, a single small RCT (N=46)[11] found no difference in fusion rates with a specific brand of demineralized bone matrix compared to iliac crest bone graft. One large administrative data study (n=133,572)[12] of IONM in patients undergoing elective posterolateral fusion for DLD did not provide adequate detail regarding clinical outcomes to reliably assess how benefits of IONM varied across fusion procedures for DLD.

Nonsurgical Procedure Questions (Key Questions 5, 6)

For Key Question 5, the greatest volume of evidence compared continuous RFA with sham (9 RCTs) in patients with presumed facet joint pain.[13-21] Three RCTs excluded patients with spondylolisthesis;[13,15,16] others did not provide information on DLS. Our review found that there was some heterogeneity in results across RCTs regarding improvement in pain and function with continuous RFA versus sham. Several trials found similar improvements for continuous RFA and sham at 3 months with other trials suggesting some benefit for RFA at 6 months (**Table ES-2**). A single trial of pulsed RFA applied to the dorsal root ganglion reported improvements with the RFA in pain and function but estimates were imprecise.[16] No difference was seen in pain or function between treatments in another trial which targeted the medial branch of the dorsal ramus, however.[18] A single open-label RCT[22] found that continuous RFA was associated with improvement in pain and function versus usual care. Similar improvements in pain, function and quality of life were seen in trials comparing ESI to placebo in patients with spinal stenosis (primarily central canal stenosis). Similar improvements in these outcomes were also seen in trials comparing facet joint (intraarticular) injection, and medial branch block (MBB) to placebo in patients with presumed facet joint pain (**Table ES-2**).

Table ES-2. Summary of evidence of interventional procedures versus sham placebo (Key Question 5) in patients with chronic low back pain

Outcome (Effect Size/SOE)[a]	Time Point	Continuous RFA vs. Sham[b]	ESI vs. Placebo[c]	FJI vs. Placebo[b]	MBB vs. Placebo[b]
Successful pain outcome[d]	3 months	Similar +	Similar ++	No evidence	Similar +
	6 months	Large +	Similar ++	No evidence	Similar +
Mean improvement in back pain	3 months	Similar ++	No evidence	Similar +	Similar +
	6 months	Small[e] +	No evidence	Similar +	Similar +
Mean improvement in leg pain	3 months	Similar +	Similar ++	No evidence	No evidence
	6 months	Large +	Similar ++	No evidence	No evidence
Successful function outcome[f]	3 months	Large +	Similar ++	No evidence	Similar +
	6 months	No evidence	Similar ++	No evidence	Similar +
Mean improvement in function	3 months	Similar +	Similar ++	Similar +	Similar +
	6 months	Similar +	Similar ++	Similar +	Similar +
Quality of Life	6 weeks	No evidence	Similar ++	No evidence	No evidence
	6 months	Similar +	No evidence	No evidence	No evidence
Serious AE	Any time	Similar +	Similar +	No evidence	No evidence
Any AEs	Any time	Similar +	Similar +	Insufficient evidence	Insufficient evidence

Abbreviations: AE = adverse event; FJI = facet joint (intraarticular) injection; MBB = medial branch block; RFA = radiofrequency ablation; SOE = strength of evidence.

[a] Effect size: Similar (no effect), small, moderate, or large difference favoring RFA; SOE: + = low, ++ = moderate, +++ = high.

[b] Population = presumed facet joint pain.

[c] Population = spinal stenosis (3 RCTs of central canal stenosis, 1 RCT of foraminal stenosis and degenerative scoliosis).

[d] A successful pain outcome was defined as ≥50% improvement in pain on the Visual Analog Scale or Numerical Rating Scale.

[e] Small effect is based off analysis excluding Moussa 2016 and Moussa 2020. In these trials, patients received post treatment injection of anesthetic and steroids; the effects of these additional treatments on outcomes were unclear.
[f] A successful function outcome was defined as ≥10 point improvement on Roland Morris Disability Questionnaire (RDQ) (continuous RFA); ≥40% or ≥50% improvement in Oswestry Disability Index or RDQ (ESI, MBB).

Comparative Effectiveness

Limited evidence indicates pain and functional improvement are similar with RFA and facet joint injection.[23-26] RFA was associated with improved pain and function compared with medial branch block.[27-29] Similar improvements in pain and function were seen for comparisons of facet joint (intraarticular) injection with systemic steroids[30] and with MBB[31] (see full report).

Strengths and Limitations

Our review includes new RCT evidence on fusion for grade I (low-grade) DLS that was not available for the reviews conducted for 2014 guidelines. A strength of our review is expansion of inclusion criteria to include patients with both stable DLS (<3 mm slip) and unstable DLS for Key Question 1 and including patients with CLBP, rather than restricting trials of interventional procedures to patients with DLS. This enabled us to capture more potentially relevant evidence to inform clinical recommendations. Our categorization of the magnitude of effects for function and pain outcomes uses the system described in our previous reviews to facilitate consistent interpretation across trials and interventions.

Many limitations to this review are related to the limitations of the available evidence. Studies of fusion for DLS did not consistently provide information on diagnostic criteria or severity of DLS. A variety of interbody fusion procedures were used. In trials of interventional procedures there was heterogeneity in how the procedures were done (e.g., RFA) and there was insufficient evidence to evaluate how technical factors and types of sham or comparator might impact findings. There was insufficient data to do subgroup analyses or evaluate modification of effects by factors such as patient characteristics, DLS grade, concurrent spine pathology, or procedural factors or to assess publication bias.

Implications and Conclusions

Our findings may facilitate shared decision-making and balancing of benefits and harms for the development of evidence-based recommendations. The addition of fusion following decompression was not associated with improved pain or function, although some improvement in quality of life and slightly lower likelihood of reoperation at 1 year was seen versus decompression alone. TLIF plus PLF was associated with a moderately higher likelihood of fusion versus PLF by 2 years. There is insufficient evidence to determine effects of IONM during fusion surgery for degenerative lumbar disease on neurological outcomes or other clinical outcomes. There was insufficient evidence for the use of interventional procedures for predicting outcomes of fusion. None of the trials of special procedures were specifically in patients with DLS. Overall, there was not consistent pain and function improvement between continuous RFA and sham for presumed facet joint pain, however, there was heterogeneity across trials with some reporting improvements in pain and function. Continuous RFA may be an option in patients who would like to avoid surgery or who may not be good candidates for surgery. ESI, facet joint injection and medial bundle branch block were not associated with improved pain or function versus placebo in patients with CLBP due to spinal stenosis. Research is needed to clarify optimal approaches to fusion surgery for stable DLS and alternatives to surgery.

References

1. North American Spine Society (NASS). Clinical Guidelines for Multidisciplinary Spine Care. Diagnosis and Treatment of Degenerative Lumbar Spondylolisthesis 2nd Edition. 2014. https://www.spine.org/Portals/0/Assets/Downloads/ResearchClinicalCare/Guidelines/Spondylolisthesis.pdf

2. Kaiser MG, Eck JC, Groff MW, et al. Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 1: introduction and methodology. J Neurosurg Spine. 2014;21(1):2-6. doi: 10.3171/2014.4.SPINE14257. PMID: 24980578.

3. Matz PG, Meagher RJ, Lamer T, et al. Guideline summary review: An evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spondylolisthesis. Spine J. 2016;16(3):439-48. doi: 10.1016/j.spinee.2015.11.055. PMID: 26681351.

4. Chan AK, Sharma V, Robinson LC, et al. Summary of Guidelines for the Treatment of Lumbar Spondylolisthesis. Neurosurgery Clinics of North America. 2019;30(3):353-64. doi: https://dx.doi.org/10.1016/j.nec.2019.02.009 . PMID: 31078236.

5. Ghogawala Z, Dziura J, Butler WE, et al. Laminectomy plus Fusion versus Laminectomy Alone for Lumbar Spondylolisthesis. New England Journal of Medicine. 2016;374(15):1424-34. doi: 10.1056/NEJMoa1508788. PMID: 27074067.

6. Kgomotso EL, Hellum C, Fagerland MW, et al. Decompression alone or with fusion for degenerative lumbar spondylolisthesis (Nordsten-DS): five year follow-up of a randomised, multicentre, non-inferiority trial. BMJ. 2024;386:e079771. doi: 10.1136/bmj-2024-079771. PMID: 39111800.

7. Karlsson T, Forsth P, Ohagen P, et al. Decompression alone or decompression with fusion for lumbar spinal stenosis: five-year clinical results from a randomized clinical trial. Bone & Joint Journal. 2024;106-B(7):705-12. doi: 10.1302/0301-620X.106B7.BJJ-2023-1160.R2. PMID: 38945544.

8. Inose H, Kato T, Sasaki M, et al. Comparison of decompression, decompression plus fusion, and decompression plus stabilization: a long-term follow-up of a prospective, randomized study. Spine Journal: Official Journal of the North American Spine Society. 2022;22(5):747-55. doi: 10.1016/j.spinee.2021.12.014. PMID: 34963630.

9. Fariborz S, Gharedaghi M, Khosravi A, et al. Comparison of results of 4 methods of surgery in grade 1 lumbosacral spondylolisthesis. Neurosurgery. 2016;26:14-8.

10. Challier V, Boissiere L, Obeid I, et al. One-Level Lumbar Degenerative Spondylolisthesis and Posterior Approach: Is Transforaminal Lateral Interbody Fusion Mandatory?: A Randomized Controlled Trial With 2-Year Follow-Up. Spine. 2017;42(8):531-9. doi: 10.1097/BRS.0000000000001857. PMID: 27548582.

11. Kang J, An H, Hilibrand A, et al. Grafton and local bone have comparable outcomes to iliac crest bone in instrumented single-level lumbar fusions. Spine. 2012;37(12):1083-91. doi: 10.1097/BRS.0b013e31823ed817. PMID: 22076647.

12. Austerman RJ, Sulhan S, Steele WJ, et al. The utility of intraoperative neuromonitoring on simple posterior lumbar fusions-analysis of the National Inpatient Sample. The Journal of Spine Surgery. 2021;7(2):132-40. doi: 10.21037/jss-20-679. PMID: 34296025.

13. Gallagher J PDVP, Wedley JR. Radiofrequency facet joint denervation in the treatment of low back pain: a prospective controlled double-blind study to assess its efficacy. Pain Clinic. 1994;7:193-8.

14. Leclaire R, Fortin L, Lambert R, et al. Radiofrequency facet joint denervation in the treatment of low back pain: a placebo-controlled clinical trial to assess efficacy. Spine. 2001;26(13):1411-6; discussion 7. doi: 10.1097/00007632-200107010-00003. PMID: 11458140.

15. Moussa WM, Khedr W. Percutaneous radiofrequency facet capsule denervation as an alternative target in lumbar facet syndrome. Clinical Neurology & Neurosurgery. 2016;150:96-104. doi: https://dx.doi.org/10.1016/j.clineuro.2016.09.004. PMID: 27618781.

16. Moussa WM, Khedr W, Elsawy M. Percutaneous pulsed radiofrequency treatment of dorsal root ganglion for treatment of lumbar facet syndrome. Clinical Neurology & Neurosurgery. 2020;199:106253. doi: 10.1016/j.clineuro.2020.106253. PMID: 33045627.

17. Nath S, Nath CA, Pettersson K. Percutaneous lumbar zygapophysial (Facet) joint neurotomy using radiofrequency current, in the management of chronic low back pain: a randomized double-blind trial. Spine. 2008;33(12):1291-7; discussion 8. doi: 10.1097/BRS.0b013e31817329f0. PMID: 18496338.

18. Tekin I, Mirzai H, Ok G, et al. A comparison of conventional and pulsed radiofrequency denervation in the treatment of chronic facet joint pain. Clinical Journal of Pain. 2007;23(6):524-9. doi: 10.1097/AJP.0b013e318074c99c. PMID: 17575493.

19. Truong K, Meier K, Ahrens LC, et al. Cryoneurolysis versus radiofrequency ablation outcome on pain experience in chronic low back pain (COPE): a single-blinded randomised controlled trial. RMD Open. 2024;10(2):09. doi: 10.1136/rmdopen-2024-004196. PMID: 38724261.

20. van Tilburg CW, Stronks DL, Groeneweg JG, et al. Randomised sham-controlled double-blind multicentre clinical trial to ascertain the effect of percutaneous radiofrequency treatment for lumbar facet joint pain. Bone & Joint Journal. 2016;98-B(11):1526-33. doi: 10.1302/0301-620X.98B11.BJJ-2016-0379.R2. PMID: 27803230.

21. van Wijk RM, Geurts JW, Wynne HJ, et al. Radiofrequency denervation of lumbar facet joints in the treatment of chronic low back pain: a randomized, double-blind, sham lesion-controlled trial. Clinical Journal of Pain. 2005;21(4):335-44. doi: 10.1097/01.ajp.0000120792.69705.c9. PMID: 15951652.

22. Chen YS, Liu B, Gu F, et al. Radiofrequency Denervation on Lumbar Facet Joint Pain in the Elderly: A Randomized Controlled Prospective Trial. Pain Physician. 2022;25(8):569-76. PMID: 36375186.

23. Do KH, Ahn SH, Cho YW, et al. Comparison of intra-articular lumbar facet joint pulsed radiofrequency and intra-articular lumbar facet joint corticosteroid injection for management of lumbar facet joint pain: A randomized controlled trial. Medicine. 2017;96(13):e6524. doi: 10.1097/MD.0000000000006524. PMID: 28353611.

24. Düger C, ÖzdemİR Kol İ, Kaygusuz K, et al. Effects of facet joint nerve block addition to radiofrequency in the treatment of low back pain. HealthMED. 2012;6(6).

25. Lakemeier S, Lind M, Schultz W, et al. A comparison of intraarticular lumbar facet joint steroid injections and lumbar facet joint radiofrequency denervation in the treatment of low back pain: a randomized, controlled, double-blind trial. Anesthesia & Analgesia. 2013;117(1):228-35. doi: 10.1213/ANE.0b013e3182910c4d. PMID: 23632051.

26. McCormick ZL, Conger A, Kendall R, et al. A pragmatic randomized prospective trial of cooled radiofrequency ablation of the medial branch nerves versus facet joint injection of corticosteroid for the treatment of lumbar facet syndrome: 12 month outcomes. Pain Medicine. 2023;24(12):1318-31. doi: 10.1093/pm/pnad107. PMID: 37578437.

27. Civelek E, Cansever T, Kabatas S, et al. Comparison of effectiveness of facet joint injection and radiofrequency denervation in chronic low back pain. Turkish Neurosurgery. 2012;22(2):200-6. doi: https://dx.doi.org/10.5137/1019-5149.JTN.5207-11.1. PMID: 22437295.

28. Hashemi M, Hashemian M, Mohajerani SA, et al. Effect of pulsed radiofrequency in treatment of facet-joint origin back pain in patients with degenerative spondylolisthesis. European Spine Journal. 2014;23(9):1927-32. doi: 10.1007/s00586-014-3412-x. PMID: 24997616.

29. Zhou Q, Zhou F, Wang L, et al. An investigation on the effect of improved X-rays-guided radiofrequency thermocoagulation denervation on lumbar facet joint syndrome. Clinical Neurology & Neurosurgery. 2016;148:115-20. doi: 10.1016/j.clineuro.2016.07.018. PMID: 27442000.

30. Ribeiro LH, Furtado RN, Konai MS, et al. Effect of facet joint injection versus systemic steroids in low back pain: a randomized controlled trial. Spine. 2013;38(23):1995-2002. doi: 10.1097/BRS.0b013e3182a76df1. PMID: 23921331.

31. Marks RC, Houston T, Thulbourne T. Facet joint injection and facet nerve block: a randomised comparison in 86 patients with chronic low back pain. Pain. 1992;49(3):325-8. doi: 10.1016/0304-3959(92)90239-8. PMID: 1408298.

1. Introduction

1.1 Background and Objectives for the Systematic Review

1.1.1 Nature and Burden of Degenerative Lumbar Spine Disease and Spondylolisthesis

Degenerative spine disease is an umbrella term that refers to a set of conditions related to gradual structural changes of the spine and associated functional difficulties. It is frequently associated with age, but some other common causes may include tumors, infection, and arthritis. The estimated prevalence of spinal degenerative disease from 2005 to 2017, in people 65 and older, based on Medicare data of approximately 1.7 million individuals, is 27.3 percent, with the highest prevalence for degenerative disc disease (12.2%).[1] Degenerative spine changes include intervertebral disc degeneration (e.g., desiccation and/or herniation of discs), osteoarthritis, and facet joint degeneration, which can compromise normal structural support and create foraminal narrowing, malfunction of the spinal ligaments, inadequate muscle stabilization, and joint laxity. The anatomic and functional changes in the lumbar spine may result in a range of symptoms including low back pain, nerve impingement causing radiculopathy (e.g., pain, weakness, numbness, tingling), and/or neurogenic claudication (pain in the legs with prolonged walking or standing), as well as limits to spinal motion and may impact the patient's quality of life. Degenerative changes may or may not cause vertebral slippage, a condition called spondylolisthesis. Spondylolisthesis is often classified as one of five types (I-V), namely dysplastic, isthmic, degenerative, traumatic, and pathologic.[2] Types I and II commonly apply to children and adolescents, and type III (degenerative) is most common in adults. Displacement may be anterior (anterolisthesis) or posterior (retrolisthesis) and may be asymmetrical and described as stable or unstable. The fourth and fifth lumbar (L4-L5) vertebrae are most frequently involved, with L5-S1 or L3-L4 involvement less common.[3,4] The degree and direction of displacement and presence of other degenerative findings impact patient presentation, symptoms, and management. Degenerative lumbar changes and DLS may cause low back pain (LBP), but many patients are asymptomatic.[5,6] There is heterogeneity in assessment of DLS and an optimal diagnostic approach is unclear.[4,5] Radiographic diagnostic criteria and stability assessment are subjective, controversial, and not standardized.[4,6-8] The impact of these radiographic criteria on clinical management and patient outcomes is unclear.[4,5,9] The Meyerding classification[10] is commonly used to describe the percent of vertebral slip as Grade I (0%-25%), Grade II (25%-50%), Grade III (50%-75%), Grade IV (75%-100%) and Grade V (>100%). Grades I and II are considered low-grade slip, are most common, and are least likely to progress.[3,4,11] Grades III-V are considered a high-grade slip.

1.1.2 Management of Degenerative Lumbar Spine Disease and Degenerative Spondylolisthesis

Optimal management of symptomatic degenerative lumbar disease (DLD), including symptomatic DLS, is controversial and unclear.[3,12,13] The treatment goal is to alleviate pain, improve neurologic function, and prevent progression or recurrence. Nonoperative management is usually the first line of treatment and surgical interventions are considered following failure of this.[3,5,12] Nonoperative management varies widely; it may include a combination of exercise,

medications, physical therapy, and behavioral therapy. Epidural steroid injections (ESI), median nerve blocks, and radiofrequency ablation (RFA) may confer short-term pain relief for patients with DLD-related chronic low back pain (CLBP)[12,14,15] and are considered if less invasive treatments provide inadequate relief. ESI decreases nerve root inflammation; however, it is unclear if steroids produce superior clinical effects compared with local anesthetics or saline[16] and their role remains unclear. Nerve blocks and RFA generally target the facet joints, and their use remains controversial.[17]

The role of elective fusion versus nonoperative care for DLD overall is unclear due to conflicting results from a limited number of randomized controlled trials (RCTs),[7,18] concern about reoperation rates,[19] and questions about what works best and for whom.[7] Surgery for patients with low-grade DLS is controversial.[6,12,20] Low-grade DLS is the most common[6] and may be an incidental radiographic finding. Patients are usually asymptomatic or exhibit mild symptoms and may not require treatment or are managed nonoperatively.[3,5] An estimated 10 to 15 percent of patients may eventually receive surgery if there is neurologic involvement or if symptoms progress or are not relieved with conservative management.[21] Decompression with or without the addition of fusion may be considered, however, decompression alone may create a degree of instability, making fusion necessary.[6,22] A decompressive procedure can alleviate the symptoms associated with neurological compression; however, decompression alone can result in progression of the vertebral misalignment. Fusion and decompression are usually considered for high-grade DLS (grades III-V). While the Spine Patient Outcomes Research Trial (SPORT) established a role for surgery in the setting of refractory symptoms of nerve compression, primarily in patients with spinal stenosis, it did not lead to clarity regarding the ideal candidate for decompression alone, decompression with fusion, and the role of instrumentation in fusion procedures.[23] In 2014, reviews conducted for guidelines from the Congress of Neurologic Surgeons and the National Association of Spine Surgeons confirmed that these gaps were still unresolved.[24-26] For example, while there is agreement that fusion has a role when degenerative lumbar spondylolisthesis is advanced or there is evidence of instability, it is not often indicated when there is no spondylolisthesis or evidence of instability. There is disagreement about the need for fusion in patients with spinal stenosis who have grade I DLS, as well as evidence that surgeons cannot always identify which patients will benefit from fusion with decompression versus decompression alone. There is lack of consensus among spine surgeons on whether fusion should be added to decompression, particularly for low-grade DLS,[3] and the best fusion methods, as the type of fusion and use of instrumentation are controversial.[6,27-29] Minimally invasive and open decompression techniques have been used for low-grade DLS. Fusion is intended to alleviate symptoms in part by preventing vertebral movement or is necessary when a facet joint needs to be removed for adequate decompression. Fusion methods include posterior and posterolateral fusion (PLF), which uses bone graft placed along the posterior (back) or posterolateral (back and sides) of the spine to produce a fusion, as well as lumbar interbody fusion (LIF) techniques, which involve placing an implant, such as a cage containing bone graft, spacer, or structural graft, into the intervertebral space after discectomy and endplate preparation. Pedicle screws, rods or plates may be used in lumbar fusion to enhance spinal stability, limit motion during healing, and improve alignment and are associated with higher fusion rates.[30] Bone grafts and bone substitutes are an integral part of fusion surgery and are used to promote and maintain fusion and enhance intersegmental stability. There are few high-quality prospective studies comparing fusion methods for DLS.

Improving patient fusion outcomes may depend on patient selection, however, evidence on tests that may predict outcomes is limited.[31-33] Difficulty ascertaining whether degenerative changes are the source of pain and lack of consensus on radiographic assessment or use of invasive procedures present challenges to optimal patient selection. Evidence on procedures (e.g., facet joint blocks) for predicting fusion outcomes in CLBP is limited and suggests lack of prognostic ability.[7,32-35] Intraoperative neuromonitoring (IONM) (e.g., somatosensory, motor evoked potential measurements, spontaneous and triggered electromyography) may improve spine surgery safety by providing neural structure assessment and detecting neurological injury during surgery. The utility of IONM in spine tumor and spinal deformity surgeries is established, however, its utility in routine surgery for degenerative lumbar disease is less clear.[36-38] Clinically, IONM is standard for lateral LIF, but not for other interbody fusions.

1.1.3 Decisional Dilemmas

Several uncertainties underlie variation in surgical management and outcomes in patients with DLS.[20,39] In patients who are candidates for lumbar surgery, the primary decisional dilemma is to choose among surgical treatments that are likely to result in the best outcomes for individual patients with DLS and improve symptoms. Specific decisions include (1) whether to add fusion to decompression; (2) whether to use cages and additional instrumentation (e.g., pedicle screws) and concomitant decisions regarding choice of graft materials and surgical approach (e.g., anterior or posterior); (3) the use of minimally invasive procedures (e.g., facet joint injections) to predict fusion outcomes; (4) the use of IONM during fusion (e.g., reduction in iatrogenic neurological events); and (5) how population characteristics and differences (e.g., age, severity of disease, concomitant degenerative disease) may impact patient selection and outcomes. Another decisional dilemma is determining benefits and harms of epidural steroid injections, radiofrequency ablation and medial branch blocks for CLBP. There will be risk/benefit trade-offs associated with each surgical or nonsurgical intervention (e.g., potential symptomatic improvement and decreased risk of neurological sequelae with surgical treatment, but increased risk of surgical complications).

1.2 Rationale for Evidence Review

Many of these dilemmas are reflected in 2014 CNS clinical guidelines, which are now a decade old.[12,24,25] As the population ages, updated evidence-based guidance on the management of DLD and DLS reflecting new evidence becomes increasingly important. The goals of this systematic review are to systematically obtain, synthesize, and update published evidence on the effectiveness and harms related to the use of fusion in patients with DLS to facilitate resolution of the decisional dilemmas, enhance clinical decision making, and point to evidence gaps. We evaluated the effectiveness of specific nonsurgical procedures for treatment and for predicting surgical outcomes if fusion is done. We also assessed how effectiveness and harms may differ by patient, intervention and disease characteristics as data permit.

This review is funded by the Patient Centered Outcomes Research Institute (PCORI; the sponsoring partner) through a Memorandum of Understanding with the Agency for Healthcare Research and Quality (AHRQ). The Congress of Neurological Surgeons plans to utilize this review to inform updated clinical guidelines on surgical management of DLS and the role of IONM and use of procedures such as epidural steroid injections, radiofrequency ablation and medial branch blocks that may be employed to relieve pain due to DLD and/or select patients

who may best benefit from fusion surgery. The review will be of interest to a broad range of constituents, including patients, clinicians, guideline developers and policy makers.

2. Methods

2.1 Review Approach

The methods for this systematic review followed the Agency for Healthcare Research and Quality (AHRQ) Methods Guide for Effectiveness and Comparative Effectiveness Reviews (https://effectivehealthcare.ahrq.gov/products/cer-methods-guide/overview). This systematic review is in accordance with the Preferred Items for Reporting in Systematic Reviews and Meta-Analyses (PRISMA).[40]

2.1.1 Key Questions

Key Informants and a Technical Expert Panel provided comments on the scope of the review. The following Key Questions and inclusion criteria reflect suggestions received and are in the final protocol. The final protocol was posted on the Effective Health Care website on December 20, 2024 (https://effectivehealthcare.ahrq.gov/products/lumbar-spinal-fusion/protocol) and registered on PROSPERO (CRD42024625535).

Questions on surgery (Key Questions 1-4):

In adults with symptomatic, stable degenerative lumbar spondylolisthesis (DLS) with or without radiculopathy or neurogenic claudication:

> **Key Question 1.** What are the benefits and harms of surgery with instrumentation in addition to decompression compared with decompression alone?

In symptomatic adults with unstable or stable DLS with or without radiculopathy or neurogenic claudication undergoing instrumented fusion:

> **Key Question 2.** What are the benefits and harms of the addition of an interbody cage to instrumentation (e.g., pedicle screws) compared to use of instrumentation alone (i.e., posterolateral fusion)?

> **Key Question 3.** What are the benefits and harms of the use of bone graft extenders and biologic substitutes compared to the use of autografts?

In adults with symptomatic, degenerative lumbar spine disease undergoing instrumented fusion:

> **Key Question 4.** Does the use of intraoperative neuromonitoring (IONM) decrease perioperative neurological injuries compared with not using IONM?

Questions on nonsurgical procedures for chronic low back pain due to degenerative spine disease (Key Questions 5 and 6):

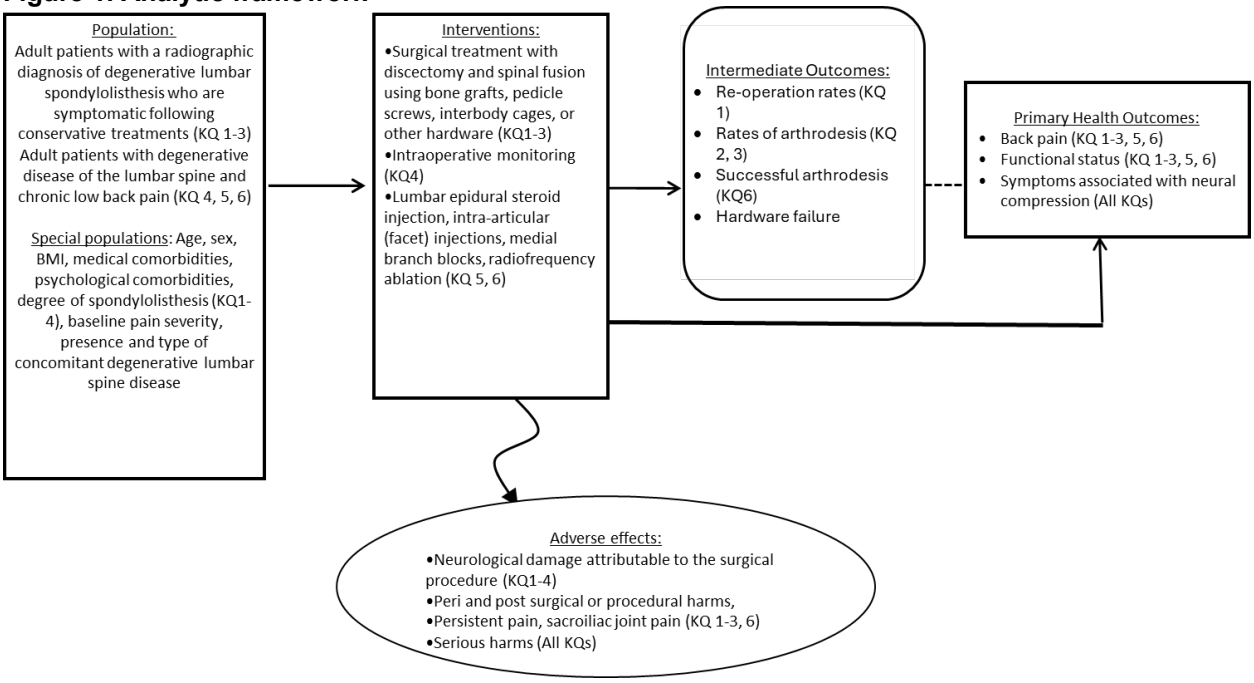
Key Question 5. In adult patients with chronic low back pain (≥3 months) resulting from degenerative disease what are the benefits and harms of lumbar epidural steroid injections, intra-articular (facet) injection, medial branch blocks, or radio frequency ablation?

Key Question 6. In adult patients with chronic low back pain (≥3 months) resulting from degenerative disease of the lumbar spine, does symptomatic improvement to therapeutic challenge with lumbar epidural steroid injections, intra-articular (facet) injection, medial branch blocks or radio frequency ablation predict positive outcomes after lumbar fusion surgery?

2.1.3 Analytic Framework

The analytic framework illustrates how the populations, interventions, and outcomes relate to the Key Questions in the review (**Figure 1**).

Figure 1. Analytic framework



Abbreviations: BMI = body mass index; KQ = Key Question

2.2 Study Selection

We searched Ovid® MEDLINE®, Cochrane Central Register of Controlled Trials, and Cochrane Database of Systematic Reviews for relevant studies published January 1, 2000 through October 23, 2024, for Key Questions 1-4; there was no start date restriction for Key

Questions 5 and 6 (nonsurgical). The year 2000 was chosen as a cut-off date after technical expert input and is due to the significant advancements in surgical techniques and instrumentation during that period (i.e., older studies may not accurately reflect current best practices or the effectiveness of newer procedures). Electronic literature searches will be updated while the draft report is posted for public comment to capture any new publications. All searches were conducted by a qualified medical librarian and were peer reviewed (See Methods, **Appendix A**).

In accordance with the *Methods Guide for Effectiveness and Comparative Effectiveness Reviews,*[41] we used the pre-established criteria in **Tables 1 and 2** to identify studies eligible for this review that were found through our searches or Supplemental Evidence and Data Submission (SEADS) entries. For all Key Questions, we focused on randomized controlled trials (RCTs), as well-conducted RCTs have the least risk of bias. Nonrandomized studies of interventions (NRSIs) in pain can be misleading, due to the subjective nature of pain, which may exacerbate effects of confounding, selection bias, and attentional and other nonspecific effects. We included comparative NRSIs that controlled for confounding to evaluate effectiveness only if no or very few RCTs were available (e.g., Key Question 4 and 6). For comparisons with sufficient RCT data, comparative NRSIs that controlled for confounding were considered for inclusion for evaluation of harms only. We excluded children/adolescents, asymptomatic patients, and patients having reoperations/revisions or repeat procedures for all Key Questions; forms of spondylolisthesis other than degenerative for Key Questions 1-3; and acute/subacute low back pain and disc herniation for Key Questions 5 and 6 (See Methods**, Appendix A, Table A-1** for detailed exclusion criteria). We used dual review to select studies. **Appendix A,** Methods, contains full details on review methods, including complete search strategies.

Table 1. Criteria for population, intervention, comparison, and outcomes of eligible studies for Key Questions 1, 2, 3, and 4 on surgery

PICOTS	Inclusion	Exclusion
Population	**Key Questions (1-3)** • Symptomatic adult patients with a radiographic diagnosis (based on dynamic (flexion and extension radiographs) of degenerative lumbar spondylolisthesis (any grade) who remain symptomatic following conservative treatment • Patients with evidence of nerve compression (radiculopathy, neurogenic claudication) **KQ 1** • Stable (non-mobile, static) DLS (<3 mm slip on extension/flexion radiographs) **KQ 2, 3** • Patients with unstable or stable DLS on radiographs **KQ 4** • Patients with symptomatic degenerative lumbar spine disease undergoing fusion of 5 or fewer levels (stratify by presence of DLS)	**ALL Key Questions** • Patients <18 years old • Asymptomatic patients • Other forms of spondylolisthesis are excluded (i.e., excluding dysplastic, isthmic, traumatic, and pathologic causes/forms) • Patients with osteoporosis, vertebral compression fractures • Exclude pts undergoing revisions or repeat procedures • Patients having reoperation/repeat procedures **KQs 1-3** • Patients without degenerative spondylolisthesis • Studies with <80% of patients have spondylolisthesis **KQ 1** • Patients with unstable (dynamic) DLS: (exclude study if stable is not specified, is unclear)

2. Methods

PICOTS	Inclusion	Exclusion
Interventions	**ALL Key Questions** • FDA approved devices or materials (or in Phase III trials) as applicable to the KQ • Open and minimally invasive (e.g., endoscopic) procedures **KQ 1** • Decompression (discectomy, indirect and direct methods) with instrumented spinal fusion (e.g., with pedicle screws, interbody cages, or other hardware) **KQ 2** • Surgical decompression and instrumented posterolateral fusion (e.g., using pedicle screws) with addition of interbody cage (expandable or static, ALIF, TLIF, LLIF). **KQ 3** • Decompression and spinal fusion using bone graft extenders or biologic substitutes (demineralized bone matrix, cadaveric allograft, cortical fibers, bone morphogenic protein, cellular allografts **KQ 4** • IONM (MEP, SSEP), Free Running EMG Direct Stimulation	**ALL Key Questions** • Devices or materials that are not FDA approved or in Phase III trials (as applicable to the question) or not available in the U.S. • Mesenchymal stem cells (MSCs) • Procedures that don't include decompression • Non-instrumented fusions • Coflex, interspinous fixation • Minimally invasive lumbar decompression (MILD) procedure • Surgical procedures not listed **KQ 4** • Other monitoring formats (e.g., imaging, computer assisted navigation systems, etc.) • Combinations of graft materials (other than with autograft) • Comparison of graft materials with each other
Comparators	**ALL Key Questions** • FDA approved devices or materials (or in Phase III trials) as applicable to the KQ • Open and minimally invasive (e.g., endoscopic) procedures **KQ 1** • Decompression alone **KQ 2** • Decompression and instrumented posterolateral spinal fusion (e.g., using pedicle screws alone) **KQ 3** • Decompression and instrumented spinal fusion using autograft **KQ 4** • No use of IONM	**ALL Key Questions** • Conservative care, non-operative care, usual care • Devices or materials that are not FDA approved or in Phase III trials (as applicable to the question) or not available in the U.S. • Mesenchymal stem cells **KQ 1** • Other surgical procedures **KQ 2, 3** • Non-instrumented fusion, • Instrumentation prior to 2000 • Coflex, interspinous fixation **KQ 3** • Combinations of graft materials with autograft

2. Methods

PICOTS	Inclusion	Exclusion

Outcomes	**ALL Key Questions:**
	• Validated measures for pain and symptoms
	  ○ Pain (e.g., VAS)
	  ○ Function (e.g., ODI)
	  ○ Quality of Life (e.g., SF-36, SF-12)
	• Peri- and post-operative harms (including serious AEs/harms[a], persistent pain, sacro-iliac joint pain, instrument failure)
	
	Additional outcomes by KQ
	KQ 1: Reoperation rates
	
	KQ 2: Fusion (arthrodesis) rates
	
	KQ 3: Fusion (arthrodesis) rates
	
	KQ 4: Persistent neurological damage based on clinical exam (e.g., foot drop)
		ALL Key Questions:
		• Measures of pain, function that are not validated
		• Measures/outcomes not listed
		• Radiographic parameters (e.g., evidence of global spinal alignment)

Timing	**Key Questions 1-3** • Pain, function, reoperation: 3, 6 and ≥ 12 months (up to 60 months) • Reoperation-any time **(KQ 2):** • Harms: any time **KQ 4** • Any time during post-operative followup	**KQ 1** • Re-operation beyond 12 months **KQs 1-3** • Outcomes measured less than 3 months (except harms) **KQ 4** • Alerts and responses to alerts during surgery
Settings	**ALL Key Questions** • Inpatient care followed by care in specialty and primary care clinics • Outpatient ambulatory surgery centers	None listed
Study designs	**ALL Key Questions** • RCTs for effectiveness/efficacy outcomes. • FDA SSED if there is inadequate information from published studies • Studies published in 2000 or later **KQ 1-3:** NRSIs will be considered for harms only and must be specifically designed to evaluate/report on AE/harms and control for confounding and focused on rare or long-term harms. **KQ 4:** NRSIs on effectiveness and harms	**ALL Key Questions** • NRSI that do not control for confounding • NRSI that include historical controls • NRSI of treatment with fewer than 50 patients per treatment arm • Case reports, case-series, single-arm and pre-post studies • **Publication types:** Conference abstracts or proceedings, editorials, letters, white papers, citations that have not been peer-reviewed, single site reports of multi-site studies • Studies published prior to 2000 • Studies not in English **KQ 1-3** • Trials with fewer than 15 patients per treatment arm

Serious adverse events are defined as events that are life-threatening or require additional medical attention.
Abbreviations: AE = adverse event; ALIF = anterior lumbar interbody fusion; DLS = degenerative lumbar spondylolisthesis; EMG = electromyography; FDA = U.S. Food and Drug Administration; IONM = intraoperative neuromonitoring; KQ = Key Question; LLIF = lateral lumbar interbody fusion; MEP = motor evoked potentials; NRS = numerical rating scale; NRSI = nonrandomized studies of interventions; ODI = Oswestry Disability Index; RCT = randomized controlled trial; SSED = Summary of Safety and Effectiveness Data; SSEP = somatosensory evoked potentials; SF-36/12 = Short Form 36 or 12 questionnaire; TLIF = transforaminal lumbar interbody fusion; VAS = visual analog scale.

Table 2. Criteria for population, intervention, comparison, and outcomes of eligible studies for Key Questions 5 and 6 on specific procedures

PICOTS	Inclusion	Exclusion
Population	**KQ 5, 6** • Adult patients with chronic non back pain (≥3 months duration) resulting from degenerative disease	**KQ 5, 6** • Patients with acute or subacute LBP • Patients with disc herniation • Patients with failed back surgery syndrome • Sacroiliac pain • Patients having reoperation
Interventions	**KQ 5, 6** • ESI • Intra-articular (facet) injections • RFA • Medial branch blocks	**KQ 5, 6** • Discoblock, provocative discography • Neuromodulation (e.g., spinal cord, dorsal column, dorsal root stimulation, peripheral nerve stimulation) • Injections: exclude other biologics (e.g., PRP), intradiscal injections • Minimally invasive lumbar decompression (MILD), percutaneous decompression • Selective nerve root blocks • Intraosseous basivertebral nerve ablation • Combinations of procedures; Studies evaluating additive benefits of one procedure with another
Comparators	**KQ 5** • Other nonoperative treatment, no treatment, sham **KQ 6** • No therapeutic challenge; (prognostic/predictive modeling study)	**KQ 5, 6** • Combinations of procedures; Studies evaluating additive benefits of one procedure to another **KQ 5** • For ESI, exclude comparison with disc procedures (e.g., discography); comparisons of medications • For RFA exclude comparisons of different types of neurotomy (conventional vs. pulsed [cooled] RFA; RFA vs. alcohol ablation)

2. Methods

PICOTS	Inclusion	Exclusion

Outcomes	**KQ 5 and 6: Harms** (e.g., serious peri-procedural and post-procedural harms)
	
	KQ 5
	• Validated measures for pain and symptoms
	  ○ Pain (e.g., VAS, NRS)
	  ○ Function (e.g., ODI)
	  ○ Quality of Life (e.g., SF-36, SF12)
	
	KQ 6
	• Response to challenge: Improvement in symptoms vs. non-improvement; [stratify other outcomes by response]
	• Validated measures for pain and symptoms following fusion surgery
	  ○ Pain (e.g., VAS, NRS)
	  ○ Function (e.g., ODI)
	  ○ Quality of Life (e.g., SF-36, SF-12)
	  ○ Symptoms associated with neural compression
	• Successful arthrodesis [as radiographically determined via x-ray/computed tomography or by proxy (e.g., lack of revision, pedicle screw loosening)]
		KQ 5, 6
		• Measures of pain, function that are not validated
		• Measures/outcomes not listed

Timing	**KQ 5 and 6** • Serious harms - periprocedural **KQ 5** • 3-month and 6-month periods following the procedure **KQ 6** • Outcomes measured at 3, 6 and ≥ 12 months after surgical procedure (up to 24 months)	None listed
Settings	**KQ 5** • Outpatient **KQ 6** • Outpatient care for therapeutic challenge. Inpatient care followed by care in specialty and primary care clinics for surgical procedure • Outpatient ambulatory surgery centers for surgery	None listed

PICOTS	Inclusion	Exclusion
Study designs	**KQ 5** • RCTs for effectiveness/efficacy outcomes • Prospective NRSIs that control for confounding will be considered for effectiveness in the absence of RCTs • NRSIs for harms must be specifically designed to evaluate/report on serious AE/harms and that control for confounding OR focused on rare or long-term harms **KQ 6** • Predictive/prognostic modeling studies evaluating the association of procedure response impact on outcomes that control for confounding	**KQ 5, 6** • NRSI that do not control for confounding • NRSI that include historical controls • NRSI with fewer than 50 patients per treatment arm • Case reports, case-series, single-arm and pre-post studies • Publication types: Conference abstracts or proceedings, editorials, letters, white papers, citations that have not been peer-reviewed, single site reports of multi-site studies • Studies not in English

Serious adverse events are defined as events that are life-threatening or anything needing additional medical attention.
Abbreviations: AE = adverse event; ESI = epidural steroid injection; FDA = Food and Drug Administration; KQ = Key Question; LBP = low back pain; NRS = numerical rating scale; NRSI = nonrandomized studies of intervention; ODI = Oswestry Disability Index; PRP = platelet rich plasma; RCT = randomized controlled trial; RFA = radiofrequency ablation; SSED = Summary of Safety and Effectiveness Data; SF-36/12 = Short Form 36 or 12 questionnaire; VAS = visual analog scale.

2.3 Data Extraction and Risk of Bias Assessment

Data were abstracted from included studies into evidence tables based on the organizational framework to include study, patient, and intervention/procedure characteristics, degree/grade of degenerative lumbar spondylolisthesis (DLS) (surgery questions), and study results (including harms), with data verified for accuracy and completeness by a second team member

The risk of bias of included studies was assessed according to established methods,[41,42] with RCTs assessed based on criteria and methods established in the *Cochrane Handbook for Systematic Reviews of Interventions (Chapter 8.5 Risk of Bias Tool)*[43] and tenets for appraisal developed by the Cochrane Back and Neck Group.[44] We assessed risk of bias using instruments tailored to observational studies[45] that consider patient selection methods (e.g., consecutive patients, use of a randomized selection of patients, an inception cohort) and appropriate control for confounding by relevant prognostic factors. We used dual review for the risk of bias assessment and individual included studies were rated as being "low," "moderate," or "high" risk of bias. We downgraded studies that did not blind participants (or providers), had a high rate of loss to followup (>20%), or demonstrated selective reporting or other bias. We focused on studies with the least potential for bias and the fewest limitations. Full details on data abstraction, data management, and risk of bias assessment (i.e., quality determination) can be found in **Appendix A**, Methods.

2.4 Data Synthesis and Analysis

We analyzed RCTs and NRSIs separately and reported them separately unless findings were very consistent across study designs and the studies were clinically homogeneous. Summary tables were constructed when appropriate to highlight the main findings.

Meta-analyses, using profile-likelihood random effects models, were conducted to summarize data and obtain more precise estimates where there were at least two studies reporting outcomes that were homogeneous enough to provide a meaningful combined estimate.[46,47] We

considered clinical and methodological diversity and assessed statistical heterogeneity using Cochran's χ^2 test and the I^2 statistic.[48] For binary outcomes, a risk ratio was used as the effect measure. For continuous outcomes, mean difference (MD) was used as the effect measure as the studies reported outcomes using the same scale, or the outcomes could be converted to the same scale (e.g., pain, converted to 0-10 scale). MD was calculated using the followup score if reported and the change score from baseline if followup scores were not reported. Sensitivity analyses were performed to explore statistical heterogeneity and differences by study quality. There were insufficient data to do subgroup analyses based on intervention differences, patient characteristics, or other factors. **Appendix A**, Methods, contains additional detail of our meta-analysis methods. We classified the magnitude of effects for continuous measures of pain and function using the same system as in prior AHRQ reviews on pain.[49-53] Effects below the threshold for small were categorized as no effect. There were too few studies (<10) included in the analyses to test for small sample effects (a marker for potential publication bias).[54]

2.5 Grading the Strength of the Body of Evidence

The strength of evidence (SOE) of primary outcome-intervention pairs were evaluated using AHRQ methods.[41] Details on the methods used are presented in the Methods **Appendix A**, and primary outcomes are delineated in **Tables 1 and 2**, above. The SOE was assigned an overall grade of high, moderate, low, or insufficient by evaluating and weighing the combined results of the following five domains: study limitations, consistency, directness, precision, and reporting bias. RCT evidence was initially considered high, with possible downgrades for any of these domains. For NRSIs, the strength started at moderate for harms outcomes, and low for benefit outcomes. While AHRQ guidance allows for upgrading NRSI evidence in certain circumstances, no upgrading was deemed appropriate. When both RCTs and NRSIs were included for a given outcome, we followed AHRQ guidance for consideration of consistency and weighing of RCTs over observational studies after evaluating each study type separately. We considered NRSI evidence to supplement RCT evidence to arrive at a final rating. We primarily used RCT evidence as that from NRSIs was of lower strength. For bodies of evidence with only a single study, we rated consistency as unknown (rather than not applicable). In these cases, we did not automatically downgrade the evidence to "insufficient" but considered the sample size or number of events available for analysis. If only poor-quality trials were available for a given outcome, SOE was considered insufficient.

3. Results

A total of 8,291 abstracts were identified, 8,278 from electronic database searches and an additional 13 from handsearching and bibliography review of included studies and systematic reviews. After dual review of titles and abstracts, 722 articles were selected for full-text review, of which 48 studies (in 58 publications) were ultimately included in this review: 35 randomized controlled trials (RCTs) (in 43 publications)[55-97] and 13 comparative nonrandomized studies of interventions (NRSIs) (in 15 publications).[98-112] For the surgical question, five RCTs (in 9 publications)[55,62,63,67,69,70,72,73,75] and seven NRSIs (in 8 publications)[101-103,106,108,109,111,112] were included in Key Question 1, but only one RCT[67] enrolled participants in a stable DSL population. The number of included studies for the remaining surgical Key Questions is as follows: Key Question 2 (2 RCTs,[57,62] 4 NRSIs[98,99,105,107,110]); Key Question 3 (1 RCT);[71] and Key Question 4 (1 NRSI).[100] For the nonsurgical/selective procedures questions, the most evidence was identified for radiofrequency ablation (17 RCTs)[58-61,66,68,78,79,87-89,91,93-97] for Key Question 5, followed by epidural steroid injections (5 RCTs in 8 publications),[64,65,76,77,81-83,90] facet joint injections (4 RCTs)[56,74,80,92] and medial branch block (2 RCTs in 3 publications).[84-86] Only one NRSI[104] was identified that met inclusion criteria for Key Question 6. Eight RCTs were assessed as low risk of bias (22.9%),[56,78,79,81,86,92,94,96] 18 as moderate risk of bias (51.4%),[55,57-60,63,67,68,70,71,82-84,88,89,91,93,95] and nine as high risk of bias (25.7%).[61,62,66,74,76,80,87,90,97] One NRSI was assessed as low risk of bias (7.7%),[100] 11 as moderate risk of bias (84.6%),[98,101,103-109,111,112] and one as high risk of bias (7.7%).[110] Search results and selection of studies are summarized in the literature flow diagram (**Appendix B, Figure B-1**). A list of included studies appears in **Appendix C** and excluded studies with reason for exclusion in **Appendix G**.

3.1 Key Question 1. Benefits and harms of surgery with instrumentation in addition to decompression compared with decompression alone for stable degenerative lumbar spondylolisthesis

3.1.1 Key Points

- Across five RCTs, there was no clear difference between decompression and fusion compared with decompression alone for improvement in back pain, leg pain, function, quality of life, frequency of serious adverse events, dural tears, deep wound infection, pulmonary embolism, deep vein thrombosis, or a multicomponent outcome of heart attack stroke, or thromboembolic event (SOE: low).
- The Spinal Laminectomy versus Instrumented Pedicle Screw (SLIP) trial, a small RCT that specifically enrolled patients with stable grade I degenerative lumbar spondylolisthesis (DLS), defined as <3 mm of slip, found slightly improved function at 4 years with decompression and fusion versus decompression alone (SOE: low)
- While the SLIP trial found slightly improved quality of life based on SF-36 at 6 months and at 2, 3, and 4 years with decompression and fusion compared with decompression alone, the Nordsten trial in predominantly stable DLS and the Swedish Spinal Stenosis study in patients with unknown DLS stability reported similar quality of life improvement on the EQ-5D across all times (SOE: low).

- Across all RCTs, the data were insufficient to determine the benefits and harms of decompression and fusion versus decompression alone on implant failure and progression of spondylolisthesis, mortality, persistent back pain, and stroke (SOE: insufficient).
- One-year reoperation rates were similar between decompression and fusion versus decompression alone at 1 month (1 NRSI) and 3 months (2 NRSIs). At 12 months, evidence from 2 NRSIs found that those who received decompression and fusion were slightly less likely to have another surgery, however it was unclear whether the reoperation occurred at the index level and/or an adjacent level (SOE: low).
- One RCT found self-reported neurological deterioration over 5 years (Nordsten trial) was slightly more likely with decompression and fusion versus decompression alone (SOE: low); data were insufficient at other timepoints.

3.1.2 Description of Included Studies

Five RCTs (reported in 9 publications)[55,62,63,67,69,70,72,73,75] and seven NRSIs (in 8 publications)[101-103,106,108,109,111,112] compared decompression plus fusion versus decompression alone for the treatment of degenerative spondylolisthesis (**Appendix D, Table D-1**). NRSIs were included for harms. Only one RCT (N=66), the SLIP trial, specifically enrolled participants with stable grade I DLS, defined as <3 mm slip on extension/flexion radiographs.[67] They report mean slippage at baseline of 6.1 mm. In the Nordsten trial (N=267),[75] about 80 percent of patients had <3 mm forward translation and mean slip at baseline was 7.6 mm however, patients with >3 mm slip were not excluded.[75] In the remaining trials, the proportion of participants who had <3 mm slippage was unclear. One trial reported a mean slip of 16.9 percent.[69] The remaining two trials did not report baseline slippage. Although four RCTs included mixed populations (stable and unstable DLS), they provide important context and information for clinicians.[55,62,63,69,70,72,73,75]

3.1.2.1 Randomized Controlled Trials

Across the RCTs, sample sizes ranged from 66 to 267 (total N=673). The average study mean age of participants was 66 years (range 63 to 68 years) among the four trials that reported a precise mean age.[55,63,67,69,70,72,73,75] One study reported that most participants were ages 30 to 40 years;[62] the average proportion of females in the trials was 69 percent (range 47% to 80%) among the four trials that reported the sex distribution.[55,62,63,67,72,73,75] No trial reported racial or ethnic distribution.

Four trials limited enrollment to participants with lumbar canal stenosis in addition to DLS;[55,63,67,69,70,72,73,75] one trial reported that some participants had lumbar stenosis and others had disk herniation in addition to spondylolisthesis but did not provide the proportions of these.[62] Three studies reported one vertebral-level intervention;[55,67,69,70,75] one trial reported participants had involvement of one or two levels;[63,72,73] and one trial reported that 81 percent (97/120) of participants had treatment at least two levels.[62]

The interventions assessed in RCTs consisted of different methods of decompression and fusion compared with fusion alone. Across RCTs, the following interventions were identified: classic posterior laminectomy and posterior instrumented fusion, transforaminal interbody fusion, and laminectomy with noninstrumented fusion. Control procedures included standard laminectomy and midline sparing laminectomies, via ipsilateral or contralateral approaches (**Appendix B, Table B-2**).

Baseline mean Oswestry Disability Index (ODI) scores (scale 0 to 100, higher is worse) ranged from 37.5 to 41.0 across three trials that reported ODI scores.[55,63,67,72,73,75] Mean back

pain and leg pain scores at baseline on a 0-10 Numerical Rating Scale (NRS, higher is worse) were 6.7 and 6.7 in one trial;[55,75] mean back and leg pain scores on a 0-100 Visual Analog Scale (VAS, higher is worse) reported in two trials were 63.5 and 57.8 for back pain and 64.5 and 71.4 for leg pain.[63,69,70,72,73]

One trial was conducted in the United States[67] and one trial each was conducted in Norway,[55,75] Sweden,[63,72,73] Japan,[69,70] and Iran.[62] Three of the five trials did not report funding; the remaining two trials reported funding from nonprofit or charitable funds and trusts with some funding coming from public grants[75] and research grants.[67]

One trial (n=120) conducted in Iran[62] was rated high risk of bias due to unclear randomization and/or allocation concealment, inadequate information on treatment group similarity at baseline and lack of reporting of attrition.[62] The remaining trials,[55,63,67,69,70,72,73,75] with fewer methodological limitations, were rated moderate risk of bias (**Appendix E, Table E-1**).

3.1.2.2 Nonrandomized Studies of Intervention

None of the NRSIs reported the proportion of participants with DLS of <3 mm slip. They are included for completeness to provide context and information for clinicians regarding harms. Two were retrospective cohorts,[105,110] two were population-based studies,[107,112] and data from one was from an RCT but only nonrandomized data for the comparison of interest here was eligible for inclusion.[99]

Across the seven NRSIs, sample sizes ranged from 102 to 75,024 (total N=78,586). The average study mean age of participants was 65 years (range 53 to 73 years); the average proportion of females in the NRSIs was 63 percent (range 50% to 72%) among the six NRSIs that reported the proportion females.[101-103,106,109,111,112] The three studies from the United States and Canada reported racial breakdown with 78 to 81 percent White participants.[103,106,112]

Almost all participants had lumbar stenosis (i.e., 99% to 100%) across the six NRSIs that reported baseline stenosis.[101-103,106,108,111,112] Two NRSIs reported one level involvement,[103,108] one reported one or two level involvement,[101,102] one reported one or two levels in participants who received bilateral partial laminectomy but one level involvement in those who received posterior lumbar interbody fusion (PLIF),[109] one study reported treatment at less than or equal to three levels,[111] and one NRSI reported involvement at one to three or more levels.[106]

Detail of intervention and control procedures varied across studies. Across NRSIs, the following interventions were identified: classic posterior laminectomy and posterior instrumented fusion, transforaminal interbody fusion, anterior lumbar interbody fusion with traditional pedicle screw constructs, and laminectomy. Control procedures included standard open laminectomy and midline sparing or minimally invasive laminectomy or laminotomy, via ipsilateral or contralateral approaches (**Appendix B, Table B-3**).

Two NRSIs were conducted in the United States[106,112] and one each was conducted in Switzerland,[111] Canada,[103] Norway,[101,102] The Netherlands,[108] and Japan.[109] Most NRSIs did not report any funding information. One NRSI reported that it received no funding,[103] one reported funding from Kaiser Permanente grant funds,[106] and the largest NRSI reported funding from the Agency for Healthcare Research and Quality (AHRQ).[112] Propensity-score matching was used by four NRSIs,[101,102,106,108,109] while the remainder conducted regression analyses with potential confounders as covariates.

3.1.3 Results, Key Question 1. Instrumentation in addition to decompression compared with decompression, Detailed Synthesis

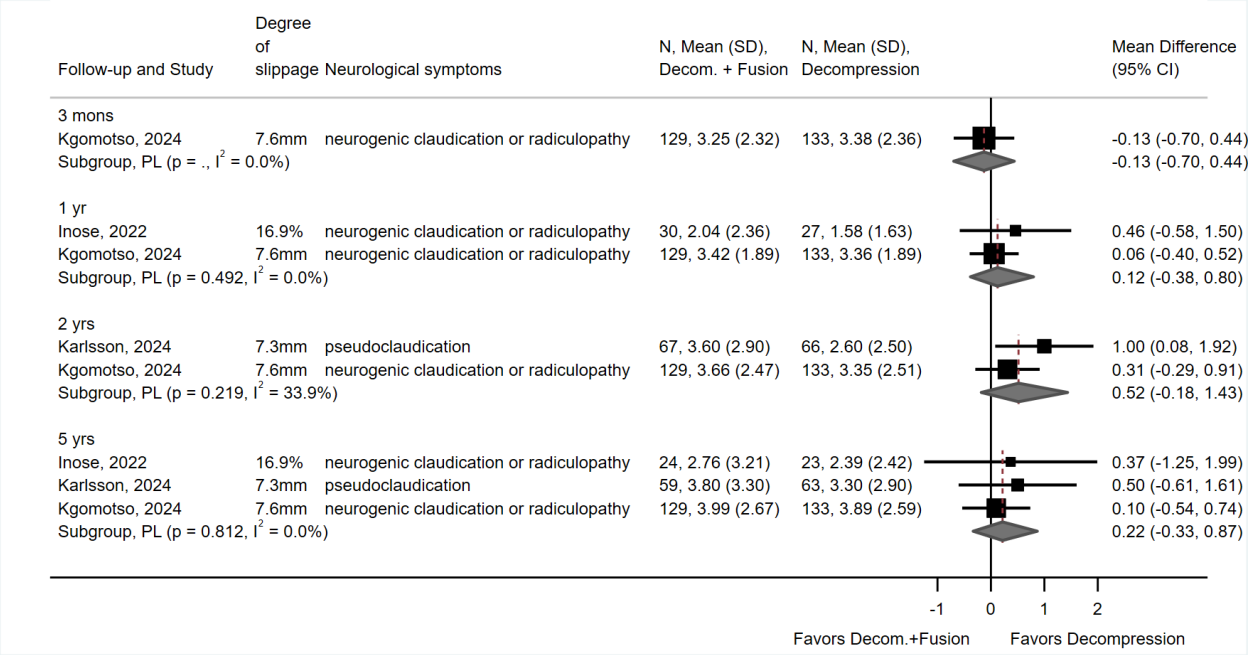
All NRSIs were rated moderate risk of bias (**Appendix E, Table E-2**). Methodological limitations included differences between groups at baseline, lack of reporting of attrition, and unclear blinding of outcome assessors and/or data analysts.

3.1.3 Detailed Synthesis

3.1.3.1 Back Pain

Three RCTs (N=452) (in 7 publications)[55,63,69,70,72,73,75] that compared decompression plus fusion versus decompression alone reported back pain from 3 months to 5 years. Back pain improvement with decompression and fusion compared with decompression on a 0 to 10 scale was similar at 3 months (1 RCT, N=262, MD -0.13, 95% CI -0.70 to 0.44), 1 year (2 RCTs, N=319, MD 0.12, 95% CI -0.38 to 0.80, I^2=0%)[75] and 5 years (3 RCTs, N=431, MD 0.22, 95% CI -0.33 to 0.87, I^2=0%) (**Figure 2**).[69,72,75] At 2 years, no difference was seen between treatments (2 RCTs, N=395, MD 0.52, 95% CI -0.18 to 1.43, I^2=33.9%).[72,75] The SLIP trial did not report measures of pain improvement.[67]

Figure 2. Decompression plus fusion versus decompression alone: back pain scores



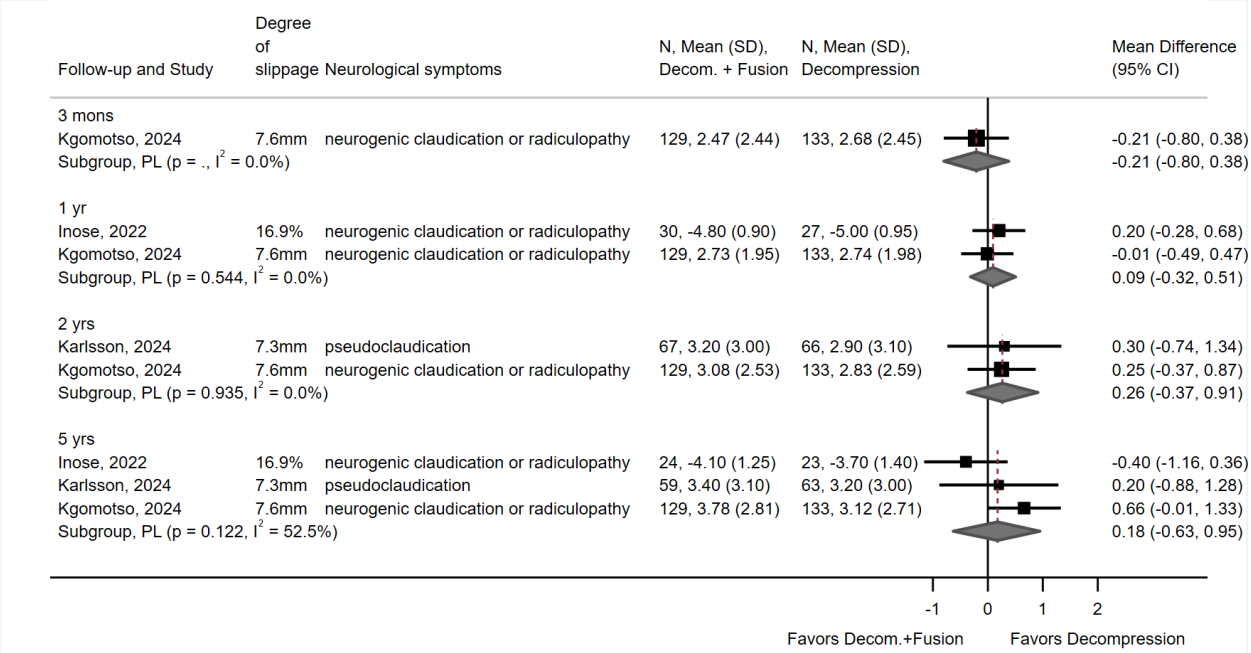
Abbreviations: CI = confidence interval; PL = profile likelihood; SD = standard deviation

3.1.3.2 Leg Pain

Three same three RCTs described above (N=452) (in 7 publications)[55,63,69,70,72,73,75] that compared decompression plus fusion versus decompression alone reported leg pain from 3 months to 5 years. The degree of leg pain with decompression and fusion compared with decompression on a 0 to 10 scale was similar at all time points: 3 months (1 RCT, N=262, MD -0.21, 95% CI -0.80 to 0.38),[75] 1 year (2 RCTs, N=319, MD 0.09, 95% CI -0.32 to 0.51, I^2=0%),[69,75] 2 years (2 RCTs, N=395, MD 0.26, 95% CI -0.37 to 0.91, I^2=0%)[72,75] and 5 years (3 RCTs, N=431, MD 0.18, 95% CI -0.63 to 0.95, I^2=52.5%) (**Figure 3**).[69,72,75]

3.1.3 Results, Key Question 1. Instrumentation in addition to decompression compared with decompression, Detailed Synthesis

Figure 3. Decompression plus fusion versus decompression alone: leg pain scores



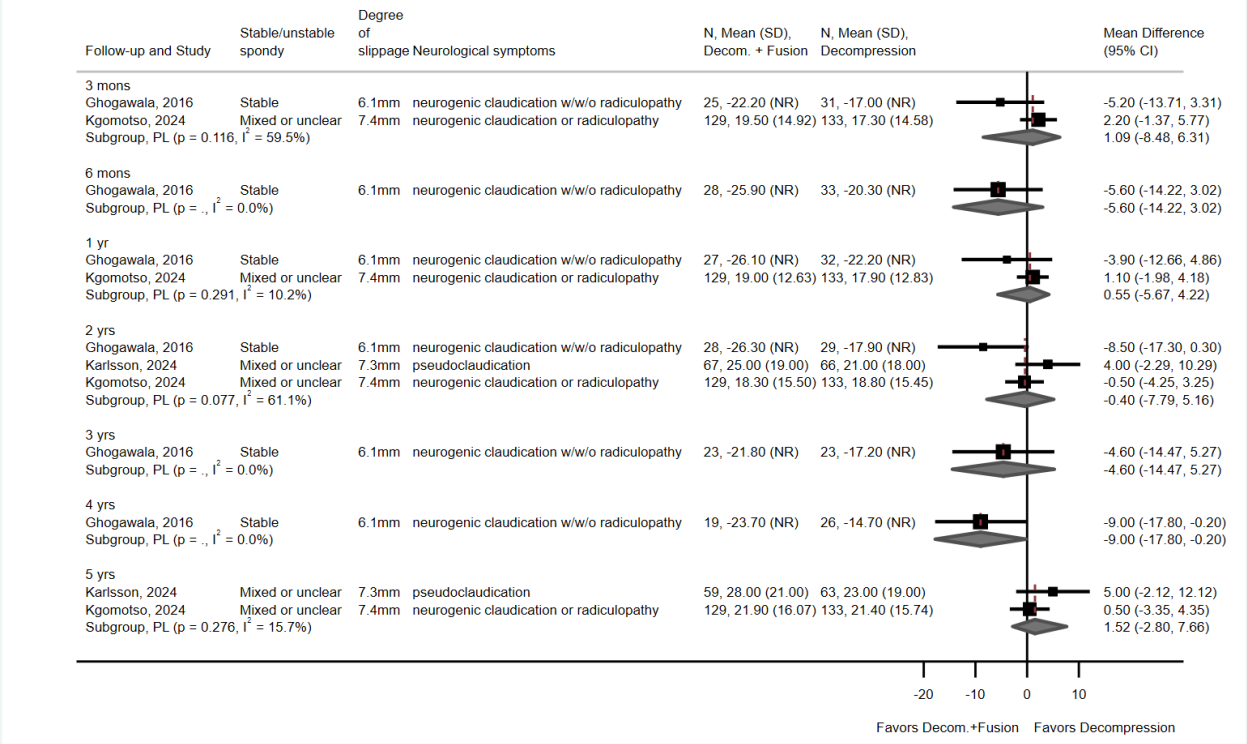
Abbreviations: CI = confidence interval; PL = profile likelihood; SD = standard deviation

3.1.3.3 Function

The SLIP (n=66)[67] and Nordsten (N=267)[55,75] trials compared decompression plus fusion versus decompression alone in patients with *stable* DLS and reported ODI scores. The Swedish Spinal Stenosis Study (SSSS) trial (3 publications)[55,63,72,73,75] that reported results in *unknown* DLS *stability* was also included in the pooled analyses (**Figure 4**). There were no differences in ODI scores (scale 0-100) between treatment arms at 3 months (2 RCTs, N=318, MD 1.09, 95% CI -8.48 to 6.31, I^2=59.5%),[67,75] 1 year (2 RCTs, N=321, MD 0.55, 95% CI -5.67 to 4.22, I^2=10.2%),[67,75] 2 years (3 RCTs, N=452, MD -0.40, 95% CI -7.79 to 5.16, I^2=61.1%),[67,72,75] 3 years (1 RCT, N=46, MD -4.60, 95% CI -14.47 to 5.27),[67] and 5 years (2 RCTs, N=196, MD 1.52, 95% CI -2.80 to 7.66, I^2=15.7%).[72,75] There was no difference in ODI scores from the single RCT in patients with *stable* DLS[67] at 3 months, 6 months, 2 years, or 3 years followup between the treatment groups; results were imprecise but tended to favor decompression with fusion (1 RCT, N=61, MD -5.60, 95% CI -14.22 to 3.02).[67] At 4 years followup, decompression and fusion were associated with a small improvement in function, but the estimate was imprecise (1 RCT, N=45, MD -9.00, 95% CI -17.80 to -0.20).[67]

3.1.3 Results, Key Question 1. Instrumentation in addition to decompression compared with decompression, Detailed Synthesis

Figure 4. Function ODI (0-100): decompression plus fusion versus decompression alone (RCTs)



Abbreviations: CI = confidence interval; NR = not reported; ODI = Oswestry Disability Index; PL = profile likelihood; RCT = randomized controlled trial; SD = standard deviation

The Nordsten trial (N=267) in patients with *stable* DLS also reported the proportion of patients who achieved a 30 percent or more reduction in baseline ODI score at 5 years that did not favor either decompression with fusion or decompression alone (64.5% vs. 63.4%, difference in percentage points: -1.1, 95% CI -12.9 to 10.9).[55,75] Another study in patients with *unknown* DLS *stability* (Fariborz, N=120) reported that a higher proportion of patients had followup ODI scores of 0 to 20 percent (mild disability) in the three fusion groups (PLIF, TLIF, PLF with pedicle screws 70, 63.4, and 66.6 percent, respectively), compared with 0 percent in patients who underwent a decompression without fusion or instrumentation ($p<0.001$) at 6 months post-operatively.[62] Results were similar at 12 months for PLIF, TLIF, and posterior lumbar fusion compared with decompression alone, where no patients scored in the 0 to 20 percent range (i.e., mild disability) on the ODI (83.4% vs. 80% vs. 83.4% vs. 0%, respectively, $p<0.001$).

One RCT (Inose, N=60) in patients with *unknown* DLS *stability* reported modified Japanese Orthopaedic Association scores which were similar between those who had decompression with fusion and those receiving decompression only at 1 and 5 years (data shown graphically in publication, $p>0.05$).[69,70]

3.1.3.4 Quality of Life

The SLIP RCT (n=66)[67] reported small improvements on the SF-36 Physical Component Summary (0-100 scale) at various timepoints with decompression and fusion compared with decompression alone, with meaningfully improved scores (at least 5 point change) at 6 months (mean difference of change scores: 1 RCT, N=61, MD 6.40, 95% CI 1.10 to 11.70), 2 years (1 RCT, N=57, MD 5.70, 95% CI 0.10 to 11.30), 3 years (1 RCT, N=46, MD 7.40, 95% CI 1.10 to 13.70), and 4 years (1 RCT, N=45, MD 6.70, 95% CI 1.20 to 12.30). Change scores at 3 months

(1 RCT, N=56, MD 4.50, 95% CI -0.70 to 9.70) and at 1 year (1 RCT, N=59, MD 3.90, 95% CI -1.50 to 9.40) were below the threshold for a small effect and not statistically significant (**Appendix B, Figure B-2**).

Two RCTs (N=397) (in 5 publications)[55,63,72,73,75] found no difference in quality of life scores between decompression and fusion and fusion alone at any timepoint on the EuroQol-5 Dimension/3 level version (EQ-5D/EQ-5D-3L): 3 months (1 RCT, N=262, standardized mean difference [SMD] -0.08, 95% CI -0.32 to 0.17),[75] 1 year (1 RCT, N=262, SMD 0.00, 95% CI -0.24 to 0.24),[75] 2 years (2 RCTs, N=395, SMD -0.02, 95% CI -0.38 to 0.26, I^2=37.9%),[72,75] and 5 years (2 RCTs, N=384, SMD -0.09, 95% CI -0.44 to 0.19).[72,75] One trial measured quality of life with the EQ-5D-3L (range -0.59 to 1.0)[72] and the other used the EQ-5D (range 0 to 1) (**Appendix B, Figure B-3**).[75]

3.1.3.5 Reoperation Rates

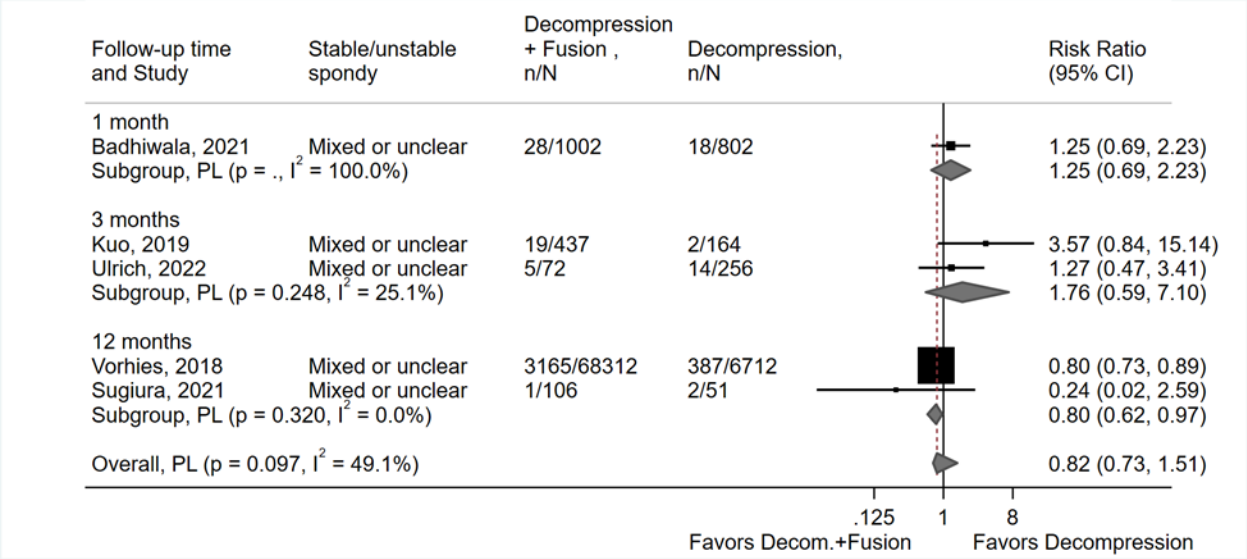
The SLIP trial (n=66) in patients with *stable,* grade I DLS[67] had similar rate of reoperation at 12 months with decompression and fusion versus decompression alone (0% vs. 0.05%, RR 0.19, 95% CI 0.01 to 3.77). Reasons for reoperation were not reported and were done at the surgeons' discretion. Within a 4-year period, authors report 14 reoperations, 10 of which occurred at the index level in patients receiving decompression alone to address clinical instability. Reoperation in the fusion group was at an adjacent segment. There was, however, differential loss to follow up at 4 years (26% vs. 39%) so results should be interpreted cautiously.

Two RCTs (N=352) (in 4 publications)[55,69,70,75] also reported reoperation rates within 12 months postoperatively. The Nordsten trial (N=267) reported that seven reoperations (6%) occurred with decompression and fusion versus two operations (2%) with decompression alone within the first 3 months after surgery, without specifying the nature of the reoperations.[55,75] Another RCT (N=85) reported that there were no revision surgeries with decompression and fusion or with decompression alone by 12 months after surgery.[69,70] Similar rates of reoperation were reported for the treatment groups in the Nordsten trial from 2 to 5 years (6 vs. 11 patients), which predominately enrolled patients with <3 mm slip,[55,75] and in the SSSS trial at 6.5 years (22% versus 21%) in patients with *unknown* DLS *stability*.[63,72]

Five NRSIs (N=77,914)[103,106,109,111,112] in patients with *unknown* DLS *stability*, reported reoperation rates in patients with decompression plus fusion compared with decompression alone at 12 months or less following index surgery. Pooled analysis of these studies found a similar risk of reoperation at 1 month postoperatively (1 NRSI, N=1,804, 2.8% vs. 2.2%, RR 1.25, 95% CI 0.69 to 2.23)[103] and a similar risk at 3 months (2 NRSIs, N=929, 4.7% vs. 3.8%, RR 1.76, 95% CI 0.59 to 7.10, I^2=25.1%)[106,111] with decompression and fusion compared with decompression alone; however the risk of reoperation at 12 months was slightly lower with decompression plus fusion (2 NRSIs, N=75,181, 4.6% vs. 5.8%, RR 0.80, 95% CI 0.62 to 0.97, I^2=0%) (**Figure 5**).[109,112] The effect estimate for reoperation at 12 months was largely driven by one large NRSI (n=75,024).[112] In this study, the second surgery for those who had decompression and fusion as the index surgery was another fusion (73%) compared with a decompression alone reoperation (27%) and for those who had decompression alone as the index surgery, 69 percent had a fusion reoperation and 31 percent had another decompression surgery. It was not specified whether additional surgeries involved the index and/or adjacent levels. Overall risk of reoperation, regardless of followup time, was similar with decompression and fusion versus decompression alone (5 NRSIs, N=77,914, 4.6% vs. 5.3%, RR 0.82, 95% CI 0.73

to 1.55, I^2=49.1%) (**Figure 5).**[103,106,109,111,112] Analyses were limited due to few studies at each followup time and estimates are imprecise.

Figure 5. Decompression plus fusion versus decompression alone: reoperation (NRSIs)



Abbreviations: CI = confidence interval; NRSI = nonrandomized studies of interventions; PL = profile likelihood

3.1.3.6 Harms and Adverse Events

3.1.3.6.1 Serious Adverse Events/Harms

At 1 month after surgery, the SLIP trial (n=66) in patients with *stable* DLS reported one patient (3.2%) with decompression plus fusion and two patients (5.7%) with decompression experienced any major complication (RR 0.57, 95% CI 0.05 to 5.93). While favoring decompression plus fusion, there were too few events with either treatment from which to draw conclusions.[67]

One NRSI (n=1,804) in patients with *unknown* DLS *stability* reported a similar risk of major complications with decompression and fusion compared with decompression alone at 1 month postoperatively, that was not statistically significant (2.9% vs. 2.0%, adjusted odds ratio [aOR] 1.51, 95% CI 0.81 to 2.81 (adjusted for age, sex, and modified Frailty Index).[103]

3.1.3.6.2 Mortality

One NRSI (N=1,804) in patients with *unknown* DLS *stability* reported that one patient treated with decompression plus fusion (0.1%) died compared with three patients (0.4%) who received decompression alone in unadjusted analysis (p=0.218) within the first 30 days following surgery.[103] The study authors reported that there were too few events to enable calculation of adjusted mortality.

One RCT, the SSSS trial (N=247) in patients with *unknown* DLS *stability* reported no deaths (0%) after 2 years in patients treated with decompression and fusion compared with two patients (1.6%) treated with decompression alone.[63] The same trial reported mortality at 5 years; there were four deaths among patients treated with decompression and fusion (3.3%) versus two deaths (1.6%) with decompression alone.[72] However, due to few deaths during the study, and the possibility that the deaths were not related to the surgery at longer followup times, it is unclear that there is any meaningful difference between treatments in mortality.

3.1.3.6.3 Persistent Pain or Sacroiliac Joint Pain

One RCT (n=247) reported that no patients treated with decompression and fusion experienced severe pain that required reoperation within the 2 years after index surgery compared with two patients (1.6%) who required reoperation for severe back pain among those treated with decompression alone.[63] Due to few patients having persistent severe pain, this difference is likely not meaningful.

No studies reported sacroiliac joint pain.

3.1.3.6.4 Implant Failure and Progression of Spondylolisthesis

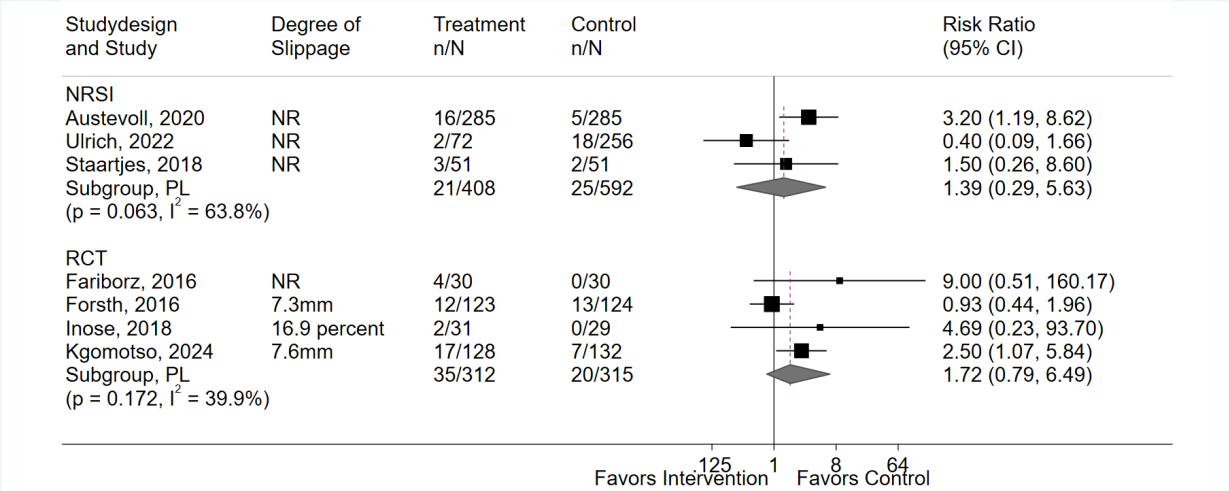
One NRSI (n=102) reported there were no implant failures with 51 decompressions and fusions.[108] This same trial also reported that one out of 51 patients who received decompression alone and none out of 51 patients who received decompression and fusion experienced progression of spondylolisthesis. Events for both outcomes were two few for estimates to be clinically meaningful.

3.1.3.6.5 Dural Tear/Puncture

Four RCTs (N=627) (in eight publications)[55,62,63,69,70,72,73,75] and three prospective NRSIs (N=1,000) in four publications[101,102,108,111] reported dural tear or puncture in patients undergoing decompression plus fusion or decompression alone. Meta-analyses of the four RCTs found similar risk of dural tear with decompression plus fusion compared with decompression alone (4 RCTs, N=627, 11.2% vs. 6.3%, RR 1.72, 95% CI 0.79 to 6.49, I^2=39.9%), with an imprecise estimate. Meta-analysis of the three NRSIs also found no difference in risk of dural tear with decompression and fusion versus decompression alone (3 NRSIs, N=1,000, 5.1% vs. 4.2%, RR 1.39, 95% CI 0.29 to 5.63, I^2=63.8%), although the largest NRSI reported similar results as the RCTs (**Figure 6**). Pooled analysis of NRSIs was limited by imprecision and moderate statistical heterogeneity, at least partly due to conflicting direction of effect estimates.

The evidence was insufficient to determine whether fusion with an interbody cage[108,111] resulted in greater risk of dural tear compared with fusion without an interbody cage[62,70,75] in patients with DLS.

Figure 6. Dural tear with decompression plus fusion versus decompression alone



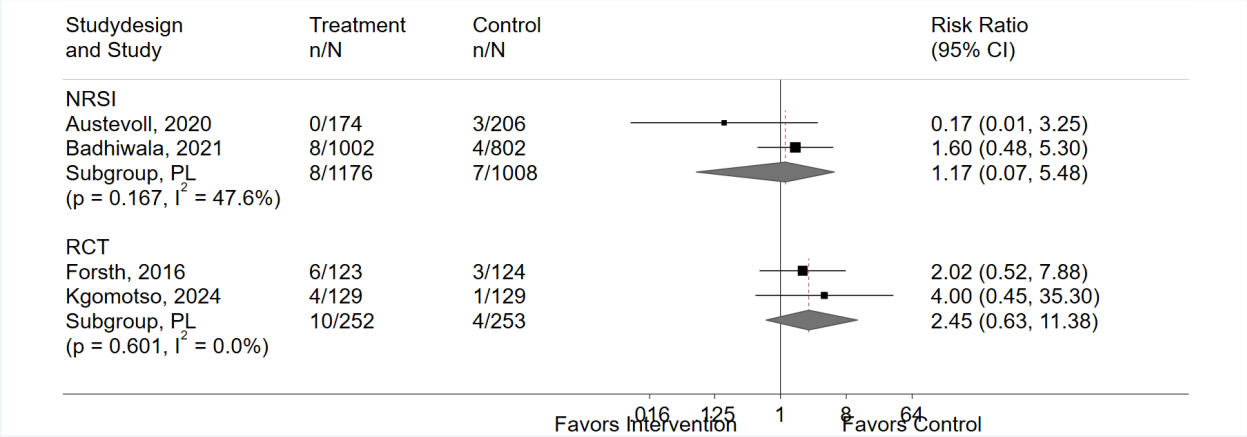
Abbreviations: CI = confidence interval; NR = not reported; NRSI = nonrandomized studies of interventions; PL = profile likelihood; RCT = randomized controlled trial

3.1.3.6.6 Deep Wound Infection and Sepsis

Two RCTs (N=505) (in 5 publications)[55,63,72,73,75] and two NRSIs (N=2,184) (in 3 publications)[101-103] reported deep wound infection and/or sepsis in patients undergoing decompression plus fusion or decompression alone in patients with *mixed or unknown stability* DLS. Pooled analysis of the two RCTs indicated a similar risk of deep wound infection with decompression and fusion compared with decompression alone (2 RCTs, N=505, 4.0% vs. 1.6%, RR 2.54, 95% CI 0.63 to 11.38, I^2=0%).[63,72,73,75]

There was also no difference in risk of deep wound infection in a pooled analysis of the two NRSIs (N=2,184, 0.70% vs. 0.70%, RR 1.17, 95% CI 0.07 to 5.48, I^2=47.6%) (**Figure 7**).[102,103] One NRSI additionally reported a two to three times greater incidence of sepsis at 30 days postoperatively with decompression and fusion versus decompression alone, although the risk of sepsis with either treatment was small (1 NRSI, N=1,804, 0.7% vs. 0.2%, RR 2.80, 95% CI 0.58 to 13.45).[103] For all analyses, estimates were imprecise and for the pooled analysis of the two NRSIs for wound infection, there was moderate statistical heterogeneity partly due to conflicting direction of effect estimates.

Figure 7. Deep wound infection decompression plus fusion versus decompression alone



Abbreviations: CI = confidence interval; NRSI = nonrandomized studies of interventions; PL = profile likelihood; RCT = randomized controlled trial

3.1.3.6.7 Cardiovascular Events

Cardiovascular events included "cardiovascular complications", venous thromboembolism, and stroke. Two RCTs (N=509) in five publications[55,63,72,73,75] and three NRSIs (N=2,975) in four publications[101-103,106] reported cardiovascular complications in patients with *unknown* DLS *stability* treated with decompression and fusion versus decompression alone.

The Nordsten trial (2 publications) reported no statistically significant differences between cardiovascular complications during the hospital stay or from hospital discharge to 3 months after discharge between decompression and fusion and decompression alone (N=260, 0% vs. 2.3%, RR 0.15, 95% CI 0.01 to 2.82; N=254, 0% vs. 0.78%, RR 0.34, 95% CI 0.01 to 8.36, respectively).[75] Of the four total cardiovascular events, two were due to atrial fibrillation and two were due to stroke (RR for atrial fibrillation and stroke, assuming all occurred during inpatient hospitalization, 0% vs. 1.5%, RR 0.21, 95% CI 0.01 to 4.25). Although the estimates of effect represent a large difference between treatment groups in risk of cardiovascular complications, there were few complications overall, and estimates were imprecise.[55,75]

3.1.3 Results, Key Question 1. Instrumentation in addition to decompression compared with decompression, Detailed Synthesis

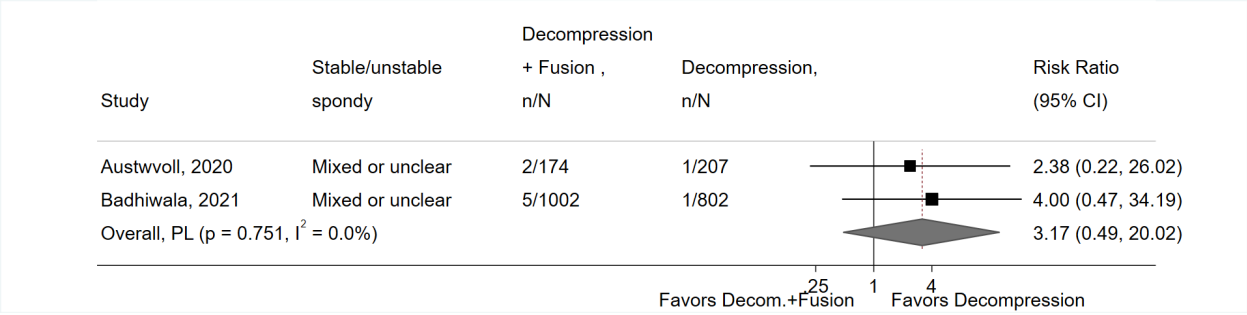
This same trial also reported venous thromboembolism and reported that no events took place in either treatment group up to 3 months from hospital discharge.[55,75]

The second RCT, the SSSS trial (N=247),[63,72,73] reported three patients treated with fusion and decompression experienced a myocardial infarction, stroke, or thromboembolic event compared with five patients who were treated with decompression alone within 2 years after surgery, and found no difference in risk of cardiac complications between treatments (3.3% vs. 4.0%, RR 0.81, 95% CI 0.22 to 2.93). This study did not report specific cardiac complications individually. Due to few events in both RCTs, estimates are imprecise and difficult to meaningfully interpret.

Three NRSIs also reported thromboembolic events[101-103,106] and one reported stroke.[103] Pooled analysis of the two NRSI that reported pulmonary embolism found greatly increased risk with decompression and fusion versus decompression alone but the number of events was small and the estimate imprecise and not statistically significant (N=2,185, 0.6% vs. 0.2%, RR 3.17, 95% CI 0.49 to 20.02, I^2=0%) (**Figure 8**). Three NRSIs reported deep venous thromboembolism (DVT) or venous thromboembolism and found similar risk of DVT with decompression and fusion and fusion alone (N=2,897, 0.7% vs. 0.8%, RR 0.92, 95% CI 0.17 to 2.31, I^2=0.1%) (**Figure 9**).[102,103] Due to few events, the overall estimate was imprecise.

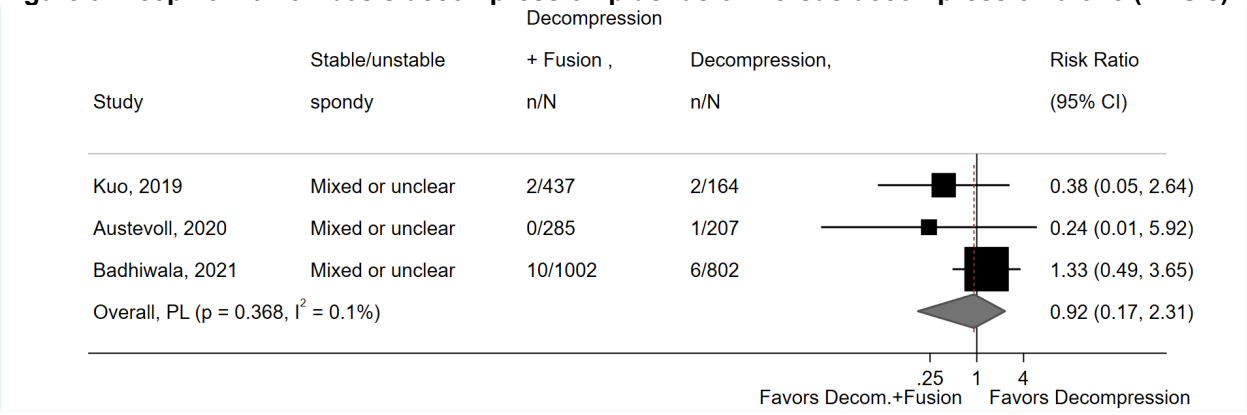
One NRSI (N=1,804) found a similar risk of stroke with decompression and fusion compared with decompression alone (0.1% vs. 0.2%, RR 0.80, 95% CI 0.11 to 5.67). However, events were few and the estimate was imprecise and not statistically significant.[103]

Figure 8. Pulmonary embolism decompression plus fusion versus decompression alone (NRSIs)



Abbreviations: CI = confidence interval; NRSI = nonrandomized studies of interventions; PL = profile likelihood

Figure 9. Deep vein thrombosis decompression plus fusion versus decompression alone (NRSIs)



Abbreviations: CI = confidence interval; NRSI = nonrandomized studies of interventions; PL = profile likelihood

3.1.3.6.8 Neurological Deterioration

The Nordsten trial (N=267) (in 2 publications)[55,75] reported neurological deterioration during multiple followup times. During the hospital stay 2 percent of patients experienced self-reported neurological deterioration with decompression and fusion compared with 1 percent with decompression alone (N=260, RR 2.06, 95% CI 0.19 to 22.47), 6 versus 2 percent from hospital discharge to 3 months (N=254, RR 2.41, 95% CI 0.64 to 9.10), 12 versus 10 percent from 3 months to 2 years postoperatively (N=241, RR 1.24, 95% CI 0.61 to 2.54), and 35 versus 32 percent up to 5 years followup (N=260, RR 1.45, 95% CI 0.99 to 2.13).[55,75] Neurological deterioration consisted of self-reported sensory, motor, or combined sensory/motor disturbance; clinical exams were not conducted. Although analyses favored decompression alone, the proportions of patients experiencing neurological deterioration were small in earlier followup times and similar between treatments at longer followup times; no analysis was statistically significant and except for the analysis at the 5-year followup. Estimates were imprecise.

When analysis was limited to those patients who reported that their condition was "much worse" or "worse than ever" on a 7-point Likert scale, there was no difference between treatments (N=227, 5% vs. 5%, RR 0.92, 95% CI 0.29 to 2.92).[55,75]

A second RCT rated high risk of bias (N=120) reported that individuals with *unknown* DLS *stability* who received transforaminal lumbar interbody fusion (TLIF), posterior lumbar interbody fusion (PLIF), posterior fusion, and decompression without fusion or instrumentation had similar ranges of sciatica, limb paresis, sphincter disorder, and sensory disorder at 12 months (data not provided).[62]

3.2 Key Question 2. In symptomatic adults with unstable or stable DLS with or without radiculopathy or neurogenic claudication undergoing instrumented fusion, what are the benefits and harms of the addition of an interbody cage to instrumentation (e.g., pedicle screws) compared to use of instrumentation alone (i.e., posterolateral fusion)?

3.2.1 Key Points

- Evidence comparing fusion using an interbody change with posterolateral fusion in patients with DLS was sparse.
- There was no difference between TLIF plus posterolateral fusion (PLF) and PLF in the proportion of patients achieving ODI scores between 0 and 20 percent (i.e., mild disability) at 2 years in one small RCT (SOE: low), however there was insufficient evidence from this RCT to draw conclusions regarding improvement in pain or quality of life (SOE: insufficient).
- TLIF plus PLF was associated with substantially higher likelihood of fusion versus PLF by 2 years based on one small RCT (SOE: low).
- Evidence on harms and adverse events was considered insufficient to draw conclusions across included studies (2 RCTs, 4 NRSIs) (SOE: insufficient).

3.2.2 Description of Included Studies

Two RCTs (N=150)[57,62] and four NRSIs (N=50,889)[99,105,107,110] compared lumbar fusion with instrumentation plus interbody cage to lumbar fusion with instrumentation alone (**Appendix D, Table D-1**). NRSIs were included for harms only. One RCT was also included as evidence for Key Question 1.[62]

3.2.2.1 Randomized Controlled Trials

Sample sizes for the two RCTs were 60 and 90. The mean age for one study, conducted in France, was 64.5 years with 70 percent female patients,[57] while the other study, conducted in Iran, enrolled participants who were mostly aged 30 to 40 years with 47 percent female patients.[62] Neither study reported racial or ethnic breakdown. The French study enrolled patients with one vertebral level involvement,[57] but most patients (81%) in the Iranian study had two or more levels needing treatment.[62] The proportion of patients with lumbar stenosis or the number of patients with unstable degenerative spondylolisthesis was not reported in either study, however, in one trial all patients had disk herniation or lumbar canal stenosis.[62]

The French study was a monocentric open-label study that compared TLIF plus PLF with PLF alone.[57] The Iranian study randomized patients to one of four groups: (1) PLIF, (2) TLIF, (3) pedicle screw and posterior fusion, and (4) decompression without fusion and instrumentation.[62] For this Key Question, findings from group 4 were not reported here as these patients did not undergo a fusion procedure. Followup was 24 months in the French study and 12 months in the Iranian study.

The French study reported that it received no funding to conduct the study;[57] funding was not reported in the Iranian study.[62] The risk of bias was rated high for the Iranian study[62] as randomization, allocation concealment, and similarity of groups at baseline were unclear, and attrition and blinding of outcomes assessors were not reported. Risk of bias was rated moderate for the French study, which had fewer methodological limitations (unclear randomization techniques and lack of blinding) (**Appendix E, Table E-1**).[57]

3.2.2.2 Nonrandomized Studies of Intervention

Across the four NRSIs (N=50,889),[99,105,107,110] the average study mean age was 62 years (range 59.7 to 63.7 years), the average study mean proportion of female participants was 60.9 percent (range 53.6% to 68.1%), and most participants were White (average study proportion White 84.4% [range 77.6% to 87.7%]). Two NRSIs enrolled participants with one vertebral level involvement;[105,110] one study enrolled participants with 1 to 3 or more levels of involvement;[99] and one NRSI did not report levels involved.[107] One NRSI reported that 92.8 percent of participants had central stenosis,[107] while the other three studies did not report stenosis. NRSIs relied upon regression analysis to control for potential confounding variables.

One study (the SPORT trial, N=376) randomized individuals to surgery or no surgery, but is included here as an NRSI since the surgery groups we are comparing are not based on randomization to those groups (decompression with interbody fusion plus posterolateral fusion with pedicle screws [360 degrees], decompression with PLF, and decompression with instrumented posterolateral fusion with pedicle screws [PPS]).[99] Another NRSI (N=48,911) utilized the Nationwide Inpatient Sample (NIS) of the Healthcare Cost and Utilization Project (HCUP) database from 2001 to 2010 and also compared findings from several surgical approaches (PLF, ALIF, PLF plus ALIF, P/TLIF, and PLS plus P/TLIF).[107] The remaining two

NRSIs were retrospective cohort studies.[105,110] Followup duration for the NRSIs ranged from 3 months to 8 years in the three NRSIs that reported followup. The study of the NIS database did not report followup duration, but it is presumed to be during the inpatient experience.[107] Two NRSIs reported that they received no funding to conduct the study,[107,110] one reported government funding,[99] and one did not report funding.[105] All NRSIs were conducted in the United States. One NRSI was rated high risk of bias as patient selection methods were unclear, differences between groups in baseline prognostic factors, some of which were not statistically adjusted for in analyses, and unclear attrition,[110] with the remainder of studies rated moderate risk of bias due to fewer methodological limitations (**Appendix E, Table E-2**).

3.2.3 Detailed Synthesis

3.2.3.1 Back and Leg Pain

Only one RCT (N=60) reported back and leg pain.[57] Back pain and leg pain scores were similar for TLIF plus PLF and PLF at 24 months followup (VAS 0-10 scale, difference in change scores: MD 0.5, p=0.646; MD -0.6, p=0.650, respectively, data for full evaluation of effect size were not reported).

3.2.3.2 Function

One RCT (N=60),[57] discussed above, found no difference between TLIF plus PLF and PLF alone on the 0-100 ODI scale (difference in change scores: MD -9, p=0.078). Additionally, 62.1 percent of patients reported scores between 0 to 20 percent on the ODI (i.e., mild disability) at the 24-month followup with TLIF plus PLF versus 44.4 percent with PLF alone (RR 1.50, 95% CI 0.89 to 2.54), whereas at baseline 3.3 and 0 percent of patients reported minimal disability, respectively.

A second RCT (N=90),[62] rated high risk of bias, reported the proportion of patients who scored between 0 and 20 percent on the ODI (i.e., mild disability) at the 6-month and 12-month followup with PLIF, TLIF, and use of pedicle screws and posterior fusion. At the 6-month followup the proportions of patients with ODI scores of 0 to 20 percent (i.e., mild disability) were similar across groups: 70 percent with PLIF versus 66.6 percent with PLF (RR 1.05, 95% CI 0.74 to 1.48), and 63.4 percent with TLIF versus 66.6 percent with PLF (RR 0.95, 95% CI 0.66 to 1.38). At the 12-month followup, the proportion remained similar when comparing PLIF with PLF (83.4% vs. 83.4%) and when comparing TLIF with PLF (80.0% vs. 83.4%, RR 0.96, 95% CI 0.76 to 1.22). The proportion of patients with 21 to 60 percent on the ODI were also similar across groups with no patient in any group reporting ODI scores greater than 60 percent.

3.2.3.3 Quality of Life

Only one RCT (N=60) reported quality of life.[57] SF-36 Physical Component Summary (PCS) and Mental Component Summary (MCS) 0-100 scores were similar between TLIF plus PLF versus PLF alone at 24-month followup (MD 2, p=0.118; MD 6, p=0.089, respectively).

3.2.3.4 Fusion

Two RCTs (N=150) reported fusion with an interbody cage versus PLF. One was rated moderate risk of bias (N=60)[57] and the other was rated high risk of bias (N=90).[62]

To determine if fusion occurred, one small RCT (N=60)[57] reported movement of less than 5 degrees in flexion or extension at 24 months as 93.3 percent with TLIF plus PLF versus 43.3

percent with PLF alone (RR 2.15, 95% CI 1.42 to 3.28) and a larger proportion achieved a Lenke A or B (i.e., Lenke A: solid trabeculated transverse process and facet fusions bilaterally; Lenke B: thick fusion mass on one side, difficult to visualize on other side) at 24 months (96.7% vs. 56.7%, RR 1.71, 95% CI 1.24 to 2.35), indicating potentially increased likelihood of fusion using an interbody. However, the number of participants who received each treatment was small (n=30).

The second, a high risk of bias RCT (N=90)[62] reported fusion results by radiographic Grade—Grade 1: obvious radiographic pseudoarthrosis; Grade 2: probably radiographic pseudoarthrosis; Grade 3: radiographic status uncertain; Grade 4: probably radiographic fusion; and Grade 5: radiographic fusion. At 6 months, Grade 5 radiographic fusion was achieved by 33.4 percent with PLIF versus 40 percent with PLF (RR 0.83, 95% CI 0.43 to 1.63) and was achieved by 27 percent with TLIF (27% vs. 40%, RR 0.67, 95% CI 0.32 to 1.39). Neither analysis was statistically significant, and estimates were imprecise. At 12 months radiographic fusion was achieved by 87 percent in those who underwent PLIF and PLF and by 77 percent of those who underwent TLIF (77% vs. 87%, RR 0.89, 95% CI 0.69 to 1.13), indicating no differences between fusion rates between PLIF, TLIF, and PLF at 12 months.

3.2.3.5 Reoperation Rates

One RCT (N=60)[57] and three NRSIs (N=1,895) reported reoperation rates.[99,105,110] Reoperation occurred for multiple reasons; some studies limited reporting of reoperations to revision surgeries[57,105] and others studies included various reasons for reoperation such as hardware pain and wound infection.[99,110] One NRSI reported annual reoperation rates for 8 years following surgery.[99]

One small RCT (N=60)[57] reported similar surgical revision rates due to failure of the initial intervention in one patient who had TLIF plus PLF versus three patients having PLF alone at 24 months followup (3.3% vs. 10%, RR 0.33, 95% CI 0.04 to 3.03). Other reoperations with TLIF were due to dural tear (n=2) and pedicle fracture with screw loosening (n=1); in patients who were treated with PLF alone, other reoperations were due to dural tear (n=2), wound infection (n=2), and adjacent segment disease (ASD) (n=1).

One NRSI (n=1,056) compared revision surgery for TLIF versus PLF by type of clinical and radiographic degenerative spondylolisthesis (CARDS) subtypes based on disk space height, anterior vertebral translation and kyphotic alignment.[105] CARDS type A: advanced disk space collapse without kyphosis; CARDS B: disk space partially preserved with translation of 5 mm or less; CARDS C: similar to CARDS B but translation more than 5 mm; CARDS D: kyphotic alignment.[113] There were no differences in the rate of revision surgeries within one year between TLIF and PLF among those with CARDS A (n=148, 4.28% vs. 11.5%, p=0.060), CARDS B (n=323, 8.97% vs. 9.55%, p=1.00), CARDS C patients (n=525) (n=525, 11.5% vs. 6.44%, p=0.061), or CARDS D (n=60, 10.3% vs. 16.1%, p=0.708). Our analysis combining the four CARDS groups indicates a similar risk of revision surgery with TLIF than with PLF (10.7% vs. 7.6%, RR 1.37, 95% CI 0.76 to 2.48). The most common reasons for the 96 revision surgeries were ASD 51.6 percent (n=50), recurrent stenosis 26.8 percent (n=26), and symptomatic pseudarthrosis 17.5 percent (n=17). ASD was more common with TLIF (63.3% vs. 35.7%), while pseudarthrosis was more common with PLF (28.6% vs. 9.1%, p=0.014); authors do not provide sufficient data to calculate effect size.

A second NRSI (N=546), rated high risk of bias, reported reoperation rates up to and beyond 5 years after TLIF or PLF and found a much lower likelihood of reoperation with TLIF

compared with PLF (9.9% vs. 15.6%, aOR 0.23, 95% CI 0.54 to 0.99) after adjusting for age, body mass index (BMI), flexion-extension difference, flexion-extension difference distance, and disk height.[110] The most common reason for reoperation was ASD for both TLIF and PLF (48.6% vs. 51.9%, p=0.800); the next most common reason for reoperation was adjacent-level spondylolisthesis (13.5% vs. 11.1%, p=0.517) followed by pseudarthrosis (8.1% vs. 11.1%, p=0.684), and implant-related reoperations (5.4% vs. 11.1%, p=0.401). After adjusting for potential confounders listed above, TLIF was associated with decreased likelihood of any reoperation rates (aOR 0.23, 95% CI 0.54 to 0.99). This study also reported that one patient who underwent TLIF (2.7%) underwent reoperation due to recurrent spondylolisthesis compared with no patients who underwent PLF.

In a third NRSI (N=292), where 96 percent of patients had spinal stenosis and 11 percent had spinal instability, few additional spine surgeries occurred within 1 year of the index surgery with interbody fusion (not otherwise specified) versus PLF (7% vs. 5%, RR 1.30, 95% CI 0.48 to 3.57).[99] The frequency of additional reoperations remained similar between treatments for subsequent years (2 years—10% vs. 10%; 3 years—13% vs. 13%; 4 years—13% vs. 14%, RR 0.91, 95% CI 0.45 to 1.81; 5 years—15% vs. 14%, RR 1.06, 95% CI 0.56 to 1.99; 6 years—17% vs. 17%; 7 years—23% vs. 18%, RR 1.25, 95% CI 0.75 to 2.09). Eight years after the index surgery, there was a similar risk of additional spine surgery with interbody fusion (not otherwise specified) versus PLF (24% vs. 20% with PLF; RR 1.24, 95% CI 0.45 to 2.03). Reasons for reoperation within 8 years following the index surgery were not different between interbody fusion and PLF. The most common reasons for reoperation were recurrent stenosis/progressive listhesis (6% vs. 11%, RR 0.54, 95% CI 0.20 to 1.52), unspecified complications (11% vs. 6%, RR 1.92, 95% CI 0.83 to 4.45), new herniations/stenosis (3% vs. 2%, RR 1.25, 95% CI 0.25 to 6.31), and pseudarthrosis/fusion exploration (3% vs. 1%, RR 2.09, 95% CI 0.36 to 12.23). Authors did not provide frequency data separately for the different indications (e.g., progressive stenosis only), however.

3.2.4 Harms and Adverse Events

3.2.4.1 Serious Adverse Events/Harms

One RCT (N=60)[57] reported major complications within 24 months of surgery and one NRSI (N=1,056)[105] reported serious adverse events within 3 months of surgery.

The only RCT meeting inclusion criteria reported half the risk of major complications with TLIF plus PLF than with PLF alone (10% vs. 20%, RR 0.50, 95% CI 0.14 to 1.82), corresponding to a large increase in the likelihood of a major complication.[57] However, the study was small, and the estimate was very imprecise. Major complications were defined as those leading to surgical revision and included dural tear repair, infection, rod failure, pedicle fracture due to screw loosing, and for adjacent segment disease (one patient) and are described under re-operation.

One NRSI (n=1,056)[105] reported 90-day complications including "Clavien-Dindo grade III or IV complications requiring invasive intervention and organ failure" by CARDS subtype in patients who underwent TLIF or PLF after adjusting for age, sex, smoking status, Charlson comorbidity index, and BMI. There were no differences between treatments at any CARDS level. Our analysis combining the results by CARDS levels also found no differences between TLIF and PLF in the risk of any complication (1.4% vs. 1.5%, RR 0.93, 95% CI 0.34 to 2.54).

3.2.4.2 Adverse Events (Serious and Nonserious Adverse Events)

Three NRSIs (N=49,750) reported the risk of adverse events/complications following interbody fusion versus PLF.[99,107,110]

The largest NRSI (n=48,911) reported the adjusted risk of any complication that occurred in at least 1 percent of the study population during the acute care phase following surgery and found a small increase in complication rate with most interbody fusions than with PLF alone; however, detail of specific complications or their seriousness is not described and it is unclear whether the differences between treatments are clinically important. Using PLF as a reference, the risk of any complication was 24 versus 21.2 percent (aOR 1.44, 95% CI 1.14 to 1.82) with ALIF plus PLF; 22.4 versus 21.2 percent (aOR 1.24, 95% CI 1.11 to 1.39) with TLIF plus PLF; 24.2 versus 21.2 percent (aOR 1.49, 95% CI 1.40 to 1.59) with ALIF alone; and 22.6 versus 21.2 percent (aOR 1.12, 95% CI 1.08 to 1.16) with TLIF alone.[107] This study adjusted for age, sex, race/ethnicity, Charlson Comorbidity Index, and payment type.

A second NRSI (N=546), rated high risk of bias, found no difference in risk of any complications within 90 days with TLIF compared with PLF in adjusted analysis that controlled for age, BMI, flexion-extension difference, flexion-extension difference distance, and disk height (aOR 1.64, 95% CI 0.15 to 18.02).[110]

A third NRSI (N=293) reported complications after 8 weeks postoperatively with interbody fusion (not otherwise specified) versus PLF and found a similar risk of any complications with the two treatments (26% vs. 35%, RR 0.74, 95% CI 0.49 to 1.13).[99] Study authors reported that there were no complications with bone grafts, or instances of cerebrospinal fluid leaks, paralysis, pseudarthrosis, or cauda equina injury in this study.

3.2.4.3 Persistent Pain

One RCT (N=60)[57] found a similar rate of persistent lumbar pain and radicular pain (not otherwise defined) 24 months postoperatively between TLIF plus PLF and PLF alone (24% vs. 28%, RR 1.14, 95% CI 0.47 to 2.75; 13% vs. 14%, RR 1.00, 95% CI 0.26 to 3.63, respectively).

3.2.4.4 Sacroiliac Joint Pain

No study reported postoperative sacroiliac joint pain comparing interbody fusion and fusion without an interbody for DLS.

3.2.4.5 Device-Related Complications

One NRSI (N=48,911)[107] reported a large increase in likelihood of device-related complications with ALIF plus PLF versus PLF (3% vs. 1.3%, aOR 4.81, 95% CI 3.44 to 6.73) and ALIF versus PLF (6.7% vs. 1.3%, aOR 5.11, 95% CI 4.36 to 5.99) during the acute phase of treatment.

3.2.4.6 Mortality

Two NRSIs (N=49,287)[99,107] reported mortality. Forty-seven patients died (0.10%) during the acute care phase of treatment in the largest NRSI (N=48,911).[107] Study authors reported a large increased likelihood of mortality with TLIF plus PLF (0.42% vs. 0.10%, aOR 5.34, 95% CI 2.57 to 11.08) and ALIF (0.14% vs. 0.10%, aOR 2.23, 95% CI 1.18 to 4.24) compared with PLF alone. However, it is unclear if this small absolute risk difference is clinically meaningful. Authors report no differences in mortality between ALIF plus PLF (0%) and PLF alone (0.10%) and between posterior TLIF (0.09%) and PLF (0.10%). This study adjusted for age, sex,

race/ethnicity, Charlson Comorbidity Index, and payment type. The second NRSI (n=376)[99] reported no deaths during the 6 weeks after surgery with interbody fusion (not otherwise specified) versus PLF. Death within 3 months of surgery was rare for interbody fusion (0/71) and PLF (1/222).

3.2.4.7 Dural Tear or Puncture

One RCT (N=90), rated high risk of bias,[62] and two NRSIs (N=922), one rated moderate risk of bias[99] and one rated high risk of bias,[110] reported dural tear or puncture or cerebrospinal fluid leak.

The RCT reported dural puncture in two patients who underwent PLIF (6.7%), in one patient who underwent TLIF (3.3%) and in no patients who underwent PLF (RR 5.00, 95% CI 0.25 to 100; RR 3.00, 95% CI 0.13 to 70.83, respectively).[62] Although the relative risks are large, due to few events and related imprecision, the differences between treatments were not reliable. One NRSI (n=293) found a substantially lower likelihood of dural tear or cerebrospinal fluid leak with interbody fusion (not otherwise specified) than with PLF (1% vs. 11%, RR 0.13, 95% CI 0.02 to 0.91).[99] The second NRSI (N=546) also reported a substantially lower likelihood of cerebrospinal fluid leak with TLIF vs. PLF (0% vs. 7.4%, RR 0.09, 95% CI 0.004 to 1.93).[110] However, the estimates was very imprecise due to few events and differences reported here may not be reliable.

3.2.4.7 Deep Wound Infection and Sepsis

One NRSI (N=546),[110] rated high risk of bias, reported reoperation due to deep wound infection and found similar risks with TLIF and PLF (5.4% vs. 3.7%, RR 0.93, 95% CI 0.09 to 10.16). However, the estimate was imprecise.

3.2.4.8 Cardiovascular Events

Two NRSIs (N=49,457), one rated moderate risk of bias[107] and one rated high risk of bias,[110] reported cardiovascular complications or events, DVT, and pulmonary embolism.

The largest NRSI (n=48,911)[107] reported a moderately increased risk of cardiac complications with ALIF compared with PLF (0.8% vs. 1.1%, aOR 1.50, 95% CI 1.15 to 1.97) after adjustment for age, sex, race/ethnicity, Charlson Comorbidity Index, and payment type. However, the risk of cardiac complications was small for both treatments and the difference may not be clinically meaningful. Risk of cardiac complications was similar between ALIF plus PLF (1.0%), posterior TLIF plus PLF (0.6%), or posterior TLIF alone (1.0%) versus PLF alone (1.1%) (adjusted comparisons not provided). This study also reported that 0.4 percent of patients who underwent posterior TLIF plus PLF, ALIF, posterior TLIF, and PLF alone experienced venous thromboembolism. This compared with 1 percent of patients who underwent ALIF plus PLF (adjusted comparisons not provided; unadjusted RR 0.25, 95% CI 0.04 to 1.75).

A second NRSI (N=546), rated high risk of bias, reported a substantial increased risk of myocardial infarction with TLIF compared with PLF (6.6% vs. 0%) within 90 days of surgery (RR 0.43, 95% CI 0.02 to 8.91).[110] However, due to the low frequency of heart attacks, the estimate is very imprecise. This study also reported that no patient who underwent TLIF or PLF experienced deep vein thrombosis or a pulmonary embolism.

3.2.4.9 Neurological Deterioration

One RCT rated high risk of bias (N=90)[62] and two NRSIs (N=49,457), one rated moderate risk of bias[107] and one rated high risk of bias,[110] reported neurological deterioration or complications.

The RCT reported that individuals who received TLIF, PLIF, and posterior fusion had similar ranges of sciatica, limb paresis, sphincter disorder, and sensory disorder at 12 months (data not provided).[62]

The largest NRSI (N=48,911) reported a moderately greater risk of neurological complications with posterior TLIF plus PLF compared with PLF alone (1.3% vs. 0.9%, aOR 1.58, 95% CI 1.05 to 2.38) and a moderately lower risk of neurological complications with ALIF compared with PLF (0.6% vs. 0.9%, aOR 0.61, 95% CI 0.43 to 0.85). However, it is unclear if these differences are clinically meaningful. This study adjusted for age, sex, race/ethnicity, Charlson Comorbidity Index, and payment type. The proportion of patients who experienced neurological complications were not different between ALIF plus PLF versus PLF (1.0% vs. 0.9%, unadjusted RR 1.10, 95% CI 0.15 to 7.89), or between posterior TLIF versus PLF (1.0% vs. 0.9%, unadjusted RR 1.11, 95% CI 0.82 to 1.51, adjusted comparison not provided for latter comparisons).

The second NRSI (N=546), rated high risk of bias, reported a similar likelihood of neurological deficit at 90 days with TLIF (13.3%) compared with PLF (9.1%), RR 1.85, 95% CI 0.21 to 16.48, with a very imprecise estimate.[110]

3.3 Key Question 3. Benefits and harms of the use of bone graft extenders and biologic substitutes compared to the use of autografts for stable or unstable degenerative lumbar spondylolisthesis

3.3.1 Key Points

- Demineralized bone matrix (DBM) was associated with similar fusion rates compared with iliac crest bone graft (ICBG) in one small RCT (SOE: low).
- Evidence from one small RCT was insufficient to draw conclusions regarding improvements in pain, function, quality of life, intervention related harms or need for reoperation for DBM versus ICBG.

3.3.2 Description of Included Study

One small RCT (N=46)[71] compared Grafton DBM (n=30) plus local bone versus ICBG (n=16) in patients with a primary diagnosis of DLS (severity not reported) who were eligible for lumbar or lumbosacral single-level fusion with instrumentation (**Appendix D, Table D-1**). Grafton DBM consisted of allograft DBM fibers with a carrier which theoretically supports both osteoconductivity and osteoinductivity. It was combined with locally harvested bone from lamina and facets. ICBG was augmented with local autograft at the surgeon's discretion. The study mean age was 64.6 years, 56.5 percent were female, and race/ethnicity were not reported. The trial was conducted the United States. The trial was rated at moderate risk of bias as randomization and methods of allocation concealment were unclear and differences in baseline

characteristics and patient reported measures across all randomized patients were not described (**Appendix E, Table E-1**). In addition, loss to followup after 6 months for patient-reported outcomes was <80 percent and differential loss to followup was noted at 24 months. This trial was funded by industry with one or more authors having conflicts of interest.

3.3.3 Detailed Synthesis

3.3.3.1 Pain, Function, and Quality of Life

The single included trial (N=46)[71] comparing DBM and ICBG did not provide sufficient data to calculate mean differences between groups and confidence intervals for outcomes of interest. There was no difference between treatments on *left* low back pain (LBP) at any timepoint (data not reported). *Right* lower back pain improvement with DBM and ICBG on a 10-point scale (instrument not specified) was reported as similar at 3 months (p=0.3809) and 12 months (p=0.5428) but DBM was associated with greater improvement versus ICBT at later timepoints (6 months, p=0.015; and 24 months, p=0.007). Reasons for the inconsistency in the association across time frames are unclear.[71] ODI scores (0 to 100 scale) were similar at 6 months (scores 20 vs. 23, p=0.808), 12 months (score for both groups 20, p=0.920), and 24 months (score 16 vs. 22.5, p=0.235) for DBM and ICBG. Data are estimated from graphs and authors do not provide sufficient data to calculate effect sizes with confidence intervals for any of these measures.[71]

Scores on the SF-36 Physical Component Summary were similar at 6 months (39 for both groups, p=0.935), 12 months (41 vs. 42, p=0.743), and 24 months (40 vs. 43, p=0.392).

3.3.3.2 Fusion

Fusion was defined as "evidence of bilateral continuous bridging trabecular bone as well as less than 3 mm of translation and less than 5 degrees of angular motion" in the single trial that provided evidence for Key Question 3.[71] If a patient was determined to be fused at the 12-month followup or 24-month followup by computed tomographic scan, then the individual was considered to have successful fusion in the final, modified intent-to-treat analysis. Fusion rates were similar between groups: 86 percent with DBM versus 92 percent with ICBG at 2 years (RR 0.93, 95% CI 0.75 to 1.16).

3.3.3.3 Reoperation Rates, Harms, and Adverse Events

The included trial did not report reoperation rates, stating only that no revision surgeries for nonunion were done in either group over 24 months of follow-up. Authors state there were no adverse events directly attributed to the DBM product but provided no information on other adverse events or on adverse events for the ICBG group.[71]

3.4 Key Question 4. Does the use of intraoperative neuromonitoring (IONM) decrease perioperative neurological injuries compared with not using IONM for patients with degenerative lumbar spine disease undergoing instrumented fusion?

3.4.1 Key Points

- There was low-strength evidence of no difference between IONM versus no IONM in odds of developing postoperative neurological complications based on one NRSI (N=133,572) (SOE: low).

3.4.2 Description of Included Studies and Detailed Synthesis

One NRSI (N=133,572)[100] compared lumbar fusion with IONM versus fusion without IONM using data from the NIS for years 2012 to 2015 in individuals undergoing first-time elective PLF (**Appendix D, Table D-1**). The analyses controlled for age, sex, race, income, payer, comorbidities, hospital teaching status, and size. Ten percent of patients were age 18 to 40 years, 38 percent were 41 to 60 years, and 52 percent were age greater than 60 years; 56 percent were female, and 78 percent were White. The primary International Classification of Diseases, Ninth Revision (ICD-9) diagnosis codes in the group who did not receive IONM in descending order were: Lumbar spinal stenosis (17.88%), lumbar/lumbosacral disc degeneration (15.35%), acquired spondylolisthesis (14.75%), lumbar disc displacement (13.57%), and lumbosacral spondylosis (10.39%). The primary ICD-9 codes in the group who received IONM in descending order were: Lumbar/lumbosacral disk degeneration (18.91%), acquired spondylolisthesis (16.28%), lumbar disk displacement (13.37%), lumbar spinal stenosis (13.28%), and lumbosacral spondylosis (11.11%). The study was rated as low risk of bias for a NRSI (**Appendix E, Table E-2**).

Neurological complications in the hospital discharge record included ICD-9 codes for: nervous system complication, unspecified; central nervous system complication; iatrogenic cerebrovascular infarction or hemorrhage; and other nervous system complications. These codes are very general; details regarding specific surgery-related neurological events were not provided.

The study[100] found similar odds of developing neurological complications with IONM compared with no IONM (aOR 0.87, 95% CI 0.70 to 1.07). Information by type of complication was not reported.

3.5 Key Question 5. Benefits and harms of lumbar epidural steroid injections, intra articular (facet) injection, medial branch blocks, or radio frequency ablation for chronic low back pain (≥3 months) resulting from degenerative disease

3.5.1 Key Findings

Radiofrequency Ablation in patients with CLBP of suspected facet joint origin with positive diagnostic block

Radiofrequency ablation versus sham

- Continuous radiofrequency ablation (RFA) was associated with a similar likelihood of a successful pain outcome (3 RCTs) and similar improvement in back (6 RCTs) and leg (3 RCTs) pain scores at 3 months compared with sham RFA; at 6 months RFA was associated with a large increased in the likelihood of a successful pain outcome (2 RCTs) and a moderate improvement in back pain scores (6 RCTs) and a large improvement in leg pain scores (3 RCTs) (SOE: low for all except back pain scores at 3 months which is moderate).
- Continuous RFA was associated with a large increase in the likelihood of a successful function outcome at 3 months (1 RCT) but similar improvement in function scores at 3 months (3 RCTs) and 6 months (4 RCTs) and quality of life score at 6 months (1 RCT) compared with sham RFA (SOE: low for all).
- Pulsed RFA targeting the dorsal root ganglion was associated with a large increase in the likelihood of a successful pain outcome and a large improvement in back pain at 3 and 6 months and in leg pain at 6 months (improvement was similar at 3 months) compared with sham RFA in one trial. Pulsed RFA targeting the medial branch nerves was associated with similar improvement in back pain scores at 6 months compared with sham in a second trial (SOE: low for all).
- Pulsed RFA targeting the dorsal root ganglion was associated with a large improvement in function scores at 3 and 6 months compared with sham RFA in one trial. Pulsed RFA targeting the medial branch nerves was associated with similar improvement in function scores at 6 months versus sham in a second trial (SOE: low for all).
- The risk of serious adverse events (continuous RFA, 3 RCTs) and any adverse events (continuous and pulsed RFA, 5 RCTs) was similar for RFA and sham RFA (SOE: low for all).

Radiofrequency ablation versus usual care

- At 3 and 6 months, continuous RFA was associated with a large increase in the likelihood of a successful pain outcome and a moderate improvement in pain scores compared with usual care in one trial (SOE: low for all).
- Continuous RFA was associated with a small increase in the likelihood of a successful function outcome at 3 months, but the effect did not persist to 6 months, and a similar (below the threshold for a small effect) improvement in function scores at 3 and 6 months compared with usual care in one trial (SOE: low for all).
- No procedure-related complications were reported. Fewer patients who received continuous RFA underwent other lumbar surgery by 6 months (SOE: low).

Radiofrequency ablation versus facet (intraarticular) injection

- Continuous RFA was associated with a similar improvement in pain and function scores at 6 months versus intraarticular steroid injection in one trial (SOE: low).
- Pulsed RFA was associated with a similar likelihood of a successful pain outcome at 3 months and a similar improvement in pain scores at 3 and 6 months compared with intraarticular steroid injection in one trial assessed as moderate risk of bias (SOE: low).
- There was insufficient evidence from one trial assessed as high risk of bias to determine the effects of cooled RFA versus intraarticular steroid injection.
- Evidence for serious or any adverse events for RFA versus intraarticular steroid injection was insufficient.

Radiofrequency ablation versus medial branch blocks

- Continuous RFA was associated with a small increase in the likelihood of a successful pain outcome (1 RCT), a moderate improvement in back pain scores (2 RCTs) and a small increase in the likelihood of a successful quality of life outcome (1 RCT) at 6 months compared with medial branch blocks; the risk of any adverse events was similar between groups (2 RCTs) (SOE: low).
- Pulsed RFA was associated with large improvements in back pain scores and function scores at 3 and 6 months in one trial (SOE: low).

Epidural Steroid Injection in patients with chronic, nondiscogenic low back

Epidural steroid injection versus placebo (lidocaine) in patients with spinal stenosis

- ESI was associated with a similar likelihood of a successful pain and function outcome at 1.5 to 3 months (3 RCTs) and 6 months (2 RCTs), a similar improvement in pain and function scores at 3 months (4 RCTs) and 6 months (3 RCTs) and in quality-of-life scores at 1.5 months (1 RCT) compared with placebo injection (SOE: moderate).
- ESI was associated with a similar, small risk of serious adverse events (4 RCTs) and a similar risk of any adverse event (2 RCTs) compared with placebo injection (SOE: low).

Epidural steroid injection versus usual care and versus inpatient physical therapy

- There was insufficient evidence from one small trial assessed as high risk of bias to determine the effects of ESI versus usual care and versus inpatient physical therapy.

Facet Joint Injections in patients with CLBP of suspected facet joint origin with positive diagnostic block

Facet joint injection versus placebo (saline)

- Facet joint (intraarticular or pericapsular) injection was associated with similar improvement in pain scores at 3 months (2 RCTs) and 6 months (1 RCT) and function scores at 3 and 6 months (1 RCT) compared with placebo injection (SOE: low).
- There was insufficient evidence from two trials to determine the harms and adverse effects of ESI versus placebo injection.

Facet joint injection versus intramuscular steroid injection

- Facet joint (intraarticular) injection, compared with intramuscular (IM) injection, was associated with a similar improvement in pain scores at 3 and 6 months and a moderate improvement in function scores at 3 months, but the effect did not persist to 6 months, in one trial (SOE: low).

- There was insufficient evidence from one trial to determine the harms and adverse effects of facet joint (intraarticular) injection versus IM steroid injection.

Facet joint injection versus physiotherapy

- There was insufficient evidence from one trial assessed as high risk of bias to determine the effectiveness and harms of facet joint (intraarticular) injection versus physiotherapy.

Medial Branch Blocks in patients with CLBP of suspected facet joint origin with positive diagnostic block

Medial branch blocks versus placebo (lidocaine)

- Medial branch block (MBB) was associated with a similar likelihood of a successful pain and function outcome and similar improvement in back pain scores and function scores compared with placebo (lidocaine only) injection in one trial (SOE: low).
- There was insufficient evidence from one trial to determine the harms and adverse effects of MBB versus placebo injection.

Medial branch block versus facet joint injection

- MBB was associated with a similar likelihood of a successful pain outcome and similar risk of any adverse events or any procedure-related adverse events in one trial (SOE: low). There was insufficient evidence for serious adverse events.

3.5.2 Radiofrequency Ablation

3.5.2.1 Radiofrequency Ablation Versus Sham

3.5.2.1.1 Description of Included Studies

Nine RCTs (total N=702; N range, 40 to 150)[66,79,88,89,91,93-96] compared RFA to sham for the treatment of chronic low back pain (CLBP) of suspected facet origin and a positive diagnostic block (**Appendix D, Table D-2**).

The average study mean age of patients was 54 years (range 45 to 62 years) and 64.3 percent were female (range 52.5% to 72.5%) across eight RCTs (age and sex not reported in 1 RCT).[66] Mean BMI was 30.9 (range 29.6 to 31.3) in three trials that reported this variable.[88,89,95] One trial[95] reported that all patients were White, otherwise trials did not report on race or ethnicity. Average baseline pain severity on a 0-10 scale was 4.9 (range, 3.1 to 7.3). Mean duration of pain was 34 months in one trial;[93] the other trials did not provide mean pain duration, but minimum durations required for inclusion ranged from 3 to 24 months with five trials specifying 3 to 6 months[66,79,94-96] and three trials 12 to 24 months.[88,89,91] Comorbidities and neurologic symptoms were not well-reported. Most trials excluded patients with major comorbidities (e.g., uncontrolled diabetes, cardiac disease, malignancy) as well as those with signs of nerve root compression, neurologic deficit or radicular syndrome (but three included patients with or without radiating pain).[88,93,96] Four trials excluded patients with a history of prior lumbar surgery[66,79,88,89] and four excluded patients who had received previous RFA treatment.[88,89,93,96] All trials required a positive diagnostic block for study entry; the way blocks were performed, the type of block and the criteria used to determine diagnostic success varied. Three trials required ≥50 percent pain relief,[93,94,96] one ≥80 percent pain relief,[91] and one ≥2 point reduction on the numeric rating scale (NRS)[95] following medical branch blocks and four required

"complete or near complete" or "significant" pain relief following facet joint injections[66,79,88,89] of local anesthetics (primarily lidocaine or bupivacaine); one trial used a combination of lidocaine plus a steroid (triamcinolone).[79]

All nine trials used continuous RFA[66,79,88,89,91,93-96] and two trials included an additional arm that received pulsed RFA.[89,93] In trials using continuous RFA, electrode tip sizes ranged from 1.2 mm to 10 mm, temperature ranged from 80 to 85 degrees Celsius, and duration of ablation ranged from 60 to 90 seconds. In the trials using pulsed RFA, tip size ranged from 2 mm to 10 mm, temperature was 42 degrees Celsius and ablation duration ranged from 4 minutes of 2 Hz pulses to four 2-minute cycles. All nine trials targeted medial branch nerves and used fluoroscopic guidance; one trial[88] included a third arm that received RFA targeted at the facet joint capsule. Sham procedures generally utilized the same technique as the RFA procedures with the electrode tip not turned on. Five trials[88,89,93,95,96] allowed crossover from sham to RFA at 3 months if patients failed to improve, two[79,94] did not allow crossover and two did not report whether crossover was allowed or occurred.[66,91] Concomitant treatments included physical therapy or graded active physiotherapy (2 RCTs),[94,95] nonsteroidal anti-inflammatory drugs (NSAIDs) (1 RCTs)[93] and antibiotics, analgesics and muscle relaxants (2 RCTs).[88,89] The latter two trials also gave all patients an post-RFA injection of steroids (20 mg methylprednisolone 0.5 ml) and anesthetic (0.5 ml of 0.5% bupivacaine) through the electrode needle. Of the remaining trials two[66,91] did not report concomitant treatments and two[79,93] told patients to limit concurrent interventions and medications

Five trials were conducted in Europe,[66,91,94-96] two in Egypt,[88,89] and one each in Canada[79] and Turkey.[93] Three trials[79,94,96] reported government funding; the other trials did not report funding.

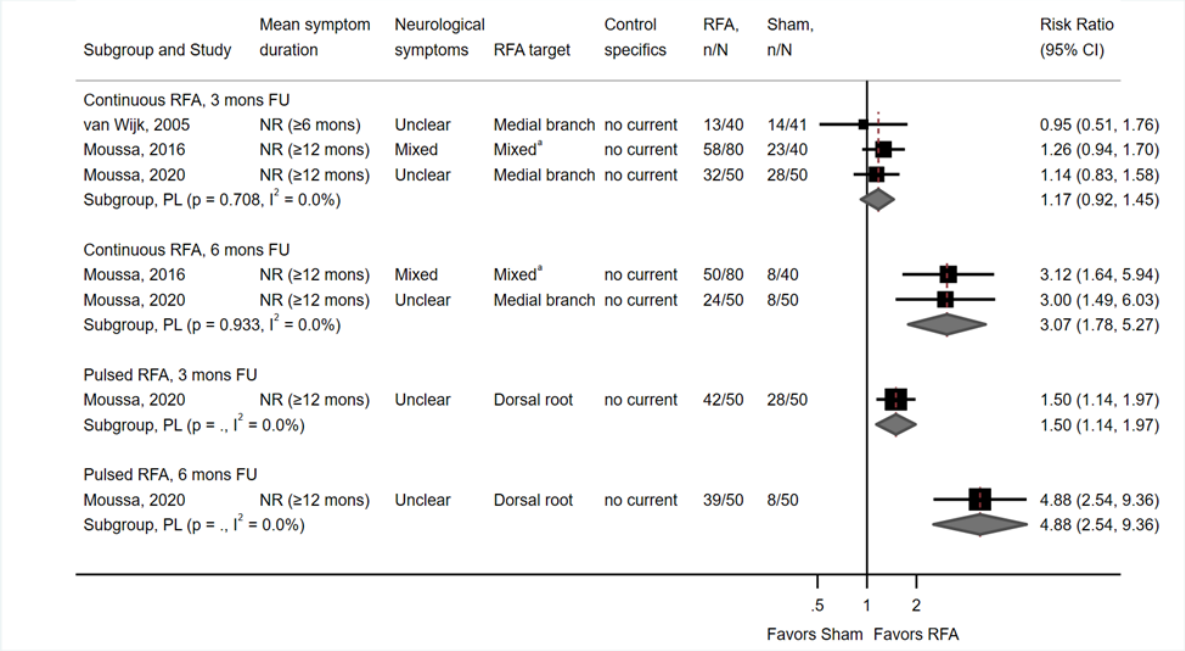
Three trials were assessed as low risk of bias,[79,94,96] five as moderate risk of bias,[88,89,91,93,95] and one as high risk of bias.[66] Common limitations included dissimilar groups at baseline, lack of blinding or unclear blinding status of care providers and outcome assessors, and unclear allocation concealment (**Appendix E, Table E-1**). Additionally, the high risk of bias trial had unclear randomization and did not report attrition.

3.5.2.1.2 Detailed Synthesis

Pain response. Three trials reported the proportion of patients who achieved ≥50 percent improvement in pain on VAS (**Figure 10**).[88,89,96] Continuous RFA was associated with a similar likelihood of achieving ≥50 percent improvement in VAS back pain at 3 months across three RCTs (N=301, 60.6% vs. 49.6%, RR 1.17, 95% CI 0.92 to 1.45, I^2=0%)[88,89,96] and a large increase in the likelihood at 6 months across two RCTs (N=220, 56.9% vs. 17.8%, RR 3.07, 95% CI 1.78 to 5.27, I^2=0%).[88,89] The RFA target was the medial branch of the dorsal ramus in all but one trial,[88] which included both the medial branch and the facet joint capsule; sensitivity analyses of RFA to the medial branch only provided results similar to the primary analyses (**Appendix B, Figure B-4**). One trial[96] reported alternative cut-offs for improvement in VAS back pain (≥2 points and ≥25% improvement) and improvement was similar between treatment groups (**Appendix D, Table D-2**). In one RCT,[89] pulsed RFA of the dorsal root ganglion was associated with a moderate increase at 3 months (N=100, 84.0% vs. 56.0%, RR 1.50, 95% CI 1.14 to 1.97) and a large increase at 6 months (N=100, 78.0% vs. 16.0%, RR 4.88, 95% CI 2.54 to 9.36) in the likelihood of achieving ≥50 percent improvement in VAS back pain. Of note, two trials by the same author group (Moussa et al.)[88,89] gave patients a post-treatment injection of steroids and anesthetic through the electrode needle and it is unclear how this may have impacted results. These trials tended to show a larger effect than the other trials.

3.5.2 Results, Key Question 5. Nonsurgical interventions for chronic low back pain, Radiofrequency Ablation

Figure 10. RFA versus sham: pain response, ≥50% improvement in back pain on 0-10 VAS



Abbreviations: CI = confidence interval; FU = followup; NR = not reported; PL = profile likelihood; RFA = radiofrequency ablation

[a] Medial branch and joint capsule.

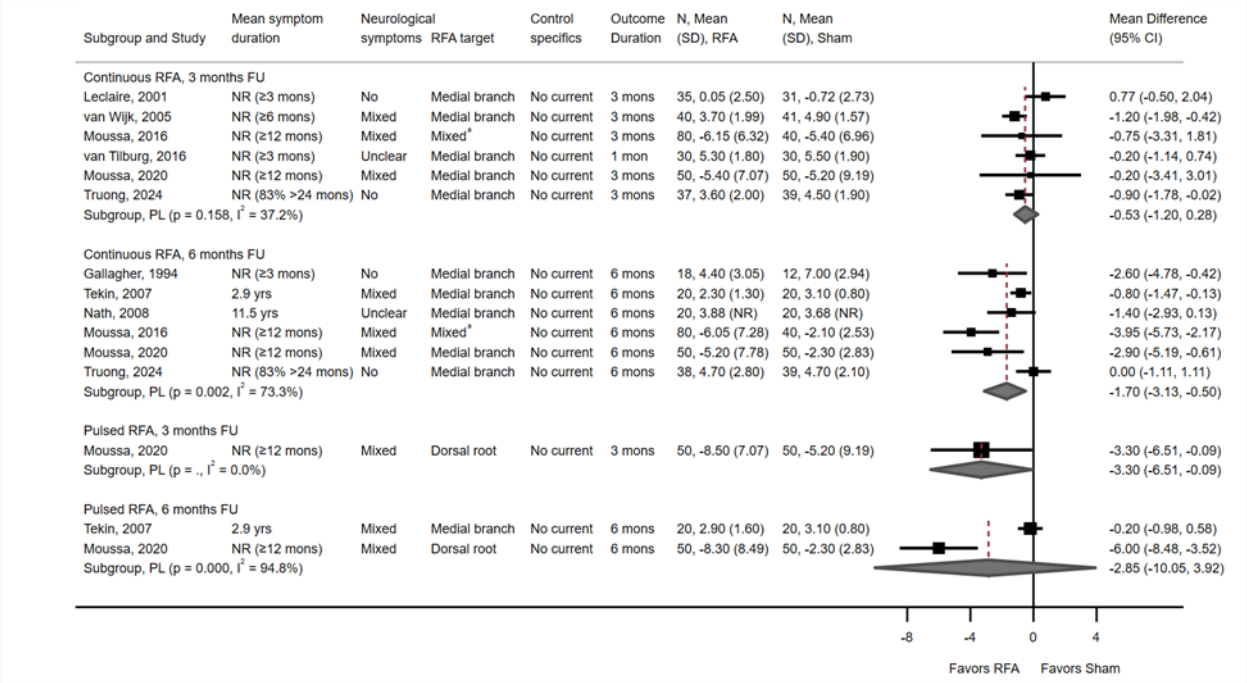
Pain scores. Nine RCTs reported pain scores on a 0-10 VAS or NRS scale.[66,79,88,89,91,93-96]

Continuous RFA was associated with moderate improvement in back pain (6 RCTs, N=407, MD -1.70, 95% CI -3.13 to -0.50, I^2=73.3%)[66,88,89,91,93,94] and large improvement in leg pain (3 RCTs, N=260, MD -2.76, 95% CI -4.83 to -0.88, I^2=65.4%)[88,89,91] at 6 months compared with sham RFA (**Figures 11 and 12**); at the earlier timepoint (3 months), improvement in both back (6 RCTs, N=503, MD -0.53, 95% CI -1.20 to 0.28, I^2=37.2%)[79,88,89,94-96] and leg (3 RCTs, N=301, MD -0.31, 95% CI -1.14 to 0.51, I^2=0%)[88,89,96] pain was similar between treatment groups. The RFA target was the medial branch in all but one trial,[88] which included both RFA of the medial branch and the joint capsule; sensitivity analyses of RFA to the medial branch only provided results similar to the primary analyses (**Appendix B, Figures B-5 and B-6**). Of note, two trials by the same author group (Moussa et al.)[88,89] gave patients a post-treatment injection of steroids and anesthetic through the electrode needle and it is unclear how this may have impacted results. These trials tended to show a larger effect than the other trials at 6 months. Analyses excluding these two trials at 6 months showed an attenuated effect for both back pain (small as opposed to moderate improvement: 4 RCTs, N=187, MD -0.79, 95% CI -1.82 to -0.16, I^2=42.3%)[66,91,93,94] (**Appendix B, Figure B-7**) and leg pain (moderate as opposed to large improvement: 1 RCT, N=40, MD -1.50, 95% CI -2.89 to -0.11)[91] (**Figure 12**). For back pain, imprecision and heterogeneity were also reduced after exclusion of these trials.

Pulsed RFA that targeted the dorsal root ganglion was associated with large improvements in back pain (N=100, MD -3.30, 95% CI -6.51 to -0.09), but similar improvement in leg pain (N=100, MD -3.00, 95% CI -6.21 to 0.21), at 3 months and large improvements in both back and leg pain at 6 months (N=100, MD -6.00, 95% CI -8.48 to -3.52 for both back and leg pain) in one RCT[89] (**Figures 11 and 12**). All estimates from this trial were imprecise. A second, small RCT[93] reported similar improvement in back pain at 6 months with pulsed RFA that targeted the medial branch nerve versus sham RFA (N=40, MD -0.20, 95% CI -0.98 to 0.58). Given the substantial

heterogeneity (I^2=95%) in the pooled estimate across the two pulsed RFA trials, we analyzed them separately; the RFA target (dorsal root vs. medial branch) may explain some of the variability in results.

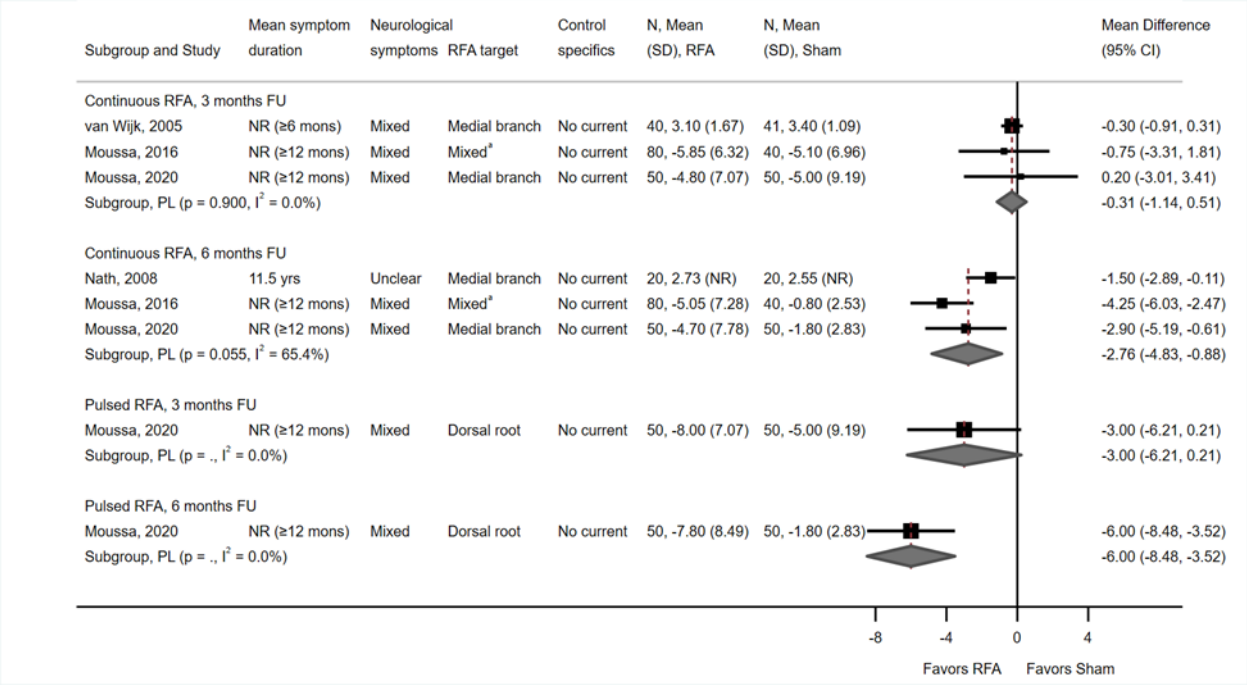
Figure 11. RFA versus sham: back pain scores on 0-10 VAS/NRS



Abbreviations: CI = confidence interval; FU = followup; NR = not reported; NRS = numeric rating scale; PL = profile likelihood; RFA = radiofrequency ablation; SD = standard deviation; VAS = visual analogue scale

[a] Medial branch or joint capsule.

Figure 12. RFA versus sham: leg pain scores on 0-10 VAS/NRS



3.5.2 Results, Key Question 5. Nonsurgical interventions for chronic low back pain, Radiofrequency Ablation

Abbreviations: CI = confidence interval; FU = followup; NR = not reported; NRS = numeric rating scale; PL = profile likelihood; RFA = radiofrequency ablation; SD = standard deviation; VAS = visual analogue scale
[a] Medial branch or joint capsule.

Function response. One RCT (N=66)[79] found continuous RFA associated with a large increase in the odds of achieving a >10-point improvement on the RDQ (0-24 scale) at 3 months compared with sham (adjusted OR 2.79, 95% CI 0.98 to 8.80) in an analysis that controlled for age, gender, number of children, and physical activities.

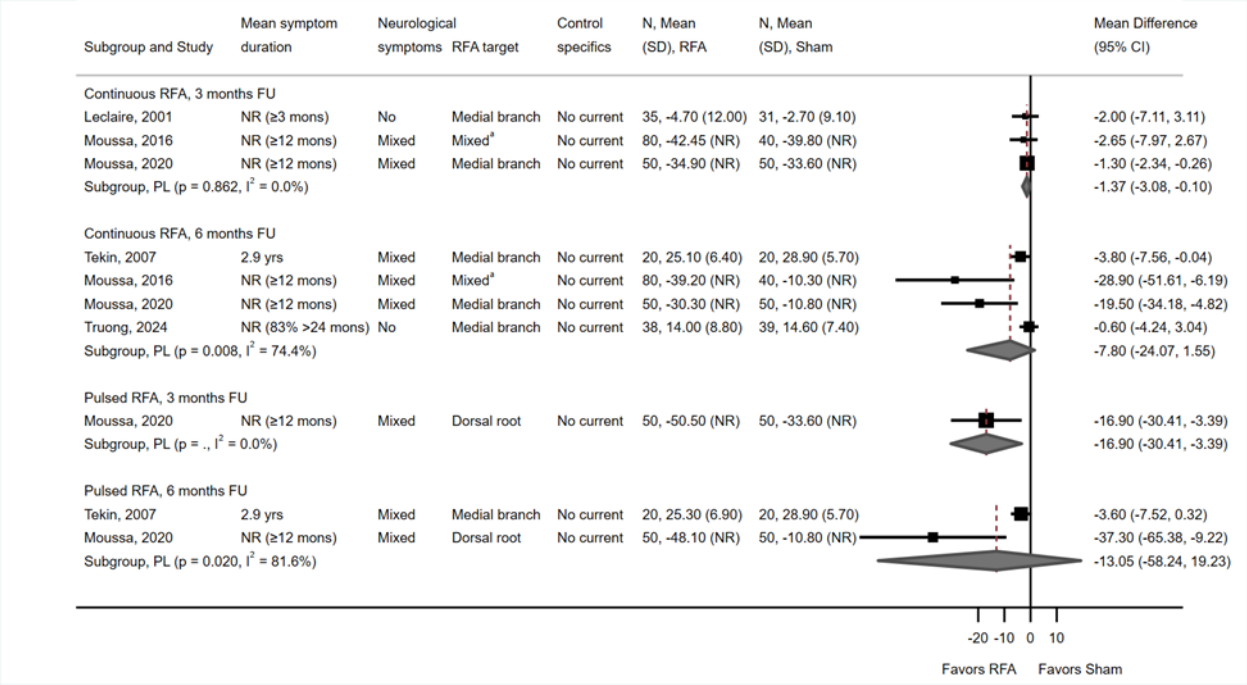
Function scores. Six RCT reported function scores using the ODI (0-100 scale) (5 RCTs)[79,88,89,93,94] or the Physical Activities Scale (0-30) (1 RCT).[96]

Continuous RFA was associated with similar improvement (below the threshold for a small effect) in ODI scores at 3 months (3 RCTs, N=286, MD -1.37, 95% CI -3.08 to -0.10, I^2=0%)[79,88,89] and 6 months (4 RCTs, N=337, MD -7.80, 95% CI -24.07 to 1.55, I^2=74.4%)[88,89,93,94] (**Figure 13**). There was heterogeneity in the pooled estimate at 6 months with three trials showing moderate to large improvement and one trial at low risk of bias[94] showing similar improvement in function with continuous RFA; however, the estimates were imprecise in all trials, especially two trials by the same author group that showed substantial improvement (MDs -28.9 and -19.5). Of note, these two trials by the same author group (Moussa et al.)[88,89] gave patients a post-RFA injection of steroids and anesthetic through the electrode needle and it is unclear how this may have impacted results. In addition, authors did not provide a measure of variance (e.g., SD) for ODI scores in these trials and SDs were estimated using SDs of the other included trials. Exclusion of these two trials at 6 months did not change the overall conclusion, however the effect estimate was attenuated and heterogeneity was reduced (2 RCTs, MD -2.15, 95% CI -6.16 to 1.78, I^2=30.5%),[93,94] and the estimate remained imprecise (**Appendix B, Figure B-8**) The RFA target was the medial branch in all but one trial,[88] which included both RFA of the medial branch and the joint capsule; sensitivity analyses that included only the patients who received RFA to the medial branch provided results similar to the primary analysis (**Appendix B, Figure B-9).** One of these trials (N=66) also reported function using the RDQ and found similar improvement in scores between continuous RFA and sham at 3 months (**Appendix D, Table D-2**).[79] In one RCT, continuous RFA associated with a large improvement in function on the Physical Activities Scale (higher score is better) at 3 months (MD 2.80, 95% CI 0.90 to 4.70).[96]

Pulsed RFA of the dorsal root ganglion was associated with a large improvement in function at 3 months (N=100, MD -16.90, 95% CI -33.44 to -0.36) and 6 months (N=100, MD -37.30, 95% CI -70.91 to -3.69) compared with sham RFA in one trial,[89] however, the estimates were very imprecise. A second, small RCT[93] reported similar improvement function at 6 months with pulsed RFA that targeted the medial branch nerve versus sham RFA (N=40, MD -3.60, 95% CI -7.52 to 0.32).

3.5.2 Results, Key Question 5. Nonsurgical interventions for chronic low back pain, Radiofrequency Ablation

Figure 13. RFA versus sham: function scores on 0-100 ODI



Abbreviations: CI = confidence interval; FU = followup; NR = not reported; ODI = Oswestry Disability Index; PL = profile likelihood; RFA = radiofrequency ablation; SD = standard deviation.
[a] Medial branch or joint capsule.

Quality of life. Compared with sham, continuous RFA was associated with similar improvement in EQ5D scores (-1 to 1 scale) at 6 months (N=77, MD -0.04, 95% CI -0.13 to 0.05) in one trial.[94] A second trial[96] reported the eight individual subscales of the SF-36; only one subscale, vitality, showed a difference favoring continuous RFA versus sham at 3 months.

Harms and adverse events. Seven RCTs reported harms;[66,79,91,93-96] all trials used continuous RFA except for one trial[93] that included both a pulsed and a continuous RFA arm. In one trial,[96] compared with sham RFA, continuous RFA was not associated with an increased risk of serious or severe adverse events, including treatment-related pain necessitating analgesics (N=78, 35.9% vs. 25.6%; RR 1.40, 95% CI 0.71 to 2.76), evident dysaesthesia or allodynia (N=79, 2.6% vs. 0%), or evident motor loss (N=79, no cases in either group) or of mild to moderate adverse events including treatment related pain (N=78, 33.3% vs. 20.5%, RR 1.63, 95% CI 0.76 to 3.48), discrete or irritating sensory (N=79, 2.6% vs. 2.5%, RR 1.03, 95% CI 0.07 to 15.83) or discrete or irritating motor function changes (N=79, 5.3% vs. 4.8%, RR 1.08, 95% CI 0.16 to 7.28). A second RCT (N=80)[94] reported that two patients (5.0%) who received continuous RFA complained of new local pain and paresthesia in untreated parts of lower back that resolved within 3 months after a short-term, low-dosage treatment with gabapentin; no adverse effects were noted in the sham group. The remaining RCTs provided statements that no serious adverse events (1 RCT, N=60),[95] no procedure-related adverse events (1 RCT, N=60, pulsed and continuous RFA)[93] and no (any) adverse events (3 RCTs, N=151)[66,79,91] occurred.

Differential effectiveness and safety. The trials included did not report on differential effectiveness and safety for RFA versus sham.

3.5.2.2 Radiofrequency Ablation Versus Usual Care

3.5.2.2.1 Description of Included Studies and Detailed Synthesis

One RCT (N=270)[58] conducted in China compared RFA to usual conservative care for the treatment of CLBP (mean duration 8.7 months) of confirmed facet origin in patients older than 60 years (mean 72 years) (**Appendix D, Table D-2**). Sixty-three percent were female. Concomitant diagnoses included: disc degeneration (15%), ligament calcification (9%), failed back surgery syndrome (21%), and radiating pain (31%). No other characteristics or comorbidities were reported. A positive diagnostic block (≥50% pain relief from 1% lidocaine 0.5 ml) was required for inclusion.

Continuous RFA was performed at 75 degrees Celsius for 120 seconds under fluoroscopic guidance (tip/electrode size was not reported). Usual care consisted of NSAIDs, glucosamine sulfate, physical therapy, and exercise rehabilitation, but no further information was provided.

The trial was rated as moderate risk of bias due to the inability to blind the patient and the provider (**Appendix E, Table E-1**). The trial did not report funding.

Pain response and pain scores. Continuous RFA was associated with a large increase in the likelihood of achieving ≥2-point improvement on the NRS (0-10) at 3 months (N=254, 61.1% vs. 26.0%, RR 2.35, 95% CI 1.69 to 3.26) and 6 months (N=232, 52.9% vs. 22.5%, RR 2.35, 95% CI 1.60 to 3.45) as well as a moderate improvement in NRS pain scores at 3 months (N=254, MD -1.10, 95% CI -1.62 to -0.58) and 6 months (N=232, MD -1.20, 95% CI -1.82 to -0.58) compared with usual conservative care.[58]

Function response and function scores. Continuous RFA was associated with a small increase in the likelihood of achieving a ≥15-point improvement on the ODI (0-100) at 3 months (N=254, 61.1% vs. 26.0%, RR 2.35, 95% CI 1.69 to 3.26) compared with usual conservative care. At 6 months, the authors reported a statistically significant difference between groups favoring RFA for this outcome (p=0.031) but there was no treatment effect according to our calculations (N=232, 36.4% vs. 27.0%, RR 1.35, 95% CI 0.91 to 1.98).[58]

RFA resulted in significantly greater improvement in ODI scores at 3 months (N=254, MD -4.40, 95% CI -8.57 to -0.23) and 6 months (N=232, MD -4.30, 95% CI -7.57 to -1.03) compared with usual care, however the estimates at both timepoints were below the threshold for a small effect.[58]

Harms and adverse events. The authors stated that there were no surgical complications such as infection, bleeding, numbness, nerve injury, or muscle strength weakness in the RFA group in the perioperative period.[58] Adverse events were not reported in the usual care group. Fewer patients who received RFA underwent other lumbar surgery by 6 months (N=270, 6.7% vs. 14.8%, RR 0.45, 95% CI 0.22 to 0.95, RD -8.2%, 95% CI -15.5% to -0.8%).

Differential effectiveness and safety. The trial did not report on differential effectiveness and safety for RFA versus usual care.

3.5.2.3 Radiofrequency Ablation Versus Facet Joint (Intraarticular) Injection

3.5.2.3.1 Description of Included Studies

Four RCTs (total N=235; N range, 39 to 80)[60,61,78,87] compared RFA to facet joint injection for the treatment of CLBP of suspected facet origin based on positive diagnostic blocks (**Appendix D, Table D-2**).

The average study mean age of participants was 57 years (range 50 to 65 years) and the average proportion of females was 52.1 percent (range 36.5% to 60.0%). One trial[87] reported patient BMI (mean 27.6). No trials reported race or ethnicity. Mean baseline pain severity was 6.2 (range 5.0 to 7.2). Mean duration of pain was 23.9 months (range 11 to 61.2 months) in three trials; one trial[78] required greater than 24 months of pain for inclusion (mean not reported). Two trials[61,87] excluded participants with radicular pain, but neurologic symptoms and comorbidities were otherwise not well-reported. Two trials excluded patients with a history of previous lumbar surgery[61,87] and one excluded patients who had received prior RFA.[87] Two trials required 50 percent or more[78] and 80 percent or more[87] pain relief following medical branch blocks with bupivacaine and/or lidocaine and one required 50 percent pain relief after intraarticular injection of lidocaine[60]; the fourth trial did not describe diagnostic injections.

One trial employed continuous RFA of the medial branch nerve at 80 degrees Celsius for 90 seconds with a 100 mm electrode tip.[78] Two trials used pulsed RFA targeting the facet joint space at 40 to 42 degrees Celsius for 6 minutes.[60,61] In one of these trials,[60] pulsed RFA was administered at 5 Hz and a 5-millisecond pulsed width at 55 V using a 100 mm electrode tip while in the other trial,[61] there was no information on the pulsed RFA parameters. The fourth trial used water-cooled RFA at 60 degrees Celsius for 165 seconds (electrode tip size not reported).[87] Intraarticular steroid injections were performed via posterior approach under fluoroscopic guidance. In one trial,[61] it was unclear if the injection target was the facet joint or the medial branch nerve; after clinical input the intervention was determined to be most consistent with a facet joint injection. A different steroid was used in all four trials: betamethasone 3 mg (1 mL),[78] methylprednisolone 20 mg (0.5 mL),[61] dexamethasone 10 mg (0.25 mL)[60] and triamcinolone 20 mg (0.5 mL);[87] three trials used bupivacaine (0.125% to 0.5%)[60,61,78] and one used lidocaine 2 percent[87] for anesthetic (the latter trial offered additional sedation beyond local anesthetic on a case-by-case basis). One trial of continuous RFA[78] performed sham RFA as part of the procedure.

One trial each was conducted in the United States,[87] Germany,[78] Korea,[60] and Turkey.[61] One trial[87] reported mixed industry and university funding; the other trials did not report a funding source.

One trial was assessed as low risk of bias,[78] one as moderate risk of bias,[60] and two as high risk of bias.[61,87] Methodological limitations in the trial considered moderate risk of bias included unclear allocation concealment methods, unclear if patients and/or providers were blinded and unclear or unacceptable differential loss to followup (**Appendix E, Table E-1**). Additionally, the trials assessed as high risk of bias had unclear randomization methods, imbalances in baseline characteristics between groups, no reporting of attrition or crossover, and unclear or higher than acceptable loss to followup, and lack of or unclear use of intention-to-treat analyses.

3.5.2.3.2 Detailed Synthesis

Pain response. Two RCTs[60,87] reported the proportion of patients who achieved ≥50 percent improvement in pain on VAS/NRS (**Figure 14**). Continuous RFA was associated with a similar likelihood of achieving pain response at 6 months compared with an intraarticular injection of dexamethasone 10 mg in one trial (N=60, 50.0% vs. 46.7%, RR 1.07, 95% CI 0.63 to 1.81).[60] In the second, smaller trial (N=32),[87] cooled RFA was associated with a large increase in the likelihood of achieving pain response at 3 months compared with an intraarticular injection of triamcinolone 20 mg (70.0% vs. 25.0%, RR 2.80, 95% CI 1.01 to 7.77) but the difference was no longer statistically significant at 6 months (55.0% vs. 25.0%, RR 2.20, 95% CI 0.76 to 6.33). Results were the same in the latter trial when a cut-off of 2-point or more improvement on NRS was used to measure pain response.

Figure 14. RFA versus facet joint injection: pain response, ≥50% improvement in back pain on 0-10 VAS/NRS

Abbreviations: CI = confidence interval; FU = followup; IA = intraarticular; NRS = numeric rating scale; PL = profile likelihood; RFA = radiofrequency ablation; VAS = visual analogue scale

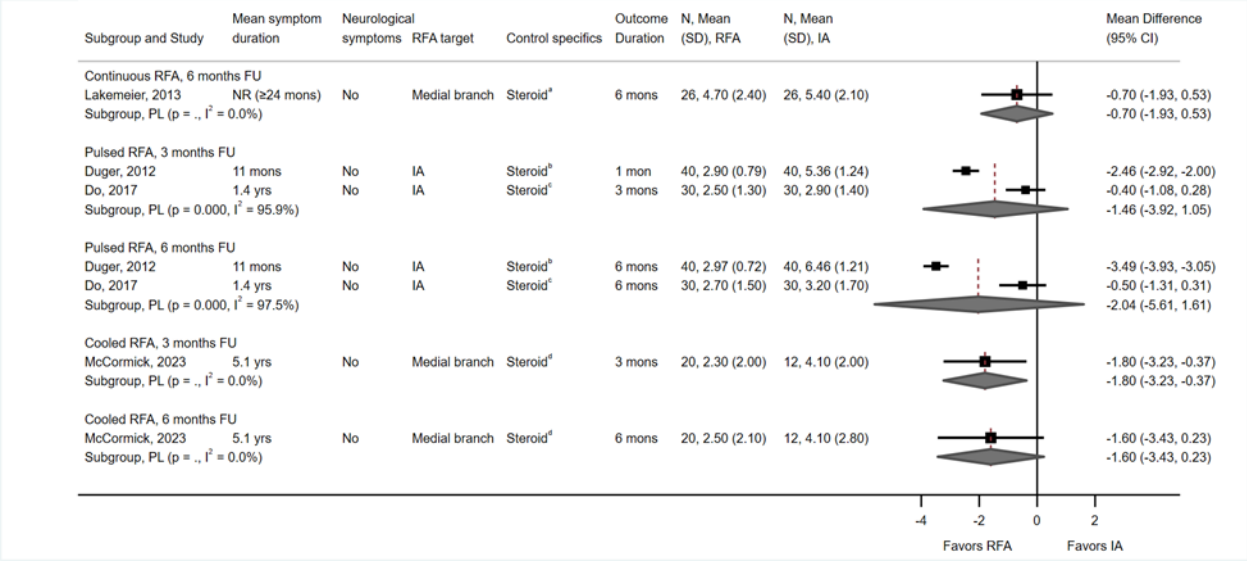
[a] Dexamethasone 10 mg + bupivacaine 0.125%.

[b] Triamcinolone 20 mg + lidocaine 2%.

Pain scores. Four RCTs reported pain scores on a 0-10 VAS or NRS scale (**Figure 15**).[60,61,78,87] One trial evaluated continuous RFA and found similar improvement in pain at 6 months compared with an intraarticular steroid injection (N=52, MD -0.70, 95% CI -1.93 to 0.53).[78] Two trials evaluated pulsed RFA[60,61] and reported conflicting results; the pooled analyses were marked by substantial heterogeneity (I^2>95%). The better quality trial (i.e., moderate risk of bias)[60] found RFA associated with similar improvement in pain scores at 3 months (N=60, MD -0.40, 95% CI -1.08 to 0.28) and 6 months (N=60, MD -0.50, 95% CI -1.31 to 0.31) compared with an intraarticular steroid injection, while the trial at high risk of bias[61] found RFA associated with large improvements in pain at 1 month (N=80, MD -2.46, 95% CI -2.92 to -2.00) and 6 months (N=80, MD -3.49, 95% CI -3.93 to -3.05) compared with steroid injection. The fourth trial evaluated cooled RFA[87] which was associated with moderate improvement in pain scores at 3 months (N=32, MD -1.80, 95% CI -3.23 to -0.37) but similar improvement at 6 months (N=32, MD -1.60, 95% CI -3.43 to 0.23) compared with an intraarticular steroid injection. Each trial used a different steroid for the intraarticular injection (methylprednisolone, betamethasone, triamcinolone, dexamethasone).

3.5.2 Results, Key Question 5. Nonsurgical interventions for chronic low back pain, Radiofrequency Ablation

Figure 15. RFA versus facet joint injection: back pain scores on 0-10 VAS/NRS



Abbreviations: CI = confidence interval; FU = followup; IA = intraarticular; NRS = numeric rating scale; PL = profile likelihood; RFA = radiofrequency ablation; SD = standard deviation; VAS = visual analogue scale

[a] Betamethasone 3 mg + bupivacaine 0.5%.

[b] Methylprednisolone 20 mg + bupivacaine 0.5%.

[c] Dexamethasone 10 mg + bupivacaine 0.125%.

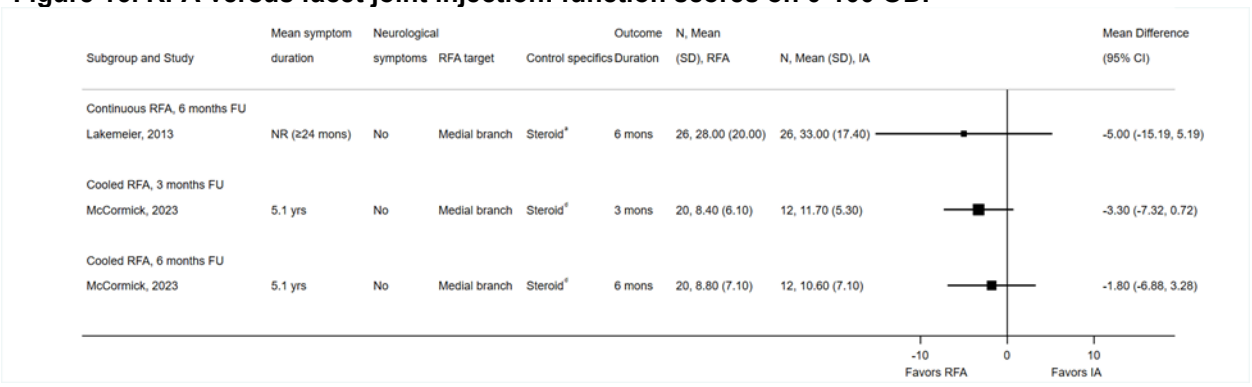
[d] Triamcinolone 20 mg + lidocaine 2%.

Function response. In one small RCT (N=32),[87] cooled RFA was associated with a similar likelihood of achieving function response on the ODI (0-100 scale) compared with an intraarticular injection of triamcinolone 20 mg at 3 months (≥15-point improvement: 15.0% vs. 8.3%, RR 1.80, 95% CI 0.21 to 15.41; ≥30% improvement: 60.0% vs. 33.3%, RR 1.80, 95% CI 0.75 to 4.32) and 6 months (≥15-point improvement: 20.0% vs. 16.7%, RR 1.20, 95% CI 0.26 to 5.59; ≥30% improvement: 40.0% vs. 25.0%, RR 1.60, 95% CI 0.52 to 4.89). Estimates were imprecise.

Function scores. In two trials, continuous RFA (1 RCT)[78] and cooled RFA (1 RCT)[87] were associated with similar improvement in ODI scores at all timepoints compared with intraarticular steroid injections (**Figure 16**). Estimates were imprecise. One trial assessed as high risk of bias found pulsed RFA associated with a large improvement in function based on the daily activity score (1 [poor] to 4 [very good] scale) compared with intraarticular injection of methylprednisolone 20 mg (N=80, MD 0.84, 95% CI 0.64 to 1.04);[61] however, this does not appear to be a validated measure.

3.5.2 Results, Key Question 5. Nonsurgical interventions for chronic low back pain, Radiofrequency Ablation

Figure 16. RFA versus facet joint injection: function scores on 0-100 ODI



Abbreviations: CI = confidence interval; FU = followup; ODI = Oswestry Disability Index; RFA = radiofrequency ablation; SD = standard deviation

[a] Betamethasone 3 mg + bupivacaine 0.5%.

[b] Triamcinolone 20 mg + lidocaine 2%.

Harms and adverse events. Two trials (N=95), one of continuous RFA[78] and one of cooled RFA,[87] reported that no serious or major adverse events occurred in either group during followup. Across two trials of pulsed RFA, one RCT (N=60)[60] reported a single case (3.3%) of hyperglycemia (minor event) in a patient randomized to intraarticular injection and the other RCT (N=80) reported that no complications were observed in either group.[61]

Quality of life and differential effectiveness and safety. The trials included did not report on quality of life outcomes or differential effectiveness and safety for RFA versus facet joint injections.

3.5.2.3 Radiofrequency Ablation Versus Medial Branch Block

3.5.2.3.1 Description of Included Studies

Three RCTs (total N=260; N range, 80 to 100)[59,68,97] compared RFA to medial branch block (MBB) for the treatment of CLBP of suspected facet origin based on diagnostic blocks (**Appendix D, Table D-2**). One trial[68] also included patients with a concomitant diagnosis of grade I DLS.

The average study mean age of participants was 60 years (range 56 to 64 years) and 44 percent (range 43.0% to 45.0%) were female in two trials (one trial did not report age or sex[59]). Mean BMI was 23 in one trial (not reported in the other trials).[68] Average baseline pain severity was 7.6 on a 0 to 10 scale (range 6.8 to 8.4). Mean duration of pain was 34.4 months (range 25.5 months to 43.2 months) in two trials; in one trial[59] a minimum pain duration of 6 weeks was an inclusion criterion (mean not reported). All trials excluded patients with radicular pain or symptomatic radiculopathy or neurological deficits and most trials excluded patients with major comorbidities. History of previous lumbar surgery or prior treatment with RFA or intraarticular injections was not reported. One trial[97] required ≥80 percent pain relief after an intraarticular or medial branch block of lidocaine and two trials required "pain relief" (not further described) after a medial branch block of lidocaine[68] or a facet joint injection[59] for inclusion.

Two trials used continuous RFA[59,97] and one trial used pulsed RFA[68] targeting the medial branch nerves. Continuous RFA was performed at 80 degrees Celsius for 120 seconds using a 5 mm active tip under fluoroscopic guidance in one trial[59] or for 90 seconds under digital subtraction angiography guidance (tip size not reported) in the other trial.[97] Pulsed RFA was

performed in two 20 millisecond pulses per second at 42 degrees Celsius for 120 seconds under fluoroscopic guidance.[68] MBB injections were performed via a posterior approach under fluoroscopic guidance in one trial,[59] a sagittal approach under digital subtraction angiography tube angiographic guidance in one trial[97] and the third trial[68] did not provide this information. A different steroid was used in each trial in conjunction with an anesthetic (bupivacaine or lidocaine, 1-2 ml): methylprednisolone 40 mg (1 ml),[59] triamcinolone 40 mg (1 ml)[68] and betamethasone (1 ml) (dose unspecified).[97] Concomitant treatment include NSAIDs or other analgesics for pain control in two trials.[59,68] One trial reported referring patients who responded favorably after 1 week to a spine rehabilitation program for 4 to 6 weeks and offering those who did not respond favorably surgery or physical therapy. However, it is unclear whether or when patients may have received these interventions.

One trial each was conducted in Turkey,[59] Iran,[68] and China.[97] Funding was not reported.

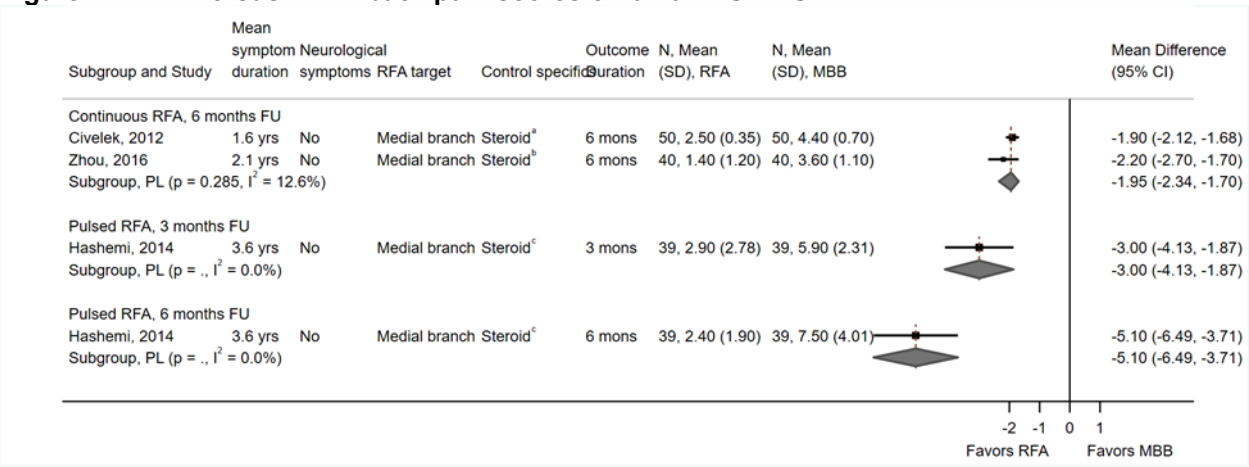
Two trials were assessed as moderate risk of bias,[59,68] and one as high risk of bias.[97] Common limitations included lack of or unclear blinding status of care providers and participants and unclear allocation concealment (**Appendix E, Table E-1**). Additionally, the high risk of bias trial contained unclear reporting of attrition or crossover, unclear or unacceptable loss to followup, and lack of or unclear use of intention-to-treat analyses.

3.5.2.3.2 Detailed Synthesis

Pain response and pain scores. Continuous RFA was associated with a small increase in the likelihood of achieving pain response (i.e., ≥50% improvement on VAS back pain) at 6 months compared with a MBB (N=100, 90.0% vs. 67.7%, RR 1.32, 95% CI 1.07 to 1.64) in one RCT.[59]

Three RCTs reported pain scores on a 0-10 VAS or NRS.[59,68,97] Compared with MBB, two trials[59,97] found continuous RFA associated with a moderate improvement in VAS pain scores at 6 months (N=180, MD -1.95, 95% CI -2.34 to -1.70, I^2=12.6%) and the third trial[68] found pulsed RFA associated with a large improvement in NRS pain scores at 3 months (N=78, MD -3.00, 95% CI -4.13 to -1.87) and 6 months (N=78, MD -5.10, 95% CI -6.49 to -3.71) (**Figure 17**). Each trial used a different steroid for the MBB (methylprednisolone, betamethasone, and triamcinolone).

Figure 17. RFA versus MBB: back pain scores on 0-10 VAS/NRS



Abbreviations: CI = confidence interval; FU = followup; MBB = medial branch block; NRS = numeric rating scale; PL = profile likelihood; RFA = radiofrequency ablation; SD = standard deviation; VAS = visual analogue scale

[a] Methylprednisolone 40 mg (1 ml) + 0.25%-0.5% bupivacaine (1.5-2 ml)

[b] Betamethasone (dose NR, 1ml) + 2% lidocaine (1 ml)

[c] Triamcinolone 40 mg (1 ml) + 0.5% bupivacaine (0.5 ml)

Function scores. In one RCT,[68] pulsed RFA was associated with a large improvement in ODI scores (0-100 scale) at 3 months (N=78, MD -21.00, 95% CI -33.02 to -8.98) and 6 months (N=78, MD -38.70, 95% CI -51.16 to -26.24) compared with MBB. None of the other trials reported on function.

Quality of life. Continuous RFA was associated with a small increase in the likelihood of achieving quality of life success, defined as a score <9 on the EQ-5D (5-15 scale), at 6 months compared with MBB (N=100, 92.0% vs. 75.4%, RR 1.21, 95% CI 1.01 to 1.44) in one RCT.[59] Mean EQ-5D scores were similar between the groups in this trial (6.5 vs. 7.2, p=0.22; data in graph).

Harms and adverse events. In one RCT (N=100)[59] two patients (4%) who received RFA experienced superficial burns (i.e., burning like sensation in the lesion-performed region and increase in the severity of LBP) which resolved after 6 to 8 weeks with medication for neuropathy. This same trial reported no cases of infection or new motor or sensory deficits in either group. A second trial (N=80)[97] reported no cases of nerve root injury or back skin anesthesia.

Differential effectiveness and safety. The trials included did not report on differential effectiveness and safety for RFA versus MBB.

3.5.3 Epidural Steroid Injections

3.5.3.1 Epidural Steroid Injection Versus Placebo Injection

3.5.3.1.1 Description of Included Studies

Four RCTs (reported in 7 publications) (total N=656; N range, 36 to 400)[64,65,77,82,83,90] compared epidural steroid injections (ESI) with placebo injection for the treatment of chronic

back pain due to central spinal stenosis (3 RCTs)[64,65,77,82,83] and degenerative scoliosis with foraminal stenosis (1 RCT) (**Appendix D, Table D-2**).[90] Additionally, all patients had neurological symptoms including radicular pain in three trials[82,83,90] and neurogenic claudication in one trial.[64,65,77]

The average age of participants was 62 years (range 52 to 73) and 62 percent (range 55% to 75%) were female. Two trials reported BMI (mean 27; range 23 to 30).[65,90] One trial reported race (white 69%, black 26%, Hispanic 4%, other 5%).[65] Mean duration of symptoms across three trials was 6.2 years (range 0.6 to 9.6 years);[82,83,90] the fourth trial[64,65,77] reported symptom durations of 3 months to 1 year (30%), 1 to 5 years (27%) and >5 years (26%). Most trials excluded major medical comorbidities (e.g., uncontrolled diabetes, cardiac disease, malignancy). In one trial,[64,65,77] 8 percent of patients were diabetic and receiving insulin. In three trials,[64,65,77,82,83] many patients (21% in one trial) were taking opioids at baseline.

ESI was performed using an interlaminar approach in two trials,[64,65,82] a transforaminal approach in two trials[64,65,90] and a caudal approach in one trial.[83] Three trials used the corticosteroid betamethasone (6 to 12 mg),[64,65,82,83] two trials used triamcinolone (20 mg to 120 mg),[64,65,90] and one trial also used dexamethasone (nonparticulate) or methylprednisolone (range 8 to 10 mg),[64,65,77] The steroids were a mix of particulate (triamcinolone, methylprednisolone, betamethasone) and nonparticulate (dexamethasone and betamethasone) formulations; one trial[83] specified use of a nonparticulate formulation of betamethasone. The total steroid injectate volume was 1 to 3 ml in all trials. All patients received 0.25 to 1 percent lidocaine (1 to 9 ml injectate volume) for a total injectate volume that ranged from 2 to 10 ml. The procedure for the placebo injection was identical to that of the steroid injection except that only the lidocaine was injected. All procedures were performed under fluoroscopic guidance. Concomitant therapies included a structured therapeutic exercise program in two trials[82,83] and continuation of any previous pharmacological therapy, including opioids, in three trials[64,65,77,82,83] One trial[64,65,77] allowed patients to crossover to the other intervention at 6 weeks (30% ESI vs. 45% placebo crossed over).[64] All four trials allowed patients to receive repeat injections at specific intervals or if there was no progress.[64,65,77,82,83,90]

Three RCTs[64,65,77,82,83] were conducted in the United States and one[90] in South Korea. The source of funding was university[90] and government[64,65,77] in one trial each; two trials did not report the funding source.[82,83]

One trial was assessed at low risk of bias,[64,65,77] two at moderate risk of bias[82,83] and one at high risk of bias (**Appendix E, Table E-1**).[90] Common methodological limitations in the moderate risk of bias trials included unclear allocation concealment methods and imbalances in some baseline characteristics. Additionally, in the trial rated high risk of bias, patients and clinicians were not blinded, attrition was higher than acceptable, and there were concerns about violation of the intent-to-treat principle due to unclear exclusions post randomization.

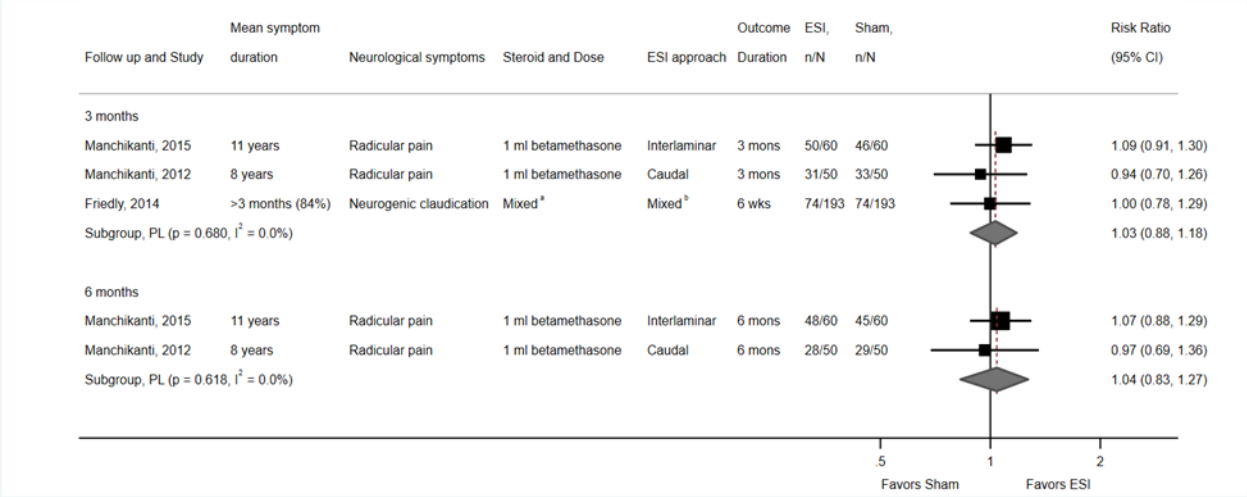
3.5.3.1.3 Detailed Synthesis

Pain response. Three RCTs reported leg pain response defined as an improvement of ≥50 percent on a 0-10 VAS or NRS (**Figure 18**)[64,82,114] ESI was associated with a similar likelihood of achieving pain response at 6 weeks to 3 months (3 RCTs, N=606, 51.2% vs. 50.5%, RR 1.03, 95% CI 0.88 to 1.18, I^2=0%)[64,82,114] and 6 months (2 RCTs, N=220, 69.1% vs. 67.3%, RR 1.04, 95% CI 0.83 to 1.27, I^2=0%)[82,114] compared with placebo in patients with central canal stenosis. One trial (N=386)[64] used an alternative cut-off for response (≥30% improvement in VAS leg pain) at 6 weeks and found similar results (49.2% vs. 49.7%, RR 0.99, 95% CI 0.81 to 1.21).

3.5.3 Results, Key Question 5. Nonsurgical interventions for chronic low back pain, Epidural Steroid Injection

Each trial used a different approach to ESI (interlaminar, caudal, or mixed transforaminal/interlaminar); two trials included patients with radicular pain[82,114] and one included patients with neurogenic claudication[64] at presentation.

Figure 18. ESI versus placebo: ≥50% improvement in leg pain on 0-10 VAS or NRS



Abbreviations: CI = confidence interval; ESI = epidural steroid injection; NRS = numeric rating scale; PL = profile likelihood; VAS = visual analogue scale

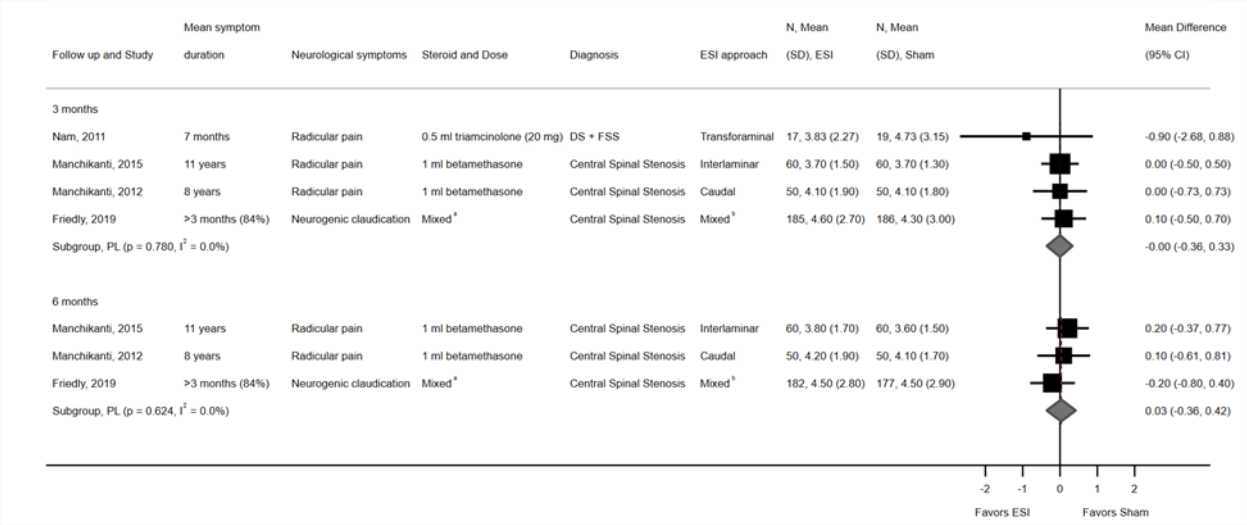
[a] 1 to 3 ml of triamcinolone (60 to 120 mg), betamethasone (6 to 12 mg), dexamethasone (8 to 10 mg), or methylprednisolone (60 to 120 mg).

[b] Transforaminal or interlaminar.

Pain scores. Four RCTs reported leg pain on a 0-10 VAS or NRS (**Figure 19**).[64,82,90,114] ESI was associated with similar improvement in leg pain at 3 months (4 RCTs, N=627, MD -0.00, 95% CI -0.36 to 0.33, I^2=0%)[64,82,90,114] and 6 months (3 RCTs, N=579, MD 0.03, 95% CI -0.36 to 0.42, I^2=0%)[64,82,114] compared with placebo. Results were similar after exclusion of the trial at high risk of bias[90] (**Appendix B, Figure B-10**) and across ESI approaches (**Appendix B, Figures B-11 and B-12**), patient diagnosis, and patient neurological symptoms at presentation. In the trial that included both an interlaminar (71%) or transforaminal (39%) approach a prespecified secondary subgroup analysis with stratification according to type of approach likewise showed no significant differences after 3 weeks.[64,65]

3.5.3 Results, Key Question 5. Nonsurgical interventions for chronic low back pain, Epidural Steroid Injection

Figure 19. ESI versus placebo: leg pain scores on a 0-10 VAS/NRS



Abbreviations: CI = confidence interval; DS = degenerative scoliosis; ESI = epidural steroid injection; FSS = foraminal spinal stenosis; NRS = numeric rating scale; PL = profile likelihood; SD = standard deviation; VAS = visual analogue scale

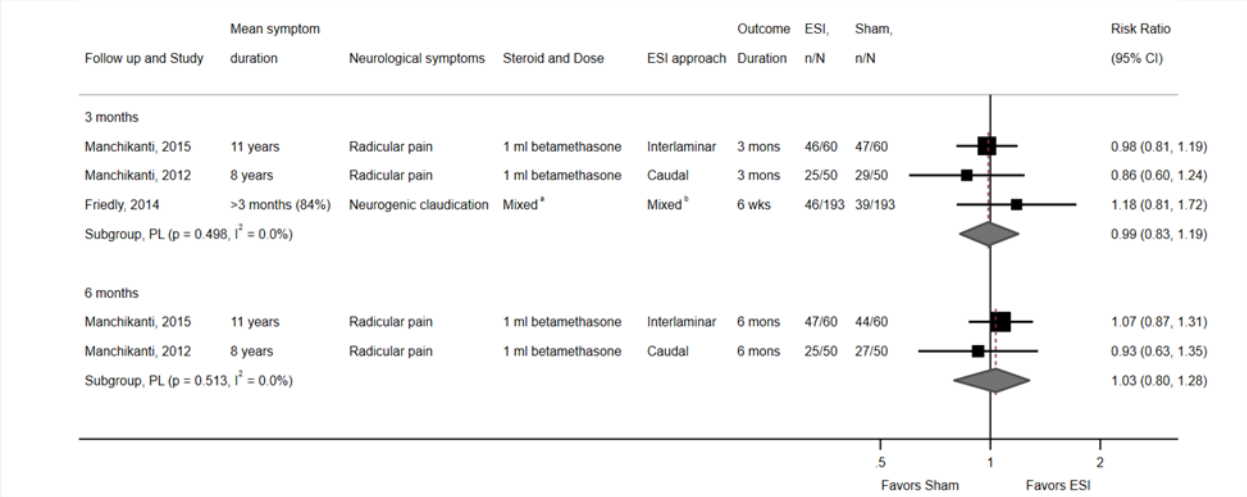
[a] 1 to 3 ml of triamcinolone (60 to 120 mg), betamethasone (6 to 12 mg), dexamethasone (8 to 10 mg), or methylprednisolone (60 to 120 mg).

[b] Transforaminal or interlaminar.

Function response. Three RCTs reported function response defined as an improvement of ≥50 percent on the ODI (0-100)[82,114] or the RDQ (0-24)[64] (**Figure 20**). ESI was associated with a similar likelihood of achieving function response at 6 weeks to 3 months (3 RCTs, N=606, 38.6% vs. 38.0%, RR 0.99, 95% CI 0.83 to 1.19, I^2=0%)[64,82,114] and 6 months (2 RCTs, N=220, 65.5% vs. 64.5%, RR 1.03, 95% CI 0.80 to 1.28, I^2=0%)[82,114] compared with placebo in patients with central canal stenosis. One trial (N=386)[64] used an alternative cut-off for response (≥30% improvement on the RDQ) at 6 weeks and found similar results (37.3% vs. 31.6%, RR 1.18, 95% CI 0.90 to 1.56). Each trial used a different approach for the ESI (interlaminar, caudal, or mixed interlaminar/transforaminal); two trials included patients with radicular pain,[82,114] and one included patients with neurogenic claudication[64] at presentation.

3.5.3 Results, Key Question 5. Nonsurgical interventions for chronic low back pain, Epidural Steroid Injection

Figure 20. ESI versus placebo: ≥50% improvement in function on ODI or RDQ



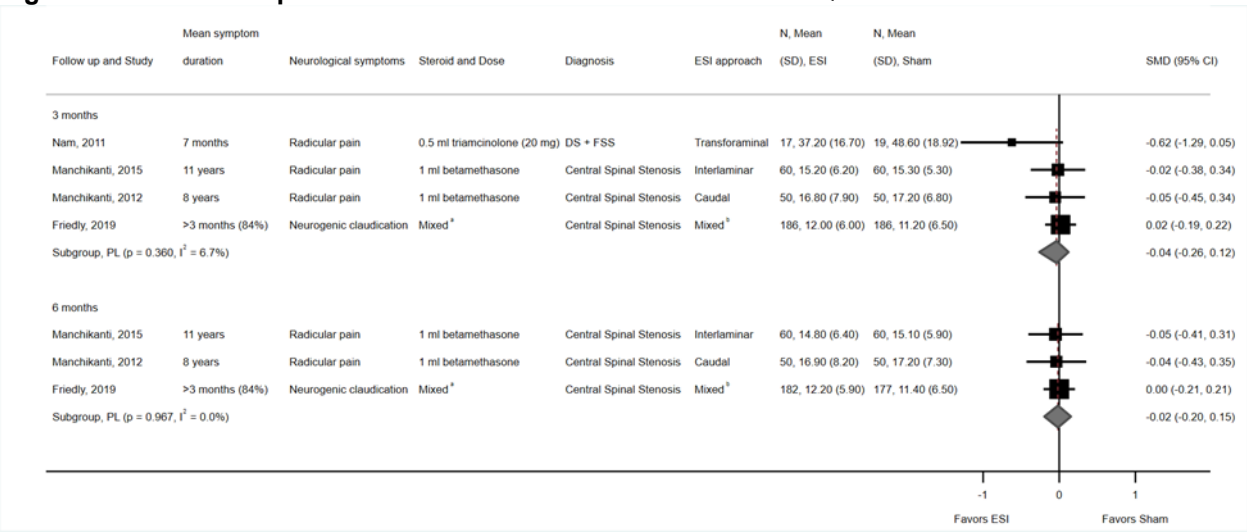
Abbreviations: CI = confidence interval; ESI = epidural steroid injection; ODI = Oswestry Disability Index; PL = profile likelihood; RDQ = Roland Morris Disability Questionnaire

[a] 1 to 3 ml of triamcinolone (60 to 120 mg), betamethasone (6 to 12 mg), dexamethasone (8 to 10 mg), or methylprednisolone (60 to 120 mg).

[b] Transforaminal or interlaminar.

Function scores. Four RCTs reported function scores using the ODI (0-50 or 0-100) or the RDQ (0-24) (**Figure 21**).[64,82,90,114] ESI was associated with similar improvement function at 3 months (4 RCTs, N=628, standardized mean difference [SMD] -0.04, 95% CI -0.26 to 0.12, I^2=6.7%)[64,82,90,114] and 6 months (3 RCTs, N=579, SMD -0.02, 95% CI -0.20 to 0.15, I^2=0%)[64,82,114] compared with placebo. Results were similar after exclusion of the trial at high risk of bias[90] (**Appendix B, Figure B-13**) and across ESI approaches (**Appendix B, Figures B-14 and B-15**), patient diagnosis, and patient neurological symptoms at presentation. In the trial that included both an interlaminar (71%) or transforaminal (39%) approach a prespecified secondary subgroup analysis with stratification according to type of approach likewise showed no significant differences after 3 weeks.[64,65]

Figure 21. ESI versus placebo: function scores on the ODI or RDQ



3.5.3 Results, Key Question 5. Nonsurgical interventions for chronic low back pain, Epidural Steroid Injection

Abbreviations: CI = confidence interval; DS = degenerative scoliosis; ESI = epidural steroid injection; FSS = foraminal spinal stenosis; ODI = Oswestry Disability Index; PL = profile likelihood; RDQ = Roland Morris Disability Questionnaire; SD = standard deviation

[a] 1 to 3 ml of triamcinolone (60 to 120 mg), betamethasone (6 to 12 mg), dexamethasone (8 to 10 mg), or methylprednisolone (60 to 120 mg).

[b] Transforaminal or interlaminar.

Quality of life. ESI was associated with a similar improvement in EQ-5D scores (0-1 scale) at 6 weeks in one RCT (N=386, MD 0.03, 95% CI -0.01 to 0.07).[65] Quality of life was not reported at later timepoints in this trial or at any time by the other trials.

Harms and adverse events. The risk of serious adverse events, i.e., hospitalization, surgery, or both was similar for ESI versus placebo (lidocaine) injection in one trial (N=400, 2.5% vs. 2.0%, RR 1.25, 95% CI 0.34 to 4.59); only one event (0.5%), which occurred in the placebo group, was considered likely procedure related (no further information provided).[65] Two trials (N=220) reported that no serious or major adverse events occurred in any patient.[82,114]

The risk of at least one adverse event (any) (21.5% vs. 15.5%, RR 1.39, 95% CI 0.91 to 2.11) and at least one procedure-related adverse event (15.0% vs. 9.5%, RR 1.58, 95% CI 0.92 to 2.71) through 6 weeks was somewhat greater for ESI versus placebo (lidocaine) injection, but the differences between groups were not statistically significant in one trial (N=400).[65] However, there were more adverse events (but not procedure-related adverse events, p=0.12) on average per person following ESI vs. placebo (p=0.02) (**Appendix D, Table D-2**). The most common procedure-related adverse events were excessive pain, headache, fever and/or infection, numbness and/or tingling, dizziness and/or light-headedness and skin irritation. There was one case (0.5%) in each group of leg swelling and dural puncture. A second trial that evaluated interlaminar injections reported a total of 14 cases of subarachnoid entry and one case each of nerve root irritation and pain and swelling at the site of injection (out of 644 procedures in 120 patients); authors did not indicate the number of events by treatment group.[82]

Differential effectiveness and safety. One RCT (N=400) (in 3 publications)[64,65,77] that compared ESI (various steroids) with placebo injection (lidocaine alone) provided pain, function and quality of life outcomes stratified by race (White vs. Nonwhite), injection approach (interlaminar vs. transforaminal) and baseline opioid use (yes or no) (**Appendix B, Table B-4**). Authors report that there were no significant interactions between race and treatment in analyses of VAS leg pain scores (p=0.99 for interaction) or RDQ scores (p=0.73 for interaction) at 6 weeks (data not provided).[65] Authors do not provide p-values for interaction for injection approach or opioid use. However, visual inspection of the confidence intervals indicates that injection approach may modify the effect of treatment, such that patients who receive ESI via an interlaminar approach reported better function scores as evaluated by the RDQ (0-24) (small improvement) compared with those who received ESI via a transforaminal approach (adjusted MD -1.4, 95% CI -2.8 vs. -0.1 vs. adjusted MD 0.3, 95% CI -1.9 to 1.8, respectively) at 6 weeks. Authors also report that the adverse event rate (i.e., number of events out of total number of patients) was lower among patients who received interlaminar (ESI, 32/143 [0.22] vs. placebo, 14/139 [0.10], p=0.02) versus transforaminal (ESI, 26/57 [0.46] vs. placebo, 20/61 [0.33], p=0.27) at 6 weeks, but no formal test for interaction by treatment group was performed. For injection approach at later timepoints (3 and 6 months) and for opioid use across outcomes and timeframes, there was considerable overlap in the confidence intervals and likely no modification of treatment effect (**Appendix D, Table D-2**).

3.5.3.2 Epidural Steroidal Injection Versus Usual Care or Inpatient Physical Therapy

3.5.3.2.1 Description of Included Studies and Detailed Synthesis

One RCT (N=29)[76] conducted in Turkey compared epidural steroid injections to usual care (n=19) and to an inpatient physical therapy program (n=20) in patients with lumbar spinal stenosis (**Appendix D, Table D-2**). The average mean age of patients was 59 years, 32 percent were female, and the mean symptom duration was 5.5 years. No other demographic information was reported. ESI consisted of two injections of 1.5 mL triamcinolone plus 3 mL of 0.5 percent bupivacaine hydrochloride via an intralaminar approach under fluoroscopic guidance. Inpatient physical therapy consisted of ultrasound for 10 minutes, hot pack for 20 minutes, and transcutaneous electrical nerve stimulation for 20 minutes to the lumbar region 5 days a week for 2 weeks. The usual care intervention was not well described but authors mention the use of exercise and NSAIDs. All patients were trained in a home exercise program and were provided oral diclofenac twice daily for 2 weeks.

This RCT was assessed as high risk of bias due to unclear randomization and allocation concealment methods, imbalances in age and sex across groups at baseline (though sample size is small), lack of blinding and unclear attrition (**Appendix E, Table E-1**). Authors reported that they did not receive funding support.

Pain and function scores and adverse events. ESI associated with similar improvement in VAS pain scores and RDQ function scores at 3 and 6 months versus usual care (i.e., exercise, NSAIDs) and versus inpatient physical therapy (**Table 3**) in patients with lumbar spinal stenosis.[76] Authors state that two patients withdrew from the trial due to adverse events (gastric complaints and angina pectoris) but do not indicate to which group they were randomized.

Table 3. ESI versus usual care and versus inpatient physical therapy: pain and function scores

Comparison	Outcome and Timing	MD (95% CI)[a]
ESI vs. usual care (N=19)	VAS pain (0-10), 3 months	-1.7 (-4.5 to 1.1)
	VAS pain (0-10), 6 months	-0.8 (-3.9 to 2.3)
	RDQ (0-24), 3 months	1.2 (-5.6 to 8.0)
	RDQ (0-24), 6 months	4.4 (-2.8 to 11.6)
ESI vs. inpatient physical therapy (N=20)	VAS pain (0-10), 3 months	-0.1 (-1.9 to 1.7)
	VAS pain (0-10), 6 months	0.6 (-1.6 to 2.8)
	RDQ (0-24), 3 months	0.2 (-5.8 to 6.2)
	RDQ (0-24), 6 months	2.2 (-4.4 to 8.8)

Abbreviations: CI = confidence interval; ESI = epidural steroid injection; MD = mean difference; PL = profile likelihood; RDQ = Roland Morris Disability Questionnaire; VAS = visual analogue scale

[a] Means used to calculate MDs and 95% CIs were estimated from graphs in article.

Differential effectiveness and safety. The trial did not report on differential effectiveness and safety for ESI versus inpatient physical therapy.

3.5.4 Facet Joint Injections

3.5.4.1 Facet Joint Injection Versus Placebo

3.5.4.1.1 Description of Included Studies and Detailed Synthesis

Two RCTs (total N=210; N range 101 to 109)[56,80] compared facet joint injections versus placebo injections for the treatment of CLBP of confirmed facet origin (**Appendix D, Table D-2**). The average patient age was 44 years (range, 43 to 44 years) and 51 percent (range, 45% to 56%) were female. Median duration of symptoms was 21 months in one trial[56] and the other trial[80] required a pain duration of at least 3 months for inclusion (mean not reported). Both trials included patients with or without radicular symptoms. One trial excluded patients with a history of previous facet joint injections or lumbar surgery;[56] in the other trial, 25 percent of patients had had previous disc surgery but authors state that their pain was like that of patients with no previous surgeries. Patients were required to show at least 50 percent pain improvement after diagnostic facet joint injections of lidocaine for inclusion.

Both trials used the corticosteroid methylprednisolone (20 to 80 mg, 1-2 ml); one trial also injected bupivacaine (6 ml)[80] while the other only used saline (1 ml).[56] The total volume of injectate ranged from 2 ml to 8 ml and all procedures were performed under fluoroscopic guidance. The approach was intraarticular in one trial[80] and either intraarticular or pericapsular (two treatment groups) in the other trial.[80] The procedure for the placebo injections was identical to that of the active injections except only saline was used. All patients in one trial were provided with acetaminophen and ask to limit all other concurrent interventions; 18 percent of patients ended up having physiotherapy, antidepressants, or peridural injections during the trial period.[56] Concomitant treatment were not described in the other trial.[80]

One trial[56] was conducted in Canada and receive government funding, and the other[80] in Finland with no funding source.

One trial[56] was assessed as low risk of bias and the other trial[80] was assessed as high risk of bias due to unclear randomization and allocation concealment methods, unclear blinding status of patients and providers and lack of information on baseline demographics (**Appendix E, Table E-1**).

Pain Scores. At 3 months, both RCTs[56,80] found facet steroid injections associated with similar improvement in VAS pain scores (0-10 scale) compared with placebo (saline) injection. In one trial (N=106)[80] mean pain scores were 4.60 for intraarticular injections, 4.25 for pericapsular injections and 4.45 for placebo injections (no other data provided). In the other trial,[56] authors did not provide 3-month data but stated that it was very similar to the 1-month results (N=96, adjusted MD -0.2, 95% CI -1.1 to 0.8). Intraarticular steroid injections were associated with a small improvement in pain scores at 6 months compared with placebo injection in the latter trial (N=95, adjusted MD -1.0, 95% CI -2.0 to -0.1); however, in a sensitivity analysis that accounted for the use of concurrent interventions (most commonly physical therapy, antidepressant medication and peridural injection) by carrying forward the last data recorded prior to receiving concurrent interventions, which were more common in the intraarticular injection group, there was no longer a treatment effect (MD -0.7, 95% CI -1.6 to 0.2). This trial also reported pain at 3 and 6 months using the McGill Pain Questionnaire; results were generally consistent with those of VAS pain scores (**Appendix D, Table D-2**).

Function scores. Intraarticular steroid injections were associated with a similar improvement in Sickness Impact Profile (SIP; 0-100) scores compared with placebo injections in one trial[56]: SIP overall score (N=95, adjusted MD -3.0, 95% CI -6.2 to 0.2) and SIP physical dimension score (N=95, adjusted MD -3.5, 95% CI -6.2 to -0.9) at 6 months. Although the SIP physical dimension results statistically favored intraarticular injections at 6 months, the estimate was below the threshold for a small effect. The sensitivity analysis accounting for receipt of concurrent interventions provided similar results. Authors did not report the 3-month data but stated that it was very similar to the 1-month results: SIP overall score (N=96, adjusted MD -0.5, 95% CI -2.8 to 1.7) and SIP physical dimension score (N=96, adjusted MD -1.1, 95% CI -2.9 to 0.8).

Adverse events. One trial (N=95) reported that no adverse events occurred in either group other than transient local pain at the injection site.[56] The other trial (N=106)[80] reported no differences in adverse events between groups, but did not report data by treatment group (6.4% for the entire sample).[80]

Differential effectiveness and safety. The trials included did not report on differential effectiveness and safety for facet joint injections versus placebo.

3.5.4.2 Facet Joint Injection Versus Systemic Steroid Injections

3.5.4.2.1 Description of Included Studies and Detailed Synthesis

One RCT (N=60)[92] conducted in Brazil compared facet joint injections with systemic steroids for the treatment of patients with facet joint syndrome of a mean 4.3 years duration (**Appendix D, Table D-2**). Mean patient age was 64 years, 82 percent were female, and mean BMI was 29. Patients with prior lumbar surgery, neurological deficits and major comorbidities were excluded.

Facet joint injections consisted of 1 ml triamcinolone and 1 ml lidocaine injected bilaterally into the facet joints under fluoroscopic guidance. Systemic steroids, 1 ml triamcinolone acetonide and 1 mL of lidocaine, were injected into 6 surface points of the lumbar paravertebral musculature bilaterally. Patients were instructed to rest for 48 hours and to take acetaminophen (maximum 750 mg 4 times per day) or diclofenac (maximum 50 mg 3 times per day).

The trial was rated as low risk of bias (**Appendix E, Table E-1**) and received university funding.

Pain, function and quality of life scores. Facet joint injections were associated with similar improvement in VAS back pain scores (0-10) at 3 months (N=60; MD -1.40, 95% CI -2.91 to 0.11) and 6 months (N=60; MD -0.50, 95% CI -1.98 to 0.98) and with a moderate improvement in RDQ (0-24) scores at 3 months compared with IM steroid injections (N=60; MD -4.10, 95% CI -7.47 to -0.73) in one trial,[92] however, the effect did not persist to 6 months for RDQ scores (N=60; MD -2.50, 95% CI -6.26 to 1.26). For quality of life, authors reported the eight individual domains of the SF-36 and found facet joint injections associated with greater improvement on the role physical domain at 3 and 6 months compared with IM steroid injection but there was no difference between groups on all other domains (**Appendix D, Table D-2**).

Harms and adverse events. One patient (3.2%) who received facet joint injections experienced aggravation of back pain after a fall resulting in spinal arthrodesis and one patient (3.4%) in the

IM steroid injection group experienced serious gastrointestinal bleeding; both events occurred between 3 and 6 months post-procedure. In terms of local and systemic events, there was no difference between groups for post-procedure pain, cutaneous hypochromia, blood glucose levels, vaginal bleeding, dizziness, or nausea (data not reported by group) (**Appendix D, Table D-2**).[92]

Differential effectiveness and safety. The trial did not report on differential effectiveness and safety for facet joint injections versus systemic steroid injections.

3.5.4.3 Facet Joint Injection Versus Physiotherapy

3.5.4.3.1 Description of Included Studies and Detailed Synthesis

One small RCT (N=18)[74] conducted in Nigeria compared intraarticular facet joint injections (0.5 mL methylprednisolone and 0.5 mL of 0.25% bupivacaine under X-ray guidance) with physiotherapy (i.e., McKenzie regimen, not otherwise described) in patients with facet joint arthropathy (**Appendix D, Table D-2**). The average age was 45, 44 percent of patients were female, and duration of symptoms were >3 months (mean not reported). Other patient characteristics and comorbidities were not reported. Patients with neurological symptoms were excluded.

The trial was assessed as high risk of bias due to unclear randomization and allocation concealment methods, inability to blind patients and assessors, and lack of information regarding attrition (**Appendix E, Table E-1**). The trial did not receive funding.

Pain and function scores, harms, and adverse events. Facet joint injection was associated with improvement VAS pain scores (0-10 scale: mean 4.3 vs. 5.5 at 3 months and 4.0 vs. 5.0 at 6 months, p=0.032) and ODI scores (0-100 scale: mean 40 vs. 53 at 3 months and 38 vs. 50, p=0.013) compared with physiotherapy, respectively in one very small trial.[74] Means for the overall population were estimate from plots and the variance was not provided. Intraarticular injection was associated with a similar likelihood of achieving clinical success (undefined) at 6 months (N=18, 90% vs. 75%, RR 1.33, 95% CI 0.89 to 1.99). Authors stated that there were no complications related to facet join injection.

Differential effectiveness and safety. Authors' primary results were stratified by male and female, but they did not do a formal subgroup analysis or provide a p-value for interaction. Authors simply stated that the female patients in both the intraarticular steroid injection and physiotherapy groups showed more improvement in pain and function than the males. Visual inspection of the effect estimates for males and females between treatments groups showed considerable overlap in the confidence intervals and likely no modification of treatment effect (**Appendix D, Table D-2**). This trial was likely too small to find an effect.

3.5.5 Medial Branch Block

3.5.5.1 Medial Branch Block Versus Placebo Block

3.5.5.1.1 Description of Included Study and Detailed Synthesis

One RCT (N=120; reported in 2 publications)[84,85] compared medial branch blocks (MBB) to placebo nerve block for treatment of lumbar facet joint pain of a mean duration of 9 years (**Appendix D, Table D-2**). Mean patient age was 47 years and 60 percent were female. Patients with a history of lumbar surgery in the past 3 months, radicular pain and major comorbidities (e.g., major psychiatric disorders, opioid use disorder, uncontrolled medical illnesses) were excluded. Patients were required to have at least 80 percent pain improvement with diagnostic block of bupivacaine on separate occasions for inclusion.

Patients in the MBB groups received injections of nonparticulate betamethasone (0.15 mg per mL, total unclear) and 0.5 mL lidocaine under fluoroscopic guidance at L1 to L4 levels and L5 dorsal ramus. Patients in the placebo group received an injection of 0.5 mL lidocaine without steroids to the same locations. All patients were further randomized to receive sarapin (50% in each group), but preliminary analyses found no difference between sarapin and nonsarapin groups and authors only reported combined results. Repeat injections of nerve blocks were provided during the trial period if pain increased significantly (not defined). Patients continued their previous pharmacological therapy (opioids and/or nonopioid analgesics) and therapeutic exercise programs as needed.

This trial was conducted in the United States and reported no funding source. Risk of bias was rated as moderate due to unclear allocation concealment and imbalances in some baseline characteristics between groups (**Appendix E, Table E-1**).

Pain response and pain scores. MBB was associated with a similar likelihood of achieving pain response (≥50% improvement on 0-10 NRS) at 3 months (N=120, 82% vs. 83%, RR 0.98, 95% CI 0.83 to 1.16) and 6 months (N=120, 93% vs. 83%, RR 1.12, 95% CI 0.98 to 1.28) and similar improvement in NRS pain scores (0-10) at 3 months (N=120; MD -0.30, 95% CI -0.74 to 0.14) and 6 months (MD -0.30, 95% CI -0.74 to 0.14) versus placebo nerve block (lidocaine only).[84]

Function response and function scores. MBB was associated with a similar likelihood of achieving function response (≥40% improvement on 0-50 ODI) at 3 months (N=120, 72% vs. 82%, RR 0.88, 95% CI 0.72 to 1.07) and 6 months (N=120, 78% vs. 83%, RR 0.94, 95% CI 0.79 to 1.12).[85] and similar improvement in ODI scores (0-50) at 3 months (N=120, MD 0.80, 95% CI -1.07 to 2.70) and 6 months (MD -0.59, 95% CI -2.24 to 1.24) compared with placebo nerve block (lidocaine only).[84]

Harms and adverse events. Authors reported that there were no major adverse events during the study period.[84,85]

Differential effectiveness and safety. The trial did not report on differential effectiveness and safety for MBB versus placebo.

3.5.5.2 Medial Branch Block Versus Facet Joint Injection

3.5.5.1.2 Description of Included Study and Detailed Synthesis

One RCT (N=86) conducted in the United States compared MBB with facet joint injection for the treatment of severe, CLBP with a median symptom duration of 8.5 years.[86] Median patient age was 43 years, 47 percent were female, 41 percent had referred pain below the knee, and 12 percent of patients had had previous spinal surgery. Other patient characteristics or comorbidities were not reported.[86] Patients in both treatment groups received injections of 20 mg methylprednisolone acetate and 1 percent lignocaine under fluoroscopic guidance between L1 and L4 and/or at L5. See **Appendix D, Table D-2** for details related to the interventions. Repeat procedures and concomitant therapies were not described. This RCT was assessed as having a low risk of bias (**Appendix E, Table E-1**).

Pain response, harms, and adverse events. MBB was associated with a similar likelihood of a good or excellent pain response, defined as complete relief of the most dominant pain or of all pain, at 3 months compared with facet joint injection (N=83, 14.3% vs. 22.0%, RR 0.64, 95% CI 0.25 to 1.63).[86] The trial did not report any other effectiveness outcomes.

No patient experienced serious complications. The risk of headache (6.8% vs. 9.5%, RR 0.73, 95% CI 0.17 to 3.01), paresthesia of one leg below the knee without motor signs (2.3% vs. 2.4%, RR 0.95, 95% CI 0.06 to 14.77), and nausea (2.3% vs. 2.4%, RR 0.95, 95% CI 0.06 to 14.77) was similar for MBB and facet joint injection (N=86); all events were transient and had vanished by the morning following infiltration. Through 1-month followup, similar proportions of patients in both groups complained of worsening pain (29.5% vs. 21.4%, RR 1.38, 95% CI 0.66 to 2.88).

Differential effectiveness and safety. The trial did not report on differential effectiveness and safety for MBB versus facet joint injection.

3.6 Key Question 6. Does symptomatic improvement to therapeutic challenge with lumbar epidural steroid injections, intra-articular (facet) injection, medial branch blocks or radio frequency ablation predict positive outcomes after lumbar fusion surgery for patients with chronic low back pain (≥3 months) resulting from degenerative disease?

3.6.1 Key Points

- There was insufficient evidence from one study to determine whether symptomatic response to therapeutic challenge with facet joint injection predicted outcomes after lumbar fusion surgery.

3.6.2 Description of Included Studies and Detailed Synthesis

One prospective NRSI (N=126)[104] evaluated the association between symptomatic response to bilateral facet joint injections (0.5 ml of lidocaine and 1 ml of bupivacaine) and outcomes after lumbar fusion surgery in patients with chronic (mean 7.6 years) mechanical, activity-related LBP (**Appendix D, Table D-2**). Most facet joint injections were performed at two levels (67%) (range, 1-3 levels). Mean patient age was 48 years, and 60 percent were female. Thirty-seven percent of patients were involved in Workers' Compensation claims and/or litigation related to their LBP. This study was rated at moderate risk of bias (**Appendix E, Table E-2**).

Authors did not describe a specific threshold for symptomatic response following facet joint injection but based response on the patient's subjective pain relief defined as full, partial or none. For the analysis, patients were divided into two groups, those who experienced full pain relief and those with partial or no relief; however, authors do not report how many patients were in each group. Of the patients who were followed at 4.5 years, 82 (65%) underwent single (24%) or multilevel (76%) lumbar fusion; 37 percent had had previous lumbar surgery (i.e., discectomy, laminectomy or fusion at other levels). Based on a system modified from Prolo et al., 46 percent of patients had a good outcome (decrease in pain, increase in function), 24 percent were unimproved (no significant change in pain or function), and 30 percent had a poor outcome (increased pain and/or decreased function) following fusion surgery. Authors report that there was no relationship between facet joint injection response and surgical outcome even after stratification by the following factors: workers' compensation and litigation status (p=0.217), history of prior (p=0.886) or no prior (p=0.184) lumbar surgery, and number of levels of facet joint injection (p=0.185). No data other than p-values and chi-squares were provided.

4. Discussion

4.1 Findings in Relation to Decisional Dilemmas

Optimal management of symptomatic degenerative lumbar disease (DLD) and degenerative lumbar spondylolisthesis (DLS), remains controversial.[26,115,116] Decisional dilemmas include the unclear benefits and harms of fusion in addition to decompression, optimal fusion methods (e.g., use of interbody cages), and use of different graft materials to promote fusion. Other decisional dilemmas include routine use of intraoperative neuromonitoring (IONM) during fusion surgery to reduce neurological complications and of the benefits and harms of epidural steroid injections (ESI), radiofrequency ablation (RFA), and medial branch blocks (MBB) in people with DLD and DLS who are being considered for lumbar fusion and whether these procedures may help identify patients who may benefit most from lumbar fusion. Key findings and the strength of evidence are summarized in **Tables 4-7** and the summary below. The summary tables reflect intervention comparisons for which there was at least low strength of evidence. An expanded version is found in **Appendix F** after the full strength of evidence tables. Where an effect size is provided in the tables, it favors the intervention unless otherwise noted.

Thirty-five randomized controlled trials (RCTs) (in 43 publications) and 13 comparative nonrandomized studies of interventions (NRSIs) (in 15 publications) provided evidence for this review. For the surgical Key Questions, there were five RCTs (in 9 publications)[55,62,63,67,69,70,72,73,75] and seven NRSIs (in 8 publications)[101-103,106,108,109,111,112] for Key Question 1; two RCTs[57,62] and four NRSIs (in 5 publications)[98,99,105,107,110]) for Key Question 2; one RCT[71] for Key Question 3; and one NRSI[100] for Key Question 4. For the nonsurgical Key Questions, there were 28 RCTs (in 32 publications)[56,58-61,64-66,68,74,76-97] for Key Question 5, most of which was for RFA (17 RCTs), and one NRSI[104] for Key Question 6. Includable evidence on the use of lumbar fusion specifically for symptomatic DLS management was sparse. Evidence for use of IONM during instrumented lumbar fusion for DLD was sparse even though we expanded the population to include any person with symptomatic degenerative lumbar spine disease undergoing fusion at five or fewer levels.

Because we anticipated that the evidence specifically for DLS due to degenerative disease would be sparse for interventional procedures, we broadened the population for the evaluation of them to include all persons with chronic low back pain (CLBP). The largest body of includable evidence was for RFA for treatment of chronic low back pain. For ESI versus placebo, the evidence was most robust for patients with CLBP due to spinal stenosis. There was insufficient evidence to draw conclusions on the benefits of procedures such as RFA, ESI, or medial branch blocks for predicting patient outcomes following lumbar fusion. Across Key Questions, reporting of serious harms and adverse events was generally poor and definitions and classifications of adverse events were variably reported.

Decompression plus fusion versus decompression alone for grade I DLS (Table 4). High-quality evidence on the benefits and risks of adding fusion to decompression specific to patients with grade I stable DLS defined as <3 mm slip based on flexion/extension radiographs was limited to two trials.[67,75] The bulk of excluded studies that otherwise would have informed this Key Question were excluded because they did not clearly address stable or unstable degenerative spondylolisthesis and/or because they included high proportions of patients with isthmic spondylolisthesis. The SLIP trial (n=60)[67] found that fusion plus decompression was associated with a small functional improvement at 4 years and in quality of life at most time points versus

decompression alone, but no differences between treatments at other time points; pain was not reported as a separate measure (strength of evidence [SOE]: low). This finding of no difference in functional outcomes was consistent with findings from three other RCTs.[69,72,75] Across these RCTs, improvements in pain, function, and quality of life were generally similar between groups at all timepoints where sufficient evidence was available.

The likelihood of reoperation at 1 year was slightly lower when fusion was added to decompression versus decompression alone (SOE: low) based on two RCTs[67,69] and evidence pooled across two NRSI[109,112] in mixed populations (i.e., those with stable or unstable DLS). The absolute difference in reoperation was small (4.5% vs. 5.8%) for the NRSIs and indications for and levels involved in reoperation (e.g., index level or adjacent level) were not specified. Estimates were imprecise. Authors did not report whether procedures were performed in a minimally invasive or open fashion in the RCTs or NRSIs. These factors, combined with the unknown impact of including a mixture of patients with stable and unstable DLS, and the general limitations of administrative data studies, suggested that results should be interpreted cautiously. Risk of reoperation was similar between treatment groups at earlier time frames for RCTs and NRSIs; RCTs may have been underpowered to identify differences. Serious adverse events were often not specified and were sparsely reported, and no differences in risk were reported in RCTs[67,75] or NRSIs[106,111] by treatment group. There was no difference in treatments groups for the likelihood of dural tears in pooled estimates for RCTs or NRSIs, however estimates were imprecise. There were too few RCTs to explore whether the use of cages modified the risk of dural tear or not.

Table 4. Summary of evidence of decompression plus fusion versus decompression alone in populations with degenerative spondylolisthesis of stable, mixed, or unknown stability (Key Question 1; pain, function, QOL, harms)

Outcome	Time Point	Effect Size/SOE[a]
Mean improvement in back and leg pain	3 months, 1, 2, 5 years	Similar +
Successful function outcome	5 years	Similar +
Mean improvement in function[b]	3, 6 months, 1, 2, 3, 4, 5 years	Similar +
Mean improvement in QOL	3 months, 1, 5 years	Similar +
	6 months, 3, 4 years	Small +
	2 years	EQ-5D: Similar + SF-36 PCS: Small +
Reoperation	1, 3 months	Similar +
	1 year	Small decrease +
Serious AE, dural tear, deep infection, PE, DVT, heart attack, or stroke	Any time	Similar +
Neurological deterioration	5 years	Small increase +

Abbreviations: AE = adverse event; DVT = deep vein thrombosis; EQ-5D = EuroQol 5 dimensions; KQ = Key Question; PE = pulmonary embolism; QOL = quality of life; SF-36 PCS= Short-form 36 questionnaire Physical Component Score; SOE = strength of evidence.

[a] Effect size: Similar (no effect), small, moderate, or large difference favoring the intervention; SOE: + = low, ++ = moderate, +++ = high.

[b] Based on Oswestry Disability Index at 1 and 5 years; evidence for Japanese Orthopaedic Association scores at these time points was insufficient.

Interbody fusion versus PLF for DLS. Evidence was very sparse to address this question. Two RCTs provided insufficient evidence to draw reliable conclusions for interbody fusion versus posterolateral fusion (PLF) in patients with DLS for most outcomes; one RCT was considered at high risk of bias[62] and the other did not provide sufficient data to calculate effect sizes and confidence intervals. One small RCT[57] found no difference between transforaminal interbody fusion (TLIF) plus PLF and PLF in the proportion of patients achieving ODI scores between 0 and 20 percent (i.e., mild disability) at 2 years but found TLIF plus PLF was associated with substantially higher likelihood of fusion versus PLF by 2 years (SOE low for both outcomes). This question was not limited to people with stable DLS. Most of the evidence identified for harms is from NRSIs and was considered insufficient due to downgrades for risk of bias, imprecision, and unknown consistency.

Graft materials for fusion in patients with DLS. Evidence on the use of bone graft extenders and biologic substitutes for fusion in patients with DLS was confined to a single, small (N=46) RCT,[71] which found no difference in fusion rates with a specific brand of demineralized bone matrix (DBM) compared with iliac crest bone graft (ICBG). Evidence was insufficient for other outcomes of interest. Most of the studies identified were not in patients with DLS, were not randomized or did not compare graft materials with autograft and were therefore excluded.

Intraoperative neuromonitoring during lumbar fusion in DLD. One large administrative data study (N=133,572)[100] reported no difference in the development of a broad range of postoperative neurologic complications prior to hospital discharge when IONM was used and when it wasn't used in patients undergoing elective PLF for DLD (SOE: low). This finding should be interpreted with caution as the primary outcome of interest included adverse events not isolated to the intraoperative setting or immediate postoperative setting or to surgical site complications (e.g., hematoma). Additionally, because this study relied upon ICD-9 codes, it was not possible to ascertain the specific neuromonitoring modalities used (e.g., somatosensory evoked potentials, motor-evoked potentials, electromyography). IONM is primarily used to assure appropriate patient positioning and nerve localization and to detect intraoperative neurologic changes that could result in persistent postoperative neurologic deficits (e.g., new radiculopathy, motor deficits such as foot drop, or those related to spinal cord damage), allowing the surgeon to reverse or correct an action to prevent permanent damage. Thus, to answer Key Question 4, information on specific modalities and how they are used intraoperatively for specific fusion procedures in patients with DLD is needed. Results were also confined to patients receiving PLF.

Radiofrequency ablation (Table 5). The greatest volume of included evidence compares continuous RFA with sham (9 RCTs) in patients with facet joint pain.[66,79,88,89,91,93-96] Three RCTs excluded patients with spondylolisthesis;[66,88,89] others did not provide information on DLS. Trials did not specify if patients had degenerative lumbar disease. For RFA versus sham, all trials required a positive diagnostic block to confirm facet joint pain for study entry, however the way blocks were performed, the type of block and the criteria used to determine diagnostic success varied. Techniques for RFA application also varied across trials. Our confidence in the evidence across most frames and outcomes was low primarily due to methodological limitations of the

studies and imprecision in effect estimates. Overall, RCTs of continuous RFA indicated similar improvement in back pain and function compared with sham, however results were variable across trials. At 3 months, pain and function improvement were similar for continuous RFA versus sham across pooled trials (SOE: moderate). There was substantial heterogeneity in findings across trials at 6 months, possibly due to difference in RFA intervention delivery and/or criteria for the diagnostic blocks across studies. Reporting of baseline neurologic symptoms also varied across trials. Two trials by the same author group[88,89] reported improvements in pain and function with continuous RFA versus sham. Patients in these two trials received a post-RFA injection of steroids and anesthetic through the electrode needle. Patients in these trials were offered RFA if they did not experience relief following the sham procedure and were no longer blinded. These and other factors may partially explain differences in results across trials.

Evidence for pulsed RFA versus sham and for continuous RFA versus usual care was limited and our confidence in the findings was low. The primary evidence for pulsed RFA was from a trial in which patients received steroids and anesthetic via the electrode needle to the dorsal root ganglion, which reported improvements in pain and function but estimates were imprecise.[89] No difference was seen in pain or function between treatments in another trial which targeted the medial branch, however.[93] Evidence comparing continuous RFA with usual care was limited to a single open-label RCT.[58]

Table 5. Summary of evidence of RFA versus sham or usual care in patients with chronic low back pain of presumed facet joint origin (Key Question 5; pain, function, QOL, harms)

Outcome	Time Point	Continuous RFA vs. sham[a]	Pulsed RFA vs. sham[a]	Continuous RFA vs. UC[a]
Successful pain outcome (Effect Size/SOE)[b]	3 months	Similar +	Large +	Large +
	6 months	Large +	Large +	Large +
Mean improvement in back pain (Effect Size/SOE)[b]	3 months	Similar ++	Large +	Moderate +
	6 months	Small[c] +	DRG: Large + MB: Similar +	Moderate +
Mean improvement in leg pain (Effect Size/SOE)[b]	3 months	Similar +	Similar +	No evidence
	6 months	Large +	Large +	No evidence
Successful function outcome (Effect Size/SOE)[b]	3 months	Large +	No evidence	Small +
	6 months	No evidence	No evidence	Similar +
Mean improvement in function (Effect Size/SOE)[b]	3 months	Similar +	Large +	Similar +
	6 months	Similar +	DRG: Large + MB: Similar +	Similar +
Quality of Life (Effect Size/SOE)[b]	6 months	Similar +	No evidence	No evidence
Serious AEs (Effect Size/SOE)[b]	Any time	Similar +	No evidence	No evidence
Any AEs (Effect Size/SOE)[b]	Any time	Similar +	Similar +	Similar +

Abbreviations: AE = adverse event; DRG = dorsal root ganglion; MB = medial branch nerve; QOL = quality of life; RFA = radiofrequency ablation; SOE = strength of evidence; UC = usual care.

[a] Evidence base is as follows: Continuous RFA vs. sham (9 RCTs; at most 6 RCTs contributed to pooled analyses); pulsed RFA vs. sham (2 RCTs; most analyses based on a single trial); continuous RFA vs. UC (1 RCT).
[b] Effect size: Similar (no effect), small, moderate, or large difference favoring RFA; SOE: + = low, ++ = moderate, +++ = high.
[c] Small effect is based off analysis excluding Moussa 2016 and Moussa 2020. In these trials, patients received post treatment injection of anesthetic and steroids; the effects of these additional treatments on outcomes were unclear.

Epidural steroid injection (Table 6). Trials of ESI versus placebo were in patients with spinal stenosis and did not specifically include DLS. One excluded patients with spinal instability requiring surgery,[64] one excluded DLS,[90] and two others did not specify exclusion of DLS.[82,83] There was moderate evidence of similar improvements in leg pain, function, and quality of life with ESI and placebo injections of lidocaine only (SOE: moderate) based on four RCTs (7 publications)[64,65,77,82,83,90] in patients with chronic back pain and neurologic symptoms due to spinal stenosis. Adverse events were also similar between groups. Data were insufficient to evaluate impact of approach (e.g., caudal, interlaminar, transforaminal) or type of steroid used (particulate, nonparticulate), or to evaluate how effects varied by type of placebo injection. There was insufficient evidence to compare the effectiveness of ESI with other treatments in our population of interest. One RCT (in 3 publications)[64,65,77] reported no significant modification of effects of ESI by race, injection approach (interlaminar vs. transforaminal), and baseline opioid use (yes or no) on pain or function scores; authors do not report on harms.

Table 6. Summary of evidence of ESI versus sham in patients with spinal stenosis (Key Question 5; pain, function, QOL, harms)

Outcome	Time Point	(Effect Size/SOE)[a]
Successful pain and function outcome	6 weeks to 3 months	Similar ++
	6 months	Similar ++
Mean improvement in leg pain and function	3 months	Similar ++
	6 months	Similar ++
Quality of life	6 weeks	Similar ++
Serious AE, any AEs, procedure-related AEs	Any time	Similar +

Abbreviations: AE = adverse event; ESI = epidural steroid injection; QOL = quality of life; SOE = strength of evidence.
[a] Effect size: Similar (no effect), small, moderate, or large difference favoring ESI; SOE: + = low, ++ = moderate, +++ = high.

Facet joint injection and medial branch blocks. Improvement in both pain and function were similar for facet joint (intraarticular) injection versus placebo (2 RCTs)[56,80] for MBB versus placebo (1 RCT [2 publications)].[84,85]

Use of procedures to predict fusion outcomes. There was insufficient evidence from one prospective NRSI (N=126)[104] to determine whether symptomatic response to therapeutic challenge with facet joint injection predicted outcomes after lumbar fusion surgery.

Comparative effectiveness of select procedures in patients with chronic low back pain (Table 7). Evidence for comparative effectiveness of different procedures was limited as only single trials were available for various comparisons. The strength of evidence was low for most comparators given the methodologic limitation of the trials, unknown consistency, and lack of precision. Limited evidence indicates pain and functional improvement are similar with RFA and facet joint injection.[60,61,78,87] RFA was associated with improved pain and function compared

with medial branch block.[59,68,97] Similar improvements in pain and function were seen for comparisons of facet joint (intraarticular) injection with systemic steroids[92] and with MBB.[86] Evidence from high risk of bias trials was insufficient for benefits and harms for ESI versus inpatient physical therapy[76] and facet joint (intraarticular) injection versus physiotherapy.[74]

Table 7. Summary of evidence for RFA, FJI, and MBB versus other active treatments in patients with chronic low back pain of presumed facet joint origin (Key Question 5; pain, function, QOL, harms)

Outcome	Time Point	Continuous RFA vs. FJI	Pulsed RFA vs. FJI	Continuous RFA vs. MBB	Pulsed RFA vs. MBB	FJI vs. IM steroid injection	MBB vs. FJI
Successful pain outcome (Effect size/SOE)[a]	3 months	No evidence	Similar +	No evidence	No evidence	No evidence	Similar +
	6 months	No evidence	No evidence	Small +	No evidence	No evidence	No evidence
Mean improvement in back pain (Effect size/SOE)[a]	3 months	No evidence	Similar +	No evidence	Large +	Similar +	No evidence
	6 months	Similar +	Similar +	Moderate +	Large +	Similar +	No evidence
Mean improvement in function (Effect size/SOE)[a]	3 months	No evidence	No evidence	No evidence	Large +	Similar +	No evidence
	6 months	Similar +	No evidence	No evidence	Large +	Similar +	No evidence
Successful QOL outcome (Effect size/SOE)[a]	6 months	No evidence	No evidence	Small +	No evidence	No evidence	No evidence
Any AEs (Effect size/SOE)[a]	Any time	No evidence	Insufficient evidence	Similar +	No evidence	No evidence	Insufficient evidence

Abbreviations: AE = adverse event; FJI = facet joint (intraarticular) injection; IM = intramuscular; MBB = medial branch block; QOL = quality of life; RFA = radiofrequency ablation; SOE = strength of evidence.
[a] Effect size: Similar (no effect), small, moderate, or large difference favoring RFA; SOE: + = low, ++ = moderate, +++ = high.

4.2 Findings in Relation to What Is Already Known

Contemporary systematic reviews all note the paucity of high-quality evidence on lumbar fusion specifically in patients with DLS. Consistent with two systematic reviews of RCTs,[117,118] our review found that the addition of fusion to decompression was not associated with improved function or pain versus decompression alone in patients with DLS. One review comparing decompression alone with decompression plus instrumented fusion (with pedicle screws) concluded that there was no benefit to adding fusion and no decrease in reoperation rates.[117] The other review in elderly patients with lumbar spinal stenosis and low-grade DLS found no difference in function or pain with the addition of fusion to decompression.[118] Both reviews describe fewer complications overall with decompression alone versus decompression with fusion.[117,118] One systematic review reported that any interbody fusion was associated with higher fusion rates versus PLF in patients with DLS but that clinical outcomes were similar based on retrospective studies.[119] The one RCT[71] comparing DBM with autograft met our inclusion criteria, found no difference in fusion rates between these. This is consistent with a 2020 review[120]of allograft and synthetic graft materials for lumbar fusion which concluded there was no significant difference and/or superiority for these materials versus autograft. Most studies were nonrandomized, at moderately high risk of bias and were in patients with various DLD conditions including degenerative disc disease. No studies of recombinant BMP-2 (rhBMP) for lumbar fusion met our inclusion criteria. Its use has been controversial. In our experience conducting a SR on Bone Morphogenetic Protein-2 (BMP-2) use in spinal fusion,[121,122] treatment

harms were underreported and benefits overstated through the use of improper comparison groups, reporting harms as "no unanticipated adverse events," and reporting bias (e.g., use of multiple publications of same data). One study of IONM met our inclusion criteria and reported no benefit to its use for elective PLF for DLD.[100] This is consistent with a recent review that found that IONM during LLIF did not improve neurologic outcomes,[123] however the methodological quality of included studies was generally poor. The utility of IONM in spine tumor and spinal deformity surgeries is established, however, its utility in routine surgery for DLD remains less clear.[36,124] Our findings of similar improvement for pain and function following ESI versus sham are consistent with those from a recent network meta-analysis and associated clinical guideline which included a broader spectrum of patients and clinical indications than our report, including those with disc herniation.[125,126] Our review excluded patients with disc herniation or discogenic pain.

Systematic reviews of RCTs comparing RFA with sham (and with other active treatments)[127-132] report that RFA may provide a benefit, typically a small improvement in pain and function primarily at longer followup (i.e., 6 months, 1 year). The positive effects of RFA at earlier times (e.g., 3 months) were less consistent. These reviews use different inclusion and exclusion criteria and methodology that we did and were variable in quality. Some based conclusions on pooled results across various comparators while our report analyzed sham intervention and other active treatments (e.g., steroid injection, facet injection) separately.

4.3 Implications

Although evidence was limited for many of the Key Questions addressed by this review, there are some potential clinical or policy implications. We found that across RCTs in a mixed population of patients with stable or unstable DLS and concomitant stenosis that improvement in pain and function was similar as was the risk of serious adverse events with the addition of fusion to decompression versus decompression alone (SOE: low). Also in this population, we found similar functional improvement between interbody fusion and PLF but potentially higher rates of fusion (SOE: low) however evidence on harms was insufficient. There was a paucity of evidence regarding the benefits and harms of different graft materials for fusion specifically in patients with DLS. One review[120] of graft materials in broader populations concluded that there was no significant difference for use of allograft and synthetic graft materials versus autograft alone, reviews of BMP specifically found lack of benefits and increased risk of harm in general with its use.[121,122] These findings may facilitate shared decision-making and balance benefits and harms for the development of evidence-based recommendations regarding fusion in patients with DLS. Controversies regarding the use of fusion in DLS may continue, however, until additional high-quality evidence is available. Although evidence on clinical benefits of IONM in patients undergoing fusion for DLD is lacking, this is routinely used in some settings and may be of value for detecting intraoperative neurological events.[36,124]

Our findings on procedures which may be considered prior to or in lieu of fusion surgery in patients with CLBP may also inform patient choices and clinical recommendations. In patients with chronic low back pain and spinal stenosis, we found moderate evidence of no benefit for ESI compared with placebo injection, which is consistent with recent systematic reviews and guidelines[125,126] that included a broader range of patients. Improvement in pain and function were also similar for comparisons of facet joint injection and medial bundle branch block versus placebo. Our review found that there was some heterogeneity in results across RCTs regarding improvement in pain and function with continuous RFA versus sham in patients with facet-

related CLBP. Several trials found similar improvements in these outcomes versus sham at 3 months with other trials suggesting some benefit at 6 months. Continuous RFA may be an option in patients who would like to avoid surgery or who may not be good candidates for surgery. Unfortunately, there was insufficient evidence to guide patient selection for fusion with these procedures.

4.4 Strengths and Limitations

Strengths. Our review has several strengths. First, we addressed clinically meaningful and controversial Key Questions regarding lumbar fusion for the management of DLS and we include new RCT evidence[67,75 62,69,72] on fusion for grade I DLS that was not available for the reviews conducted for 2014 guidelines.[24-26] We also addressed questions related to interventional procedures that may be considered prior to or in lieu of fusion in patients with chronic low back pain.. A strength of our review is our expansion of Key Question 1 to include patients with both stable and unstable DLS and not restricting trials of interventional procedures to patients with DLS but including patients with chronic low back pain. This enabled us to capture more potentially relevant evidence to inform clinical recommendations. Our review sheds additional light on ongoing evidence gaps specific to fusion for management of symptomatic DLS, use of IONM and special procedures for the management of chronic low back pain that may stimulate additional research.

Another strength of our review is our categorization of the magnitude of effects for function and pain outcomes using the system described in our previous reviews.[49-53,133] This categorization facilitates transparent and consistent interpretation across trials and interventions by providing an indication of the relative size of effects. We classified effects below the threshold for small as no effect. Based on this system, beneficial effects identified were usually considered small to moderate. We acknowledge that effects that we classify as small (e.g., 5 to 10 points on a 0 to 100 scale) may be below some proposed thresholds for minimum clinically important differences for some measures, however values for minimum clinically important difference vary based on populations and methods used to determine them. Mean differences in outcomes represent "average" effects, and some patients will experience larger effects; in addition, patients differ in how they value small effects. Where evidence was available, we reported the proportion of patients who experienced a clinically important improvement in pain or function to provide better insight into patient treatment response.

Limitations. Many of the limitations to this review are related to the limitations of the available evidence. Studies of fusion for DLS did not generally report on diagnostic methods or criteria used, the severity, grade of slippage and whether it was stable or unstable nor were definitions of stable and unstable generally provided. Radiographic diagnostic criteria and stability assessment are subjective, controversial, and not standardized.[4,6-8] The impact of these radiographic criteria on clinical management and patient outcomes is unclear.[4,5,9]

Many studies did not report on how DLS was determined or graded (e.g., use of flexion-extension radiographs, computed tomography alone) possibly adding to the heterogeneity in enrolled populations. For Key Question 1, the SLIP RCT[67] was specifically in patients with stable DLS defined as <3 mm slip on flexion-extension radiographs and the Nordsten trial[75] included ~80% of patients with <3 mm slip. We expanded our criteria to include trials which either did not specify stability or included patients with stable and unstable DLS. Although the findings across the trials which included patients with stable or unstable DLS were generally

concordant with the SLIP and Nordsten trials the evidence specific to stability is somewhat limited to draw firm conclusions about whether or not fusion should be added to decompression in patients stable DLS as defined for this review. There is, however, not a clear consensus regarding definitions stability and instability.[134] There also substantial heterogeneity with regard to procedures used for fusion and decompression. Evidence comparing interbody fusion and PLF for DLS was mostly insufficient. Studies of fusion and decompression did not specify whether procedures were open or if minimally invasive techniques were used and use of graft material was poorly described. Failure to specify such details may result in misunderstanding of reported benefits and harms. The variety of approaches and procedures used in the NRSIs in particular and lack of detail regarding the specific procedures used, precludes comparison of the benefits and harms across them. For studies of interventional procedures, heterogeneity in populations and in how the procedures were done (e.g., RFA) was noted. For all key questions where meta-analysis was possible, we used random effects models for pooling given the heterogeneity in patient populations and procedures used. We also conducted sensitivity analyses to explore heterogeneity and the impact of study quality when possible. There was insufficient data from includable studies to do subgroup analyses or evaluate modification of benefits or harms by factors such as patient characteristics, DLS grade/degree or stability, concurrent spine pathology or other factors. Stratified analyses were limited by small numbers of studies and for some interventions (e.g. RFA, ESI) there is insufficient evidence to evaluate how technical factors and types of sham or comparator might impact findings. We were also unable to assess for publication bias as there were too few studies for graphical or statistical tests for small sample effects.

Although we restricted to English language studies, we did not identify relevant non-English language studies in searches or reference lists.

4.5 Applicability

The applicability of our findings may be impacted by several factors. First, many studies of fusion excluded from this review were in populations of patients with DLS *or* isthmic spondylolisthesis which likely reflect the mixture of patients seen in clinical practice; our focus was specific to DLS. For Key Question 1, we expanded our criteria to include trials which either did not specify stability or included patients with stable and unstable DLS, given the paucity of trials in stable DLS as defined for this review. To the extent that this expansion better reflects the mix of patients seen clinically, our findings may be applicable to real world practice. Across included studies of fusion, a variety of approaches and techniques were used. Similarly, there was a variability in the interventional procedures that were employed. The variability seen in the included studies may reflect the broader range of fusion procedures and interventions used clinically. Given the paucity of evidence on various graft materials for fusion and for the use of IONM, it is not possible to evaluate the applicability of our findings to clinical practice.

4.6 Future Research

Our review identified important areas for future research on DLS management as well as use of IONM in lumbar fusion surgery for DLD. RCTs of fusion plus decompression versus decompression alone in patients with stable low-grade DLS are needed as are RCTs comparing interbody fusion with PLF in the broader population of patients with stable or unstable DLS. Clarification and standardization of diagnostic criteria and definitions of DLS stability would facilitate comparison of findings across studies. RCTs should report on specific diagnostic

criteria for determining DLS stability, the proportions of patients meeting specific criteria, severity of the slippage or grade of DLS and presence of concurrent pathology. RCTs employing standardized surgical procedures are needed. RCTs comparing bone graft extenders and graft materials other than BMP to local autologous bone graft specially in patients with DLS are needed. Studies that help identify which patients with DLS may benefit most from surgery are needed. All trials need to be sufficiently powered to evaluate and describe procedure-related harms and to effectively evaluate potential differential effectiveness and harms of treatment options based on patient characteristics or other factors. Well-designed NRSIs could also be informative if RCTs are implausible. High quality studies of the accuracy and utility of IONM during fusion procedures for degenerative lumbar disease that evaluate the association between monitoring and clinical outcomes such as persistence of neurological events are needed. Robust prospective studies that assess whether a symptomatic response to therapeutic challenge using interventions (e.g., facet injection, median branch block) may predict outcomes after lumbar fusion surgery may assist clinicians in identifying which patients may benefit most from fusion surgery.

4.7 Conclusions

High quality evidence on optimal approaches to fusion surgery in patients with stable DLS is limited. The addition of fusion following decompression was not associated with improved pain or function up to 3 years, although some improvement in quality of life and slightly lower likelihood of reoperation at 1 year was seen versus decompression alone. TLIF plus PLF was associated with a moderately higher likelihood of fusion versus PLF by 2 years. There is inadequate detail about specific surgery-related neurological events to determine benefits of IONM during fusion surgery for degenerative lumbar disease. There was insufficient evidence on use interventional procedures for predicting outcomes of fusion. None of the trials of special procedures were specifically in patients with DLS. Overall, there was not consistent pain and function improvement with continuous RFA versus sham for presumed facet joint pain, however, there was heterogeneity across trials with some reporting improvements in pain and function. ESI, facet joint injection and medial bundle branch block were not associated with improved pain or function versus placebo in patients with chronic low back pain due to spinal stenosis. Research is needed to clarify optimal approaches to fusion surgery for stable DLS and alternatives to surgery.

References

1. Parenteau CS, Lau EC, Campbell IC, et al. Prevalence of spine degeneration diagnosis by type, age, gender, and obesity using Medicare data. Sci Rep. 2021;11(1):5389. doi: 10.1038/s41598-021-84724-6. PMID: 33686128.
2. Wiltse LL, Newman PH, Macnab I. Classification of spondylolisis and spondylolisthesis. Clin Orthop Relat Res. 1976 (117):23-9. PMID: 1277669.
3. Bydon M, Alvi MA, Goyal A. Degenerative Lumbar Spondylolisthesis: Definition, Natural History, Conservative Management, and Surgical Treatment. Neurosurg Clin N Am. 2019;30(3):299-304. doi: 10.1016/j.nec.2019.02.003. PMID: 31078230.
4. Wang YXJ, Kaplar Z, Deng M, et al. Lumbar degenerative spondylolisthesis epidemiology: A systematic review with a focus on gender-specific and age-specific prevalence. J Orthop Translat. 2017;11:39-52. doi: 10.1016/j.jot.2016.11.001. PMID: 29662768.
5. Evans N, McCarthy M. Management of symptomatic degenerative low-grade lumbar spondylolisthesis. EFORT Open Rev. 2018;3(12):620-31. doi: 10.1302/2058-5241.3.180020. PMID: 30697442.
6. Koslosky E, Gendelberg D. Classification in Brief: The Meyerding Classification System of Spondylolisthesis. Clin Orthop Relat Res. 2020;478(5):1125-30. doi: 10.1097/CORR.0000000000001153. PMID: 32282463.
7. Reid PC, Morr S, Kaiser MG. State of the union: a review of lumbar fusion indications and techniques for degenerative spine disease. J Neurosurg Spine. 2019;31(1):1-14. doi: 10.3171/2019.4.SPINE18915. PMID: 31261133.
8. Chatprem T, Puntumetakul R, Kanpittaya J, et al. A diagnostic tool for people with lumbar instability: a criterion-related validity study. BMC Musculoskelet Disord. 2021;22(1):976. doi: 10.1186/s12891-021-04854-w. PMID: 34814879.
9. Reeves NP, Cholewicki J, van Dieen JH, et al. Are Stability and Instability Relevant Concepts for Back Pain? J Orthop Sports Phys Ther. 2019;49(6):415-24. doi: 10.2519/jospt.2019.8144. PMID: 31021689.
10. Meyerding HW. Low backache and sciatic pain associated with spondylolisthesis and protruded intervertebral disc: incidence, signficance, and treatment. J Bone Joint Surg Am. 1941;23(2):461-70.
11. Fredrickson BE, Baker D, McHolick WJ, et al. The natural history of spondylolysis and spondylolisthesis. J Bone Joint Surg Am. 1984;66(5):699-707. PMID: 6373773.
12. Chan AK, Sharma V, Robinson LC, et al. Summary of Guidelines for the Treatment of Lumbar Spondylolisthesis. Neurosurg Clin N Am. 2019;30(3):353-64. doi: 10.1016/j.nec.2019.02.009. PMID: 31078236.
13. Jia H, Zhang Z, Qin J, et al. Management for degenerative lumbar spondylolisthesis: a network meta-analysis and systematic review basing on randomized controlled trials. Int J Surg. 2024;110(5):3050-9. doi: 10.1097/JS9.0000000000001228. PMID: 38446872.
14. Conger A, Schuster NM, Cheng DS, et al. The Effectiveness of Intraosseous Basivertebral Nerve Radiofrequency Neurotomy for the Treatment of Chronic Low Back Pain in Patients with Modic Changes: A Systematic Review. Pain Med. 2021;22(5):1039-54. doi: 10.1093/pm/pnab040. PMID: 33544851.
15. Resnick DK, Choudhri TF, Dailey AT, et al. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 1: introduction and methodology. J Neurosurg Spine. 2005;2(6):637-8. doi: 10.3171/spi.2005.2.6.0637. PMID: 16028729.
16. Lee JH, Kim DH, Kim DH, et al. Comparison of Clinical Efficacy of Epidural Injection With or Without Steroid in Lumbosacral Disc Herniation: A Systematic Review and Meta-analysis. Pain Physician. 2018;21(5):449-68. PMID: 30282390.

17. Cohen SP, Bhaskar A, Bhatia A, et al. Consensus practice guidelines on interventions for lumbar facet joint pain from a multispecialty, international working group. Reg Anesth Pain Med. 2020;45(6):424-67. PMID: 32245841.

18. Chou R, Baisden J, Carragee EJ, et al. Surgery for low back pain: a review of the evidence for an American Pain Society Clinical Practice Guideline. Spine (Phila Pa 1976). 2009;34(10):1094-109. doi: 10.1097/BRS.0b013e3181a105fc. PMID: 19363455.

19. Martin BI, Mirza SK, Comstock BA, et al. Reoperation rates following lumbar spine surgery and the influence of spinal fusion procedures. Spine (Phila Pa 1976). 2007;32(3):382-7. doi: 10.1097/01.brs.0000254104.55716.46. PMID: 17268274.

20. Azad TD, Vail D, O'Connell C, et al. Geographic variation in the surgical management of lumbar spondylolisthesis: characterizing practice patterns and outcomes. Spine J. 2018;18(12):2232-8. doi: 10.1016/j.spinee.2018.05.008. PMID: 29746964.

21. Postacchini F, Cinotti G, Perugia D. Degenerative lumbar spondylolisthesis. II. Surgical treatment. Ital J Orthop Traumatol. 1991;17(4):467-77. PMID: 1816152.

22. Blumenthal C, Curran J, Benzel EC, et al. Radiographic predictors of delayed instability following decompression without fusion for degenerative grade I lumbar spondylolisthesis. J Neurosurg Spine. 2013;18(4):340-6.

23. North American Spine Society (NASS). Clinical Guidelines for Multidisciplinary Spine Care. Diagnosis and Treatment of Degenerative Lumbar Spondylolisthesis 2nd Edition. 2014. https://www.spine.org/Portals/0/Assets/Downloads/ResearchClinicalCare/Guidelines/Spondylolisthesis.pdf

24. Kaiser MG, Eck JC, Groff MW, et al. Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 1: introduction and methodology. J Neurosurg Spine. 2014;21(1):2-6. doi: 10.3171/2014.4.SPINE14257. PMID: 24980578.

25. Matz PG, Meagher RJ, Lamer T, et al. Guideline summary review: An evidence-based clinical guideline for the diagnosis and treatment of degenerative lumbar spondylolisthesis. Spine J. 2016;16(3):439-48. doi: 10.1016/j.spinee.2015.11.055. PMID: 26681351.

26. Chan AK, Sharma V, Robinson LC, et al. Summary of Guidelines for the Treatment of Lumbar Spondylolisthesis. Neurosurgery Clinics of North America. 2019;30(3):353-64. doi: https://dx.doi.org/10.1016/j.nec.2019.02.009 . PMID: 31078236.

27. Rathbone J, Rackham M, Nielsen D, et al. A systematic review of anterior lumbar interbody fusion (ALIF) versus posterior lumbar interbody fusion (PLIF), transforaminal lumbar interbody fusion (TLIF), posterolateral lumbar fusion (PLF). Eur Spine J. 2023;32(6):1911-26.

28. Herkowitz HN, Kurz LT. Degenerative lumbar spondylolisthesis with spinal stenosis. A prospective study comparing decompression with decompression and intertransverse process arthrodesis. J Bone Joint Surg Am. 1991;73(6):802-8. PMID: 2071615.

29. Sengupta DK, Herkowitz HN. Degenerative spondylolisthesis: review of current trends and controversies. Spine (Phila Pa 1976). 2005;30(6 Suppl):S71-81. doi: 10.1097/01.brs.0000155579.88537.8e. PMID: 15767890.

30. Hirase T, Ling JF, Haghshenas V, et al. Instrumented Versus Noninstrumented Spinal Fusion for Degenerative Lumbar Spondylolisthesis: A Systematic Review. Clin Spine Surg. 2022;35(5):213-21. doi: 10.1097/BSD.0000000000001266. PMID: 35239288.

31. Achttien RJ, Powell A, Zoulas K, et al. Prognostic factors for outcome following lumbar spine fusion surgery: a systematic review and narrative synthesis. Eur Spine J. 2022;31(3):623-68. doi: 10.1007/s00586-021-07018-5. PMID: 34705106.

32. Chou R. Commentary: Successful spinal fusion surgery: can we improve the odds? Spine J. 2013;13(2):110-2. doi: 10.1016/j.spinee.2012.11.020. PMID: 23452566.

33. Willems PC, Staal JB, Walenkamp GH, et al. Spinal fusion for chronic low back pain: systematic review on the accuracy of tests for patient selection. Spine J. 2013;13(2):99-109. doi: 10.1016/j.spinee.2012.10.001. PMID: 23127364.

34. Willems P. Decision making in surgical treatment of chronic low back pain: the performance of prognostic tests to select patients for lumbar spinal fusion. Acta Orthop Suppl. 2013;84(349):1-35. doi: 10.3109/17453674.2012.753565. PMID: 23427903.

35. Resnick DK, Choudhri TF, Dailey AT, et al. Guidelines for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 13: injection therapies, low-back pain, and lumbar fusion. J Neurosurg Spine. 2005;2(6):707-15. doi: 10.3171/spi.2005.2.6.0707. PMID: 16028741.

36. Alvi MA, Kwon BK, Hejrati N, et al. Accuracy of Intraoperative Neuromonitoring in the Diagnosis of Intraoperative Neurological Decline in the Setting of Spinal Surgery-A Systematic Review and Meta-Analysis. Global Spine J. 2024;14(3_suppl):105S-49S. PMID: Pmc10964897.

37. Austerman RJ, Sulhan S, Steele WJ, et al. The utility of intraoperative neuromonitoring on simple posterior lumbar fusions—analysis of the National Inpatient Sample. Journal of Spine Surgery. 2021;7(2):132-40.

38. Sharan A, Groff MW, Dailey AT, et al. Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 15: electrophysiological monitoring and lumbar fusion. J Neurosurg Spine. 2014;21(1):102-5.

39. Resnick DK, Schmidt BT, Momin E, et al. Interobserver variance and patient heterogeneity influencing the treatment of grade I spondylolisthesis. Spine J. 2020;20(12):1934-9. doi: 10.1016/j.spinee.2020.06.001. PMID: 32534135.

40. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Bmj. 2021;372:n71. doi: 10.1136/bmj.n71. PMID: 33782057.

41. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Content last reviewed October 2022. Effective Health Care Program, Agency for Healthcare Research and Quality. Rockville, MD. https://effectivehealthcare.ahrq.gov/products/collections/cer-methods-guide2022.

42. Viswanathan M, Ansari MT, Berkman ND, et al. Chapter 9: Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions. In: Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2014. Chapters available at: www.effectivehealthcare.ahrq.gov.

43. Higgins JP, Thomas J, Chandler J, et al. Cochrane handbook for systematic reviews of interventions version 6.2: John Wiley & Sons; 2021.

44. Furlan AD, Malmivaara A, Chou R, et al. 2015 Updated Method Guideline for Systematic Reviews in the Cochrane Back and Neck Group. Spine (Phila Pa 1976). 2015;40(21):1660-73. doi: 10.1097/BRS.0000000000001061. PMID: 26208232.

45. Viswanathan M, Berkman ND. Development of the RTI item bank on risk of bias and precision of observational studies. J Clin Epidemiol. 2012;65(2):163-78. doi: 10.1016/j.jclinepi.2011.05.008. PMID: 21959223.

46. Fu R, Gartlehner G, Grant M, et al. Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. J Clin Epidemiol. 2011;64(11):1187-97. doi: 10.1016/j.jclinepi.2010.08.010. PMID: 21477993.

47. Morton SC, Murad MH, O'Connor E, et al. Quantitative Synthesis—An Update. Methods Guide for Comparative Effectiveness Reviews. (Prepared by the Scientific Resource Center under Contract No. 290-2012-0004-C). AHRQ Publication No. 18-EHC007-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2018. doi: 10 .23970/AHRQEPCMETHGUIDE3.

48. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. Bmj. 2003;327(7414):557-60. doi: 10.1136/bmj.327.7414.557. PMID: 12958120.

49. Chou R, Wagner J, Ahmed AY, et al. Treatments for Acute Pain: A Systematic Review. Comparative Effectiveness Review No. 240 (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. HHSA 290-2015-00009-I.) AHRQ Publication No. 20(21)-EHC006. Rockville, MD: Agency for Healthcare Research and Quality; December 2020. doi: https://doi.org/10.23970/AHRQEPCCER240.

50. Skelly AC, Chou R, Dettori JR, et al. Noninvasive Nonpharmacological Treatment for Chronic Pain: A Systematic Review Update. Comparative Effectiveness Review No. 227. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2015-00009-I.) AHRQ Publication No. 20-EHC009. Rockville, MD: Agency for Healthcare Research and Quality; April 2020. https://effectivehealthcare.ahrq.gov/products/noninvasive-nonpharm-pain-update/researchdoi: https://doi.org/10.23970/AHRQEPCCER227.

51. Skelly AC, Chou R, Dettori JR, et al. Noninvasive Nonpharmacological Treatment for Chronic Pain: A Systematic Review. Comparative Effectiveness Review No. 209. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2015-00009-I.) AHRQ Publication No 18-EHC013-EF. Rockville, MD: Agency for Healthcare Research and Quality; June 2018. doi: https://doi.org/10.23970/AHRQEPCCER209. .

52. Chou R, Deyo R, Friedly J, et al. Noninvasive Treatments for Low Back Pain. Comparative Effectiveness Review No. 169. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2012-00014-I.) AHRQ Publication No. 16-EHC004-EF. Agency for Healthcare Research and Quality. Rockville, MD: 2016. PMID: 26985522.

53. Chou R, Hartung D, Turner J, et al. Opioid Treatments for Chronic Pain. Comparative Effectiveness Review No. 229. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290 2015-00009-I.) AHRQ Publication No. 20 EHC011 Agency for Healthcare Research and Quality. Rockville, MD: 2020. doi: 10.23970/AHRQEPCCER229. PMID: 32338848.

54. Sterne JAC, Sutton AJ, Ioannidis JPA, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011;343:d4002. doi: 10.1136/bmj.d4002.

55. Austevoll IM, Hermansen E, Fagerland MW, et al. Decompression with or without Fusion in Degenerative Lumbar Spondylolisthesis. New England Journal of Medicine. 2021;385(6):526-38. doi: 10.1056/NEJMoa2100990. PMID: 34347953.

56. Carette S, Marcoux S, Truchon R, et al. A controlled trial of corticosteroid injections into facet joints for chronic low back pain. New England Journal of Medicine. 1991;325(14):1002-7. doi: 10.1056/NEJM199110033251405. PMID: 1832209.

57. Challier V, Boissiere L, Obeid I, et al. One-Level Lumbar Degenerative Spondylolisthesis and Posterior Approach: Is Transforaminal Lateral Interbody Fusion Mandatory?: A Randomized Controlled Trial With 2-Year Follow-Up. Spine. 2017;42(8):531-9. doi: 10.1097/BRS.0000000000001857. PMID: 27548582.

58. Chen YS, Liu B, Gu F, et al. Radiofrequency Denervation on Lumbar Facet Joint Pain in the Elderly: A Randomized Controlled Prospective Trial. Pain Physician. 2022;25(8):569-76. PMID: 36375186.

59. Civelek E, Cansever T, Kabatas S, et al. Comparison of effectiveness of facet joint injection and radiofrequency denervation in chronic low back pain. Turkish Neurosurgery. 2012;22(2):200-6. doi: https://dx.doi.org/10.5137/1019-5149.JTN.5207-11.1. PMID: 22437295.

60. Do KH, Ahn SH, Cho YW, et al. Comparison of intra-articular lumbar facet joint pulsed radiofrequency and intra-articular lumbar facet joint corticosteroid injection for management of lumbar facet joint pain: A randomized controlled trial. Medicine. 2017;96(13):e6524. doi: 10.1097/MD.0000000000006524. PMID: 28353611.

61. Düger C, ÖzdemİR Kol İ, Kaygusuz K, et al. Effects of facet joint nerve block addition to radiofrequency in the treatment of low back pain. HealthMED. 2012;6(6).

62. Fariborz S, Gharedaghi M, Khosravi A, et al. Comparison of results of 4 methods of surgery in grade 1 lumbosacral spondylolisthesis. Neurosurgery. 2016;26:14-8.

63. Forsth P, Olafsson G, Carlsson T, et al. A Randomized, Controlled Trial of Fusion Surgery for Lumbar Spinal Stenosis. New England Journal of Medicine. 2016;374(15):1413-23. doi: 10.1056/NEJMoa1513721. PMID: 27074066.

64. Friedly JL, Bauer Z, Comstock B, et al. Comparing the Effects of Two Types of Epidural Shots on Pain and Physical Ability in Older Adults with Lumbar Spinal Stenosis. Patient Centered Outcomes Research Institute. 2019;04:04. doi: 10.25302/4.2019.CE.12114469. PMID: 38315786.

65. Friedly JL, Comstock BA, Turner JA, et al. A randomized trial of epidural glucocorticoid injections for spinal stenosis. New England Journal of Medicine. 2014;371(1):11-21. doi: 10.1056/NEJMoa1313265. PMID: 24988555.

66. Gallagher J PDVP, Wedley JR. Radiofrequency facet joint denervation in the treatment of low back pain: a prospective controlled double-blind study to assess its efficacy. Pain Clinic. 1994;7:193-8.

67. Ghogawala Z, Dziura J, Butler WE, et al. Laminectomy plus Fusion versus Laminectomy Alone for Lumbar Spondylolisthesis. New England Journal of Medicine. 2016;374(15):1424-34. doi: 10.1056/NEJMoa1508788. PMID: 27074067.

68. Hashemi M, Hashemian M, Mohajerani SA, et al. Effect of pulsed radiofrequency in treatment of facet-joint origin back pain in patients with degenerative spondylolisthesis. European Spine Journal. 2014;23(9):1927-32. doi: 10.1007/s00586-014-3412-x. PMID: 24997616.

69. Inose H, Kato T, Sasaki M, et al. Comparison of decompression, decompression plus fusion, and decompression plus stabilization: a long-term follow-up of a prospective, randomized study. Spine Journal: Official Journal of the North American Spine Society. 2022;22(5):747-55. doi: 10.1016/j.spinee.2021.12.014. PMID: 34963630.

70. Inose H, Kato T, Yuasa M, et al. Comparison of Decompression, Decompression Plus Fusion, and Decompression Plus Stabilization for Degenerative Spondylolisthesis: A Prospective, Randomized Study. Clinical Spine Surgery : A Spine Publication. 2018;31(7):E347-E52. doi: 10.1097/BSD.0000000000000659. PMID: 29877872.

71. Kang J, An H, Hilibrand A, et al. Grafton and local bone have comparable outcomes to iliac crest bone in instrumented single-level lumbar fusions. Spine. 2012;37(12):1083-91. doi: 10.1097/BRS.0b013e31823ed817. PMID: 22076647.

72. Karlsson T, Forsth P, Ohagen P, et al. Decompression alone or decompression with fusion for lumbar spinal stenosis: five-year clinical results from a randomized clinical trial. Bone & Joint Journal. 2024;106-B(7):705-12. doi: 10.1302/0301-620X.106B7.BJJ-2023-1160.R2. PMID: 38945544.

73. Karlsson T, Forsth P, Skorpil M, et al. Decompression alone or decompression with fusion for lumbar spinal stenosis: a randomized clinical trial with two-year MRI follow-up. Bone & Joint Journal. 2022;104-B(12):1343-51. doi: 10.1302/0301-620X.104B12.BJJ-2022-0340.R1. PMID: 36453045.

74. Kawu AA, Olawepo A, Salami AO. Facet joints infiltration: a viable alternative treatment to physiotherapy in patients with low back pain due to facet joint arthropathy. Nigerian Journal of Clinical Practice. 2011;14(2):219-22. doi: https://dx.doi.org/10.4103/1119-3077.84021. PMID: 21860143.

75. Kgomotso EL, Hellum C, Fagerland MW, et al. Decompression alone or with fusion for degenerative lumbar spondylolisthesis (Nordsten-DS): five year follow-up of a randomised, multicentre, non-inferiority trial. BMJ. 2024;386:e079771. doi: 10.1136/bmj-2024-079771. PMID: 39111800.

76. Koc Z, Ozcakir S, Sivrioglu K, et al. Effectiveness of physical therapy and epidural steroid injections in lumbar spinal stenosis. Spine. 2009;34(10):985-9. doi: 10.1097/BRS.0b013e31819c0a6b. PMID: 19404172.

77. Kolodge G, Gold LS, Jarvik JG, et al. Treatment With Opioids Is Not Associated With Poor Outcomes Among Older Adults With Lumbar Spinal Stenosis Receiving Epidural Injections. Spine. 2023;48(7):445-51. doi: 10.1097/BRS.0000000000004558. PMID: 36729991.

78. Lakemeier S, Lind M, Schultz W, et al. A comparison of intraarticular lumbar facet joint steroid injections and lumbar facet joint radiofrequency denervation in the treatment of low back pain: a randomized, controlled, double-blind trial. Anesthesia & Analgesia. 2013;117(1):228-35. doi: 10.1213/ANE.0b013e3182910c4d. PMID: 23632051.

79. Leclaire R, Fortin L, Lambert R, et al. Radiofrequency facet joint denervation in the treatment of low back pain: a placebo-controlled clinical trial to assess efficacy. Spine. 2001;26(13):1411-6; discussion 7. doi: 10.1097/00007632-200107010-00003. PMID: 11458140.

80. Lilius G, Laasonen EM, Myllynen P, et al. Lumbar facet joint syndrome. A randomised clinical trial. Journal of Bone & Joint Surgery - British Volume. 1989;71(4):681-4. doi: 10.1302/0301-620X.71B4.2527856. PMID: 2527856.

81. Makris UE, Edwards TC, Lavallee DC, et al. Patient Priority Weighting of the Roland Morris Disability Questionnaire Does Not Change Results of the Lumbar Epidural Steroid Injections for Spinal Stenosis Trial. Spine. 2017;42(1):42-8. doi: 10.1097/BRS.0000000000001647. PMID: 27105466.

82. Manchikanti L, Cash KA, McManus CD, et al. A randomized, double-blind controlled trial of lumbar interlaminar epidural injections in central spinal stenosis: 2-year follow-up. Pain Physician. 2015;18(1):79-92. PMID: 25675062.

83. Manchikanti L, Cash KA, McManus CD, et al. Results of 2-year follow-up of a randomized, double-blind, controlled trial of fluoroscopic caudal epidural injections in central spinal stenosis. Pain Physician. 2012;15(5):371-84. PMID: 22996849.

84. Manchikanti L, Singh V, Falco FJ, et al. Lumbar facet joint nerve blocks in managing chronic facet joint pain: one-year follow-up of a randomized, double-blind controlled trial: Clinical Trial NCT00355914. Pain Physician. 2008;11(2):121-32. PMID: 18354721.

85. Manchikanti L, Singh V, Falco FJ, et al. Evaluation of lumbar facet joint nerve blocks in managing chronic low back pain: a randomized, double-blind, controlled trial with a 2-year follow-up. International Journal of Medical Sciences. 2010;7(3):124-35. doi: 10.7150/ijms.7.124. PMID: 20567613.

86. Marks RC, Houston T, Thulbourne T. Facet joint injection and facet nerve block: a randomised comparison in 86 patients with chronic low back pain. Pain. 1992;49(3):325-8. doi: 10.1016/0304-3959(92)90239-8. PMID: 1408298.

87. McCormick ZL, Conger A, Kendall R, et al. A pragmatic randomized prospective trial of cooled radiofrequency ablation of the medial branch nerves versus facet joint injection of corticosteroid for the treatment of lumbar facet syndrome: 12 month outcomes. Pain Medicine. 2023;24(12):1318-31. doi: 10.1093/pm/pnad107. PMID: 37578437.

88. Moussa WM, Khedr W. Percutaneous radiofrequency facet capsule denervation as an alternative target in lumbar facet syndrome. Clinical Neurology & Neurosurgery. 2016;150:96-104. doi: https://dx.doi.org/10.1016/j.clineuro.2016.09.004. PMID: 27618781.

89. Moussa WM, Khedr W, Elsawy M. Percutaneous pulsed radiofrequency treatment of dorsal root ganglion for treatment of lumbar facet syndrome. Clinical Neurology & Neurosurgery. 2020;199:106253. doi: 10.1016/j.clineuro.2020.106253. PMID: 33045627.

90. Nam HS, Park YB. Effects of transforaminal injection for degenerative lumbar scoliosis combined with spinal stenosis. Annals of Rehabilitation Medicine. 2011;35(4):514-23. doi: 10.5535/arm.2011.35.4.514. PMID: 22506167.

91. Nath S, Nath CA, Pettersson K. Percutaneous lumbar zygapophysial (Facet) joint neurotomy using radiofrequency current, in the management of chronic low back pain: a randomized double-blind trial. Spine. 2008;33(12):1291-7; discussion 8. doi: 10.1097/BRS.0b013e31817329f0. PMID: 18496338.

92. Ribeiro LH, Furtado RN, Konai MS, et al. Effect of facet joint injection versus systemic steroids in low back pain: a randomized controlled trial. Spine. 2013;38(23):1995-2002. doi: 10.1097/BRS.0b013e3182a76df1. PMID: 23921331.

93. Tekin I, Mirzai H, Ok G, et al. A comparison of conventional and pulsed radiofrequency denervation in the treatment of chronic facet joint pain. Clinical Journal of Pain. 2007;23(6):524-9. doi: 10.1097/AJP.0b013e318074c99c. PMID: 17575493.

94. Truong K, Meier K, Ahrens LC, et al. Cryoneurolysis versus radiofrequency ablation outcome on pain experience in chronic low back pain (COPE): a single-blinded randomised controlled trial. RMD Open. 2024;10(2):09. doi: 10.1136/rmdopen-2024-004196. PMID: 38724261.

95. van Tilburg CW, Stronks DL, Groeneweg JG, et al. Randomised sham-controlled double-blind multicentre clinical trial to ascertain the effect of percutaneous radiofrequency treatment for lumbar facet joint pain. Bone & Joint Journal. 2016;98-B(11):1526-33. doi: 10.1302/0301-620X.98B11.BJJ-2016-0379.R2. PMID: 27803230.

96. van Wijk RM, Geurts JW, Wynne HJ, et al. Radiofrequency denervation of lumbar facet joints in the treatment of chronic low back pain: a randomized, double-blind, sham lesion-controlled trial. Clinical Journal of Pain. 2005;21(4):335-44. doi: 10.1097/01.ajp.0000120792.69705.c9. PMID: 15951652.

97. Zhou Q, Zhou F, Wang L, et al. An investigation on the effect of improved X-rays-guided radiofrequency thermocoagulation denervation on lumbar facet joint syndrome. Clinical Neurology & Neurosurgery. 2016;148:115-20. doi: 10.1016/j.clineuro.2016.07.018. PMID: 27442000.

98. Abdu WA, Lurie JD, Spratt KF, et al. Degenerative spondylolisthesis: does fusion method influence outcome? Four-year results of the spine patient outcomes research trial. Spine (Phila Pa 1976). 2009;34(21):2351-60. doi: 10.1097/BRS.0b013e3181b8a829. PMID: 19755935.

99. Abdu WA, Sacks OA, Tosteson ANA, et al. Long-Term Results of Surgery Compared With Nonoperative Treatment for Lumbar Degenerative Spondylolisthesis in the Spine Patient Outcomes Research Trial (SPORT). Spine. 2018;43(23):1619-30. doi: 10.1097/BRS.0000000000002682. PMID: 29652786.

100. Austerman RJ, Sulhan S, Steele WJ, et al. The utility of intraoperative neuromonitoring on simple posterior lumbar fusions-analysis of the National Inpatient Sample. The Journal of Spine Surgery. 2021;7(2):132-40. doi: 10.21037/jss-20-679. PMID: 34296025.

101. Austevoll IM, Gjestad R, Brox JI, et al. The effectiveness of decompression alone compared with additional fusion for lumbar spinal stenosis with degenerative spondylolisthesis: a pragmatic comparative non-inferiority observational study from the Norwegian Registry for Spine Surgery. European Spine Journal. 2017;26(2):404-13. doi: 10.1007/s00586-016-4683-1. PMID: 27421276.

102. Austevoll IM, Gjestad R, Solberg T, et al. Comparative Effectiveness of Microdecompression Alone vs Decompression Plus Instrumented Fusion in Lumbar Degenerative Spondylolisthesis. JAMA Network Open. 2020;3(9):e2015015. doi: 10.1001/jamanetworkopen.2020.15015. PMID: 32910195.

103. Badhiwala JH, Leung SN, Jiang F, et al. In-hospital Course and Complications of Laminectomy Alone Versus Laminectomy Plus Instrumented Posterolateral Fusion for Lumbar Degenerative Spondylolisthesis: A Retrospective Analysis of 1804 Patients from the NSQIP Database. Spine. 2021;46(9):617-23. doi: 10.1097/BRS.0000000000003858. PMID: 33290365.

104. Esses SI, Moro JK. The value of facet joint blocks in patient selection for lumbar fusion. Spine. 1993;18(2):185-90. doi: 10.1097/00007632-199302000-00003. PMID: 8441932.

105. Issa TZ, Lee Y, Lambrechts MJ, et al. Comparing Posterior Lumbar Decompression and Fusion and Transforaminal Lumbar Interbody Fusion in Lumbar Degenerative Spondylolisthesis as Assessed by the CARDS Classification System. World Neurosurgery. 2023;175:e861-e75. doi: 10.1016/j.wneu.2023.04.036. PMID: 37075895.

106. Kuo CC, Merchant M, Kardile MP, et al. In Degenerative Spondylolisthesis, Unilateral Laminotomy for Bilateral Decompression Leads to Less Reoperations at 5 Years When Compared to Posterior Decompression With Instrumented Fusion: A Propensity-matched Retrospective Analysis. Spine. 2019;44(21):1530-7. doi: 10.1097/BRS.0000000000003121. PMID: 31181016.

107. Norton RP, Bianco K, Klifto C, et al. Degenerative Spondylolisthesis: An Analysis of the Nationwide Inpatient Sample Database. Spine. 2015;40(15):1219-27. doi: 10.1097/BRS.0000000000000987. PMID: 26020842.

108. Staartjes VE, Schroder ML. Effectiveness of a Decision-Making Protocol for the Surgical Treatment of Lumbar Stenosis with Grade 1 Degenerative Spondylolisthesis. World Neurosurgery. 2018;110:e355-e61. doi: 10.1016/j.wneu.2017.11.001. PMID: 29133000.

109. Sugiura T, Okuda S, Takenaka S, et al. Comparing Investigation Between Bilateral Partial Laminectomy and Posterior Lumbar Interbody Fusion for Mild Degenerative Spondylolisthesis. Clinical Spine Surgery : A Spine Publication. 2021;34(7):E403-E9. doi: 10.1097/BSD.0000000000001109. PMID: 33290326.

110. Tang AR, Chanbour H, Steinle AM, et al. Transforaminal Lumbar Interbody Fusion Versus Posterolateral Fusion Alone in the Treatment of Grade 1 Degenerative Spondylolisthesis. Neurosurgery. 2023;93(1):186-97. doi: 10.1227/neu.0000000000002402. PMID: 36848669.

111. Ulrich NH, Burgstaller JM, Valeri F, et al. Incidence of Revision Surgery After Decompression With vs Without Fusion Among Patients With Degenerative Lumbar Spinal Stenosis. JAMA Network Open. 2022;5(7):e2223803. doi: 10.1001/jamanetworkopen.2022.23803. PMID: 35881393.

112. Vorhies JS, Hernandez-Boussard T, Alamin T. Treatment of Degenerative Lumbar Spondylolisthesis With Fusion or Decompression Alone Results in Similar Rates of Reoperation at 5 Years. Clin Spine Surg. 2018;31(1):E74-e9. doi: 10.1097/bsd.0000000000000564. PMID: 28671881.

113. Kepler CK, Hilibrand AS, Sayadipour A, et al. Clinical and radiographic degenerative spondylolisthesis (CARDS) classification. Spine J. 2015;15(8):1804-11. doi: 10.1016/j.spinee.2014.03.045. PMID: 24704503.

114. Manchikanti L, Cash KA, McManus CD, et al. Fluoroscopic caudal epidural injections with or without steroids in managing pain of lumbar spinal stenosis: one-year results of randomized, double-blind, active-controlled trial. Journal of Spinal Disorders & Techniques. 2012;25(4):226-34. doi: 10.1097/BSD.0b013e3182160068. PMID: 22652990.

115. Bydon M, Alvi MA, Goyal A. Degenerative Lumbar Spondylolisthesis: Definition, Natural History, Conservative Management, and Surgical Treatment. Neurosurgery Clinics of North America. 2019;30(3):299-304. doi: https://dx.doi.org/10.1016/j.nec.2019.02.003. PMID: 31078230.

116. Jia H, Zhang Z, Qin J, et al. Management for degenerative lumbar spondylolisthesis: a network meta-analysis and systematic review basing on randomized controlled trials. International Journal Of Surgery. 2024;110(5):3050-9. doi: 10.1097/JS9.0000000000001228. PMID: 38446872.

117. Kaiser R, Kantorova L, Langaufova A, et al. Decompression alone versus decompression with instrumented fusion in the treatment of lumbar degenerative spondylolisthesis: a systematic review and meta-analysis of randomised trials. Journal of Neurology, Neurosurgery & Psychiatry. 2023;94(8):657-66. doi: 10.1136/jnnp-2022-330158. PMID: 36849239.

118. Abdel-Fattah AR, Bell F, Boden L, et al. To fuse or not to fuse: The elderly patient with lumbar stenosis and low-grade spondylolisthesis. Systematic review and meta-analysis of randomised controlled trials. Surgeon Journal of the Royal Colleges of Surgeons of Edinburgh & Ireland. 2023;21(1):e23-e31. doi: https://dx.doi.org/10.1016/j.surge.2022.02.008. PMID: 35305933.

119. Dantas F, Dantas FLR, Botelho RV. Effect of interbody fusion compared with posterolateral fusion on lumbar degenerative spondylolisthesis: a systematic review and meta-analysis. Spine Journal: Official Journal of the North American Spine Society. 2022;22(5):756-68. doi: 10.1016/j.spinee.2021.12.001. PMID: 34896611.

120. Salamanna F, Tschon M, Borsari V, et al. Spinal fusion procedures in the adult and young population: a systematic review on allogenic bone and synthetic grafts when compared to autologous bone. J Mater Sci Mater Med. 2020;31(6):51. doi: 10.1007/s10856-020-06389-3. PMID: 32451687.

121. Fu R, Selph S, McDonagh M, et al. Effectiveness and harms of recombinant human bone morphogenetic protein-2 in spine fusion: a systematic review and meta-analysis. Ann Intern Med. 2013;158(12):890-902. doi: 10.7326/0003-4819-158-12-201306180-00006. PMID: 23778906.

122. Laurie AL, Chen Y, Chou R, et al. Meta-analysis of the Impact of Patient Characteristics on Estimates of Effectiveness and Harms of Recombinant Human Bone Morphogenetic Protein-2 in Lumbar Spinal Fusion. Spine (Phila Pa 1976). 2016;41(18):E1115-E23. doi: 10.1097/BRS.0000000000001580. PMID: 27031772.

123. Nie JW, Hartman TJ, Zheng E, et al. Neuromonitoring in Lateral Approaches for Lumbar Interbody Fusion: A Systematic Review. World Neurosurg. 2022;168:268-77.e1. doi: 10.1016/j.wneu.2022.10.031. PMID: 36243359.

124. Sharan A, Groff MW, Dailey AT, et al. Guideline update for the performance of fusion procedures for degenerative disease of the lumbar spine. Part 15: electrophysiological monitoring and lumbar fusion. Journal of Neurosurgery Spine. 2014;21(1):102-5. doi: https://dx.doi.org/10.3171/2014.4.SPINE14324. PMID: 24980592.

125. Busse JW, Genevay S, Agarwal A, et al. Commonly used interventional procedures for non-cancer chronic spine pain: a clinical practice guideline. BMJ. 2025;388:e079970. doi: 10.1136/bmj-2024-079970. PMID: 39971339.

126. Wang X, Martin G, Sadeghirad B, et al. Common interventional procedures for chronic non-cancer spine pain: a systematic review and network meta-analysis of randomised trials. BMJ. 2025;388:e079971. doi: 10.1136/bmj-2024-079971. PMID: 39971346.

127. Bernardes A, Correa RF, Trajano L, et al. Lumbar Facet Syndrome and the Use of Radiofrequency Ablation Technique as an Alternative Therapy: A Systematic Review. Revista Brasileira de Ortopedia. 2023;58(2):199-205. doi: 10.1055/s-0043-1768695. PMID: 37252302.

128. Du R, Gao J, Wang B, et al. Percutaneous radiofrequency ablation and endoscopic neurotomy for lumbar facet joint syndrome: are they good enough? European Spine Journal. 2024;33(2):463-73. doi: 10.1007/s00586-023-08078-5. PMID: 38141106.

129. Falco FJ, Manchikanti L, Datta S, et al. An update of the effectiveness of therapeutic lumbar facet joint interventions. Pain Physician. 2012;15(6):E909-53. PMID: 23159980.

130. Lee CH, Chung CK, Kim CH. The efficacy of conventional radiofrequency denervation in patients with chronic low back pain originating from the facet joints: a meta-analysis of randomized controlled trials. Spine Journal: Official Journal of the North American Spine Society. 2017;17(11):1770-80. doi: 10.1016/j.spinee.2017.05.006. PMID: 28576500.

131. Leggett LE, Soril LJ, Lorenzetti DL, et al. Radiofrequency ablation for chronic low back pain: a systematic review of randomized controlled trials. Pain Research & Management. 2014;19(5):e146-53. doi: https://dx.doi.org/10.1155/2014/834369. PMID: 25068973.

132. Maas ET, Ostelo WJGR, Niemisto L, et al. Radiofrequency denervation for chronic low back pain. Cochrane Database of Systematic Reviews. 2015 (10) PMID: 00075320-100000000-07055.

133. Skelly AC, Chou R, Dettori JR, et al. Integrated and Comprehensive Pain Management Programs: Effectiveness and Harms. Comparative Effectiveness Review No. 251. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 75Q80120D00006.) AHRQ Publication No. 22-EHC002. Rockville, MD: Agency for Healthcare Research and Quality; October 2021. doi: https://doi.org/10.23970/AHRQEPCCER251.

134. Spina N, Schoutens C, Martin BI, et al. Defining Instability in Degenerative Spondylolisthesis: Surgeon Views. Clin Spine Surg. 2019;32(10):E434-E9. doi: 10.1097/BSD.0000000000000874. PMID: 31490244.

Abbreviations and Acronyms

AHRQ	Agency for Healthcare Research and Quality
ALIF	anterior lumbar interbody fusion
aOR	adjusted odds ratio
BMI	body mass index
CARDS	clinical and radiographic degenerative spondylolisthesis
CI	confidence interval
CLBP	chronic low back pain
DBM	demineralized bone matrix
DLD	degenerative lumbar disease
DLS	degenerative lumbar spondylolisthesis
DS	degenerative scoliosis
EMG	electromyography
FDA	U.S. Food and Drug Administration
EPC	Evidence-based practice center
EQ-5D	EuroQol- 5 Dimension
EQ-5D-3L	EuroQol- 5 Dimension, 3 level version
ESI	epidural steroid injection
FJI	facet joint (intraarticular) injection
FSS	foraminal spinal stenosis
ICBG	Iliac crest bone graft
ICD-9	International Classification of Diseases, Ninth Revision
IONM	intraoperative neuromonitoring
IM	intramuscular
KQ	Key Question
LBP	low back pain
LIF	lumbar interbody fusion
LLIF	lateral lumbar interbody fusion
MEP	motor evoked potentials
MBB	medial branch block
MD	mean difference
NR	not reported
NRS	numeric rating scale
NRSI	nonrandomized study of intervention
NSAID	nonsteroidal anti-inflammatory drug
ODI	Oswestry Disability Index
OLIF	oblique lateral interbody fusion

PL	profile likelihood
PLF	posterolateral fusion
PLIF	posterior lumbar interbody fusion
QOL	quality of life
RCT	randomized controlled trial
RDQ	Roland Morris Disability Questionnaire
RFA	radiofrequency ablation
RR	relative risk
SD	standard deviation
SIP	Sickness Impact Profile Score
SSED	Summary of Safety and Effectiveness Data
SSEP	somatosensory evoked potentials
SF-36/12	Short Form 36 or 12 questionnaire
SMD	standardized mean difference
SOE	strength of evidence
TLIF	transforaminal interbody fusion
VAS	visual analogue scale