Washington State Health Care Authority

Surgery for Lumbar Radiculopathy/ Sciatica

Final evidence report

April 13, 2018

Health Technology Assessment Program (HTA)

Washington State Health Care Authority

PO Box 42712
Olympia, WA 98504-2712
(360) 725-5126
www.hca.wa.gov/hta
shtap@hca.wa.gov

Prepared by:

RTI International–University of North Carolina Evidence-based Practice Center
Research Triangle Park, NC 27709
www.rti.org

This evidence report is based on research conducted by the RTI-UNC Evidence-based Practice Center through a contract between RTI International and the State of Washington Health Care Authority (HCA). The findings and conclusions in this document are those of the authors, who are responsible for its contents. The findings and conclusions do not represent the views of the Washington HCA and no statement in this report should be construed as an official position of Washington HCA.

The information in this report is intended to help the State of Washington's independent Health Technology Clinical Committee make well-informed coverage determinations. 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).

None of the individuals involved in producing this report reported any financial or non-financial conflicts of interest regarding the topic presented in this report.

Acknowledgments

The following individuals from the RTI-UNC Evidence-based Practice Center contributed to this report:

Lead Investigator: Leila Kahwati, MD, MPH
Co-Investigator: Rachel Palmieri Weber, PhD
Clinical Advisor: Moe Lim, MD
Analysts: Rachel Clark, BA; Katrina Burson, RN, MS
Scientific Reviewer: Meera Viswanathan, PhD
Library/Document Preparation: Mark Howell, MLS; Loraine Monroe; Laura Small, BA

The following individuals independently peer-reviewed the Draft Report; these individuals did not receive any compensation in exchange for their review:

Timothy Carey, MD, MPH; Distinguished Professor of Medicine, School of Medicine, University of North Carolina
Roger Chou, MD; Professor of Medicine, School of Medicine, Oregon Health & Science University

Contents

List of Appendices

List of Figures

List of Tables

List of Abbreviations

AMD	Absolute mean difference
CI	Confidence interval
CPG	Clinical practice guideline
CQ	Cost question
EQ	Efficacy question
HTA	Health technology assessment
MID	Minimally important difference
NR	Not reported
NS	Not significant
RCT	Randomized controlled trial
RM	Repeated measures
RR	Relative risk ratio
SQ	Safety question
U.K.	United Kingdom
U.S.	United States

Executive Summary

Structured Abstract

Purpose: To conduct a health technology assessment (HTA) on the efficacy, safety, and cost of surgery for the treatment of symptomatic lumbar radiculopathy, also referred to as sciatica.

Data Sources: PubMed from January 2007 through November 9, 2017; clinical trial registry; government, payor, and clinical specialty organization websites; hand searches of bibliographies, relevant clinical practice guidelines, and systematic reviews to identify studies published prior to 2007.

Study Selection: Using a priori criteria, we selected English-language primary research studies published in any year that were conducted in very highly developed countries that enrolled adults with symptomatic lumbar radiculopathy and compared surgery for radiculopathy (primarily discectomy or microdiscectomy) to nonsurgical interventions, or that compared alternative surgical procedures, for example minimally-invasive procedures performed percutaneously or with endoscopy, compared with open procedures. We selected trials that reported efficacy outcomes (pain, functioning and disability, quality of life, neurological symptoms, return to work), safety outcomes (mortality, surgical morbidity, reoperations, persistent opioid use), or cost-analyses that reported costs or cost per quality-adjusted life year. We also selected relevant clinical practice guidelines (CPG) for quality appraisal and synthesis.

Data Extraction: One research team member extracted data and a second checked for accuracy. Two investigators independently assessed risk of bias of included primary research studies and conducted a quality assessment of included CPGs.

Data Synthesis: We included 25 primary research studies published between 1983 and 2017. Twenty-four randomized controlled trials (RCTs) provided findings related to efficacy and safety and 7 cost analyses provided findings related to cost-effectiveness. One RCT was rated as low risk of bias, 12 were rated as having some concerns for bias, and 12 were rated as high risk of bias.

Seven RCTs (total number of participants (N) = 1,158) compared microdiscectomy or discectomy to nonsurgical interventions. In these trials, surgery reduced leg pain by 6 to 26 points more than nonsurgical interventions as measured on a 0 to 100-point visual analog scale of patient-reported pain at up to 26 weeks follow-up; differences between groups did not persist at 1 year or later. The evidence was mixed for functioning and disability as measured by the Oswestry Disability Index, Roland-Morris Disability Questionnaire, and Short Form 36 (SF-36) Physical Functioning subscale in follow-up through 26 weeks, but no between-group differences were observed at 1 year or later. Surgery and nonsurgical interventions produced similar improvements in quality of life, neurologic symptoms, and return to work. No surgical deaths occurred in any study and surgical morbidity was infrequent. The incidence of reoperations among participants who underwent surgery ranged from 0% to 10%. Studies reported higher

quality-adjusted life years for participants who underwent surgery compared to nonsurgical interventions, but similar or higher costs. The average cost per quality-adjusted life year gained from a health care payor perspective ranged from $51,156 to $83,322 in 2010 United States (U.S.) dollars.

Thirteen RCTs (total N = 1,288) compared minimally-invasive surgical procedures to open microdiscectomy or discectomy. In general, minimally-invasive surgery produced similar improvements in pain, function/disability, quality of life, and neurologic symptoms, but resulted in return to work 4 to 15 weeks sooner. No surgical deaths occurred in any trials and with few exceptions, surgical morbidity was similar between groups. The incidence of reoperations across study groups ranged from 2% to 65%; 2 of the 10 trials reporting this outcome reported a statistically significant higher incidence of reoperation among participants who underwent minimally-invasive procedures but the other 8 RCTs reported a similar incidence between groups (pooled relative risk 1.37 [95% CI, 0.74 to 2.52]; 10 RCTs; 1,172 participants; I^2=60.8%). Three RCTs (total N = 282) compared microdiscectomy to discectomy and reported similar improvements pain at 26 weeks and later. Two RCTs (total N = 160) reported efficacy and safety outcomes of revision surgery for recurrent radiculopathy; findings were mixed.

We identified 14 clinical practice guidelines; the 4 higher quality clinical practice guidelines were in general agreement about recommending discectomy or microdiscectomy (and related decompressive procedures) as acceptable treatment for radiculopathy based on evidence that it improves outcomes in the short- to medium-term.

Limitations: The included RCTs were limited by methodologic designs that increased risk for bias, including extensive participant crossover, lack of participant and outcome assessor blinding, and inadequate randomization and allocation concealment in some studies. Many RCTs either did not report outcomes of interest or were underpowered, leading to imprecision for many effect estimates reported. This HTA was limited to English-language studies; it did not include observational studies or 'as-treated' analyses reported by some RCTs.

Conclusions: Most findings are based on a body of RCT evidence graded as low to very low certainty. Compared with nonsurgical interventions, surgery reduces pain and improves function more up to 26 weeks follow-up, but this difference does not persist at 1 year or longer. Minimally-invasive surgery, microdiscectomy, and discectomy are generally comparable with respect to efficacy and surgical morbidity; findings are mixed for reoperations. Surgery may be cost-effective when compared with nonsurgical interventions, depending on a decision maker's willingness to pay threshold, but the evidence is inconclusive about the cost-effectiveness of minimally-invasive surgery.

ES-1. Background

We designed this health technology assessment (HTA) to assist the State of Washington's independent Health Technology Clinical Committee with determining coverage for selected surgical interventions to treat symptomatic lumbar radiculopathy, also known as sciatica.

ES-1.1 Condition Description

Lumbar radiculopathy is a clinical syndrome characterized by radiating leg pain, with or without motor weakness, and sensory disturbances in a myotomal or dermatomal distribution. Lumbar radiculopathy is a heterogenous condition that may present acutely (as in the case of an acute disc herniation) or more insidiously (as in the case of spondylosis).[1-3] The objective of treatment for radiculopathy is symptom relief through nonsurgical management of symptoms, or surgical intervention to address the underlying causative mechanism, or both.

ES-1.2 Disease Burden

Estimates of the incidence and prevalence of lumbar radiculopathy vary widely because of variation in definitions and differences between self-reported and clinically assessed symptoms.[4] A 2008 systematic review of 23 studies assessing sciatica prevalence estimates reported a lifetime prevalence ranging from 3% to 43% (5 studies), a period prevalence over 1 year ranging from 2.2% to 34% (15 studies), and a point prevalence ranging from 1.6% to 13.4%. (4 studies).[4]

ES-1.3 Technology Description

The choice of surgical procedure to treat symptomatic lumbar radiculopathy in part depends on etiology and extent of nerve root compression. Discectomy or microdiscectomy may be used to address radiculopathy resulting from disc herniation, whereas laminectomy and other decompressive procedures (e.g., foraminotomy) may be used to address radiculopathy resulting from spondylolysis. ***Table ES-1*** lists the surgical procedures used to treat lumbar radiculopathy.

Table ES-1. Surgical interventions used to treat lumbar radiculopathy

Category of Intervention	Examples
Disc removal procedures	• Discectomy • Microdiscectomy
Decompression procedures	• Laminectomy • Microlaminectomy • Laminotomy • Foraminotomy
Minimally-invasive procedures	• Percutaneously or endoscopically performed discectomy, discoplasty, nucloetomy, nucleoplasty; performed manually or with automated devices, includes laser-assisted procedures.

Standard, open surgical interventions remove parts of the intervertebral disc, with or without additional decompression of spinal nerve root(s) through removal of parts of the bony vertebrae, facet joints (e.g., laminectomy or partial facetectomy) and/or other soft tissues impinging on the nerve root(s). Decompression and disc removal interventions are often performed with a microscope or other magnifying instrument ("micro" approaches). Such an approach makes it

possible to minimize the length of incision and area of dissection, thereby reducing the degree of structural alteration to surrounding tissues. Both standard open and "micro" approaches allow for direct visualization of the disc and surrounding structures.

In contrast to open procedures, interventions that use either an endoscopic approach to allow direct visualization of the surgical field and anatomy, or that use a percutaneous approach, which does not allow direct visualization of the disc and surrounding tissue, are also available. These procedures use mechanical (manual or automated), radiofrequency thermal, coblation (also known as plasma), or laser-assisted techniques for disc removal, destruction, or decompression. Although these procedures may vary, for this report, we refer to these procedures as 'minimally-invasive" surgical procedures. Minimally-invasive procedures allow for a smaller-incision and less tissue damage relative to open procedures.

ES-1.4 Regulatory Status

The U.S. Food and Drug Administration (FDA) regulates some surgical instruments and devices used in spine-related surgery. The FDA has cleared several devices for cutting, grinding, and aspiration of disc material during discectomy and for ablation and coagulation. FDA has also cleared laser instruments for incision, excision, resection, ablation, vaporization, and coagulation of tissue during surgical procedures including but not limited to discectomy. See the Full Report for a detailed description. All devices referred to in the Full Report were cleared by the FDA through the 510(k) process, which is based on evidence that the device is 'substantially equivalent' to a device that the FDA has already cleared or that was marketed before 1976. None were approved through the premarket approval process, which requires manufacturers to demonstrate that the device is safe and effective, a higher standard than the 510(k) process.

ES-1.5 Policy Context

Numerous surgical and nonsurgical approaches to the management of lumbar radiculopathy are routinely used within current clinical practice. In addition to standard surgical techniques (e.g., discectomy with or without laminectomy), minimally-invasive surgical techniques that use percutaneous, endoscopic, or laser-assisted approaches are now available. The State of Washington Health Care Authority selected surgery for lumbar radiculopathy as a topic for an HTA based on medium concerns for efficacy, medium concerns for safety, and high concerns for cost.

The State of Washington Health Care Authority provided data on the use of surgical procedures for the treatment of radiculopathy for the time period 2015 to 2017. This data is provided in ***Appendix A*** of the Full Report. Data is provided for Medicaid (fee for service and managed care organization), Department of Labor and Industries Workers' Compensation Program, and the Public Employee Benefit Board Uniform Medical Plan, including Medicare.

ES-2. Methods

This HTA includes two separate, but related components. The first component is a systematic review of primary research studies and the second component is a quality appraisal and synthesis of relevant clinical practice guidelines.

ES-2.1 Research Questions and Analytic Framework for Systematic Review of Primary Research Studies

We developed the following research questions and analytic framework (***Figure ES-1***) to guide the systematic review of primary research studies:

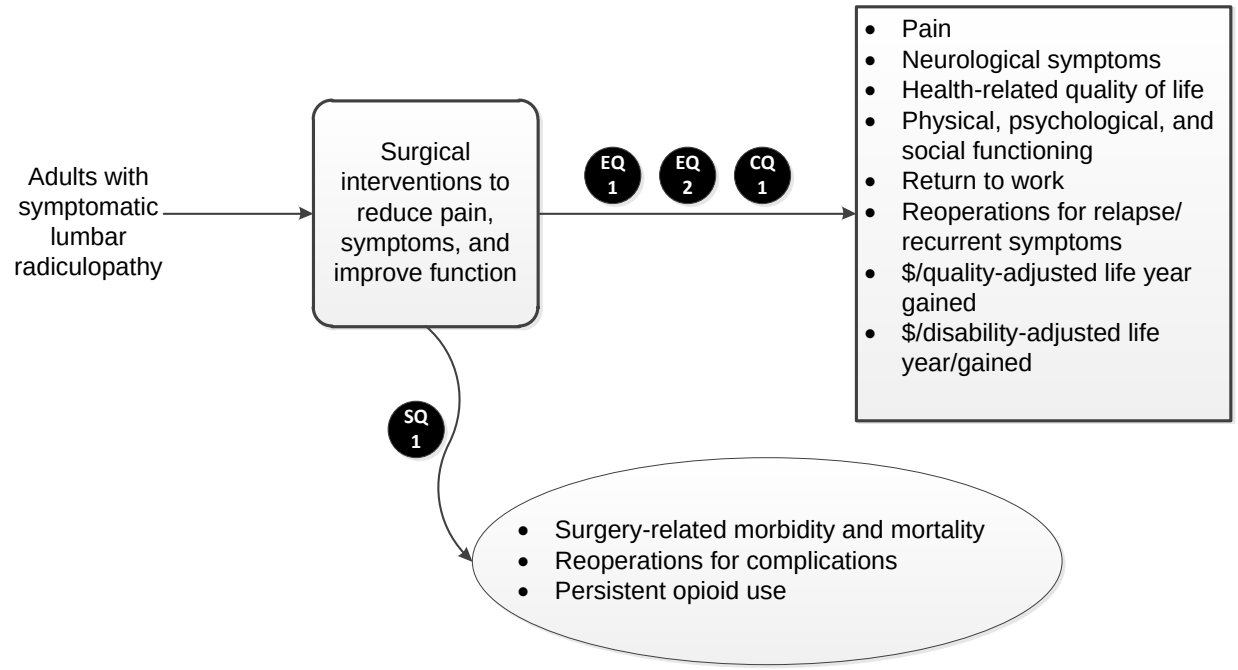
Efficacy Question 1 (EQ1). In adults with symptomatic lumbar radiculopathy, what is the effectiveness and comparative effectiveness of surgical interventions?

Efficacy Question 2 (EQ2). In adults with symptomatic lumbar radiculopathy, does effectiveness or comparative effectiveness of surgical interventions vary for patients who are not employed because of disability or patients who are undergoing recurrent surgery for relapse?

Safety Question 1 (SQ1). In adults with symptomatic lumbar radiculopathy, what are the adverse events associated with surgical interventions?

Cost Question 1 (CQ1). In adults with symptomatic lumbar radiculopathy, what is the cost-effectiveness of surgical interventions?

Figure ES-1. Analytic framework for HTA on surgery for lumbar radiculopathy



Abbreviations: CQ=cost question; EQ=efficacy question; SQ=safety question

ES-2.1.1 Data Sources and Search

The search strategy is detailed in ***Appendix B***. We searched MEDLINE® (via PubMed) from 2007, the Cochrane Library, a clinical trials registry (clinicaltrials.gov) and relevant government, payer, and health care professional society websites for relevant English-language articles. In addition, we reviewed the reference lists of relevant studies, practice guidelines, and other HTAs to identify any relevant articles not found through the electronic search and to identify studies published prior to 2007. We used medical subject headings (MeSH terms) and text words associated with the surgical interventions of interest combined with MeSH terms for radiculopathy and lumbar disc disease.

ES-2.1.2 Study Selection

Table ES-2 summarizes the study selection criteria related to the population, intervention, comparator, outcomes, time period, and setting that defined the scope of this HTA. We screened titles and abstracts and full-text articles based on these study selection criteria.

Table ES-2. Population, intervention, comparator, outcome, timing, setting and other study selection criteria for HTA on surgery for lumbar radiculopathy

Domain	Included	Excluded
Population	Adults age 18 years and over with symptomatic lumbar radiculopathy (i.e., sciatica) unrelated to infection, cancer, inflammatory, congenital, or traumatic etiologies. For studies of mixed populations, results must be stratified and reported separately for patients with lumbar radiculopathy.	Adults with: • Cervical or thoracic radiculopathy • Cauda equina syndrome • Neurogenic claudication or low back and leg symptoms related primarily to central spinal stenosis • Spondylolisthesis • Traumatic or congenital structural spinal abnormalities • Nonradicular leg or low back pain (i.e., discogenic or other nonspecific low back pain)
Intervention	Surgical interventions for the treatment of radiculopathy, for example: • Discectomy • Laminectomy, laminotomy • Foraminotomy • Nucleotomy Includes "micro" approaches to the above procedures, which involve smaller incisions and/or areas of dissection and/or use of microscope or loupe magnification. Minimally-invasive surgical procedures designed for treating radicular pain: percutaneous discectomy, discoplasty, nucleotomy, or nucleoplasty that are manual, automated, endoscopic, or laser-assisted, or use radiofrequency heat or coblation technology. Interventions involving combinations of the above interventions are eligible.	Surgical interventions primarily designed to treat neurogenic claudication and central spinal stenosis, spinal instability, or nonradicular low back pain, for example: • Spinal fusion • Arthroplasty • Artificial disc replacement • Interspinous process decompression (e.g., X-STOP® IPD System,[5] Coflex® Interlaminar Technology)[6] • Minimally-invasive lumbar decompression (mild® procedure)[7] • Other minimally-invasive procedures designed for treating discogenic (i.e., nonradicular) low back pain Epidural, spinal, or disc injections of enzymatic (e.g., chymopapain), chemical, or biologic (e.g., stem cells, mesenchymal cells) agents. Interventions involving combinations of procedures that include an above intervention are ineligible.

(continued)

Table ES-2.Population, intervention, comparator, outcome, timing, setting and other study selection criteria for HTA on surgery for lumbar radiculopathy (continued)

Domain	Included	Excluded
Comparator	*Placebo or no treatment comparators:* sham surgery, expectant management, no treatment *Active treatment comparators:* • Nonsurgical management (e.g., physical therapy, exercise, pharmacologic treatment of symptoms, spinal manipulation, chiropractic treatment, epidural steroid or pain injections, other noninvasive treatments) • Surgical interventions as listed under "intervention"	• No comparator • Chemonucleolysis Studies using "usual care" comparator groups will not be excluded but will be synthesized separately if no information was provided about the components of "usual care".
Outcomes	*Efficacy (at 4 weeks post-op or later):* • Pain • Physical functioning • Social functioning • Psychological/emotional distress • Health-related quality of life • Neurologic symptoms (e.g., weakness, sensory alteration) • Return to work *Safety:* • Surgery-related morbidity (e.g., venous thromboembolism, paralysis, new neurological symptoms, dural tear, epidural hematoma) • Surgical mortality (30 day) • Reoperations • Persistent opioid use *Costs and cost-effectiveness:* • Direct medical costs • Cost per quality-adjusted life year gained • Cost per disability-adjusted life year gained	Other outcomes not specifically listed as eligible. Pain, quality of life, and functional outcomes not measured using valid and reliable instruments or scales.[8,9]
Setting	Inpatient or outpatient settings in countries categorized as "very high" on United Nations Human Development Index.[10]	Studies conducted in countries not categorized as "very high" on United Nations Human Development index.[10]
Study Design and Risk of Bias Rating	*For all Efficacy and Safety Research Questions:* CCTs, RCTs, and SRs of CCTs or RCTs with similar scope as this HTA. For studies using surgical interventions as active comparators, only RCTs or SRs of RCTs will be included. *For Cost-Effectiveness Questions*: CEA, CUA, or CBA performed from the societal or payer perspective *For All Studies:* Intent-to-treat analyses. Studies with any risk of bias rating will be included, but high risk of bias studies will only be used in quantitative syntheses if fewer than 3 studies total are available.	Editorials, comments, letters, narrative reviews, case reports, case series, cohort studies, case-control studies. As treated or per-protocol analyses reported by RCTs.

(continued)

Table ES-2.Population, intervention, comparator, outcome, timing, setting and other study selection criteria for HTA on surgery for lumbar radiculopathy (continued)

Domain	Included	Excluded
Language and Time Period	English language, any time period.	Languages other than English.

Abbreviations: CCT = controlled clinical trial; CEA = cost-effectiveness analysis; CUA = cost-utility analysis; CBA = cost-benefit analysis; HTA = health technology assessment; RCT = randomized controlled trial; SR = systematic review.

ES-2.1.3 What is Excluded from This HTA

This review did not include studies published in languages other than English or conducted in countries that are not very highly developed based on the United Nations Human Development Index.[10] This review did not include studies conducted among children or adolescents. This review was designed to focus primarily on surgery to treat symptomatic radiculopathy, and we excluded studies evaluating surgical interventions performed primarily to manage central spinal canal stenosis (e.g., neurogenic claudication), spondylolisthesis, traumatic or congenital abnormalities. Further, this review did not cover surgical interventions for low back pain that was not radicular in nature (e.g., chronic discogenic pain). We refer readers to the State of Washington's 2015 Health Technology Assessment Final Evidence Report on Lumbar Fusion for Patients with Degenerative Disc Disease Uncomplicated by Comorbid Spinal Conditions available at the Program website.[11] Lastly, this review did not include observational study designs (e.g., case series, comparative cohort studies) or 'as treated' or 'per protocol' analyses from randomized controlled trials (RCTs) because these analyses have a high risk of bias relative to intent-to-treat analyses from RCTs.

ES-2.1.4 Data Abstraction and Risk of Bias Assessment

One team member extracted relevant study data into a structured abstraction form and the lead investigator checked it for accuracy. We used the Cochrane Risk of Bias (RoB 2.0) tool to assess the risk of bias for each included trial.[12] Domains assessed with this tool include: bias arising from randomization process, bias due to deviations from intended interventions, bias due to missing outcome data, bias in measurement of the outcome, and bias in selection of the reported result. Risk of bias was assessed as "high," "some concerns," or "low" at the study level unless different outcomes within a single study required different risk of bias ratings. We used the Quality of Health Economic Studies Instrument to assess the quality of included cost analyses.[13] Two team members conducted independent risk of bias or quality assessments on all included studies.

ES-2.1.5 Data Synthesis and Analysis

Study characteristics and results from intent-to treat analyses were qualitatively synthesized for each research question in tabular and narrative formats. We synthesized studies comparing the surgical interventions to nonsurgical interventions separately from studies comparing alternative surgical interventions. We summarized continuous outcome measures as absolute mean differences (AMD) between treatment groups wherever possible. We summarized categorical outcomes using differences in proportions between groups. When studies did not report statistical significance testing, we calculated it when possible. We identify all values that we calculated in the text and tables as "calculated" values. We transformed cost outcomes reported

in foreign currency to U.S. dollars based on the U.S. Department of Treasury mid-year exchange rate for the year reported by study authors and then used the chain-weighted consumer price index (CPI) to adjust to 2010 U.S. dollars (***Appendix C***).[14,15]

We required three or more studies with similar intervention and comparator with same outcome measure at approximately the same follow-up time point to calculate a pooled treatment effect. We considered outcomes reported at less than 12 weeks to be short-term, outcomes reported between 12 weeks up to 52 weeks as medium-term, and outcomes reported at 52 weeks or later as long-term. We estimated pooled effects using a random effects model with the 'metafor' package in R using the DerSimonian and Laird method.[16] We assessed statistical heterogeneity with the I^2 statistic,[17] and investigated heterogeneity qualitatively based on factors such as risk of bias, type of intervention, and type of comparator.

We graded the strength of evidence for each research question and outcome measure using a modification to GRADE, which assesses the strength of evidence based on domains relating to risk of bias, inconsistency, imprecision, indirectness, and other considerations, such as publication bias.[18] For each outcome measure, we rated the evidence for between-group differences in short-, medium-, and long-term outcomes separately when required because of differences in GRADE domains at different follow-up time periods. With GRADE, the strength of evidence can be graded as "very low," "low," "moderate," or "high", and this rating represents the overall certainty of the findings. ***Table ES-3*** defines these levels of certainty.[19] We modified the GRADE approach to allow for a rating of "insufficient" for single-study bodies of evidence with very serious concerns in one or more domains, or when we are unable to draw a conclusion about the treatment effect because of inconsistent findings.

Table ES-3. Strength of evidence grades and definitions[19]

Grade	Definition
High	**We are very confident that the estimate of effect lies close to the true effect for this outcome.** The body of evidence has few or no deficiencies. We believe that the findings are stable, that is, another study would not change the conclusions.
Moderate	**We are moderately confident that the estimate of effect lies close to the true effect for this outcome.** The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
Low	**We have limited confidence that the estimate of effect lies close to the true effect for this outcome.** The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
Very Low	**We have very limited confidence that the estimate of effect lies close to the true effect for this outcome.** The body of evidence has numerous major deficiencies. We believe that substantial additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
Insufficient	**We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome**. No evidence is available or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

When multiple outcome measures within a clinical domain were reported (e.g., pain, function), we only graded strength of evidence for the clinical measures with known validity and reliability

and that were reported by at least 2 studies. To draw overall conclusions about a clinical domain reporting multiple measures we considered all strength of evidence ratings within the domain.

ES-2.2 Clinical Practice Guideline Synthesis

In addition to the systematic evidence review portion of this HTA, we also identified relevant clinical practice guidelines and conducted a quality assessment of each guideline using the Appraisal of Guidelines for Research & Evaluation II (AGREE) instrument.[20,21] With this instrument, six domains are assessed and an overall score of between 1 (lowest possible) and 7 (highest possible) are assigned to reflect the overall quality of the guideline. We synthesized clinical practice guidelines in a tabular format.

ES-3. Results

ES-3.1 Literature Yield

We identified and screened 1,861 unique citations. We excluded 1,638 citations after title and abstract review. We reviewed the full-text of 223 articles and included a total of 25 studies reported in 39 articles published between 1983 and 2017. Twenty-two RCTs provided evidence on efficacy or comparative effectiveness (EQ1), two RCTs provided evidence on the effectiveness or comparative effectiveness of revision surgical interventions for relapse (EQ2), 24 RCTs provided evidence on safety (SQ1), and seven studies (six RCTs and one cost-effectiveness analysis) provided evidence on costs or cost-effectiveness (CQ1). The Full Report describes individual study and population characteristics and findings for all included studies (***Appendix D***), the list of articles we screened but excluded at the full-text stage (***Appendix E***), and risk of bias assessments for included studies (***Appendix F)***.

ES-3.2 Efficacy

ES-3.2.1 Efficacy Question 1

In adults with symptomatic lumbar radiculopathy, what is the effectiveness and comparative effectiveness of surgical interventions?

We included 22 RCTs. Four were conducted in the U.S.;[22] the rest were conducted in Canada (k=1),[23] Taiwan (k=1),[24] Japan (k=1)[25] or various European countries (k=15).[26-40] Seven RCTs provided evidence for the efficacy of surgery compared with nonsurgical treatment;[22,23,26,32,33,37,41] 15 RCTs provided evidence for the comparative effectiveness of alternative surgical interventions.[24,25,27-31,35,36,38-40,42-44] The interventions and comparators evaluated are summarized in ***Table ES-4***. The total number of participants randomized ranged from as few as 21 to as many as 501. The mean age of participants generally ranged from mid-30s to mid-40s. All studies enrolled both men and women with signs or symptoms of lumbar radiculopathy and confirmatory imaging, usually computerized tomography (CT) or magnetic resonance imaging (MRI).

Across the included RCTs, studies reported outcomes at various time points spanning from immediately postoperative to up to 10 years postoperative; no single efficacy measure was used

consistently across all included studies. We rated one RCT as low risk of bias,[40] 10 RCTs as some concerns for bias,[23,24,27-29,35,36,38,39,43] and 10 RCTs as high risk of bias.[22,25,26,30-34,37,42] One RCT was rated as some concerns for bias for 6 week outcomes and high risk of bias for outcomes at 12 weeks and later.[41] All but one study[40] did not blind participants to treatment allocation, and since nearly all studies relied on patient-reported outcomes, most studies had at least some concerns for bias since knowledge of the assigned treatment may impact such outcomes. Studies rated as high risk of bias generally used inadequate randomization and allocation concealment (e.g., use of even /odd [37]) or had moderate to extensive levels of crossover between treatment arms. For example, in the Weinstein et al. RCT (Spine Patient Outcomes Research Trial [SPORT]), 46.1% of participants allocated to surgery did not receive surgery by 26 weeks follow-up, and 36.3% of participants allocated to conservative management received surgery.[22]

Table ES-4. Surgical and comparator interventions used among 22 included studies for EQ1

	Surgical Intervention[a]	Comparator Intervention[b]
Efficacy RCTs (k=7)	Microdiscectomy	Spinal manipulation *(McMorland 2010)*[23]; Physiotherapy *(Osterman 2003)*[33]
	Percutaneous disc decompression with coblation technology *(Gerszten 2003)*[41]	Epidural steroid injection
	Percutaneous disc decompression *(Erginousakis 2011)*[37] Discectomy *(Weber 1983)*[26,32] Discectomy/microdiscectomy *(Weinstein 2006 [SPORT])*[22] Microdiscectomy *(Peul 2007)*[32]	Conservative management
Comparative effectiveness RCTs (k=15)	Tubular/trocar discectomy *(Arts 2011, Ryang 2008)*[30,40] Automated percutaneous lumbar discectomy *(Chatterjee 1995)*[38] Percutaneous endoscopic discectomy *(Mayer 1993)*[34] Endoscopic discectomy *(Ruetten 2008)*[31] Microendoscopic discectomy *(Sasaoka 2006, Teli 2010[c])*[25,29] Sequestrectomy *(Thome 2005)*[28] Percutaneous laser disc decompression *(Brouwer 2015)*[39] Microscopically-assisted percutaneous nucleotomy *(Franke 2009)*[36]	Microdiscectomy
	Automated/endoscopic percutaneous discectomy *(Haines 2002)*[42] Video-assisted arthroscopic microdiscectomy *(Hermantin 1999)*[43] Microendoscopic discectomy *(Huang 2005, Teli 2010[c])*[24,29]	Discectomy
	Microdiscectomy *(Henriksen 1996, Teli 2010[c], Tullberg 1993)*[27,29,35]	Discectomy

[a] In the Appendix D Evidence Tables, these interventions are considered the surgical group and are denoted as SG1.

[b] In the Appendix D Evidence Tables, these interventions are considered the comparator groups; nonsurgical comparator groups are denoted as NS1 and surgical comparators are denoted as SG2 or SG3.

[c] This study was a three-arm RCT that allocated participants to microendoscopic discectomy, microdiscectomy, and standard discectomy; thus, it contributes to three comparisons of interest for this HTA.

Abbreviations: k = number of studies; RCT = randomized controlled trial; SPORT = Spine Patient Outcomes Research Trial

A. Surgery compared with nonsurgical interventions

Seven RCTs that compared surgery to nonsurgical interventions reported at least one efficacy outcome.[22,23,26,32,33,37,41] Five were rated as high risk of bias,[22,26,32,33,37] one was rated as some concerns for bias,[23] and one was rated as high risk of bias for outcomes later than 12 weeks and some concerns for bias for outcomes less than 12 weeks.[41] The nonsurgical interventions to which surgery was compared included medications, physical therapy, patient education/counseling, spinal manipulation, and epidural steroid injection. We were unable to

conduct quantitative synthesis for any outcomes because of outcome measure and reporting heterogeneity or because some studies did not report measures of variance needed to conduct a meta-analysis. ***Table ES-5*** summarizes findings and strength of evidence ratings.

Table ES-5. Summary of efficacy outcome findings and strength of evidence ratings comparing surgery to nonsurgical interventions in persons with symptomatic lumbar radiculopathy (EQ1)

Outcomes Length of follow-up	No. RCTs (k) No. participants (N)	Summary of effect ⇨ favors surgery; ⇦ favors nonsurgical intervention; ⇔ no difference, ? unable to determine		Certainty[a]
Pain[b]	k=7; N=1,158			
Up to 26 weeks	k=5; N=970	⇨	Surgery reduced pain more than nonsurgical interventions by an amount considered a minimally-important difference for most measures reported.	⊕⊕◯◯ LOW
Between 1 and 8 years	k=3; N=840	⇔	Surgery and nonsurgical interventions decreased pain by about the same amount.	⊕◯◯◯ VERY LOW
Function/Disability[c]	k=5; N=970			
Up to 26 weeks	k=5; N=970	?	Findings were mixed across studies. Surgery generally improved function and reduced disability more than nonsurgical interventions but the magnitude of differences were not consistently above the minimally-important difference for the measures reported or were not statistically significant between-groups.	◯◯◯◯ INSUFFICIENT
Between 1 and 8 years	k=3; N=840	⇔	Surgery and nonsurgical interventions improve function and reduce disability by about the same amount.	⊕◯◯◯ VERY LOW
Quality of life[d]	k=2; N=96			
Up to 12 weeks	k=2; N=96	⇔	Surgery and nonsurgical interventions improve quality of life by about the same amount.	⊕◯◯◯ VERY LOW
Between 52 weeks and 2 years	k=1; N=56	?	No significant between-group differences observed in quality of life measure at 52 weeks or 2 years.	◯◯◯◯ INSUFFICIENT
Neurologic symptoms[e] 6 weeks to 52 weeks	k=2; N=146	⇔	Surgery and nonsurgical interventions improve neurologic symptoms by about the same amount.	⊕◯◯◯ VERY LOW
Return to work[f] Between 12 weeks and 10 years	k=5; N=835	⇔	Return to work outcomes are similar for surgery and nonsurgical interventions.	⊕◯◯◯ VERY LOW

[a] We assessed certainty using a modified GRADE approach, which assesses the evidence base for each outcome measure based on risk of bias, inconsistency, imprecision, indirectness, and other considerations; certainty is rated as "insufficient", "very low", "low", "moderate", or "high". For domains with more than 1 measure reported (e.g., pain, function), we rated each measure separately but this summary table reflects our overall assessment across measures reported in more than one study. See Full Report for individual outcome measure strength of evidence ratings.

[b] As measured by visual analog scale (VAS 100 mm) for leg pain and for back pain, SF-36 Bodily Pain subscale, and Sciatica Index.

[c] As measured by Oswestry Disability Index, Roland-Morris Disability Questionnaire, and SF-36 Physical Functioning subscale.

[d] As measured by SF-36 and 15D health-related quality of life measures.

[e] As measured by physical exam or patient-report.

[f] As measured by actual return to work, self-reported ability to work, receipt of disability benefits, or other related measures.

Abbreviations: k = number of studies; N = number of participants; RCT = randomized controlled trial

Pain

All seven RCTs reported as least one pain outcome. Pain outcomes reported included the Visual Analog Scale (VAS) 100 mm or 10 cm for leg pain, the VAS 100 mm or 10 cm for back pain,

the Short Form 36 (SF-36) Bodily Pain subscale, the Sciatica index, the McGill Pain Questionnaire, and the Abderdeen back pain scale. A few studies also reported the frequency and proportion of participants reporting reduced pain, no pain, or relief from pain. Peul et al.[32] Osterman et al.[33] and Gerszten et al.[41] reported decreased VAS scores (i.e., improvement) for leg pain in participants allocated to both the surgical treatment (range 41 to 57 point decrease) and nonsurgical comparator (range 20 to 36.5 decrease) from baseline through short-term (6- and 8-week follow-up); these scores decreased by 6 to 26 points more at short- and medium-term follow-up among participants allocated to surgery compared with those allocated to nonsurgical intervention.[32,33,41] VAS scores for back pain followed a similar pattern. Of the two RCTs reporting VAS leg and back pain scores between 52 weeks and 5 years (Peul et al.[32] and Osterman et al.[33]) between-group differences were generally smaller than the minimally important difference (MID) and not statistically significant suggesting no difference in treatment effect. The four RCTs reporting pain outcomes using the SF-36 reported mixed findings at various follow-up time points.[22,23,32,41] Weinstein et al. [SPORT][22] and Peul et al.[32] reported pain outcomes using the Sciatic Index; scores decreased by 2.1 to 4.0 points more among participants allocated to surgery in the short- and medium-term. Scores reported at timepoints between 52-weeks and 8 years observed attenuation in the between-group difference such that these differences were less than the MID, though these differences remained statistically significant at multiple long-term follow-up timepoints in Weinstein et al. [SPORT].[22]

Summary: We concluded with low certainty that surgery improves pain more in the short- to medium-term compared with nonsurgical treatment, but by one year the surgery and nonsurgical interventions are similar with respect to improving pain (very low certainty).

Function and Disability

Five RCTs reported various measures of physical, mental, emotional, and social functioning or disability.[22,23,32,33,41] The most commonly reported functional outcomes were the Oswestry Disability Index, the Roland-Morris Disability Questionnaire, and the Physical Functioning subscale of the SF-36. Across studies, function improved in both participants allocated to surgery and in participants allocated to nonsurgical treatment; however, between-group differences varied depending on follow-up time point.

Gerszten et al.,[41] Osterman et al.,[33] and Weinstein et al. [SPORT][22] observed between-group differences on the Oswestry Disability Index that favored surgery (range of AMDs -4.7 to -10) through 26 weeks. Two of these trials also reported long-term follow-up for this measure but the between-groups differences did not persist.[33,22]

Peul et al,[32] observed larger short-term (through 8 weeks) improvements in the Roland-Morris Disability Questionnaire and Physical Functioning subscale of the SF-36 among participants allocated to surgery; these differences did not persist in the medium- and long-term. McMorland et al.[23] observed no between-group differences at 6 weeks and 12 weeks on this measure. Lastly, Weinstein et al. [SPORT] [22] observed larger improvements among participants allocated to surgery at 12 weeks on one measure of functional status (Oswestry Disability Index) as noted previously, but this difference was not observed by the other measure of functioning (SF-36 Physical Functioning subscale) reported by this trial.[22] Peul et al,[32] and Weinstein et al. [SPORT]

[22] also reported long-term outcomes (through 5 and 8 years) with the SF-36 Physical Functioning subscale and observed no between-group differences.

Summary: We concluded the evidence was insufficient to draw a conclusion about treatment effect on functioning/disability for short- and medium-term outcomes, but concluded with very low certainty that surgery and nonsurgical treatments are similar with respect to long-term improvements in function/disability.

Quality of life

McMorland et al.[23] and Osterman et al.,[33] reported health-related quality of life (QOL) outcomes. These studies reported outcomes using the total SF-36 score[23] (sum of all normed subscales, possible range 0 to 800) and the 15D QOL measure (range 0 to 1.0).[33] In both studies, QOL improved from baseline to follow-up in the surgery and nonsurgical groups; no significant differences between groups were observed by either study.

Summary: We concluded with very low certainty that surgery and nonsurgical treatments are similar with respect to improvements in QOL in the short- and medium-term, but the evidence was insufficient to draw a conclusion regarding long-term impact on QOL because of a single-study body of evidence.

Neurological symptoms

Osterman et al.,[33] and Gerszten et al.,[41] reported outcomes related to neurological symptoms, specifically sensory or motor deficits.[33,41] With few exceptions, no between-group differences were observed. For example, Osterman et al. reported a similar proportion of participants with muscle weakness among those allocated to microdiscectomy compared with those allocated to physiotherapy at 6 weeks (53.8% vs. 46.2%), 12 weeks (42.3% vs. 46.2%), and 52 weeks (28.6% vs. 30.0%).[33]

Summary: We concluded with very low certainty that surgery and nonsurgical treatments are similar with respect to improvements in neurological symptoms at all follow-up time points.

Return to work

Five RCTs reported various outcomes related to "return to work."[22,26,33,37,41] Some measures captured actual return to work, whereas others reflected somewhat indirect measures, such as self-reported ability to work, receipt of disability benefits, or pain affecting occupational status. Except for Erginousakis et al.[37]; no between-group differences in return to work outcomes were observed. For example, Weinstein et al. [SPORT][22] reported a difference of -2.2% (95% confidence interval (CI), -10.6% to 6.2%) in the proportion of participants working full time at 2 years follow-up between the participants allocated to surgery compared with the participants allocated to conservative management.

Summary: We concluded with very low certainty that surgery and nonsurgical treatments are similar with respect to "return to work" outcomes.

Other efficacy outcomes

Four RCTs reported other efficacy outcomes related to perceived recovery, overall time to recovery, overall result, and patient satisfaction with symptoms.[22,26,32,33] These outcomes were consistent with previously reported efficacy outcomes that suggest more favorable outcomes for participants who are allocated to surgery in the short- or medium-term. Peul et al.[32] reported a significant difference in median time to recovery (4.0 weeks [95% CI, 3.7 to 4.4] vs. 12.1 weeks [95% CI, 9.5 to 14.9]; AMD not reported (NR), P< 0.001) among participants allocated to microdiscectomy compared with participants allocated to conservative management. Weber et al.[26] reported a significant difference in proportion of category of result (good, fair, poor, bad) at 52 weeks (P=0.0015) but not at 4 years or 10 years. Weinstein et al. [SPORT][22] reported a higher proportion of patients reporting being satisfied with symptoms at 12 weeks (AMD 11.3% [95% CI 1.6% to 20.9%]) but no significant difference at single follow-up time points through 8 years; however, the repeated measure for this outcome was significant (P=0.013).

Summary: We did not use these additional efficacy outcomes in our strength of evidence ratings because of heterogeneity in outcome definition; however, the findings are consistent with the pattern of previously reported efficacy outcomes.

B. Minimally-invasive surgery compared with microdiscectomy or discectomy

Ten RCTs comparing minimally-invasive surgical interventions (tubular/trocar discectomy,[30,40] percutaneous endoscopic discectomy,[34] endoscopic interlaminar or transforaminal discectomy,[31] microendoscopic discectomy,[25,29] sequestrectomy,[28] percutaneous laser disc decompression,[45] microscopically assisted percutaneous nucleotomy,[36] and video-assisted microdiscectomy[43]) reported at least one efficacy outcome. Four were rated as high risk of bias,[25,30,31,34] five were rated as having some concerns for bias,[28,29,36,39,43] and one was rated as low risk for bias.[40] ***Table ES-6*** summarizes findings and strength of evidence ratings.

Table ES-6. Summary of efficacy outcome findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery in persons with symptomatic lumbar radiculopathy (EQ1)

Outcomes Length of follow-up	No. RCTs (k) No. participants (N)	Summary of effect ⇨ favors minimally-invasive surgery; ⇦ favors standard surgery; ⇔ no difference		Certainty[a]
Pain[b]	k=10; N=1,155			
Up to 26 weeks	k=5; N=869	⇔	Improvements in pain are similar between minimally-invasive surgery and standard surgery. Pooled between-group mean difference in VAS 100 mm for leg pain at 12 w to 26w was 0.3 (95% CI, -2.2 to 2.9, 4 RCTs, 642 participants, I^2=0%).	⊕⊕⊕◯ MODERATE
52 weeks to 2 years	k=5; N=869	⇔	Improvements in pain are similar between minimally-invasive surgery and standard surgery. Pooled between-group mean difference in VAS 100 mm for leg pain at 52w to 1.5y was 1.6 (95% CI, -1.5 to 4.6, 4 RCTs, 640 participants, I^2=28.1%).	⊕⊕◯◯ LOW

(continued)

Table ES-6. Summary of efficacy outcome findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery in persons with symptomatic lumbar radiculopathy (EQ1) (continued)

Outcomes Length of follow-up	No. RCTs (k) No. participants (N)	Summary of effect ⇨ favors minimally-invasive surgery; ⇦ favors standard surgery; ⇔ no difference		Certainty[a]
Function/Disability[c]	k=8; N=1,063			
Up to 26 weeks	k=6; N=903	⇔	Improvements in function are similar between minimally-invasive surgery and standard surgery.	⊕⊕◯◯ LOW to ⊕◯◯◯ VERY LOW
52 weeks to 2 years	k=8; N=1,063	⇔	Improvements in function are similar between minimally-invasive surgery and standard surgery.	⊕◯◯◯ VERY LOW
Quality of life[d] 12 weeks to 3 years	k=3; N=286	⇔	Quality of life improvements are similar between minimally-invasive surgery and standard surgery.	⊕◯◯◯ VERY LOW
Neurologic symptoms[e] 12 weeks to 2 years	k=6; N=602	⇔	Neurologic symptom improvements are similar between minimally-invasive surgery and standard surgery.	⊕◯◯◯ VERY LOW
Return to work[f]	K=6; N= 555	⇨	Minimally-invasive surgery reduces the duration of postoperative work disability by a range of 4 weeks to 15 weeks compared to standard surgery.	⊕◯◯◯ VERY LOW

[a] We assessed certainty using a modified GRADE approach, which assesses the evidence base for each outcome measure based on risk of bias, inconsistency, imprecision, indirectness, and other considerations; certainty is rated as "insufficient", "very low", "low", "moderate", or "high". For domains with more than 1 measure reported (e.g., pain, function), we rated each measure separately but this summary table reflects our overall assessment across measures reported in more than one study. See Full Report for individual outcome measure strength of evidence ratings.

[b] Only studies that assessed pain as measured by visual analog scale (VAS) for leg pain and for back pain, SF-36 Bodily Pain subscale, and Sciatica Index were included in strength of evidence ratings.

[c] Only studies that assessed pain as measured by Oswestry Disability Index, Roland-Morris Disability Questionnaire, and SF-36 Physical Functioning subscale, and Prolo scale were included in strength of evidence ratings.

[d] As measured by the physical health and mental health component summary scores of the SF-36.

[e] As measured by physical exam or patient-report.

[f] As measured by mean duration of post-operative disability or self-reported "work impairment".

Abbreviations: CI = confidence interval; k = number of studies; N = number of participants; RCT = randomized controlled trial; VAS = visual analog scale; w = week(s); y = year(s);

Pain

All 10 RCTs reported at least one pain outcome. For VAS 100 mm leg pain, two RCTs [39,40] reported outcomes at 4 weeks and 8 weeks follow-up and five RCTs [28,29,31,39,40] reported medium-term outcomes. The range of decreases in VAS leg pain scores at the earliest follow-up in each study (4 weeks to 26 weeks) was 42.5 to 69 points among participants allocated to minimally-invasive surgery and 29.8 to 62 points for standard surgery. Between-group differences in short-term outcomes were not significant and with one exception (Brouwer et al.[39] VAS 100 mm back pain at 26 weeks, favored microdiscectomy), between-group differences in medium-term outcomes were also not significant. For example, Arts et al.[40] reported between group differences of 4.5 (95% CI, -0.3 to 9.3) at 4 weeks and 4.5 (95% CI, -0.4 to 9.3) at 8 weeks for VAS 100 mm leg pain scores. The pooled between-group difference in VAS 100 mm leg pain at 12 to 26 weeks was 0.3 (95% CI, -2.2 to 2.9; 4 RCTs; 642 participants; I^2=0%; ***Appendix G***, ***Figure G-1***). The pooled between-group difference in VAS 100 mm back pain scores at 12 to 26 weeks was 1.3 (95% CI, -3.5 to 6.2; 4 RCTs; 642 participants; I^2=61.7%, ***Appendix G***, ***Figure G-2***).

Five RCTs[28,29,31,39,40] also reported long-term outcomes with VAS 100 mm for leg and back pain at 52 weeks and 2 years and the reductions observed in the short and medium-term generally persisted. The pooled between-group difference in VAS 100 mm leg pain at 52 weeks to 1.5 years was 1.6 (95% CI, -1.5 to 4.6; 4 RCTs; 640 participants; I^2=28.1%) as shown in ***Appendix G, Figure G-1*** and the pooled between-group difference in VAS 100 mm back pain was 1.5 (95% CI, -3.0 to 5.9; 4 RCTs; 640 participants; I^2=57.6%) as shown in ***Appendix G, Figure G-2***. Pooled findings at 2 years also observed no significant between-group differences in leg or back pain scores.

Four RCTs[28,30,39,40] reported pain outcomes using the Bodily Pain subscale of the SF-36. In all studies, pain scores improved from baseline to short-, medium-, and long-term follow-up among participants allocated to both surgical groups. Increases in scores at the earliest follow-up in each study (range 4 weeks to 26 weeks) ranged from 6.7 to 46.5 among participants allocated to minimally-invasive surgery and from 5.9 to 51.1 among participants allocated to standard surgery. With one exception (Brouwer et al. [39] 26 weeks, favored microdiscectomy), no between-group differences were observed at any follow-up time. The pooled mean difference in SF-36 Bodily Pain scores at 12 to 26 weeks was -3.0 (95 % CI, -12.8 to 6.8; 3 RCTs; 500 participants; I^2=75.4%; ***Appendix G, Figure G-3***).

Arts et al. [39] and Brower et al. [40] reported pain outcomes using the Sciatica Index at follow-up time points from 4 weeks to 2 years.[39,40] Scores decreased from baseline to follow-up in both surgical groups, and no significant between-group differences were observed at any single time point or in repeated measures analysis.

Summary: We concluded that improvements in pain were similar between minimally-invasive surgery and standard surgery in the short- and medium-term (moderate certainty) and in the long-term (low certainty).

Function and disability

Eight RCTs reported at least one outcome related to functioning or disability[28-31,36,39,40,42] Functional outcomes reported included the Oswestry Disability Index, the Roland-Morris Disability Questionnaire, the Physical Functioning subscale of the SF-36, the Prolo Scale, and various other subscales of the SF-36. With few exceptions, between-group differences were minimal.

Ruetten et al.[31] Teli et al.,[29] Franke et al.,[36] and Ryang et al.,[30] reported decreases on the Oswestry Disability Index of 28 to 53 points among participants allocated to minimally-invasive surgery and 29 to 47 points among participants allocated to standard surgery at the earliest follow-up (12 weeks to 26 weeks). Improvements persisted in the long-term and no between-groups differences were observed.

Arts et al.,[40] Brouwer et al.,[39] and Haines et al.,[42] reported decreases on the Roland-Morris Disability Questionnaire of 4.9 to 9.7 points among participants allocated to minimally-invasive surgery and 2.3 to 10.6 points among participants allocated to standard surgery at the earliest follow-up (4 weeks to 26 weeks). Between group differences were not observed at 26 weeks by

Haines et al.[42] Between group differences were significant at 4 weeks in Brouwer et al. (AMD -2.5 [95% CI, -4.7 to-0.2]) but not at 8 weeks or at any other follow-up time through 2 years. The only significant between-group differences reported by Arts et al. through 5 years of follow-up was at 52 weeks (AMD 1.3 [95% CI 0.03 to 2.6]); however, this difference was below the minimally-important difference.

Arts et al.,[40], Haines et al.,[42] Ryang et al.,[30] , Thome et al.,[28] and Brouwer et al.,[39] reported increases on the SF-36 Physical Functioning subscale of 27.2 to 41.8 points among participants allocated to minimally-invasive surgery and 2.6 to 51.9 points among participants allocated to standard surgery at earliest follow-up (4 weeks to 26 weeks). Similar to the Roland-Morris Disability outcome previously reported, between-group differences favoring minimally-invasive surgery at 4 weeks were observed by one study,[39] and this difference did not persist at 8 weeks or any subsequent follow-up time point. The pooled between-group mean difference in SF-36 Physical Functioning subscale was -2.4 (95 % CI, -6.1 to 1.2; 4 RCTs; 527 participants; I^2=0.0%; ***Appendix G, Figure G-4***) at 12 weeks to 26 weeks. Some between-group differences were observed at 52 weeks and 2 years; but the findings were mixed.

Summary: We concluded that improvements in function/disability were similar between minimally-invasive surgery and standard surgery in the short- and medium-term (very low to low certainty) and in the long-term (very low certainty).

Quality of life

Three RCTs reported health-related QOL using physical health component summary (PCS) and mental health component summary (MCS) scores over 12 weeks to 2.8 years.[28-30] In all studies, quality of life as measured by both component scores improved over time in both intervention groups, and with one exception, no statistically significant between-group differences were observed. Ryang et al. observed a significant difference in the MCS at the 1.3 year follow-up; participants allocated to microdiscectomy had a higher score (mean 51.9 [standard deviation (SD) 7.8]) compared with participants allocated to minimal access trocar microdiscectomy (mean 44.0 [SD 13.2], P=0.03), but it is not clear whether this comparison adjusted for small differences in baseline scores.[30]

Summary: We concluded with very low certainty that improvements in quality of life were similar for minimally-invasive surgery and standard surgery at medium- and long-term follow-up.

Neurological symptoms

Six RCTs comparing minimally-invasive surgery to microdiscectomy[28,30,31,34,36] and discectomy[43] reported outcomes related to neurological symptoms. Findings were not reported by group in one RCT,[36] the remaining five studies observed no between-group differences. Three RCTs reported no statistical difference in neurological symptoms between intervention groups.[28,30,31] We calculated no significant between-group differences in the other two RCTs.[34,43] For example, Ryang et al. reported no difference in the proportion of participants with sensory deficits (40% vs. 43%) in participants allocated to minimal access trocar discectomy compared with participants allocated to microdiscectomy, respectively, over an average of 1.3 years

follow-up (P=0.31).[30] Similar findings were observed for the proportion with motor deficits (27% vs. 23%, P=0.86).

Summary: We concluded with very low certainty that improvements in neurological symptoms were similar for minimally-invasive surgery and standard surgery at medium- and long-term follow-up.

Return to work

Six RCTs reported various outcomes related to "return to work", though in some studies this outcome was not reported by group.[28,31,34,36,38,43] Of the four RCTs that reported between-group differences, three RCTs[31,34,43] suggest that participants allocated to minimally-invasive surgery return to work sooner than participants allocated to standard surgery as measured by weeks of postoperative disability. The range of this difference was 3.4 weeks to 15.2 weeks. The remaining RCT[28] reported no significant between-group differences; however; this study used a multi-level categorical measure of work impairment, which may be measuring a related, but different construct compared to the other three RCTs. In this study, Thome et al. reported specific categories of impairment of work at 12 to 26 weeks and at 2 years.[28] Thirty-one percent of participants allocated to sequestrectomy reported that their work impairment was "much better" at 12 to 26 weeks compared with 33% of participants allocated to microdiscectomy. At 2 years, the proportions were 37% and 31%, respectively. The proportion of participants endorsing various categories of work impairment were not significantly different between groups (P=0.415 at 12 to 26 weeks, P=0.112 at 2 years).

Summary: We concluded with very low certainty that participants receiving minimally-invasive surgery return to work sooner than participants who receive standard surgery. However, this finding is associated with numerous limitations detailed in the Full Report.

Other efficacy outcomes

Ten RCTs reported other efficacy outcomes, related to perceived recovery, overall time to recovery, overall result, and patient satisfaction with symptoms. With few exceptions, most observed no significant differences between groups. We did not use these outcomes in our strength of evidence ratings because of heterogeneity in outcome definition.

C. Microdiscectomy compared to discectomy

Three RCTs comparing microdiscectomy to discectomy reported efficacy outcomes.[27,29,35] All three were rated as some concerns for bias. ***Table ES-7*** summarizes findings and strength of evidence ratings.

Pain

Three trials reported pain outcomes using a VAS 10 cm scale at 4 and 6 weeks,[35] at 52 weeks,[27] and at 26 weeks, 52 weeks, and 2 years.[29] In Henriksen et al.,[35] actual VAS values were not reported but no differences were reported between groups for both VAS leg pain and VAS back pain at 4 weeks and at 6 weeks follow-up. Tullberg et al.[27] reported a mean baseline VAS 10 cm leg pain score of 7.0 (SD NR) among participants allocated to microdiscectomy and 7.0 (SD NR) among participants allocated to discectomy. The mean scores at 52 weeks were 2.1 (SD NR) and

2.3 (SD NR), respectively (AMDs and P value NR). Teli et al.[29] reported VAS 10 cm leg pain score of 8 (SD 1) at baseline decreasing to 2 (1 SD) at 26 weeks, 1 (SD 1) at 52 weeks, and 2 (SD 1) at 2 years in both surgical groups (P=0.73 for between-group differences).

Summary: We concluded with very low certainty that improvements in pain are similar for microdiscectomy compared with discectomy in the medium- and long-term but the evidence is insufficient for assessing outcomes in the short-term because of a single-study body of evidence.

Table ES-7. Summary of efficacy outcome findings and strength of evidence ratings comparing microdiscectomy to discectomy in persons with symptomatic lumbar radiculopathy (EQ1)

Outcomes Length of follow-up	No. RCTs (k) No. participants (N)	Summary of effect ⇨ favors microdiscectomy; ⇦ favors discectomy; ⇔ no difference; ? unable to determine		Certainty[a]
Pain[b]	k=3; N=282			
Up to 6 weeks	k=1; N=80	?	Pain decreased by similar amounts in both surgical groups.	◯◯◯◯ INSUFFICIENT
26 weeks to 2 years	k=2; N=202	⇔	Pain decreased by similar amounts in both surgical groups.	⊕◯◯◯ VERY LOW
Function/Disability[c] 26 weeks to 2 years	k=1; N=142	?	Function improved by similar amounts in both surgical groups.	◯◯◯◯ INSUFFICIENT
Quality of life[d] 26 weeks to 2 years	k=1; N=142	?	Health-related quality of life improved by similar amounts in both surgical groups.	◯◯◯◯ INSUFFICIENT
Neurologic symptoms[e]	k=0	?	No studies.	◯◯◯◯ INSUFFICIENT
Return to work[e]	k=1; N=60	?	Both surgeries result in similar duration of postoperative work disability; 10.4 weeks for microdiscectomy compared with 10.1 weeks for discectomy.	◯◯◯◯ INSUFFICIENT

[a] We assessed certainty using a modified GRADE approach, which assesses the evidence base for each outcome measure based on risk of bias, inconsistency, imprecision, indirectness, and other considerations; certainty is rated as "insufficient", "very low", "low", "moderate", or "high". For domains with more than 1 measure reported (e.g., pain, function), we rated each measure separately but this summary table reflects our overall assessment across measures. See Full Report for individual outcome measure strength of evidence ratings.

[b] As measured by visual analog scale (VAS) for leg pain and for back pain.

[c] As measured by Oswestry Disability Index.

[d] As measured by the physical health and mental health component summary scores of the SF-36.

[e] As measured by duration of postoperative disability and by the proportion out of work at unspecified follow-up time.

Abbreviations: k = number of studies; N = number of participants; RCT = randomized controlled trials

Function and disability

One RCT reported outcomes with the Oswestry Disability Index. Teli et al.[29] observed scores improve at 26 weeks, 52 weeks, and 2 years among participants allocated to microdiscectomy and in participants allocated to discectomy. Among participants allocated to microdiscectomy, score decreases from baseline (40 [SD 4]) ranged from 25 to 29 points; among participants allocated to discectomy score decreases from baseline (39 [SD 4]) ranged from 24 to 27 points. No significant between-group differences were observed (P=0.81).

Summary: We assessed this single study body of evidence as insufficient for drawing a conclusion about function and disability outcomes comparing microdiscectomy and discectomy.

Quality of life

One RCT reported outcomes with the SF-36 PCS score and the MCS score at 26 weeks, 52 weeks, and 2 years. For PCS, Teli et al.[29] reported increases from baseline (21 [SD 4]) ranging from 19 to 23 points among participants allocated to microdiscectomy at the various follow-up time points compared with increases from baseline (22 [SD 4]) ranging from 18 to 22 points among participants allocated to discectomy. No significant between-group differences were observed (P=0.68). Similar findings were reported for the MCS (P=0.78 for between-group differences).

Summary: We assessed this single study body of evidence as insufficient for drawing a conclusion about quality of life outcomes comparing microdiscectomy and discectomy.

Neurological symptoms

No studies reported outcomes related to neurological symptoms, thus we concluded the evidence was insufficient.

Return to work

One RCT reported on outcomes related to "return to work". Tullberg et al.[27] reported a mean duration of postoperative, full-time sick leave of 10.4 weeks (SD NR) in participants allocated to microdiscectomy compared with 10.1 weeks (SD NR) in participants allocated to discectomy (P value NR). The proportion out of work at an unspecified follow-up time point was 16.7% among those allocated to microdiscectomy and 6.7% among those allocated to discectomy (calculated P =0.42).

Summary: We assessed this single study body of evidence as insufficient for drawing a conclusion about return to work outcomes comparing microdiscectomy and discectomy.

Other efficacy outcomes

Tullberg et al.[27] reported the frequency and proportion of participants with a specified opinion on recovery at 52 weeks (total recovery, almost recovered, good, unchanged, or worse). Among participants allocated to microdiscectomy, 11 (37.9%) reported total recovery, and 8 (27.6%) reported almost recovered. Among participants allocated to discectomy these outcomes were 6 (20.7%) and 14 (28.3%), respectively (calculated P=0.25 and 0.18, respectively). We did not use these outcomes in our strength of evidence ratings because they were only reported by one study.

ES-3.2.2 Efficacy Question 2

In adults with symptomatic lumbar radiculopathy, does effectiveness or comparative effectiveness of surgical interventions vary for patients who are not employed because of disability or patients who are undergoing recurrent surgery for relapse?

We did not identify any studies that reported outcomes specifically for patients not employed because of disability.

We identified two studies focused on the efficacy[46] or comparative effectiveness[47] of revision surgery for relapse; both were rated as high risk of bias. North et al.[46] was conducted in the

United States and randomized 50 participants with persistent radicular pain despite one or more prior lumbosacral spine surgeries to either repeat lumbosacral decompression or spinal cord stimulation. Ruetten et al.[47] was conducted in Germany and randomized 100 adults who had a previous conventional discectomy with acute occurrence of radicular leg symptoms after a pain-free interval in combination with a recurrent disc herniation on magnetic resonance imaging (MRI) to either revision endoscopic discectomy or revision microdiscectomy. ***Table ES-8*** and ***Table ES-9*** summarizes findings and strength of evidence ratings for each of these two comparisons.

Table ES-8. Summary of efficacy findings and strength of evidence ratings comparing repeat lumbosacral decompression surgery with spinal cord stimulation for treatment of lumbar radiculopathy relapses (EQ2)

Outcomes Length of follow-up	No. RCTs (k) No. participants (N)	Summary of effect ⇨ favors repeat surgery; ⇦ favors spinal cord stimulation; ⇔ no difference; ? unable to determine		Certainty[a]
Pain	k=0	?	No studies.	◯◯◯◯ INSUFFICIENT
Function/disability[b] 1.8 to 5.7 years	k=1; N=50	?	Similar levels of improvement in function and disability from repeat surgery and spinal cord stimulation.	◯◯◯◯ INSUFFICIENT
Quality of life	k=0	?	No studies.	◯◯◯◯ INSUFFICIENT
Neurologic symptoms	k=0	?	No studies.	◯◯◯◯ INSUFFICIENT
Return to work[c] 1.8 to 5.7 years	k=1; N=50	?	Return to work outcomes were similar between participants receiving repeat lumbosacral decompression and participants receiving spinal cord stimulation.	◯◯◯◯ INSUFFICIENT

[a] We assessed certainty using a modified GRADE approach, which assesses the evidence base for each outcome measure based on risk of bias, inconsistency, imprecision, indirectness, and other considerations; certainty is rated as "insufficient", "very low", "low", "moderate", or "high". For domains with more than 1 measure reported (e.g., pain, function), we rated each measure separately but this summary table reflects our overall assessment across measures. See Full Report for individual outcome measure strength of evidence ratings.

[b] As measured by patient-reported impairment from pain in performing everyday activities.

[c] Specific measure used was poorly defined.

Table ES-9. Summary of efficacy findings and strength of evidence ratings comparing revision endoscopic discectomy with revision microdiscectomy for treatment of relapsed lumbar radiculopathy (EQ2)

Outcomes Length of follow-up	No. RCTs (k) No. participants (N)	Summary of effect ⇨ favors revision endoscopic surgery; ⇦ favors microdiscectomy;⇔ no difference; ? unable to determine		Certainty[a]
Pain[b] 12 weeks to 2 years	k=1; N=100	?	Similar improvements in pain over time in both surgical groups.	◯◯◯◯ INSUFFICIENT
Function/disability[c] 12 weeks to 2 years	k=1; N=100	?	Similar improvements in function and disability over time in both surgical groups.	◯◯◯◯ INSUFFICIENT
Quality of life	k=0	?	No studies.	◯◯◯◯ INSUFFICIENT
Neurologic symptoms[d] 12 weeks to 2 years	k=1; N=100	?	Similar improvement in neurologic symptoms over time in both surgical groups.	◯◯◯◯ INSUFFICIENT
Return to work[e]	k=1; N=100	?	Revision endoscopic surgery results in a shorter duration of postoperative disability (4 weeks) compared to microdiscectomy (7.4 weeks) (P < 0.01).	◯◯◯◯ INSUFFICIENT

[a] We assessed certainty using a modified GRADE approach, which assesses the evidence base for each outcome measure based on risk of bias, inconsistency, imprecision, indirectness, and other considerations; certainty is rated as "insufficient", "very low", "low", "moderate", or "high". For domains with more than 1 measure reported (e.g., pain, function), we rated each measure separately but this summary table reflects our overall assessment across measures. See Full Report for individual outcome measure strength of evidence ratings.

[b] As measured by visual analog scale 100 mm for leg and back pain, and North American Spine Society Pain score.

[c] As measured by Oswestry Disability Index.

[d] As measured by North American Spine Society Neurology score.

[e] As measured by mean duration of post-operative disability, P< 0.01.

Pain

North et al.,[46] which compared repeat lumbosacral decompression to spinal cord stimulation did not report any outcomes related to pain. Ruetten et al.,[47] which compared revision endoscopic discectomy to revision microdiscectomy reported improvement in VAS 100 mm leg pain score from baseline to 12 weeks, 26 weeks, 52 weeks, and 2 years among participants allocated to revision endoscopic discectomy and among participants allocated to revision microdiscectomy. Between-group differences were reported as not significant at any follow-up time point (AMDs were NR). A similar pattern was observed for VAS 100 mm back pain scores and North American Spine Society pain scores.

Summary: Because no studies reported pain outcomes for one comparison and only a single study body of evidence for the other comparison, we assessed the evidence as insufficient to draw a conclusion for pain outcomes.

Functioning/disability

North et al.[46] reported no significant differences in qualitative assessment of impairment among participants who received repeat lumbosacral decompression compared with participants who received spinal cord stimulation. Ruetten et al.[47] reported improvements as measured by the Oswestry Disability Index from baseline to 12 weeks, 26 weeks, 52 weeks, and 2 years among participants allocated to revision endoscopic discectomy and among participants allocated to revision microdiscectomy. The between-group differences were reported as not significant.

Summary: Because of single study bodies of evidence for each comparison, we assessed the evidence as insufficient to draw a conclusion for function and disability outcomes.

Quality of life

Neither study reported outcomes related to overall quality of life; thus, we assessed the evidence as insufficient for drawing conclusions about quality of life outcomes.

Neurological symptoms

North et al. [46] did not report any outcomes related to neurologic symptoms. Ruetten et al.[47] reported mean North American Spine Society Neurology scores at 12 weeks, 26 weeks, 52 weeks, and 2 years. Scores in both groups improved over time and the differences between groups were reported as not significant.

Summary: Because no studies reported neurological symptom outcomes for one comparison and only a single study body of evidence reported these outcomes for the other comparison, we assessed the evidence as insufficient to draw a conclusion for neurological symptom outcomes.

Return to work

North et al.[46] reported no significant differences in return to work at a mean follow-up of 2.9 years, but actual values were not reported. Ruetten et al.[47] reported a significant difference between groups in the mean duration of postoperative disability. Among participants allocated to revision endoscopic discectomy, the mean was 4 weeks (SD NR) and among participants allocated to revision microdiscectomy the mean was 7.4 weeks (SD NR, $P < 0.01$).

Summary: Because of single study bodies of evidence for each comparison, we assessed the evidence as insufficient to draw a conclusion for return to work outcomes.

Other efficacy outcomes

North et al.[46] reported on the frequency and proportion of successful treatment over a mean follow-up time of 2.9 years (range 1.8 years to 5.7 years). Success was defined as at least 50% pain relief and patient satisfaction with treatment. A significant difference in treatment success was observed ($P < 0.01$). Among those allocated to repeat lumbosacral decompression, successful treatment was observed in 3 (12%) and among those allocated to spinal cord stimulation, successful treatment was observed in 9 (47%). Ruetten et al.[47] reported on the frequency and proportion of patient satisfaction with surgery and whether participants would undergo the operation again. Among those allocated to revision endoscopic discectomy, 43 (95%) were satisfied; among those allocated to revision microdiscectomy 36 (86%) were satisfied (P value for comparison NR). We did not use these outcomes in our strength of evidence ratings because of heterogeneity in outcome definition.

ES-3.3 Safety

Safety Question 1

In adults with symptomatic lumbar radiculopathy, what are the adverse events associated with surgical interventions?

All 24 RCTs included for EQ1 and the two RCTs included for EQ2 also provided evidence for safety outcomes.

A. Surgery compared with nonsurgical interventions

Seven RCTs that compared surgery to nonsurgical interventions reported at least one safety outcome.[22,23,26,32,33,37,41] ***Table ES-10*** summarizes findings and strength of evidence ratings.

Table ES-10. Summary of safety outcome findings and strength of evidence ratings comparing surgery to nonsurgical interventions in persons with symptomatic lumbar radiculopathy (SQ1)

Outcomes Length of follow-up	No. RCTs (k) No. participants (N)	Summary of effect ⇨ favors surgery; ⇦ favors nonsurgical intervention; ⇔ no difference; ? unable to determine		Certainty[a]
Surgical mortality	k=6; N=1,096	NA[b]	Surgical mortality is rare; no deaths reported among participants allocated to surgery in any studies.	⊕⊕◯◯ LOW
All-cause mortality Up to 10 years	k=3; N=717	⇔	All-cause mortality is rare and is similar for surgery and nonsurgical interventions.	⊕⊕◯◯ LOW
Surgical morbidity	k=6; N=1,032	NA[b]	Surgical morbidity occurs with low frequency; dural tears are the most common adverse event (reported in up to 4% of cases).	⊕⊕◯◯ LOW
Reoperations Up to 5 years	k=5; N=942	NA[b]	The incidence of reoperations varies from 0% to 10%.	⊕◯◯◯ VERY LOW
Persistent opioid use Up to 26 weeks	k=1; N=90	?	Surgery and nonsurgical interventions result in similar frequency of persistent opioid use.	◯◯◯◯ INSUFFICIENT

[a] We assessed certainty using a modified GRADE approach, which assesses the evidence base for each outcome measure based on risk of bias, inconsistency, imprecision, indirectness, and other considerations; certainty is rated as "insufficient", "very low", "low", "moderate", or "high".

[b] Not applicable for comparative assessment because comparator treatment is nonsurgical intervention.

Abbreviations: k = number of studies; N = number of participants; NA = not applicable; RCT = randomized controlled trial

Mortality

Six RCTS reported on mortality.[22,23,26,32,33,41] All of these studies were rated as low risk of bias for this specific outcome. Surgical mortality is not relevant as a comparative outcome given the nonsurgical comparison group. Thus, the strength of evidence for surgical mortality reflects our certainty about the absolute incidence of surgical mortality in the surgical intervention group. Of the RCTs that reported surgical mortality, no studies reported any deaths relating to percutaneous disc decompression,[41] microdiscectomy,[23,33] discectomy,[32] or discectomy/microdiscectomy procedures.[22] Three RCTs reported all-cause mortality.[22,26,41] For example, Weber et al. reported three deaths (5.0%) among participants allocated to discectomy and no deaths among participants allocated to conservative management at 10 years.[26] One death was due to cancer and two due to heart disease.

Summary: We concluded with low certainty that surgical mortality is rare, and that no difference in all-cause mortality exists between surgical and nonsurgical interventions.

Surgical Morbidity

Six RCTs reported surgical morbidity outcomes.[22,23,32,33,37,41] All of these studies were rated as low risk of bias for this specific outcome. Surgical morbidity is not relevant as a comparative outcome given the nonsurgical comparison group. Thus, the strength of evidence for surgical morbidity reflects our certainty about the absolute incidence of surgical morbidity in the surgical intervention group. Surgical complications were generally rare among participants who underwent surgical intervention. In the largest of the trials, Weinstein et al. [SPORT] reported 10 (4.0%) dural tear or spinal fluid leaks, 4 (1.6%) superficial postoperative wound infections, 1 (0.40%) vascular injury, 2 (0.81%) other intraoperative complications, and 9 (3.6%) other unspecified postoperative complications among participants who underwent discectomy/microdiscectomy.[22]

Summary: We concluded with low certainty that surgical complications were generally rare among participants who underwent surgical interventions.

Reoperations

Five RCTs reported reoperation rates in participants that were allocated to and underwent the surgical intervention; some studies also reported reoperations among participants who crossed over from the nonsurgical intervention to surgery.[22,23,32,33,37] Reoperations is not relevant as a comparative outcome given the nonsurgical comparison group. Thus, the strength of evidence for reoperation reflects our certainty about the absolute incidence of reoperations among those who underwent surgery, whether initially allocated to the surgical group or among those who crossed over to surgery at some point during the trial. The incidence of reoperations across the five RCTs varied from 0% to 10%. For example, Peul et al.[32] reported that 7 (6%) participants allocated to microdiscectomy had reoperations for recurrent sciatic within 2 years and 9 (7%) by 5 years. Among participants allocated to conservative management who crossed over to receive surgery, 4 (6%) underwent a reoperation by 2 years and 8 (12%) by 5 years.

Summary: We concluded with very low certainty that the incidence of reoperations varies between 0 and 10%.

Persistent Opioid use

Only one RCT reported outcomes related to persistent opioid use. Gerszten et al.[41] reported that reduction in use of narcotics was not significantly different at 26 weeks between participants who underwent percutaneous disc decompression and those who underwent conservative management (actual values NR, P value NR).

Summary: Because of only one study, we assessed the evidence as insufficient to draw a conclusion about persistent opioid use outcomes.

B. Minimally-invasive surgery compared with microdiscectomy or discectomy

Twelve RCTs that compared minimally-invasive surgery to nonsurgical interventions reported at least one safety outcome.[24,28-31,34-36,38-40,43] ***Table ES-11***summarizes findings and strength of evidence ratings.

Table ES-11. Summary of safety outcome findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery in persons with symptomatic lumbar radiculopathy (SQ1)

Outcomes Length of follow-up	No. RCTs (k) No. participants (N)	Summary of effect ⇨ favors minimally-invasive surgery; ⇦ favors standard surgery;⇔ no difference; ? unable to determine		Certainty[a]
Surgical mortality	k=5; N=464	⇔	No surgery-related deaths reported in any studies.	⊕⊕◯◯ LOW
All-cause mortality	k=2; N=428	⇔	Only 1 death unrelated to surgery was reported in 1 RCT; other RCT reported 2 vs. 3 deaths by 5 years.	⊕⊕◯◯ LOW
Surgical morbidity	k=10; N=1,151	⇔	The most commonly reported complications were dural tears and spinal fluid leaks. Between-group differences were generally similar between groups with one exception.	⊕◯◯◯ VERY LOW
Reoperations	k=10; N=1,200	?	The proportion of participants that had reoperations varied extensively across study groups (from 2.5 % to 64.5%). Between-group differences were not significant in 8 RCTs, but favored standard surgery in 2 RCTs.[38,39] Pooled ARD 7% (95% CI, -2% to 17%; 10 RCTs; 1,172 participants I^2=86.1%); pooled RR 1.37 (95% CI 0.74 to 2.62; I^2=60.6%).	◯◯◯◯ INSUFFICIENT
Persistent opioid use	k=1; N=60	?	The duration of postoperative narcotic use ranged from 0.43 to 2 weeks (average 1 week) for participants who underwent video-assisted arthroscopic microdiscectomy and 1 to 8 weeks for participants who underwent discectomy.	◯◯◯◯ INSUFFICIENT

[a] We assessed certainty using a modified GRADE approach, which assesses the evidence base for each outcome measure based on risk of bias, inconsistency, imprecision, indirectness, and other considerations; certainty is rated as "insufficient", "very low", "low", "moderate", or "high".

Abbreviations: ARD = absolute risk difference; CI = confidence interval; k = number of studies; N = number of participants; RCT = randomized controlled trial; RR = relative risk ratio.

Mortality

Five RCTs reported mortality outcomes.[24,29,31,34,43] No surgery-related deaths were reported. Two RCTs reported all-cause mortality.[31,40,48] Ruetten et al.[31] reported one death (0.5%) unrelated to surgery; the authors did not specify whether this death occurred among participants allocated to the minimally-invasive surgery or among participants allocated to microdiscectomy. Arts et al.[40,48] reported two deaths among participants allocated to tubular discectomy and three deaths among participants allocated to microdiscectomy by 5 years.

Summary: We concluded with low certainty that all-cause and surgical mortality are rare and similar for minimally-invasive surgery and standard surgery.

Surgical Morbidity

Ten RCTs[24,28-31,34,36,39,40,43] reported surgical morbidity outcomes. The most common complications reported were those relating to dural tear and spinal fluid leak. In nine of the 10 RCTs, morbidity incidence was similar between groups, though few reported statistical significance testing. One RCT reported significantly fewer complications among participants

who underwent endoscopic discectomy compared to participants who underwent microdiscectomy.[31] In this study, Ruetten et al.[31] reported significantly fewer complications (P<0.05) among participants who underwent endoscopic discectomy compared to participants who underwent microdiscectomy participants. Complications included transient postoperative dysesthesia (3.3% vs 5.7%), postoperative bleeding (0% vs 2.3%), delayed wound healing (0% vs 1.1%), and soft tissue infection (0% vs 1.1%).

Summary: We concluded with very low certainty that the incidence of surgical morbidity is similar for minimally-invasive surgery and standard surgery.

Reoperations

Ten RCTs reported the incidence of reoperations.[28-31,34,36,38-40,43] The proportion of participants that had reoperations varied extensively across study groups (from 2.5 % to 64.5%). Chatterjee et al.[38] (automated percutaneous lumbar discectomy) and Brouwer et al. [39] (percutaneous laser disc decompression) reported a significantly higher frequency of reoperations among participants who underwent minimally-invasive surgery compared to standard surgery. These findings were inconsistent with findings from the other 8 RCTs, which observed a similar incidence of reoperations between surgical groups. The pooled relative risk ratio (RR) was 1.37 (95% CI, 0.74 to 2.52; 10 RCTs; 1,172 participants; I^2=60.8%)) and pooled absolute risk difference (ARD) was 7% (95% CI, -2% to 17%; I^2=86.1%). Because of unique circumstances in the Chatterjee et al [38] study, we excluded it from the pooled estimate in a sensitivity analysis. The pooled ARD without it was 2% (95% CI, -4% to 8%; 9 RCTs, 1,101 participants; I^2= 60.7%) and the pooled RR was 1.17 (95% CI, 0.70 to 1.97; I^2= 44.4%). We believe the residual inconsistency in pooled estimates are likely explained by varying definitions and ascertainment methods (e.g., timing of measurement), and because some studies may have been more or less aggressive in offering participants reoperations for residual symptoms.

Summary: We concluded the evidence on incidence of reoperations is insufficient because of inconsistent findings, likely because of study limitations (risk of bias) and varying definitions and ascertainment methods.

Persistent Opioid use

Hermantin et al.[43] reported the duration of postoperative narcotic use ranged from 0.43 to 2 weeks (average 1 week) for participants who underwent video-assisted arthroscopic microdiscectomy and 1 to 8 weeks (average 3.65 weeks) for participants who underwent discectomy (P value NR).

Summary: We condcluded the evidence was insufficient to draw a conclusion about persistent opioid use because of a single-study body of evidence.

C. Microdiscectomy compared to discectomy

Three RCTs that compared microdiscectomy to discectomy reported at least one safety outcome.[27,29,35] ***Table ES-12*** summarizes findings and strength of evidence ratings.

Table ES-12. Summary of safety outcome findings and strength of evidence ratings comparing microdiscectomy to discectomy in persons with symptomatic lumbar radiculopathy (SQ1)

Outcomes Length of follow-up	No. RCTs (k) No. participants (N)	Summary of effect ⇨ favors microdiscectomy; ⇦ favors discectomy; ⇔ no difference; ? unable to determine		Certainty[a]
Surgical mortality	k=1; N=142	?	No surgery-related deaths reported in either surgical group.	◯◯◯◯ INSUFFICIENT
All-cause mortality	k=0	?	No studies.	◯◯◯◯ INSUFFICIENT
Surgical morbidity	k=3; N=282	⇔	Surgical morbidity was infrequent and similar in both surgical groups.	⊕◯◯◯ VERY LOW
Reoperations	k=2; N=202	⇔	Incidence of reoperation similar in both surgical groups; range from 3% to 4%.	⊕◯◯◯ VERY LOW
Persistent opioid use	k=0	?	No studies.	◯◯◯◯ INSUFFICIENT

[a] We assessed certainty using a modified GRADE approach, which assesses the evidence base for each outcome measure based on risk of bias, inconsistency, imprecision, indirectness, and other considerations; certainty is rated as "insufficient", "very low", "low", "moderate", or "high".

Abbreviations: k = number of studies; N = number of participants; RCT = randomized controlled trial

Mortality

Only one RCT reported surgical mortality. Teli et al.[29] reported no surgical deaths in either group. We concluded the evidence was insufficient to draw a conclusion about mortality because of a single-study body of evidence.

Surgical Morbidity

Three RCTs reported surgical morbidity,[27,29] but one[35] did not report by group. In one RCT, the overall frequency of surgical infection was 6.3%.[35] The other two RCTs reported similar frequency of complications between groups with respect to dural tears and nerve root injury. We concluded with very low certainty that surgical morbidity is similar for microdiscectomy and discectomy.

Reoperations

Two RCTs reported on reoperations. Tullberg et al.[27] reported 1 (3.3%) reoperation by 52 weeks in each surgical group (microdiscectomy and discectomy). Teli et al.[29] reported 3 (4.2%) reoperations among participants who underwent microdiscectomy compared with 2 (3%) among participants who underwent discectomy (calculated P=1.0). We concluded with very low certainty that the incidence or reoperations is similar for microdiscectomy and discectomy.

Persistent Opioid use

No RCTs reported persistent opioid use outcomes; thus the evidence was insufficient to draw a conclusion.

D. Revision Surgery

One RCT compared repeat lumbosacral decompression with spinal cord stimulation,[46] and another RCT compared revision endoscopic discectomy with revision microdiscectomy.[47] ***Table***

ES-13 and ***Table ES-14*** summarizes findings and strength of evidence ratings. Across all outcomes reported, we considered the evidence insufficient to draw conclusions because of single-study bodies of evidence for both comparisons.

Table ES-13. Summary of safety outcome findings and strength of evidence ratings comparing repeat lumbosacral decompression to spinal cord stimulation in persons with relapsed lumbar radiculopathy (SQ1)

Outcomes Length of follow-up	No. RCTs (k) No. participants (N)	Summary of effect ⇨ favors revision surgery; ⇦ favors comparator; ⇔ no difference; ? unable to determine		Certainty[a]
Surgical mortality	k=1; N=50	?	No surgery-related deaths reported in either group.	◯◯◯◯ INSUFFICIENT
All-cause mortality	k=1; N=50	?	1 death reported among participants receiving spinal cord stimulation; the death was due to a sudden cardiac event.	◯◯◯◯ INSUFFICIENT
Surgical morbidity	k=1; N=50	?	1 site infection reported among participants who underwent spinal cord stimulations; none in surgery group. No other complications reported.	◯◯◯◯ INSUFFICIENT
Reoperations	k=1; N=50	?	No reoperations among participants who underwent repeat decompression compared with 3 who underwent spinal cord stimulation.	◯◯◯◯ INSUFFICIENT
Persistent opioid use	k=1; N=50	?	15 (58%) participants who underwent repeat decompression with stable or decreased opioid use at 1.8 to 5.7 years follow-up compared with 20 (80%) of participants who underwent spinal cord stimulation (P=0.025).	◯◯◯◯ INSUFFICIENT

[a] We assessed certainty using a modified GRADE approach, which assesses the evidence base for each outcome measure based on risk of bias, inconsistency, imprecision, indirectness, and other considerations; certainty is rated as "insufficient", "very low", "low", "moderate", or "high".

Abbreviations: k = number of studies; N = number of participants; RCT = randomized controlled trial

Table ES-14. Summary of safety outcome findings and strength of evidence ratings comparing revision endoscopic surgery to revision microdiscectomy in persons with relapsed lumbar radiculopathy (SQ1)

Outcomes Length of follow-up	No. RCTs (k) No. participants (N)	Summary of effect ⇨ favors revision surgery; ⇦ favors comparator; ⇔ no difference; ? unable to determine		Certainty[a]
Surgical mortality	k=1; N=100	?	No surgery-related deaths reported in either group.	◯◯◯◯ INSUFFICIENT
All-cause mortality	k=1; N=100	?	No deaths reported in either group.	◯◯◯◯ INSUFFICIENT
Surgical morbidity	k=1; N=100	?	Significantly fewer serious complications among participants receiving revision microendoscopic surgery, compared to revision microdiscectomy (6% vs. 21%, P < 0.05).	◯◯◯◯ INSUFFICIENT
Reoperations	k=1; N=100	?	2 reoperations among participants who underwent revision microendoscopic surgery compared with 3 who underwent revision microdiscectomy.	◯◯◯◯ INSUFFICIENT
Persistent opioid use	k=0	?	No studies.	◯◯◯◯ INSUFFICIENT

[a] We assessed certainty using a modified GRADE approach, which assesses the evidence base for each outcome measure based on risk of bias, inconsistency, imprecision, indirectness, and other considerations; certainty is rated as "insufficient", "very low", "low", "moderate", or "high".

Abbreviations: k = number of studies; N = number of participants; RCT = randomized controlled trial

ES 3.4 Cost

Cost Question 1

In adults with symptomatic lumbar radiculopathy, what is the cost-effectiveness of surgical interventions?

We identified seven eligible studies reporting cost.[29,44,49-53] Five studies reported cost-effectiveness analyses related to RCTs that we also included for efficacy and safety outcomes (Peul et al., Weinstein et al., Arts et al., Brouwer et al., and Chatterjee et al.).[49-53] One study, Teli et al.,[29] was a trial we also included for efficacy and safety outcomes that reported on surgical costs of three alternative surgical interventions. Lastly, Malter et al.[44] reported a cost-effectiveness analysis using cost and effectiveness inputs from a variety of sources. Two studies were conducted in the United States;[44,50] the rest were conducted in the Netherlands,[49,51,52] Italy,[29] and the United Kingdom.[53] The time horizon used in studies ranged from 26 weeks to 10 years. Studies reported cost findings using different currency and base years; thus, we converted all figures to 2010 U.S. dollars (see ***Appendix C*** for details on conversion) for this report. We rated one cost-effectiveness analysis as poor quality[53], one as fair quality [44], and four as good quality.[49-52] We did not assess the quality for Teli et al, as it only reported costs and not cost-effectiveness.

A. Surgery compared with nonsurgical interventions

Three studies provided evidence for the cost-effectiveness of surgery compared with nonsurgical treatment.[44,49,50] ***Table ES-15*** summarizes findings and strength of evidence ratings.

Table ES-15. Summary of cost-effectiveness findings comparing surgery to nonsurgical interventions in persons with symptomatic lumbar radiculopathy (CQ1)

Outcomes Length of follow-up	No. studies (k) No. participants (N)	Summary of effect	Certainty[a]
Cost-effectiveness Between 1 and 10 years	k=3; N=1,474[b]	Surgery results in higher quality-adjusted life years but similar or higher costs compared to nonsurgical interventions. The mean cost per quality-adjusted life year gained from the payor perspective ranged from $51,156 to $83,322 (in 2010 U.S. dollars).	⊕◯◯◯ VERY LOW

[a] We assessed certainty using a modified GRADE approach, which assesses the evidence base for each outcome measure based on risk of bias, inconsistency, imprecision, indirectness, and other considerations; certainty is rated as "insufficient", "very low", "low", "moderate", or "high".

[b] One study was a decision analysis not concurrent to a trial so no N reported; one study combined data from a trial and a concurrent observational study.

Abbreviations: k = number of studies; N = number of participants; U.S. = United States.

B. Minimally-invasive surgery compared with microdiscectomy or discectomy

Four studies[29,51-53] provided evidence for the cost-effectiveness of alternative surgical interventions, including percutaneous laser discectomy,[52] tubular discectomy,[51] and automated percutaneous discectomy[53] compared to microdiscectomy and a three-arm study comparing microendoscopic discectomy, microdiscectomy, and discectomy.[29] ***Table ES-16*** summarizes findings and strength of evidence ratings.

Table ES-16. Summary of cost-effectiveness findings comparing minimally-invasive surgery to standard surgery in persons with symptomatic lumbar radiculopathy (CQ1)

Outcomes Length of follow-up	No. studies (k) No. participants (N)	Summary of effect	Certainty[a]
Cost-effectiveness	k=4; N=656	Inconsistent findings across studies. One study[29] found higher surgical costs (AMD $722; 95% CI, $551 to $892) for minimally-invasive surgery compared to standard surgery; one study[53] calculated a higher cost ($3,573) per successful outcome for minimally-invasive surgery compared to standard surgery; one study[51] reported no significant differences in quality-adjusted life years, total costs, or health care costs between groups, but point estimates suggest minimally-invasive surgery is less effective and costs more; and the last study[52] reported no significant differences in quality-adjusted life years or total costs but some differences in health care costs suggesting minimally-invasive surgery costs less (AMD -$2,393; 95% CI, $-4,376 to $-409) but is also less effective (calculated costs/QALY gained $97,424 for microdiscectomy compared to minimally-invasive surgery).	◯◯◯◯ INSUFFICIENT

[a] We assessed certainty using a modified GRADE approach, which assesses the evidence base for each outcome measure based on risk of bias, inconsistency, imprecision, indirectness, and other considerations; certainty is rated as "insufficient", "very low", "low", "moderate", or "high".

Abbreviations: AMD = absolute mean difference; CI = confidence interval; k = number of studies; N = number of participants; QALY = quality-adjusted life years.

C. Microdiscectomy compared to discectomy

One study reported the surgical costs of microdiscectomy to discectomy, but did not report cost-effectiveness.[29] The cost of microdiscectomy was $3,156 (SD $438) and the cost of discectomy was $2,976 (SD $322) (calculated AMD $65; 95% CI, $52 to $307). ***Table ES-17*** summarizes findings and strength of evidence ratings.

Table ES-17. Summary of costs comparing microdiscectomy to discectomy in persons with symptomatic lumbar radiculopathy (CQ1)

Outcomes Length of follow-up	No. studies (k) No. participants (N)	Summary of effect	Certainty[a]
Costs	k=1; N=142	Costs are slightly higher for microdiscectomy compared to discectomy ($3,156 (SD $438) vs $2,976 (SD $322). Calculated AMD $65 (95% CI, $52 to $307).	◯◯◯◯ INSUFFICIENT

[a] We assessed certainty using a modified GRADE approach, which assesses the evidence base for each outcome measure based on risk of bias, inconsistency, imprecision, indirectness, and other considerations; certainty is rated as "insufficient", "very low", "low", "moderate", or "high". See Full Report for individual outcome measure strength of evidence ratings.

Abbreviations: AMD = absolute mean difference; CI = confidence interval; k = number of studies; N = number of participants; SD = standard deviation;

ES-3.5 Synthesis of Clinical Practice Guidelines

A synopsis of clinical practice guidelines (CPGs) and guidance related to the use of surgical procedures for lumbar radiculopathy is summarized in ***Table ES-18.*** Please refer to the Full Report for additional details. We assessed the quality of each CPG or procedure guidance using the Appraisal of Guidelines for Research & Evaluation II instrument.[20,21] With this instrument, six domains are assessed and an overall score of between 1 (lowest possible) and 7 (highest possible) are assigned to reflect the overall quality of the guideline.

Overall, the guidelines we identified were in general agreement about considering discectomy or microdiscectomy (and related decompressive procedures) as acceptable treatment based on evidence that it improves outcomes in the short- to medium-term. One guideline specifies that this surgery can be considered when symptoms have not improved with conservative therapy.[45] Another guideline suggests that conservative therapy is reasonable for patients with nonprogressive symptoms who wish to delay surgery.[54] The guideline recommendations relating to minimally-invasive spine surgery varied; one did not consider these specific procedures within their scope.[45] Three of the guidelines were developed 5 or more years ago; thus may not include the most recent evidence for these procedures.[55-57]

Table ES-18. Synopsis of clinical practice guidelines related to lumbar radiculopathy or herniated intervertebral lumbar disc

Organization ***Guideline Title (Year)*** **Guideline Quality**[a]	**Synopsis of Recommendation**[b]
National Institute for Health and Care Excellence (United Kingdom) *Low back pain and sciatica in over 16s: assessment and management-Invasive treatments (2016)*[45] Quality Rating: 6 out of 7	• Consider spinal decompression for sciatica (includes laminectomy, foraminotomy, and/or discectomy) when nonsurgical treatment has not improved pain or function and radiological findings are consistent with sciatica symptoms.
American Pain Society *Interventional Therapies, Surgery, and Interdisciplinary Rehabilitation for Low Back Pain (2009)*[55] Quality Rating: 5 out of 7	• Open discectomy or microdiscectomy recommended for radiculopathy with prolapsed disc. • Insufficient evidence for determining superiority of open vs. micro approaches, and to evaluate alternative surgical methods including laser- or endoscopic-assisted techniques.
North American Spine Society *Clinical Guidelines for Diagnosis and Treatment of Lumbar Disc Herniation with Radiculopathy (2012)*[57] Quality Rating: 5 out of 7	• Discectomy is suggested to provide more effective symptom relief than medical/interventional care for patients with lumbar disc herniation with radiculopathy whose symptoms warrant surgical intervention. In patients with less severe symptoms, surgery or medical/interventional care appear to be effective for both short- and long-term relief. • Surgical intervention prior to 6 months is suggested in patients with symptomatic lumbar disc herniation whose symptoms are severe enough to warrant surgery. Earlier surgery (within 6 months to 1 year) is associated with faster recovery and improved long-term outcomes. • Use of an operative microscope is suggested to obtain comparable outcomes to open discectomy for patients with lumbar disc herniation with radiculopathy. • Endoscopic percutaneous discectomy is suggested for carefully selected patients to reduce early postoperative disability and reduce opioid use compared with open discectomy in the treatment of patients with lumbar disc herniation with radiculopathy. • In a select group of patients automated percutaneous lumbar discectomy (APLD) may achieve equivalent results to open discectomy, however, this equivalence is not felt to be generalizable to all patients with lumbar disc herniation with radiculopathy whose symptoms warrant surgery. • Insufficient evidence for other procedures (See Full Report for details)

(continued)

Table ES-18. Synopsis of clinical practice guidelines related to lumbar radiculopathy or herniated intervertebral lumbar disc (continued)

Organization *Guideline Title (Year)* Guideline Quality[a]	Synopsis of Recommendation[b]
American Society of Interventional Pain Physicians *An Update of Comprehensive Evidence-Based Guidelines for Interventional Techniques in Chronic Spinal Pain (2013)*[21,56] Quality Rating: 4 out of 7	For lumbar disc prolapse, protrusion, and extrusion: automated percutaneous lumbar decompression (APLD), percutaneous lumbar disc decompression (PLDD), and mechanical decompression with nucleoplasty are recommended in select cases.
American College of Occupational and Environmental Medicine *Low Back Disorders. In Occupational medicine practice guidelines: evaluation and management of common health problems and functional recovery in workers (2016)*[54] Quality Rating: Unknown[c]	Patients with evidence of specific nerve root compromise confirmed by appropriate imaging studies may be expected to potentially benefit from surgery.
National Institute for Health and Care Excellence (United Kingdom) *Percutaneous transforaminal endoscopic lumbar discectomy for sciatic: Interventional procedures guidance [IPG 556] (2016)]*[58] Quality Rating: 2 out of 7	Current evidence on the safety and efficacy of percutaneous transforaminal endoscopic lumbar discectomy for sciatica is adequate to support the use of this procedure provided that standard arrangements are in place for clinical governance, consent and audit.
National Institute for Health and Care Excellence (United Kingdom) *Percutaneous interlaminar endoscopic lumbar discectomy for sciatica: Interventional procedures guidance[IPG555](2016)*[59] Quality Rating: 2 out of 7	Current evidence on the safety and efficacy of percutaneous interlaminar endoscopic lumbar discectomy for sciatica is adequate to support the use of this procedure provided that standard arrangements are in place for clinical governance, consent and audit.
National Institute for Health and Care Excellence (United Kingdom) *Percutaneous coblation of the intervertebral disc for low back pain and sciatica Interventional procedures guidance[IPG543](2016)*[60] Quality Rating: 2 out of 7	Current evidence on percutaneous coblation of the intervertebral disc for low back pain and sciatica raises no major safety concerns. The evidence on efficacy is adequate and includes large numbers of patients with appropriate follow-up periods. Therefore, this procedure may be used provided that normal arrangements are in place for clinical governance, consent and audit.
National Institute for Health and Care Excellence (United Kingdom) Percutaneous electrothermal treatment of the intervertebral disc annulus for low back pain and sciatic[a] Interventional procedures guidance[IPG544](2016)[61] Quality Rating: 2 out of 7	Current evidence on percutaneous electrothermal treatment of the intervertebral disc annulus for low back pain and sciatica raises no major safety concerns. The evidence on efficacy is inconsistent and of poor quality. Therefore, this procedure should only be used with special arrangements for clinical governance, consent and audit or research.
National Institute for Health and Care Excellence (United Kingdom) *Percutaneous intradiscal radiofrequency treatment of the intervertebral disc nucleus for low back pain. Interventional procedures guidance[IPG545] (2016)*[62] Quality Rating: 2 out of 7	Current evidence on percutaneous intradiscal radiofrequency treatment of the intervertebral disc nucleus for low back pain raises no major safety concerns. The evidence on its efficacy is limited in quantity and quality. Therefore, this procedure should only be used with special arrangements for clinical governance, consent and audit or research.
National Institute for Health and Care Excellence (United Kingdom) *Epiduroscopic lumbar discectomy through the sacral hiatus for sciatica Interventional procedures guidance[IPG570] (2016)*[63] Quality Rating: 2 out of 7	Current evidence on the safety and efficacy of epiduroscopic lumbar discectomy through the sacral hiatus for sciatica is limited in quantity and quality. Therefore, this procedure should only be used in the context of research.

(continued)

Table ES-18. Synopsis of clinical practice guidelines related to lumbar radiculopathy or herniated intervertebral lumbar disc (continued)

Organization *Guideline Title (Year)* Guideline Quality[a]	Synopsis of Recommendation[b]
National Institute for Health and Care Excellence (United Kingdom) *Percutaneous intradiscal laser ablation in the lumbar spine. Interventional procedures guidance[IPG357] (2010)*[64] Quality Rating: 2 out of 7	Current evidence on the safety and efficacy of percutaneous intradiscal laser ablation in the lumbar spine is adequate to support the use of this procedure provided that normal arrangements are in place for clinical governance, consent and audit. Patients selected for the procedure should be limited to those with severe pain refractory to conservative treatment, in whom imaging studies show bulging of an intact disc, and who do not have neurological deficit requiring surgical decompression.
National Institute for Health and Care Excellence (United Kingdom) *Automated percutaneous mechanical lumbar discectomy: Interventional procedures guidance[IPG141])(2005)*[65] Quality Rating: 2 out of 7	Current evidence suggests that there are no major safety concerns associated with automated percutaneous mechanical lumbar discectomy. There is limited evidence of efficacy based on uncontrolled case series of heterogeneous groups of patients, but evidence from small randomized controlled trials shows conflicting results. In view of the uncertainties about the efficacy of the procedure, it should not be used without special arrangements for consent and for audit or research.
National Institute for Health and Care Excellence (United Kingdom) *Endoscopic laser foraminoplasty. Interventional procedures guidance[IPG31] (2003)*[66] Quality Rating: 2 out of 7	Current evidence of the safety and efficacy of endoscopic laser foraminoplasty does not appear adequate to support the use of this procedure without special arrangements for consent and for audit or research.

[a] We assessed the quality of guideline using the Appraisal of Guidelines For Research & Evaluation II (AGREE II) Instrument, version 2017.[21] The lowest quality score possible is 1, the highest possible quality score is 7.

[b] Only recommendations from the guideline pertinent to surgical interventions for lumbar radiculopathy are summarized; see the Full Report for a more complete summary.

[c] The complete guideline is not publicly accessible; thus, a full quality appraisal and summary of the evidence base and strength of evidence ratings were not possible.

ES-4. Discussion

ES-4.1 Summary of the Evidence

Evidence maps summarizing the overall findings and strength of evidence are provided in ***Figures 3, 4, 5, 6,*** and ***7*** of the Full Report. With few exceptions, most findings that we considered sufficient for assessment were based on evidence graded as low to very low certainty, primarily because of some or high concerns for bias among included studies and imprecision in study estimates. Most outcomes assessed as having insufficient evidence were single study bodies of evidence.

ES-4.1.1 Surgery compared with nonsurgical interventions

Findings for this comparison are summarized in an evidence map (***Figure 3***) in the Full Report. Surgery reduces pain more than nonsurgical interventions in the short and medium-term (up to 26 weeks) but this difference does not persist in the long-term. Several explanations for this are possible. One explanation for the mitigation of short-term benefits observed is that the impact of

participants that crossover between groups accumulates over time. For example, Peul et al.[32] reports that of the 142 participants allocated to conservative management, 55 (39%) underwent surgery during the first year after a median of 14.6 weeks, 62 (44%) underwent surgery by 2 years, and 66 (46%) by 5 years. In an intent-to-treat analysis, any treatment effects that might exist are mitigated by these crossovers. Another explanation is that long-term outcomes simply reflect the natural history of radiculopathy, particularly radiculopathy that results from disc herniation. The evidence was insufficient to assess short- and medium-term impact on function because of inconsistent findings, but long-term impact on function suggests no difference between treatments. The impact on other outcomes including quality of life, neurological symptoms, and return to work also found no meaningful differences between treatment groups.

No surgery-related deaths were observed and surgery-related complications were rare, but these findings may not be applicable to community practice where enrolled participants may have more comorbidities than participants enrolled in RCTs. As might be expected, no difference in all-cause mortality was observed. The evidence was insufficient to assess outcomes related to persistent opioid use because of a single-study body of evidence. Surgery compared with nonsurgical interventions may be cost-effective depending on a decision-makers willingness to pay threshold. In this HTA, the cost per quality-adjusted life year (QALY) gained ranged from $51,156 to $83,322 in 2010 U.S. dollars from a healthcare payor perspective.

ES-4.1.2 Minimally-invasive surgery compared with standard surgery

Findings are summarized in an evidence map (***Figure 4***) in the Full Report. For the purposes of this HTA, we synthesized interventions under the broad term "minimally-invasive surgical procedures", however, these procedures may represent a heterogenous set of interventions. With few exceptions, minimally-invasive surgery and standard surgery similarly reduce pain and improve function. However, minimally-invasive surgery seems to result in a quicker return to work, though this finding should be interpreted with caution because of the varying definitions of return to work used by studies, differences in work culture between U.S. and European countries, and because the advice given to participants as to when to return to work may be in part based on the procedure they received.

No surgical deaths were reported and surgical morbidity was similar between both approaches. Although 10 studies reported on the incidence of reoperations, the evidence was insufficient to draw a definitive conclusion because of mixed findings and imprecision in estimates. The evidence for persistent opioid use outcomes was also insufficient because of a single-study body of evidence. The evidence on cost-effectiveness for minimally-invasive surgery compared to standard approaches was also insufficient; further, none of the cost analyses were conducted in the U.S.

ES-4.1.3 Microdiscectomy compared to discectomy

Findings are summarized in an evidence map (***Figure 5***) in the Full Report. Microdiscectomy and discectomy were comparable with respect to pain, surgical morbidity and incidence of reoperations. However, the evidence was insufficient to draw conclusions about other outcomes because no studies reported these outcomes (neurological symptoms, persistent opioid use) or

these outcomes were reported by only a single-study (function/disability, quality of life, return to work, surgical mortality). The evidence was also insufficient for drawing conclusions about cost or cost-effectiveness.

ES-4.1.4 Repeat surgery for recurrent radiculopathy

Findings are summarized in evidence maps (***Figure 6, Figure 7***) in the Full Report. Only two RCTs reported on repeat surgery for recurrent radiculopathy, and used different comparator groups resulting in a single-study body of evidence for each comparison. Thus, the evidence was insufficient to draw any conclusions about efficacy, safety, or costs.

ES-4.2 Limitations of the Evidence Base

The studies we identified for inclusion in this HTA had numerous limitations as summarized in this section. Please refer to the Full Report for a more detailed description of each of these limitations.

- **Nearly half of included studies were assessed as high risk for bias**. Common sources of bias include lack of participant and outcome assessor blinding, deviations from intended interventions (i.e., crossovers), and inadequate randomization or treatment allocation in some studies.
- **Studies were generally underpowered to detect between-group differences for most outcomes of interest in this HTA.** Only 11 of the 24 included RCTs for efficacy and safety designated a primary outcome and described the sample size required to detect an a priori effect size. Few described how this effect size was determined or whether it represented a minimally important clinical difference and whether the analysis was designed for detecting superiority or noninferiority.
- **Variation in diagnosis of radiculopathy and severity of symptoms among participants at baseline.** Most studies required participants to have a clinical diagnosis of radiculopathy with disc herniation or nerve root compression confirmed by imaging (usually CT or MRI) for enrollment. However, few studies described the criteria for clinical diagnosis. Further, the duration of symptoms and criteria related to provision of conservative therapy prior to enrollment was variable across studies. Further, all studies were focused on populations with radiculopathy caused by intervertebral disc herniation, not from radiculopathy caused by more generalized spondylosis.
- **Limited number of comparative effectiveness trials for any one minimally-invasive surgical procedure.** We identified 15 trials comparing minimally-invasive surgery to open surgery; however, most of these interventions were only evaluated by 1 to 3 RCTs at the most, and variations in the outcomes reported limited our ability to draw conclusions for specific minimally-invasive procedures. Many studies lacked a full description of the surgical intervention, including the procedure, the skill and experience of the surgeon and surgical team, and pre- and postoperative care.
- **Variation in type, timing, and completeness in reporting outcomes.** Some studies reported between-group differences at multiple follow-up time points without a priori specification of a primary time point; others more appropriately used repeated measures analysis, to account for multiple observations over time, and some reported both. Our

ability to conduct quantitative syntheses (i.e., meta-analysis) was limited by variation in specific outcomes reported and by incomplete reporting. "Return to work" outcomes are particularly challenging to interpret in this evidence report. Safety outcomes reported were very heterogenous, particularly with respect to 'reoperations' and 'persistent opioid use'. For example, some studies reported 'all-cause' reoperations, some studies only report reoperations for technical failures, and some studies did not define or explain how reoperations were defined. See the Full Report for additional details.

- **Applicability of this evidence base to current practice.** Six RCTs were conducted prior to the year 2000. Changes in surgical technique and pre- and post-operative care may limit the applicability of findings from these older studies to current practice. Further, this RCT evidence base may underestimate differences in safety outcomes as participants in trials may have fewer comorbidities than individuals within the general population. Large case series and cohort studies could provide additional information on safety outcomes.
- **Limited number of United States cost studies.** Of the seven cost studies identified, only two were conducted in the U.S., limiting the applicability of findings to U.S. settings.
- **Limitations in the AGREE guideline appraisal instrument.** The AGREE guideline appraisal instrument largely focuses on evaluating the processes through which a guideline is developed; it does not assess how well the evidence included in the guideline was evaluated and interpreted correctly, or whether the conclusions of the guideline are consistent with the evidence. Thus, some guidelines may score artificially high and explains why conclusions may differ between guidelines despite having nearly similar scores on the instrument.

ES-4.3 Other Related HTAs

The only related HTA that we identified was commissioned by the National Institute for Health Research (U.K.) Health Technology Assessment programme.[67] This HTA included both surgical and nonsurgical interventions for the management of sciatica and used a network meta-analysis to provide a measure of relative therapeutic effect across 18 different treatment categories. The findings suggest that nonopioid medication, epidural corticosteroids injections, and disc surgery are effective for reducing sciatica. This HTA also concluded that stepped care approaches to treatment are cost-effective relative to direct referral for surgery.

ES-4.4 Payer Coverage

The Centers for Medicare and Medicaid Services (CMS) does not have a national coverage determination related to open standard or microsurgical decompressive procedures (i.e., discectomy, microdiscectomy, foraminotomy, laminectomy/otomy). With respect to the use of lasers, CMS recognizes their use to alter, revise, or destroy tissues in place of more conventional techniques as part of surgical procedures.[68] Medicare administrative contractors have been advised to use discretion in determining coverage for procedures performed with a laser when the laser has been FDA-approved, the procedure is considered reasonable and necessary, and a noncoverage instruction does not exist (effective date May 1, 1997).[68] CMS does have a national coverage determination related to thermal intradiscal procedures; these procedures are not covered (effective date January 1, 2009).[68] Percutaneous disc decompression falls within the

category of procedures covered by this determination. ***Table ES-19*** provides an overview of other payer coverage policies; please see the Full Report for complete details.

Table ES-19. Overview of payer coverage policies

Procedure	Medicare	Premera	Regence	Cigna	United	Aetna	Humana	Kaiser
Laminectomy, laminotomy, discectomy, foraminotomy (open technique including microsurgical approaches)	—	✓[a]	—	—	—	✓[a]	✓[a]	—
Automated percutaneous lumbar disc decompression	✕[b]	✕	✕	✕	✕	✓[c]	✕	—
(Percutaneous) endoscopic discectomy	✕	✕	✕	✕	✕	—	No additional reimbursement.	—
(Percutaneous) laser discectomy	✕	✕	✕	✕	✕	No additional reimbursement.	✕	—
Percutaneous nucleoplasty with coblation technology	✕	✕	✕	✕	✕	—	—	—

Notes: ✓ = covered; ✕ = not covered; — = no policy identified

[a] If specific clinical criteria are met. See Table 50 in the Full Report for details.

[b] All percutaneous disc decompression procedures fall under a Medicare National Coverage Determination related to thermal intradiscal procedures.

[c] Also covers percutaneous manual discectomy, see Table 50 in the Full Report for details.

In general, payers cover decompressive procedures, including discectomy, laminectomy/otomy, foraminectomy/otomy, including microsurgical approaches, for disc herniation with radicular symptoms. Specific criteria vary by payer but often include a failed trial of conservative management for 6 to 12 weeks. Most payers also require imaging confirmation of nerve root compression that corresponds to symptoms and physical examination findings. The coverage of minimally-invasive procedures varies by payer.

ES-4.5 Limitations of this HTA

This HTA was limited to studies and other information published or publicly available in English. The electronic search was limited to three databases. For efficiency, we relied on hand searches of existing systematic reviews to identify eligible studies published prior to 2007. Although this approach may have resulted in missed studies, we think this is unlikely since we hand searched more than 40 systematic reviews. We used a single reviewer to screen titles and abstract; however, we mitigated this risk through reviewer training, quantitative assessment of interrater reliability during initial dual-review of 50 titles/abstracts, and using a low threshold for reviewers to request a second screening by another team member. We only included efficacy outcomes reported at 4 weeks or later; thus, immediate and very short-term benefits are not reflected in our synthesis.

Our grouping of minimally-invasive surgical procedures combines procedures that in fact may be heterogenous. Although the surgical approach used may be slightly different (e.g., direct vs. indirect visualization, use of different ablative techniques), the objective of the procedure (disc removal and decompression) is similar. For outcomes where quantitative synthesis was possible, we did not consistently observe heterogeneity in treatment effects, which suggests that factors

other than the specific type of minimally-invasive intervention may explain the heterogeneity of treatment effect where it was observed.

Lastly, our HTA excluded observational designs, which may provide additional information for safety outcomes that could be more generalizable than data from participants in trials, who generally have fewer comorbidities than the general population.

ES-4.6 Ongoing Research and Future Research Needs

We did not identify any ongoing trials of surgical interventions specifically for lumbar radiculopathy through our search of a clinical trials registry. Several trials related to injections of biologics (e.g., condoliase into nucleus pulposus)[69] or pharmacologics (e.g., epidural clonidine)[70] or use of adjunctive treatments (e.g., epidural steroid injections, stem cell injections, annular repair technologies) during or after discectomy to improve outcomes are ongoing. The challenges faced in conducting methodologically rigorous randomized trials of surgical interventions are well-documented.[71] However, additional trials on treatment of lumbar radiculopathy with the same methodologic flaws will be unlikely to change the certainty of findings. Additional research on patient preferences and values related to timing of treatment or surgery, and establishment of minimally important clinical differences in outcomes that are specific to sciatica would advance research in this area. Finally, advanced analytic and statistical techniques could be used within trials to quantify and mitigate the impact of crossovers on treatment effects and could be used within observational studies to mitigate biases introduced by nonrandomized study designs, potentially broadening the evidence base available to address important research questions.

ES-5. Conclusion

Most findings in this HTA are based on a body of RCT evidence graded as low to very low certainty.

Surgery (discectomy or microdiscectomy) for symptomatic lumbar radiculopathy reduces pain more in the short and medium-term (up to 26 weeks) compared to nonsurgical interventions, but these findings do not persist at one year or longer follow-up. The evidence is insufficient to assess short- and medium-term impact on function because of inconsistent findings, but long-term impact on function suggests no difference between treatments. Surgery compared with nonsurgical interventions result in similar improvements in neurologic symptoms, quality of life, and return to work. No surgery-related deaths were observed and surgery-related complications were rare. The evidence is insufficient to assess outcomes related to persistent opioid use. Surgery compared with nonsurgical interventions may be cost-effective depending on a decision-makers willingness to pay threshold.

Minimally-invasive surgery is comparable to microdiscectomy or discectomy for reducing pain and improving function, quality of life, and neurological symptoms. No surgery-related deaths were observed and surgical morbidity is similar. The evidence is insufficient for drawing conclusions about differences in incidence of reoperations, persistent opioid use, and cost-effectiveness.

Microdiscectomy compared with discectomy are similar with respect to pain reduction, surgical morbidity, and incidence of reoperations, but the evidence is insufficient for drawing conclusions about differences in other efficacy, safety, and cost outcomes.

The evidence is insufficient for drawing conclusions about repeat surgery among individuals with recurrent radiculopathy.

Full Technical Report

Structured Abstract

Purpose: To conduct a health technology assessment (HTA) on the efficacy, safety, and cost of surgery for the treatment of symptomatic lumbar radiculopathy, also referred to as sciatica.

Data Sources: PubMed from January 2007 through November 9, 2017; clinical trial registry; government, payor, and clinical specialty organization websites; hand searches of bibliographies, relevant clinical practice guidelines, and systematic reviews to identify studies published prior to 2007.

Study Selection: Using a priori criteria, we selected English-language primary research studies published in any year that were conducted in very highly developed countries that enrolled adults with symptomatic lumbar radiculopathy and compared surgery for radiculopathy (primarily discectomy or microdiscectomy) to nonsurgical interventions, or that compared alternative surgical procedures, for example minimally-invasive procedures performed percutaneously or with endoscopy, compared with open procedures. We selected trials that reported efficacy outcomes (pain, functioning and disability, quality of life, neurological symptoms, return to work), safety outcomes (mortality, surgical morbidity, reoperations, persistent opioid use), or cost-analyses that reported costs or cost per quality-adjusted life year. We also selected relevant clinical practice guidelines (CPG) for quality appraisal and synthesis.

Data Extraction: One research team member extracted data and a second checked for accuracy. Two investigators independently assessed risk of bias of included primary research studies and conducted a quality assessment of included CPGs.

Data Synthesis: We included 25 primary research studies published between 1983 and 2017. Twenty-four randomized controlled trials (RCTs) provided findings related to efficacy and safety and 7 cost analyses provided findings related to cost-effectiveness. One RCT was rated as low risk of bias, 12 were rated as having some concerns for bias, and 12 were rated as high risk of bias.

Seven RCTs (total number of participants (N) = 1,158) compared microdiscectomy or discectomy to nonsurgical interventions. In these trials, surgery reduced leg pain by 6 to 26 points more than nonsurgical interventions as measured on a 0 to 100-point visual analog scale of patient-reported pain at up to 26 weeks follow-up; differences between groups did not persist at follow-up 1 year or later. The evidence was mixed for functioning and disability as measured by the Oswestry Disability Index, Roland-Morris Disability Questionnaire, and Short Form 36 (SF-36) Physical Functioning subscale in follow-up through 26 weeks, but no between-group differences were observed at 1 year or later. Surgery and nonsurgical interventions produced similar improvements in quality of life, neurologic symptoms, and return to work. No surgical deaths occurred in any study and surgical morbidity was infrequent. The incidence of reoperations among participants who underwent surgery ranged from 0% to 10%. Studies

reported higher quality-adjusted life years for participants who underwent surgery compared to nonsurgical interventions, but similar or higher costs. The average cost per quality-adjusted life year gained from a health care payor perspective ranged from $51,156 to $83,322 in 2010 United States (U.S.) dollars.

Thirteen RCTs (total N = 1,288) compared minimally-invasive surgical procedures to open microdiscectomy or discectomy. In general, minimally-invasive surgery produced similar improvements in pain, function/disability, quality of life, and neurologic symptoms, but resulted in return to work 4 to 15 weeks sooner. No surgical deaths occurred in any trials and with few exceptions, surgical morbidity was similar between groups. The incidence of reoperations across study groups ranged from 2% to 65%; 2 of the 10 trials reporting this outcome reported a statistically significant higher incidence of reoperation among participants who underwent minimally-invasive procedures but the other 8 RCTs reported a similar incidence between groups (pooled relative risk 1.37 [95% CI, 0.74 to 2.52]; 10 RCTs; 1,172 participants; I^2=60.8%). Three RCTs (total N = 282) compared microdiscectomy to discectomy and reported similar improvements pain at outcomes 26 weeks and later. Two RCTs (total N = 160) reported efficacy and safety outcomes of revision surgery for recurrent radiculopathy; findings were mixed.

We identified 14 clinical practice guidelines; the 4 higher quality clinical practice guidelines were in general agreement about recommending discectomy or microdiscectomy (and related decompressive procedures) as acceptable treatment for radiculopathy based on evidence that it improves outcomes in the short- to medium-term.

Limitations: The included RCTs were limited by methodologic designs that increased risk for bias, including extensive participant crossover, lack of participant and outcome assessor blinding, and inadequate randomization and allocation concealment in some studies. Many RCTs either did not report outcomes of interest or were underpowered, leading to imprecision for many effect estimates reported. This HTA was limited to English-language studies; it did not include observational studies or 'as-treated' analyses reported by some RCTs.

Conclusions: Most findings are based on a body of RCT evidence graded as low to very low certainty. Compared with nonsurgical interventions, surgery reduces pain and improves function more up to 26 weeks follow-up, but this difference does not persist at 1 year or longer. Minimally invasive surgery, microdiscectomy, and discectomy are generally comparable with respect to efficacy and surgical morbidity; findings are mixed for reoperations. Surgery may be cost-effective when compared with nonsurgical interventions, depending on a decision maker's willingness to pay threshold, but the evidence is inconclusive about the cost-effectiveness of minimally-invasive surgery.

1. Background

We designed this health technology assessment (HTA) to assist the State of Washington's independent Health Technology Clinical Committee with determining coverage for selected surgical interventions to treat symptomatic lumbar radiculopathy, also known as sciatica.

1.1 Condition Description

Radiculopathy is a clinical syndrome characterized by radiating leg pain, with or without motor weakness, and sensory disturbances in a myotomal or dermatomal distribution. When radicular symptoms occur in the legs, this condition is referred to as lumbar radiculopathy or sciatica. Nerve root compression is a common cause of radiculopathy and various pathological processes may be responsible, but most often it results from disc herniation or spondylosis (i.e., degenerative joint and disc disease).[1-3] Both processes can cause stenosis of the lateral recesses or neural foramina and resulting spinal nerve root compression.[1-3] Degenerative changes can also produce spondylolisthesis, central spinal canal stenosis, and facet joint hypertrophy, which may be associated with radiculopathy and nonradicular low back pain.[1] Less common etiologies of radiculopathy include infection, inflammation, neoplasm, vascular disease, and congenital abnormalities.[1,2] Radiculopathy is a clinical diagnosis because spinal nerve root compression identified with imaging may not always be symptomatic. Thus, correlation of symptoms and physical exam with imaging is usually used to diagnose radiculopathy, with electromyography reserved for selected patients.

Lumbar radiculopathy is a heterogenous condition that may present acutely (as in the case of an acute disc herniation) or more insidiously (as in the case of spondylosis).[2] Further, radiculopathy may present only with pain or with varying degrees of sensory disturbance or motor weakness.[72] Although stenosis of the central spinal canal more commonly presents with neurogenic claudication, it can present with radicular symptoms.[3]

The objective of treatment for radiculopathy is symptom relief through nonsurgical management of symptoms, or surgical intervention to address the underlying causative mechanism, or both. Clinical trials of surgery to address causative mechanisms often specify persistent pain after 6 weeks of conservative management (i.e., medications, physical therapy, epidural steroid injections, etc.) as a patient eligibility criterion for enrollment. In one observational study of 338 patients referred by general practitioners for low back pain or sciatica, 36% had major improvements in symptoms within 2 weeks and 73% had reasonable to major improvements within 12 weeks.[73] A recent systematic review of preoperative predictors for postoperative clinical outcomes following lumbar surgery for disc herniation reported 17 predictors of a favorable outcome, 20 predictors of a nonfavorable outcome, and 15 predictors with conflicting evidence of impact on outcomes.[74] The four predictors with the highest level of supporting evidence for a favorable outcome after surgery include younger age, better mental health, more severe and dominant leg pain (vs. back pain), and absence of worker's compensation claim.

1.2 Disease Burden

Estimates of the incidence and prevalence of lumbar radiculopathy vary widely because of variation in definitions and differences between self-reported and clinically assessed symptoms.[4] A 2008 systematic review of 23 studies assessing sciatica prevalence estimates reported a lifetime prevalence ranging from 3% to 43% (5 studies), a period prevalence over 1 year ranging from 2.2% to 34 % (15 studies), and a point prevalence ranging from 1.6% to 13.4%. (4 studies).[4]

Although some studies report that radiculopathy is distributed equally between men and women,[75] others have found that men are more commonly affected.[1,26] The highest prevalence likely occurs between the ages of 45 and 64 years.[76] Men may be more likely to develop symptoms in their forties while women are at a higher risk in their fifties.[1] Previous history of axial low back pain is an established risk factor for radiculopathy.[26,75] Other risk factors include a prior history of trauma, prolonged driving, pregnancy, and jobs requiring manual labor.[75-77] Several studies have shown a genetic linkage for spinal canal size as well as disc herniation.[75,78]

1.3 Technology Description

The choice of surgical procedure to treat symptomatic lumbar radiculopathy in part depends on etiology and extent of spinal involvement. Discectomy or microdiscectomy may be used to address radiculopathy resulting from disc herniation, whereas laminectomy and other decompressive procedures may be used to address radiculopathy resulting from spondylolysis. ***Table 1*** provides descriptions of the surgical procedures used to treat lumbar radiculopathy.

Standard, open surgical interventions remove parts of the intervertebral disc, with or without additional decompression of spinal nerve root(s) through removal of parts of the bony vertebrae, facet joints (e.g., laminectomy or partial facetectomy) and/or other soft tissues impinging on the nerve root(s). Decompression and disc removal interventions are often performed with a microscope or other magnifying instrument ("micro" approaches). Such an approach makes it possible to minimize the length of incision and area of dissection, thereby reducing the degree of structural alteration to surrounding tissues. Both standard and microsurgical approaches allow for direct visualization of the disc and surrounding structures.

In contrast to open procedures, interventions that use either an endoscopic approach to allow direct visualization of the surgical field and anatomy, or that use a percutaneous approach, which does not allow direct visualization of the disc and surrounding tissue, are also available. These procedures use mechanical (manual or automated), radiofrequency thermal, coblation (also known as plasma), or laser-assisted techniques for disc removal, destruction, or decompression. Although the terms used for procedures in this HTA may vary, for this report, we refer to these procedures as 'minimally-invasive" surgical procedures. Minimally-invasive procedures allow for a smaller-incisions and less tissue damage relative to open procedures.

Table 1. Description of surgical interventions used to treat lumbar radiculopathy

Category of Intervention	Examples
Open decompression procedures	• Laminectomy is the removal of part or all the lamina portion of the vertebra bone and ligaments to reduce compression on nerve roots. • Microlaminectomy is the removal of part or all the lamina portion of the vertebra bone, but using a microscope or magnification to allow for a smaller incision and area of dissection to minimize disruption to surrounding bones, joints, ligaments, and muscles. • Laminotomy is the creation of an opening in the lamina portion of the vertebra bone to reduce compression on nerve roots. • Foraminotomy is the creation of a wider neuroforaminal opening of the vertebra to reduce compression within the neural foramina.
Open disc removal procedures	• Discectomy is the removal of some or all the intervertebral disk to reduce compression on nerve roots, and is sometimes combined with laminotomy or laminectomy or other procedures. • Microdiscectomy is removal of some or all the intervertebral disk using a microscope or magnification to allow for a smaller incision and area of dissection to minimize disruption of the surrounding bones, joints, ligaments, and muscles.
Minimally-invasive procedures	These procedures are designed to reduce nerve root compression but do not structurally alter the bony spine. Some are performed either percutaneously or endoscopically. The core objective of these procedures is the same as open procedures, specifically disc removal or ablation and nerve root decompression. • Percutaneous and/or endoscopically performed procedures using mechanical (manual or automated), radiofrequency thermal, coblation (also known as plasma), or laser-assisted techniques for disc removal, destruction, or nerve root decompression. Includes procedures termed nucleotomy, nucleoplasty, sequestrectomy, discectomy, and discoplasty. • Chemonucleolysis is the injection of enzymes (e.g., chymopapain) or other chemical substances into the nucleus pulposus to induce disc shrinkage and reduce compression. This procedure is no longer in routine clinical use in the U.S.

1.4 Regulatory Status

The United States (U.S.) Food and Drug Administration (FDA) regulates some surgical instruments and devices used in spine-related surgery. FDA-cleared electrosurgical cutting instruments, aspiration or coagulation devices, endoscopes, and other related accessories can be found under product codes GEI, GXI, HRX, BSO, and BSP. The following summary is not an all-inclusive list. All devices referred to here were cleared by the FDA through the 510(k) process, which is based on evidence that the device is 'substantially equivalent' to a device that the FDA has already cleared or that was marketed before 1976. None were approved through the premarket approval process, which requires manufacturers to demonstrate that the device is safe and effective, a higher standard than the 510(k) clearance process.

The FDA has cleared several devices for aspiration during percutaneous discectomies including the DeKompressor® Percutaneous Discectomy Probe[79] (Stryker, November 7, 2003), Herniatome Percutaneous Discectomy Device[80] (Gallini Medical Devices, December 8, 2014), Nucleotome®[81] (Clarus Medical, June 1, 2004), and Laurimed Percutaneous Discectomy System[82] (Laurimed LLC, August 28, 2008). These are all Class II devices indicated for "aspiration of disc material during percutaneous discectomies in the lumbar, thoracic and cervical regions of the spine." The SpineView ENSPIRET™[83] Debrider (SpineView Inc, April 15, 2009) and the enSpire™ Discectomy System[84] (Spine View Inc, June 26, 2012) are Class II

devices indicated for use in "cutting, grinding and aspirating intervertebral disc material during discectomy procedures in the cervical, thoracic and lumbar spine."

Numerous devices have been cleared for ablation and coagulation. A commonly noted device in several payer coverage policies is the Disc-FX™ system[85] (Ellman International Inc, February 27, 2006), a Class II device indicated for "ablation and coagulation of intervertebral disc material during discectomy procedures in the lumbar spine." Several Class II catheters, such as the SpineCATH Intradiscal Catheter[86] (ORATEC Interventions Inc, December 17, 1999) and the Nucleotomy Catheter[87] (ORATEC Interventions Inc, January 31, 2002), have been cleared for the indicated use of "coagulation and decompression of disc material." One included study in this HTA[41] used the Perc-D SpineWand™[88] (Arthrocare, August 22, 2013), a Class II device indicated for "ablation, coagulation, and decompression of disc material," to conduct percutaneous disc decompression with coblation technology.

One study included in this HTA[29] utilized the Metr'X system arthroscope (Medtronic Sofamor Danek, Inc., November 24, 2000).[89] Frequently cited in insurance coverage policies, the Yeung Endoscopic Spine System (Richard Wolf Instrument Company, March 13, 1998)[90] is a Class II device indicated for the "visualization and removal of herniated discs in the lumbar region." Though also often discussed in payer policies as a medical device, the AccuraScope™ Discectomy and Neural Decompression (DND) procedure has been marketed by North American Spine, a private provider of minimally-invasive spine surgery,[91] but is not FDA-regulated.

FDA has cleared laser instruments for incision, excision, resection, ablation, vaporization, and coagulation of tissue during surgical procedures including but not limited to discectomy. One included study in this HTA[39] utilized the Ceralas 980 Diode Laser[92] (Biolitec Inc, January 24, 2008), a Class II device indicated for "delivery of laser light to soft tissue in the contact or non-contact mode during surgical procedures, including via endoscopes, introducers, or catheters," to conduct percutaneous laser disc decompression. For more information on coverage of surgical procedures that use laser devices, see the Selected Payer Coverage Policies section of this report.

1.5 Policy Context

Numerous surgical and nonsurgical approaches to the management of lumbar radiculopathy are routinely used within current clinical practice. In addition to standard open surgical techniques (e.g., laminectomy, microdiscectomy), minimally-invasive surgical techniques that use percutaneous, endoscopic, or laser-assisted approaches are now available. The State of Washington Health Care Authority selected surgery for lumbar radiculopathy as a topic for an HTA based on medium concerns for efficacy, medium concerns for safety, and high concerns for cost. This HTA is designed to assist the State of Washington's Health Technology Clinical Committee in determining coverage for selected surgical interventions to treat symptomatic lumbar radiculopathy.

1.6 Washington State Agency Utilization Data

The State of Washington Health Care Authority provided data on the use of surgical procedures for the treatment of radiculopathy for the time period 2015 to 2017. This data is provided in

Appendix A. Data is provided for Medicaid (fee for service and managed care organization), Department of Labor and Industries Workers' Compensation Program, and the Public Employee Benefit Board Uniform Medical Plan, including Medicare.

2. Methods

This HTA includes two separate, but related components. The first component is a systematic review of primary research studies and the second component is a quality appraisal and synthesis of relevant clinical practice guidelines.

2.1 Research Questions and Analytic Framework for Systematic Review of Primary Research Studies

We developed the following research questions and analytic framework (***Figure 1***) to guide the systematic evidence review of primary research studies:

Efficacy Question 1 (EQ1). In adults with symptomatic lumbar radiculopathy, what is the effectiveness and comparative effectiveness of surgical interventions?

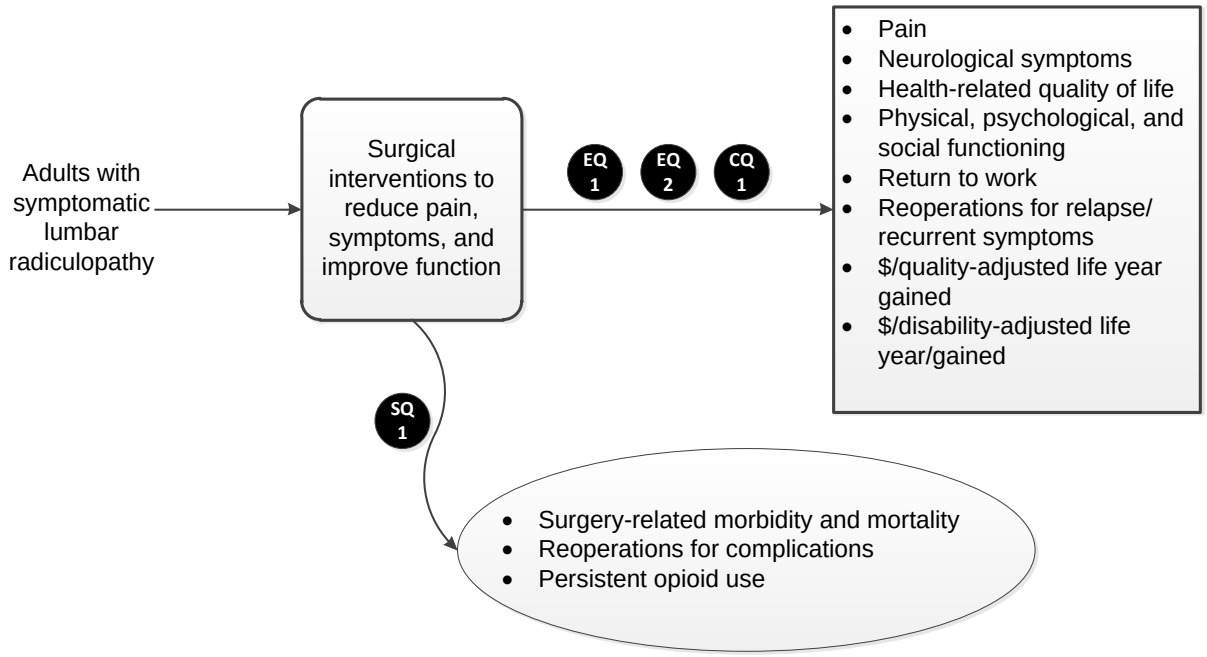
Efficacy Question 2 (EQ2). In adults with symptomatic lumbar radiculopathy, does effectiveness or comparative effectiveness of surgical interventions vary for patients who are not employed because of disability or patients who are undergoing recurrent surgery for relapse?

Safety Question 1 (SQ1). In adults with symptomatic lumbar radiculopathy, what are the adverse events associated with surgical interventions?

Cost Question 1 (CQ1). In adults with symptomatic lumbar radiculopathy, what is the cost-effectiveness of surgical interventions?

The State of Washington Health Technology Assessment Program posted a draft of these research questions with study selection criteria for public comment from November 14, 2017 to November 27, 2017. The final key questions and response to public comments on the draft key questions is available at the Program's website.[93] A draft version of this evidence report was reviewed by two independent, external peer reviewers and was also posted for public comment from February 22, 2018 until March 23, 2018. Feedback from peer reviewers and from public comments was incorporated into this Final Evidence Report; responses to public and peer review comments are summarized in a separate document also available at the Program's website.[93]

Figure 1. Analytic framework for HTA on surgery for lumbar radiculopathy



Abbreviations: CQ=cost question; EQ=efficacy question; SQ=safety question

2.1.1 Data Sources and Searches

We searched MEDLINE® (via PubMed) from 2007, the Cochrane Library, and a clinical trials registry (clinicaltrials.gov) for relevant English-language studies. We searched the Centers for Medicare and Medicaid Services and FDA websites, selected payer and health care professional society websites, and websites of other organizations that conduct and disseminate HTAs. In addition, we reviewed the reference lists of relevant studies, systematic reviews, practice guidelines, and other HTAs on this topic to identify any relevant primary research studies not found through the electronic search and to identify studies published prior to 2007. The detailed search strategy is provided in ***Appendix B.***

In brief, we used medical subject headings (MeSH terms) and text words associated with the surgical interventions of interest combined with MeSH terms for radiculopathy and lumbar disc disease. We limited the search by eliminating studies indexed using terms for infants, children, or adolescents, selected animals, and conditions that indicate excluded populations, such as cancer, tuberculosis, fracture, scoliosis, spondylolysis, and cervical vertebrae. We used MeSH terms to select studies most likely to be trials or systematic reviews and to remove editorials, letters, and publication types that do not represent primary research studies.

2.1.2 Study Selection

Table 2 summarizes the study selection criteria related to the population, intervention, comparator, outcomes, time period, and setting that defined the scope of this HTA; these are further described following the table. We screened titles and abstracts and full-text articles based on these study selection criteria. Four team members independently screened the same initial 50 titles/abstracts. Because we had substantial interrater reliability (Light's Kappa = 0.84),[94] the

remaining titles/abstracts were divided among team members and a single reviewer screened them. The lead investigator and one additional team member independently screened all full-text articles; discrepancies were resolved by discussion.

Table 2. Population, intervention, comparator, outcome, timing, setting and other study selection criteria for HTA on surgery for lumbar radiculopathy

Domain	Included	Excluded
Population	Adults age 18 years and over with symptomatic lumbar radiculopathy (i.e., sciatica) unrelated to infection, cancer, inflammatory, congenital, or traumatic etiologies. For studies of mixed populations, results must be stratified and reported separately for patients with lumbar radiculopathy.	Adults with: • Cervical or thoracic radiculopathy • Cauda equina syndrome • Neurogenic claudication or low back and leg symptoms related primarily to central spinal stenosis • Spondylolisthesis • Traumatic or congenital structural spinal abnormalities • Nonradicular leg or low back pain (i.e., discogenic or other nonspecific low back pain)
Intervention	Surgical interventions for the treatment of radiculopathy, for example: • Discectomy • Laminectomy, laminotomy • Foraminotomy • Nucleotomy Includes "micro" approaches to the above procedures, which involve smaller incisions and/or areas of dissection and/or use of microscope or loupe magnification. Minimally-invasive surgical procedures designed for treating radicular pain: percutaneous discectomy, discoplasty, nucleotomy, or nucleoplasty that are manual, automated, endoscopic, or laser-assisted, or use radiofrequency heat or coblation technology. Interventions involving combinations of the above interventions are eligible.	Surgical interventions primarily designed to treat neurogenic claudication and central spinal stenosis, spinal instability, or nonradicular low back pain, for example: • Spinal fusion • Arthroplasty • Artificial disc replacement • Interspinous process decompression (e.g., X-STOP® IPD System,[5] Coflex® Interlaminar Technology)[6] • Minimally-invasive lumbar decompression (mild® procedure)[7] • Other minimally-invasive procedures designed for treating discogenic (i.e., nonradicular) low back pain Epidural, spinal, or disc injections of enzymatic (e.g., chymopapain), chemical, or biologic (e.g., stem cells, mesenchymal cells) agents. Interventions involving combinations of procedures that include an above intervention are ineligible.
Comparator	*Placebo or no treatment comparators:* sham surgery, expectant management, no treatment *Active treatment comparators:* • Nonsurgical management (e.g., physical therapy, exercise, pharmacologic treatment of symptoms, spinal manipulation, chiropractic treatment, epidural steroid or pain injections, other noninvasive treatments) • Surgical interventions as listed under "intervention"	• No comparator • Chemonucleolysis Studies using "usual care" comparator groups will not be excluded but will be synthesized separately if no information was provided about the components of "usual care."

(continued)

Table 2. Proposed population, intervention, comparator, outcome, timing, and setting for HTA on surgical interventions for lumbar radiculopathy (continued)

Domain	Included	Excluded
Outcomes	*Efficacy (at 4 weeks post-op or later):* • Pain • Physical functioning • Social functioning • Psychological/emotional distress • Health-related quality of life • Neurologic symptoms (e.g., weakness, sensory alteration) • Return to work *Safety:* • Surgery-related morbidity (e.g., venous thromboembolism, paralysis, new neurological symptoms, dural tear, epidural hematoma) • Surgical mortality (30 day) • Reoperations • Persistent opioid use *Cost and cost-effectiveness:* • Direct medical costs • Cost per quality-adjusted life year gained • Cost per disability-adjusted life year gained	Other outcomes not specifically listed as eligible. Pain, quality of life, and functional outcomes not measured using valid and reliable instruments or scales.[8,9]
Setting	Inpatient or outpatient settings in countries categorized as "very high" on United Nations Human Development Index	Studies conducted in countries not categorized as "very high" on United Nations Human Development index.
Study Design and Risk of Bias Rating	*For all Efficacy and Safety Research Questions:* CCTs, RCTs, and SRs of CCTs or RCTs with similar scope as this HTA. For studies using surgical interventions as active comparators, only RCTs or SRs of RCTs will be included. *For Cost-Effectiveness Questions*: CEA, CUA, or CBA performed from the societal or payer perspective *For all studies:* Intent-to-treat analyses. Studies with any risk of bias rating will be included, but high risk of bias studies will only be used in quantitative syntheses if fewer than 3 studies are available.	Editorials, comments, letters, narrative reviews, case reports, case series, cohort studies, case-control studies. As-treated or per-protocol analyses reported by RCTs.
Language and Time Period	English language, any time period.	Languages other than English.

Abbreviations: CCT = controlled clinical trial; CEA = cost-effectiveness analysis; CUA = cost-utility analysis; CBA = cost-benefit analysis; HTA = health technology assessment; RCT = randomized controlled trial; SR = systematic review.

2.1.2.1 Population

Studies were selected if they enrolled adults age 18 years or over with symptomatic lumbar radiculopathy unrelated to infection, cancer, inflammatory, congenital or traumatic etiologies. Studies that enrolled participants with lumbar spinal stenosis and neurogenic claudication, spondylolisthesis, or chronic discogenic (i.e., nonradicular) low back pain were excluded.

2.1.2.2 Intervention and Comparator

For all research questions, comparative studies where at least one study group included a surgical intervention to relieve lumbar radicular pain were eligible for selection. This included standard open discectomy with or without laminotomy and laminectomy, foraminotomy, sequestrectomy, or nucleotomy. Any forms of these surgeries including minimally-invasive approaches (e.g., endoscopic or percutaneous approaches), use of microsurgical techniques (e.g., microdiscectomy), or laser-assisted procedures were eligible. Comparison groups that were placebo or no treatment comparators or active treatment comparators were eligible for selection. Active treatment comparators could include nonsurgical management (e.g., analgesics, physical therapy, spinal manipulation, epidural injection, etc.) or could include an alternative surgical intervention (e.g., endoscopic discectomy compared with standard open discectomy).

2.1.2.3 Outcomes

For the research questions on efficacy (EQ1, EQ2), studies that reported outcomes related to pain, quality of life, and functional outcomes were eligible for selection and we required studies to use valid and reliable measures of these constructs (e.g., Short Form 36 (SF-36), Roland-Morris Disability Questionnaire, etc.) for use within our strength of evidence ratings. In addition, studies that reported on change in neurologic symptoms and return to work were also eligible for selection. For the research question on safety (SQ1), studies that reported on perioperative or postoperative morbidity and mortality, reoperations, or persistent opioid use were eligible for selection. For the research question on cost (CQ1), studies that reported costs or cost-effectiveness measures, specifically cost per quality-adjusted life year gained (QALY) or cost per disability-adjusted life year gained (DALY) were eligible for selection.

2.1.2.4 Settings

Studies conducted in any inpatient or outpatient clinical setting were eligible for selection. Studies that were conducted in countries designated as "very high" by the United Nations Human Development Programme were eligible for selection as these countries (Europe, Australia, New Zealand, Japan, S. Korea, Singapore, Hong Kong, and selected Middle Eastern countries) are like the United States with respect to standards of medical practice.[10] We excluded studies conducted in countries designated as less than "very high."

2.1.2.5 Study Design

Nonrandomized controlled clinical trials, randomized controlled trials (RCTs), and systematic reviews of trials were eligible for selection. However, we required trials comparing two alternative surgeries to be randomized. Cohort studies, studies that analyzed randomized trial data with cohort study data together, case series, and case reports were not eligible for selection.

2.1.2.6 Time Period

We did not restrict included studies based on year conducted or published.

2.1.3 What is Excluded from This HTA

This review did not include studies published in languages other than English or conducted in countries that are not very highly developed based on the United Nations Human Development Index.[10] This review did not include studies conducted among children or adolescents. This review was designed to focus primarily on surgery to treat symptomatic radiculopathy, and we excluded studies evaluating surgical interventions performed primarily to manage central spinal canal stenosis (e.g., neurogenic claudication), spondylolisthesis, or traumatic or congenital abnormalities. Further, this review did not cover surgical interventions for low back pain that was not radicular in nature (e.g., chronic discogenic pain). We refer readers to the State of Washington's 2015 Health Technology Assessment Final Evidence Report on Lumbar Fusion for Patients with Degenerative Disc Disease Uncomplicated by Comorbid Spinal Conditions available at the Program website.[11] Lastly, this review did not include observational study designs (e.g., case series, comparative cohort studies) or 'as treated' or 'per protocol' analyses from RCTs because these analyses have a high risk of bias relative to intent-to-treat analyses from RCTs.

2.1.4 Data Abstraction and Risk of Bias Assessment

One team member extracted relevant study data into a structured abstraction form and the lead investigator checked it for accuracy. For consistency in reporting findings across studies, we transposed some treatment effects reported in studies to ensure all our abstracted data represented the effect of the intervention group relative to the comparator group. We used the Cochrane Risk of Bias (RoB 2.0) tool to assess the risk of bias for each included trial.[12] Domains assessed with this tool include: bias arising from randomization process, bias due to deviations from intended interventions, bias due to missing outcome data, bias in measurement of the outcome, and bias in selection of the reported result. Risk of bias was assessed as "high," "some concerns," or "low" at the study level unless different outcomes within a single study required outcom-level risk of bias ratings. We used the Quality of Health Economic Studies Instrument to assess the quality of included cost analyses.[13] We considered studies with scores on this instrument above 90 to be good quality, studies with scores between 60 and 89 to be fair quality, and studies with scores below 60 to be poor quality. Two team members conducted independent risk of bias or quality assessments on all included studies; discrepancies were resolved by discussion, in consultation with the lead investigator if needed.

2.1.5 Data Synthesis and Strength of Evidence Rating

Study characteristics and results from intent-to treat analyses were qualitatively synthesized for each research question in tabular and narrative formats. We synthesized studies comparing the surgical interventions to nonsurgical interventions separately from studies comparing alternative surgical interventions. We summarized continuous outcome measures as absolute mean differences (AMD) between treatment groups wherever possible. When studies did not report the AMD for critical outcomes, we calculated it and the 95% confidence intervals (CI) if the appropriate data were reported in the article to be able to do so (e.g., mean, standard deviation

(SD) for each group). We summarized categorical outcomes using differences in proportions between groups, and when studies did not report tests of statistical significance we calculated P values using Fisher's exact test. We identify all values that we calculated in the text and tables as "calculated" values. For cost outcomes, we adjusted all reported outcomes in foreign currency to U.S. dollars based on the U.S. Department of Treasury mid-year exchange rate for the year reported by study authors and then used the chain-weighted consumer price index (CPI) to adjust to 2010 U.S. dollars (***Appendix C***).[14,15]

To determine whether quantitative synthesis was appropriate, we assessed the number of studies and the clinical and methodological heterogeneity present based on established guidance.[95,96] We required three or more studies with similar intervention and comparator with same outcome measure at approximately the same follow-up time point to calculate a pooled treatment effect. We considered outcomes reported at less than 12 weeks to be short-term, outcomes reported between 12 weeks up to 52 weeks as medium-term, and outcomes reported at 52 weeks or later as long-term. We estimated pooled effects using a random effects model with the 'metafor' package in R using the DerSimonian and Laird method.[16] We assessed statistical heterogeneity with the I^2 statistic; an I^2 between 0% and 40 % might not be important, 30% to 60% may represent moderate heterogeneity, and 50% to 90% percent may represent substantial heterogeneity.[17] We investigated heterogeneity qualitatively based on factors such as risk of bias, type of intervention, and type of comparator.

We graded the strength of evidence for each research question and outcome measure using a modification to GRADE, which assesses the strength of evidence based on domains relating to risk of bias, inconsistency, imprecision, indirectness, and other considerations, such as publication bias.[18] For each outcome measure, we rated the evidence for between-group differences in short-, medium-, and long-term outcomes separately when required because of differences in GRADE domains at different follow-up time periods. With GRADE, the strength of evidence can be graded as "very low," "low," "moderate," or "high", and this rating represents the overall certainty of the findings. ***Table 3*** defines these levels of certainty.[19] We modified the GRADE approach to allow for a rating of "insufficient" for single-study bodies of evidence with very serious concerns in one or more domains, or when we are unable to draw a conclusion about the treatment effect because of inconsistent findings.

Table 3. Strength of evidence grades and definitions[19]

Grade	Definition
High	**We are very confident that the estimate of effect lies close to the true effect for this outcome.** The body of evidence has few or no deficiencies. We believe that the findings are stable, that is, another study would not change the conclusions.
Moderate	**We are moderately confident that the estimate of effect lies close to the true effect for this outcome.** The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
Low	**We have limited confidence that the estimate of effect lies close to the true effect for this outcome.** The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
Very Low	**We have very limited confidence that the estimate of effect lies close to the true effect for this outcome.** The body of evidence has numerous major deficiencies. We believe that substantial additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
Insufficient	**We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome**. No evidence is available or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

To assess the consistency domain within GRADE, we evaluated both the consistency in the direction and magnitude of treatment effect. We judged whether the effect was consistent based on whether the evidence consistently supported a minimally important difference (MID) between intervention and comparator groups, or consistently supported no meaningful difference. We identified the range of MIDs for key outcomes a priori based on the literature. When a quantitative synthesis was possible, we also used the I^2 statistic to assess consistency.

To assess the precision domain within GRADE, we calculated the sample size requirements to detect a range of MID from the literature based on 80% power, alpha level of 0.05, and two-tailed tests using STATA version 14.0 (Stata Corp). Bodies of evidence generally not meeting the sample size requirement to detect the lower end of the MID threshold were downgraded one level (e.g., not serious to serious concern); bodies of evidence not meeting the sample size requirement to meet the upper end of the MID threshold were downgraded by two levels (i.e., not serious to very serious concern).

When multiple outcome measures within a clinical domain were reported (e.g., pain, function), we only graded strength of evidence for the clinical measures with known validity and reliability and that were reported by at least 2 studies. To draw overall conclusions about a clinical domain reporting multiple measures we considered all strength of evidence ratings within the domain.

2.2 Clinical Practice Guideline Synthesis

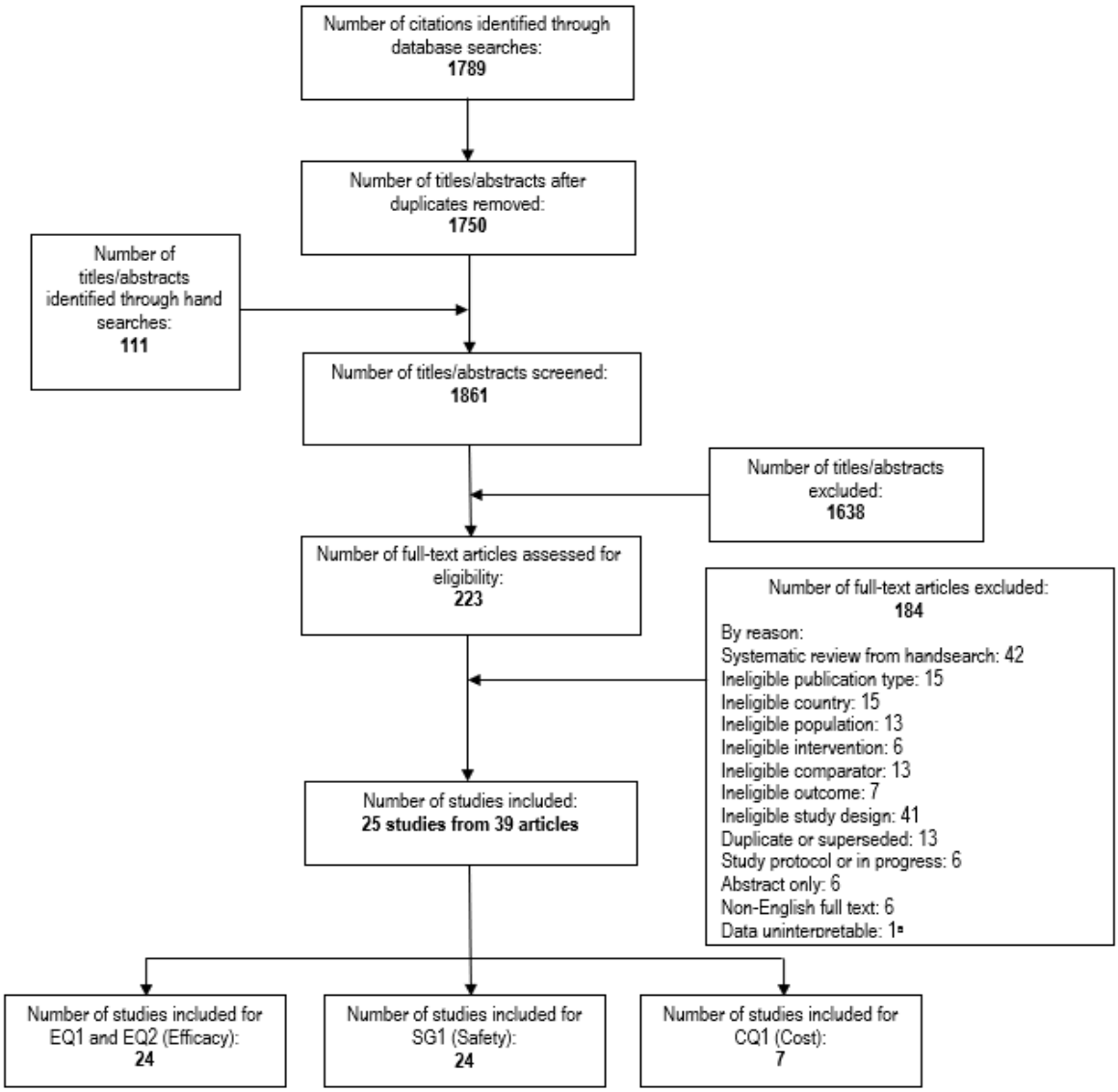
In addition to the systematic evidence review portion of this HTA, we also identified relevant clinical practice guidelines and conducted a quality assessment of each guideline using the Appraisal of Guidelines for Research & Evaluation II (AGREE) instrument.[20,21] With this instrument, six domains are assessed and an overall score of between 1 (lowest possible) and 7 (highest possible) are assigned to reflect the overall quality of the guideline. We synthesized clinical practice guidelines in a tabular format.

3. Results

3.1 Literature Search

Figure 2 depicts the study flow diagram. We identified and screened 1,861 unique citations. We excluded 1,638 citations after title and abstract review. We reviewed the full-text of 223 articles and included a total of 25 studies reported in 39 articles published between 1983 and 2017. Twenty-two RCTs provided evidence on efficacy or comparative effectiveness (EQ1), two RCTs provided evidence on the effectiveness or comparative effectiveness of revision surgical interventions for relapse (EQ2), 24 RCTs provided evidence on safety (SQ1), and seven studies (six RCTs and one cost-effectiveness analysis) provided evidence on costs or cost-effectiveness (CQ1).

Figure 2. Study flow diagram for HTA on surgery for lumbar radiculopathy



[a] We contacted the study author for clarification but did not receive a reply.

Individual study and population characteristics and findings for all included studies are summarized in ***Appendix D***. The list of articles we screened at the full-text stage, but which we excluded, is provided in ***Appendix E***. Note that articles may have been excluded based on more than one reason but we report only one reason. We report our individual study risk of bias assessments for included studies in ***Appendix F***.

3.2 Efficacy

3.2.1 Efficacy Question 1

In adults with symptomatic lumbar radiculopathy, what is the effectiveness and comparative effectiveness of surgical interventions?

We included 22 RCTs. Seven RCTs provided evidence for the efficacy of surgery compared with nonsurgical treatment; study and population characteristics are summarized in ***Table 4***.[22,23,26,32,33,37,41] Of the 15 RCTs providing evidence for the comparative effectiveness of alternative surgical interventions; 12 compared minimally-invasive surgery to standard microdiscectomy or discectomy (***Table 5***),[24,25,28,30,31,36,38-40,42-44] two compared microdiscectomy to discectomy (***Table 6***) [27,35] and one compared minimally-invasive surgery to both microdiscectomy and to discectomy.[29] ***Tables D-1*** and ***D-2*** in ***Appendix D*** provide detailed individual study and population characteristics and ***Tables D-3*** and ***D-4*** provide detailed individual study outcomes related to efficacy.

3.2.1.1 Study Characteristics

All 22 included studies were parallel-group RCTs. Four were conducted in the U.S.;[22,41-43] the rest were conducted in Canada (N=1),[23] Taiwan (N=1),[24] Japan (N=1)[25] or various European countries (N=15).[26-40] The total number of participants randomized ranged from as few as 21 to as many as 501. The mean age of participants generally ranged from mid-30s to mid-40s. All studies enrolled both men and women with signs or symptoms of lumbar radiculopathy and confirmatory imaging, usually computerized tomography (CT) or magnetic resonance imaging (MRI). Most studies excluded participants with indications for immediate surgery, for example, cauda equina syndrome. Only six studies reported the proportion of participants disabled at baseline,[22,23,27,32,39,40] and they did not use a consistent definition of disability. The mean duration of symptoms at baseline ranged from 8 weeks to 2 years among the 13 included studies reporting this population characteristic.[23,27-29,31-34,38-41,43]

Table 4. Study and population characteristics of the seven randomized controlled trials comparing surgery to nonsurgical interventions for management of lumbar radiculopathy (EQ1)

Author (Year); Country; Risk of Bias	**Population Characteristics Age, mean (SD); Women N (%); Duration of symptoms, mean (SD)**	***Surgical Intervention* (SG1); N randomized; N analyzed (% of randomized); N crossovers (% of randomized)**	***Comparator(s)* (NS1); N randomized; N analyzed (% of randomized); N crossovers (% of randomized)**	***Primary Outcome* (effect size detectable with 80% power, α = 5%)[a]; • Other outcomes**
Erginousakis (2011)[37]; Greece; High	Age SG1: 38 (4.2); NS1: 36 (5.8) Women: SG1: 12 (38.7%); NS1: 14 (45.2%) Duration of symptoms: NR	*Percutaneous disc* decompression N randomized: 31; N analyzed: 31 (100%); N crossovers: 0 (0%)	*Conservative management* N randomized: 31; N analyzed: 31 (100%); N crossovers: 0 (0%)	*VAS 10 Pain (NR);* • N (%) with category of pain reduction • N (%) reporting pain affected occupational status
Gerszten (2010)[41] United States; Some concerns (6w outcomes); High (12w and later outcomes)	Age SG1: 46 (12); NS1: 42 (11) Women SG1: 24 (53%); NS1: 19 (48%) Duration of symptoms: SG1: 52w (range 4w to 16y) NS1: 2y (range 10w to 13y)	*Plasma disc decompression with coblation technology* N randomized: 46; 45 ITT sample; N analyzed: 29 (64% of ITT sample) at 26w; N crossovers: 12 were unresolved and received a second, unspecified procedure.	*Epidural steroid injection* N randomized: 44; 40 ITT sample; N analyzed: 28 (70% of ITT sample) at 26w; N crossovers: 8 were unresolved and received a second, unspecified procedure.	*VAS 100 leg pain (15 points);* • VAS 100 back pain • SF-36 Bodily Pain • Oswestry Disability Index • SF-36 Physical Functioning • SF-36 Physical and Mental Component Summary • Other SF-36 subscales • N (%) with full muscle strength • N (%) with normal tactile sensitivity • N (%) working full or part-time
McMorland (2010)[23] Canada; Some concerns	Age SG1 Men: 42.85 (NR); SG1 Women: 40.1 (NR); NS1 Men: 36.4 (NR); NS1 Women: 48.33 (NR) Women SG1: 7 (35%); NS1: 9 (45%) N (%) with duration of complaint 12-26w SG1: 3 (15%); NS1: 6 (30%)	*Microdiscectomy* N randomized: 20; N analyzed: 20 (100%) (outcomes reported only to 12w for ITT analysis) 24w: 20 (100%) 52w: 15 (75%); N crossovers: 3 (15%) enrolled in spinal manipulation 26-34w after surgery so received both.	*Spinal manipulation* N randomized: 20; N analyzed: 20 (100%) (outcomes reported only to 12w for ITT analysis) 24w: 20 (100%) 52w: 17 (85%); N crossovers: 8 (40%) underwent microdiscectomy after 12w of spinal manipulation care.	*Aberdeen Back Pain Scale (6 points);* • McGill Pain Questionnaire • Roland-Morris Disability Questionnaire • SF-36 Bodily Pain • SF-36 Physical Functioning • Other SF-36 subscales • SF-36 Total Score

(continued)

Table 4. Study and population characteristics of the seven randomized controlled trials comparing surgery to nonsurgical interventions for management of lumbar radiculopathy (EQ1) (continued)

Author (Year); Country; Risk of Bias	Population Characteristics Age, mean (SD); Women N (%); Duration of symptoms, mean (SD)	*Surgical Intervention* (SG1); N randomized; N analyzed (% of randomized); N crossovers (% of randomized)	*Comparator(s)* (NS1); N randomized; N analyzed (% of randomized); N crossovers (% of randomized)	*Primary Outcome* (effect size detectable with 80% power, α = 5%)[a]; • Other outcomes
Osterman (2003)[33] Finland; High	Age SG1: 37 (7); NS1: 38 (7) Women; SG1: 13 (46.4%) NS1: 9 (32.1%) Duration of leg pain, mean (SD) SG1: 11.0w (4.6); NS1: 8.6w (3.0)	*Microdiscectomy* N randomized: 28; N analyzed 6w: 26 (93%) 26w: 26 (93%) 52w: 21 (75%) 2y: 26 (93%); N crossovers: 0 (0%)	*Physiotherapy* N randomized: 28; N analyzed 6w: 26 (93%) 26w: 22 (78.6%) 52w: 20 (71.4%) 2y: 24 (86%); N crossovers: 11 (39.3%)	*VAS 100 leg pain (15 points);* • VAS 100 back pain • VAS 100 Work Ability Score • Oswestry Disability Index • 15D HRQOL • N (%) with muscle weakness • N (%) reporting full recovery
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98] Netherlands; High	Age SG1: 41.7 (9.9); NS1: 43.4 (9.6) Women SG1: 52 (37%); NS1: 45 (32%) Duration of symptoms, mean (SD) SG1: 9.43w (2.37w); NS1: 9.48w (2.11w)	*Microdiscectomy* N randomized: 141; N analyzed 52w: 140 (99.3%) 2y: 130 (92.2%) 5y: 115 (81.6%); N crossovers 52w: 16 (11.3%) 2y: 16 (11.3%) 5y: 16 (11.3%);	*Conservative management* N randomized: 142; N analyzed 52w: 141 (99.3%) 2y: 130 (91.5%) 5y: 116 (81.7%); N crossovers 52w: 55 (38.7%) 2y: 62 (43.7%) 5y: 66 (46.5%);	*Roland-Morris Disability Questionnaire (90% power, 3 points)*; • VAS 100 leg pain • VAS 100 back pain • Sciatica index • SF-36 Bodily Pain • SF-36 Physical Functioning • Other SF-36 subscales • Prolo Scale • Median time to recovery • Global perception of recovery
Weber (1983)[26] Norway; High	Age SG1: 40.0 (NR); NS1: 41.7 (NR) Women SG1: 28 (46.7%); NS1: 30 (45.5%) Duration of symptoms: NR	*Discectomy* N randomized: 60; N analyzed: 60 (100%); N crossovers: 1 (1.7%)	*Conservative management* N randomized: 66; N analyzed: 66 (100%); N crossovers: 17 (25.8%)	*NR (NR);* • N (%) with categories of radiating pain • N (%) with permanent incapacitation and receiving disablement benefits • N (%) with category of result (Good, Fair, Poor Bad)

(continued)

Table 4. Study and population characteristics of the seven randomized controlled trials comparing surgery to nonsurgical interventions for management of lumbar radiculopathy (EQ1) (continued)

Author (Year); Country; Risk of Bias	Population Characteristics Age, mean (SD); Women N (%); Duration of symptoms, mean (SD)	*Surgical Intervention* (SG1); N randomized; N analyzed (% of randomized); N crossovers (% of randomized)	*Comparator(s)* (NS1); N randomized; N analyzed (% of randomized); N crossovers (% of randomized)	*Primary Outcome* (effect size detectable with 80% power, α = 5%)[a]; • Other outcomes
Weinstein (2006) [22] Weinstein (2008)[99] Lurie Jon (2014)[100] United States; High	Age SG1: 41.7 (11.8); NS1: 43.0 (11.3) Women SG1: 101 (44%); NS1: 93 (39%) Duration of symptoms: NR	*Discectomy/microdiscectomy* N randomized: 245; N analyzed: 232 in main study's primary analyses. N crossovers: Cumulative over time: 26w: 113 (46.1%) 2y: 105 (42.9%) 8y: 97 (39.6%)	*Conservative management* N randomized: 256; N analyzed: 240 included in main study's primary analyses. N crossovers: Cumulative over time: 26w: 93 36.3%) 2y: 107 (41.8% 8y: 122 (47.7%)	*SF-36 Bodily Pain, SF-36 Physical Functioning, Oswestry Disability Index (85% power, 10 points in SF-36 scales or "similar effect size in the Oswestry Disability Index");* • Sciatica index • N (%) working full or part time • N (%) satisfied with symptoms • N (%) with self-rated major improvement

[a] As specified and reported by study authors.

Abbreviations: HRQOL = health-related quality of life; ITT = intent to treat; N = number; NR = not reported; NS= nonsurgical group; SD = standard deviation; SF-36 = Short Form 36; SG = surgical group; VAS = visual analog scale; w = week(s); y = years(s).

Table 5. Study and population characteristics of the twelve randomized controlled trials comparing minimally-invasive surgery to standard surgery for management of lumbar radiculopathy (EQ1)

Author (Year); Country; Risk of Bias	Population Characteristics Age, mean (SD); Women N (%); Duration of symptoms, mean (SD)	*Surgical Intervention* (SG1); N randomized; N analyzed (% of randomized); N crossovers (% of randomized);	*Comparator(s)* (SG2, SG3); N randomized; N analyzed (% of randomized); N crossovers (% of randomized);	*Primary Outcome* (effect size detectable with 80% power, α = 5%)[a]; • Other outcomes
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48] The Netherlands; Low	Age SG1: 41.6 (9.8); SG2: 41.3 (11.7) Women SG1: 82 (49%); SG2: 71 (45%) Duration of symptoms, median (range) SG1: 21.0w (13 to 30) SG2: 21.0w (13 to 34)	*Tubular discectomy* N randomized: 167; N analyzed: 166 (99.4%); N crossovers: 2 (1.2%)	*Microdiscectomy* N randomized: 161; N analyzed: 159 (98.8%); N crossovers: 0 (0%)	*Roland-Morris Disability* (90% power, 4 points); • VAS 100 leg pain • VAS 100 back pain • SF-36 Bodily Pain • SF-36 Physical Functioning • Sciatica index • Prolo Scale • Global perception of recovery
Brouwer (2015)[39] Brouwer (2017)[102] The Netherlands; Some concerns	Age SG1: 43.2 (11.8); SG2: 43.7 (9.7) Women SG1: 19 (35%); SG2: 24 (42%) Duration of sciatica, median (range) SG1: 30.0w (9 to 182) SG2: 26.0w (8 to 260)	*Percutaneous laser disc decompression* N randomized: 57; N analyzed: 55 (96.5%); N crossovers: Unclear	*Microdiscectomy* N randomized: 58; N analyzed: 57 (98.3%); N crossovers: 0 (0%)	*Roland-Morris Disability* (90% power, 4 points); • VAS 100 leg pain • VAS 100 back pain • SF-36 Bodily Pain • SF-36 Physical Functioning • Sciatica index • Prolo Scale (functional and economic subscales) • Global perception of recovery
Chatterjee (1995)[38] United Kingdom; Some concerns	Age SG1: 38.9 (range 20 to 56); SG2: 41.3 (range 21 to 67) Women SG1: NR (51%); SG2: NR (40%) Duration of current episode of radicular pain, mean (range) SG1: 13w (6 to 30) SG2: 20w (6 to 38)	*Automated Percutaneous Lumbar Discectomy* N randomized: 31; N analyzed: 31 (100%); N crossovers: 0 (0%) in main analysis, 20 eventually underwent microdiscectomy	*Microdiscectomy* N randomized: 40; N analyzed: 40 (100%); N crossovers: 0 (0%)	*NR (NR);* • N (%) returned to work or previous level of activity • N (%) with excellent/good outcome based on MacNab criteria
Franke (2009)[36] Germany; Some concerns	Age 44 (11.7) Women 40 (40%) Duration of symptoms: NR	*Microscopically assisted percutaneous nucleotomy* N randomized: 52; N analyzed: 52 (100%); N crossovers: 0 (0%)	*Microdiscectomy* N randomized: 48; N analyzed: 48 (100%); N crossovers: 0 (0%)	*Operation duration (15 minutes)*; • VAS (sum of leg and back pain) • N (%) with motor deficits • N (%) with sensory deficits • Oswestry Disability Index • Duration of postoperative inability to work

(continued)

Table 5. Study and population characteristics of the 15 randomized controlled trials comparing alternative surgical interventions for management of lumbar radiculopathy (EQ1) (continued)

Author (Year); Country; Risk of Bias	Population Characteristics Age, mean (SD); Women N (%); Duration of symptoms, mean (SD)	*Surgical Intervention* (SG1); N randomized; N analyzed (% of randomized); N crossovers (% of randomized);	*Comparator(s)* (SG2, SG3); N randomized; N analyzed (% of randomized); N crossovers (% of randomized);	*Primary Outcome* (effect size detectable with 80% power, α = 5%)[a]; • Other outcomes
Haines (2002)[42] United States; High	Age SG1: 42.2 (12.0); SG2: 35.4 (10.1) Women SG1: 10 (47.6%); SG2: 5 (38.4%) Duration of symptoms: NR	*Automated percutaneous discectomy, endoscopic percutaneous discectomy* N randomized: 21; N analyzed: 17 (81.0%) at 26w; N crossovers: 0 (0%)	*Discectomy* N randomized: 13; N analyzed: 10 (76.9%) at 26w; N crossovers: 1 (7.69%)	*Difference in success rates (15%);* • Roland-Morris Disability Questionnaire • SF-36 Physical Functioning • N (%) with success • N (%) with success based on MacNab criteria
Hermantin (1999)[43] United States; Some concerns	Age SG1: Mean 40 (range 18-67); SG2: Mean 39 (range 15-66) Women's SG1: 13 (43.4%); SG2: 8 (26.7%) Duration of symptoms: minimum duration of nonoperative treatment prior to randomization was 14w in both groups.	*Video-assisted arthroscopic microdiscectomy* N randomized: 30; N analyzed 30 (100%); N crossovers 0 (0%)	*Discectomy, with laminotomy* N randomized: 30; N analyzed: 30 (100%); N crossovers: 0 (0%)	*NR (NR);* • VAS 10 leg pain • N (%) with postoperative reflex abnormalities • N (%) with sensory deficits • N (%) with motor weakness • Duration of postoperative disability in time lost from work or until able to resume normal activity • N (%) self-reported satisfied with operative result • N (%) with satisfactory outcomes
Huang (2005)[24] Taiwan; Some concerns	Age SG1: 39.2 (10.8); SG2: 39.8 (11.0) Women SG1: 4 (40%); SG2: 3 (25%) Duration of symptoms: NR	*Microendoscopic discectomy* N randomized: 10; N analyzed: 10 (100%); N crossovers: 0 (0%)	*Discectomy* N randomized: 12; N analyzed: 12 (100%); N crossovers: 0 (0%)	*NR (NR);* • N (%) with excellent/good outcome based on MacNab criteria
Mayer (1993)[34] Germany; High	Age SG1: 39.8 (10.4); SG2: 42.7 (10) Women SG1: 8 (40%); SG2: 6 (30%) Duration of symptoms, mean (SD) SG1: 27.6w (NR); SG2: 29.2w (NR)	*Percutaneous endoscopic discectomy* N randomized: 20; N analyzed: 20 (100%); N crossovers: 3 (15%)	*Microdiscectomy* N randomized: 20; N analyzed: 20 (100%); N crossovers: 0 (0%)	*NR (NR);* • N (%) with low back pain • N (%) with sciatica • N (%) with sensory deficit • N (%) with motor deficit • Duration of postoperative disability • N (%) returning to work • Clinical score • N (%) with specified clinical score • N (%) with self-reported success of surgery

(continued)

Table 5. Study and population characteristics of the 15 randomized controlled trials comparing alternative surgical interventions for management of lumbar radiculopathy (EQ1) (continued)

Author (Year); Country; Risk of Bias	Population Characteristics Age, mean (SD); Women N (%); Duration of symptoms, mean (SD)	*Surgical Intervention* (SG1); N randomized; N analyzed (% of randomized); N crossovers (% of randomized);	*Comparator(s)* (SG2, SG3); N randomized; N analyzed (% of randomized); N crossovers (% of randomized);	*Primary Outcome* (effect size detectable with 80% power, α = 5%)[a]; • Other outcomes
Ruetten (2008)[31] Germany; High	Age: 43 (range 20 to 68) Women: 116 (58%) Duration of symptoms, mean (range): 11.71w (0.14 to 68)	*Endoscopic (interlaminar or transforaminal) discectomy* N randomized: 100; N analyzed: 91 (91%); N crossovers: 0 (0%)	*Microdiscectomy* N randomized: 100; N analyzed: 87 (87%); N crossovers: 0 (0%)	*NR (NR);* • VAS 100 leg pain • VAS 100 back pain • NASS pain score • NASS neurology score • Oswestry Disability Index • N (%) with no leg pain • N (%) with leg pain occasionally or pain was greatly reduced • N (%) with no improvements in leg pain • N (%) with progredient back pain • Duration of postoperative work disability
Ryang (2008)[30] Gempt (2013)[103] Germany; High	Age SG1: 38.2 (9.3); SG2: 39.1 (11.3) Women; SG1: 17 (56.7%); SG2: 11 (36.7%) Duration of symptoms: NR	*Trocar microdiscectomy* N randomized: 30; N analyzed: unclear; N crossovers: 0 (0%)	*Microdiscectomy* N randomized: 30; N analyzed: unclear; N crossovers: 0 (0%)	*NR (NR);* • VAS 10 pain • VAS 10 for improvement • SF-36 Bodily Pain • SF-36 Physical Functioning • SF-36 physical component summary and mental component summary • Other SF-36 subscales • Oswestry Disability Index • N (%) with radicular pain • N (%) with sensory deficits • N (%) with motor deficits
Sasaoka (2006)[25] Japan; High	Age 42.4 (range 20 to 72) SG1: 36.5 (range 25 to 60) SG2: 37.7 (range 20 to 58) Women 14 (42.4%) SG1: 9 (60.0%) SG2: 3 (27.3%) Duration of symptoms: NR	*Microendoscopic discectomy* N randomized: 15; N analyzed: unclear; N crossovers: 0 (0%)	*Microdiscectomy* N randomized: 11; N analyzed: unclear; N crossovers: 0 (0%)	*NR (NR);* • Japanese Orthopaedic Association Score • N (%) with residual low back pain or lumbar discomfort

(continued)

Table 5. Study and population characteristics of the 15 randomized controlled trials comparing alternative surgical interventions for management of lumbar radiculopathy (EQ1) (continued)

Author (Year); Country; Risk of Bias	Population Characteristics Age, mean (SD); Women N (%); Duration of symptoms, mean (SD)	*Surgical Intervention* (SG1); N randomized; N analyzed (% of randomized); N crossovers (% of randomized);	*Comparator(s)* (SG2, SG3); N randomized; N analyzed (% of randomized); N crossovers (% of randomized);	*Primary Outcome* (effect size detectable with 80% power, α = 5%)[a]; • Other outcomes
Teli (2010)[29] Italy; Some concerns	Age: 39.3 (range 27 to 61) Women: 73 (34.4%) Duration of pain, mean (SD): SG1: 11w (5) SG2: 12w (6)	*Microendoscopic discectomy* N randomized: NR; N analyzed: 70; N crossovers: 0 (0%)	*Microdiscectomy* N randomized: NR; N analyzed: 72; N crossovers: 0 (0%)	*VAS10 leg pain and VAS 10 back pain (1.5 points);* • SF-36 physical component summary and mental component summary • Oswestry Disability Index
Thome (2005)[28] Barth (2008)[104] Germany; Some concerns	Age SG1:42 (9); SG2: 40 (10) Women SG1: 18 (42.9%); SG2: 19 (45.2%) Duration of symptoms, mean (SD) SG1: 11w (12) SG2: 8w (10)	*Sequestrectomy* N randomized: 42; N analyzed: 42 (100%); N crossovers: 0 (0%)	*Microdiscectomy* N randomized: 42; N analyzed: 42 (100%); N crossovers: 0 (0%)	*NR (NR);* • VAS10 leg pain • VAS10 back pain • SF-36 Bodily Pain • SF-36 Physical Functioning • Other SF-36 subscales • SF-36 physical component summary and mental component summary • N (%) with total Prolo score >=7 • N (%) with specified total Prolo score • N (%) with improvement in sensory deficit • N (%) with improvement in motor deficit • N (%) with specified change in sensory index • N (%) with specified change in motor deficit • N (%) with specified categories or impairment at work • N (%) with specified patient satisfaction index scores

[a] As specified and reported by study authors.

Abbreviations: N = number; NR = not reported; SD = standard deviation; SF-36 = Short Form 36; SG1 = surgical intervention group; SG2 = surgical comparator group; VAS = visual analog scale; w = week(s); y = years(s).

Table 6. Study and population characteristics of the 3 randomized controlled trials comparing microdiscectomy to discectomy for the management of lumbar radiculopathy (EQ1)

Author (Year); Country; Risk of Bias	**Population Characteristics Age, mean (SD); Women N (%); Duration of symptoms, mean (SD)**	***Surgical Intervention* (SG1); N randomized; N analyzed (% of randomized); N crossovers (% of randomized)**	***Comparator(s)* (SG2, SG3); N randomized; N analyzed (% of randomized); N crossovers (% of randomized)**	***Primary Outcome* (effect size detectable with 80% power, α = 5%)[a]; • Other outcomes**
Henriksen (1996)[35] Denmark; Some concerns	Age Median (IQR) SG1: 39.7 (30 to 46); SG2: 42.8 (36 to 48) Women; SG1: 15 (38.5%); SG2: 14 (35%) Duration of symptoms: NR	*Microdiscectomy* N randomized: 40; N analyzed: 39 (97.5%); N crossovers: 0 (0%)	*Discectomy* N randomized: 40; N analyzed: 40 (100%); N crossovers: 0 (0%)	*Hospital stay duration (one day)*; • VAS 100 leg pain • VAS 100 back pain
Teli (2010)[29] Italy; Some concerns	Age: 39.3 (range 27 to 61) Women: 73 (34.4%) Duration of pain, mean (SD): SG2: 12w (6) SG3: 11w (5)	*Microdiscectomy* N randomized: NR; N analyzed: 72; N crossovers: 0 (0%)	*Discectomy* N randomized: NR; N analyzed: 70; N crossovers: 0 (0%)	*VAS10 leg pain and VAS 10 back pain (1.5 points);* • SF-36 physical component summary and mental component summary • Oswestry Disability Index
Tullberg (1993)[27] Sweden; Some concerns	Age SG1: 40 (range 17 to 59); SG2: 38 (range 18 to 64) Women SG1: 12 (40%); SG2: 9 (30%) N (%) with specified duration of symptoms: SG1: <4w:1 (NR) SG2: <4w: 0 (NR)	*Microdiscectomy* N randomized: 30; N analyzed: 29 (97%); N crossovers: 0 (0%)	*Discectomy* N randomized: 30; N analyzed: 29 (97%); N crossovers: 0 (0%)	*NR (NR);* • VAS 10 leg pain • VAS 10 back pain • Postoperative sick leave • N (%) of patients out of work • N (%) with specified option on recovery

[a] As specified and reported by study authors.

Abbreviations: N = number; NR = not reported; SD = standard deviation; SF-36 = Short Form 36; SG1 = surgical intervention group; SG2 = surgical comparator group; VAS = visual analog scale; w = week(s).

The interventions and comparators evaluated by included studies are summarized in ***Table 7***. The comparator interventions used in the seven RCTs evaluating the efficacy of surgery compared to nonsurgical management varied. One RCT compared the efficacy of microdiscectomy with spinal manipulation (formalized protocol, mean number of treatment sessions was 21).[23] One RCT compared the efficacy of microdiscectomy with a specific physiotherapy protocol.[33] One RCT compared percutaneous disc decompression with coblation technology to epidural steroid injection.[41] The other four RCTs compared discectomy, microdiscectomy, or percutaneous discectomy to 'conservative management' as needed and directed by treating clinicians.[22,26,32,37] Conservative management may have included analgesic or anti-inflammatory medication, bed rest, physical therapy, home exercise instruction, and education or counseling about the natural course of the disease.

Table 7. Surgical and comparator interventions used among 22 included studies for EQ1

	Surgical Intervention[a]	Comparator Intervention[b]
Efficacy RCTs (k=7)	Microdiscectomy	Spinal manipulation *(McMorland 2010)*[23]; Physiotherapy *(Osterman 2003)*[33]
	Percutaneous disc decompression with coblation technology *(Gerszten 2003)*[41]	Epidural steroid injection
	Percutaneous disc decompression *(Erginousakis 2011)*[37] Discectomy *(Weber 1983)*[26,32] Discectomy/microdiscectomy *(Weinstein 2006 [SPORT])*[22] Microdiscectomy *(Peul 2007)*[32]	Conservative management
Comparative effectiveness RCTs (k=15)	Tubular/trocar discectomy *(Arts 2011, Ryang 2008)*[30,40] Automated percutaneous lumbar discectomy *(Chatterjee 1995)*[38] Percutaneous endoscopic discectomy *(Mayer 1993)*[34] Endoscopic discectomy *(Ruetten 2008)*[31] Microendoscopic discectomy *(Sasaoka 2006, Teli 2010[c])*[25,29] Sequestrectomy *(Thome 2005)*[28] Percutaneous laser disc decompression *(Brouwer 2015)*[39] Microscopically-assisted percutaneous nucleotomy *(Franke 2009)*[36]	Microdiscectomy
	Automated percutaneous discectomy/endoscopic percutaneous discectomy *(Haines 2002)*[42] Video-assisted arthroscopic microdiscectomy *(Hermantin 1999)*[43] Microendoscopic discectomy *(Huang 2005, Teli 2010[c])*[24,29]	Discectomy
	Microdiscectomy *(Henrisksen 1996, Teli 2010[c], Tullberg 1993)*[27,29,35]	Discectomy

[a] In the Appendix D Evidence Tables, these interventions are considered the surgical group and are denoted as SG1.

[b] In the Appendix D Evidence Tables, these interventions are considered the comparator groups; nonsurgical comparator groups are denoted as NS1 and surgical comparators are denoted as SG2 or SG3.

[c] This study was a three-arm RCT that allocated participants to microendoscopic discectomy, microdiscectomy, and standard discectomy; thus, it contributes to three comparisons of interest for this HTA.

Abbreviations: k = number of studies; RCT = randomized controlled trial

The surgical interventions in the 15 RCTs evaluating the comparative effectiveness of surgical interventions also varied and one RCT[29] included more than 2 study groups. Ten RCTs evaluated various minimally-invasive surgical procedures compared with microdiscectomy.[28-31,34,36,38-40] Four RCTs evaluated various minimally-invasive surgical procedures compared with standard discectomy.[24,29,42,43] Three RCTs compared microdiscectomy with discectomy.[27,29,35]

Across the 22 included RCTs, studies reported outcomes at various time points spanning from immediately postoperative to up to 10 years postoperative; no single efficacy measure was used consistently across all included studies. The most common measures of pain included patient-reported visual analog scales (VAS) of leg pain, back pain, or both; and the Bodily Pain subscale of the SF-36. Measures of functioning also varied across studies and most commonly included validated patient-reported instruments including the Physical Functioning subscale of the SF-36, the Roland-Morris Disability Questionnaire, and the Oswestry Disability Index. Some studies also used the Prolo Scale, a measure based on an observer's assessment of the patient and includes a functional subscale and an economic subscale. A few studies reported neurologic symptoms and overall health-related quality of life (QOL). Outcomes related to "return to work" were variably defined and reported by just over half of the 22 included studies. Other outcomes reported included satisfaction with symptoms and degree of recovery, both measured using Likert scales.

We rated one RCT as low risk of bias,[40] 10 RCTs as some concerns for bias,[23,24,27-29,35,36,38,39,43] and 10 RCTs as high risk of bias.[22,25,26,30-34,37,42] We rated one study as having some concerns for bias for outcomes reported at 6 weeks but high risk of bias for outcomes reported at 12 weeks or later because of high attrition at later follow-up time points (30% or greater in both study groups).[41] All but one study[40] did not blind participants to treatment allocation, and since nearly all studies relied on patient-reported outcomes, most studies had at least some concerns for bias since knowledge of the assigned treatment may impact such outcomes. Studies rated as high risk of bias generally used inadequate randomization and allocation concealment (e.g., use of even /odd [37]) or had moderate to extensive levels of crossover between treatment arms. For example, in the Weinstein et al. RCT (Spine Patient Outcomes Research Trial [SPORT]), 46.1% of participants allocated to surgery did not receive surgery by 26 weeks follow-up, and 36.3% of participants allocated to conservative management received surgery.[22]

3.2.1.2 Findings

This section is organized by outcome: pain; function/disability, quality of life, neurological symptoms, return to work, and other outcomes (e.g., satisfaction with treatment outcome, recovery). Within each outcome domain, we synthesized the seven RCTs comparing surgery to nonsurgical interventions separate from the 15 RCTs comparing alternative surgical interventions. For consistency, we reported the findings from the surgical intervention group first, followed by the nonsurgical comparator group for the efficacy RCTs or the surgical comparator groups for the comparative effectiveness RCTs. All results presented are the intent-to-treat analyses for which participants are analyzed in the groups to which they were randomized. Studies may have also reported "as treated" and "per protocol" analyses (see Discussion). Some studies did not report the actual outcome measure values, between-group differences, or statistical tests of significance; we use "Not Reported" ("NR") to indicate when this occurred or provided calculated values when possible. ***Table 8*** describes the most common outcomes reported by included studies, including how the outcome is assessed, the range of possible scores, the directionality of the score, the minimally important clinical difference reported in the literature, and the required sample sizes that we calculated to detect various between-group differences with 80% power at an alpha level of 0.05.

Table 8. Summary of efficacy outcomes reported by included studies, including score range, minimally important clinical difference, and required sample size to detect various between-group differences

Instrument	Administration	Score Range	Interpretation of Between-Group Treatment Effect[a]	Minimally Important Difference from Literature[b]	Between-Group Difference (Δ) and Corresponding Sample Size Requirements (N)[c]
VAS 100 mm[d] Leg or Back Pain	Patient reported	0 to 100 Higher scores represent more severe symptoms	Negative absolute mean difference favors intervention group	Between 7 to 11 points[6,45,105]	Δ 15: N=70 Δ 11: N=128 *Δ 7: N=314* (assuming SD of 22)
SF-36 Bodily Pain subscale	Patient reported, 2 items from the SF-36 instrument	0 to 100 (norm-based: mean 50, SD (10)) Higher scores represent less severe symptoms	Positive absolute mean difference favors intervention group	Between 3 to 4 points [45,106,107]	Δ 10: N=128 Δ 7: N=260 *Δ 3: N=1,398* (assuming SD of 20)
SF-36 Physical Functioning subscale	Patient reported, 10 items from the SF-36 instrument				
SF-36 Physical Component Summary	Patient reported, scores multiplied by subscale factor score coefficients and summed over all 8 subscales			2 points for PCS 3 points for MCS.[45,106]	Δ 5: N=84 *Δ 3: N=226* *Δ 2: N=506* (assuming SD of 8)
SF-36 Mental Component Summary					
Roland Morris Disability Questionnaire	Patient reported, 24 items with yes/no responses	1 to 24 Higher scores represent worse functional status	Negative mean difference favors intervention group	Between 2 to 5 points[6,45,107]	Δ 8: N=16 Δ 5: N=34 Δ 4: N=52 *Δ 2: N=200* (assuming SD of 5)
Oswestry Disability Index	Patient reported, 10 items with 6-point Likert Scale	0 to 100 Higher scores represent worse functional status	Negative mean difference favors intervention group	Between 30% to 50% relative difference, or absolute difference of 8 to11 points[107] (though some studies report range from 5 to 17 points)[6,45,108,109]	Δ 17: N=30 Δ 15: N=38 Δ 11: N=70 *Δ 5: N=324* (assuming SD of 16)
Sciatica Index-Bothersomeness subscale and Frequency subscale	Patient reported 6-point Likert scale 0 (not bothersome) to 6 (extremely bothersome)	0 to 24 Higher scores represent more severe symptoms	Negative mean difference favors intervention group	None established, a 10% relative difference (2.4) is probably reasonable.	Δ 4: 42 Δ 2.4: 114 Δ 1: 638 (assuming SD 4.5)

(continued)

Table 8. Summary of efficacy outcomes reported by included studies, including score range, minimally important clinical difference, and required sample size to detect various between-group differences (continued)

Instrument	Administration	Score Range	Interpretation of Between-Group Treatment Effect[a]	Minimally Important Difference from Literature[b]	Between-Group Difference (Δ) and Corresponding Sample Size Requirements (N)[c]
Prolo Score-Economic and Functional subscales	Clinician assessor using 4-point Likert scale 0 (worse) to 4 (best)	Higher score represents increased ability to work or better functioning	Positive mean difference favors intervention group	None established, a 10% relative difference (0.4) is probably reasonable.	Δ 0.8: N=52 *Δ 0.4: N=200* Δ 0.2: N=788 (assuming SD of 1.0)

Δ = between-group difference in means

[a] Treatment effect is difference in mean scores at follow-up time point, with or without adjustment for baseline scores.

[b] From the broader musculoskeletal pain literature; we identified no studies establishing between-group MIDs specific to lumbar radiculopathy.

[c] We calculated the sample size requirements based on 80% power, two-tailed test, alpha level = 0.05 using STATA version 14.0 with the standard deviation indicated in the table cell. The standard deviation used represents the median standard deviation of the measure at baseline for the studies included in this HTA. The sample size in italics is the sample size we used to assess the imprecision domain for strength of evidence ratings; bodies of evidence generally not meeting the lower end of the MID threshold were downgraded one level (not serious to serious); bodies of evidence not meeting the upper end of the MID threshold were downgraded by two levels (not serious to very serious).

[d] Also applicable to VAS 10 cm, between-group differences in VAS 100 mm can be divided by 10 to be applicable to VAS 10 cm.

Abbreviations: MCS = mental component summary score: N = number of participants required; PCS = physical component summary score; SF-36 = Short Form 36; SD = standard deviation; VAS = visual analog scale.

Pain

A. Surgery compared with nonsurgical interventions

Seven RCTs reported at least one pain outcome. Five were rated as high risk of bias,[22,26,32,33,37] one was rated as some concerns for bias,[23] and one was rated as high risk of bias for outcomes later than 12 weeks and some concerns for bias for outcomes less than 12 weeks.[41] Pain outcomes reported included the VAS 100 mm or 10 cm for leg pain, the VAS 100 mm or 10 cm for back pain, the SF-36 Bodily Pain subscale, the Sciatica index, the McGill Pain Questionnaire, the Abderdeen back pain scale. A few studies also reported the frequency and proportion of participants reporting reduced pain, no pain, or relief from pain. We were unable to conduct quantitative synthesis for pain outcomes within this comparison because of outcome measure and reporting heterogeneity and because some studies did not report measures of variance needed to conduct a meta-analysis. ***Table 9*** summarizes the findings and strength of evidence related to pain outcomes for this comparison. Overall all studies observed improvements in pain among both surgical and nonsurgical comparison groups. For most measures, studies also reported between-group differences favoring surgery in the short- to medium-term, however, between-group differences did not persist in the long-term. A detailed description of findings follows this table.

Table 9. Summary of findings and strength of evidence ratings comparing surgery to nonsurgical interventions for pain in persons with symptomatic lumbar radiculopathy (EQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Pain(leg)-VAS 100 mm (short- and medium-term) (follow-up: range 6 weeks to 26 weeks) (MID 7 to 11 points)						
3 RCTs	Very serious[a]	Not serious	Not serious	Not serious	Pain improved in both treatment groups. Scores decreased by 6 to 26 points more in surgery groups. Peul et al.[32] (N=283) 8w AMD -17.7 (95% CI -23.1 to -12.3), 26w AMD -6.1 (95% CI -10.0 to -2.2). Osterman et al.[33] (N=56) (calculated 6w AMD -17, calculated 26w AMD -13). Gerszten et al.[41] (N=90) calculated 6w AMD -21 (p=0.002), calculated 12w AMD -23 (p=0.0001), calculated 26w AMD -26 (P=0.0008).	⊕⊕◯◯ LOW Favors surgery

(continued)

Table 9. Summary of findings and strength of evidence ratings comparing surgery to nonsurgical interventions for pain in persons with symptomatic lumbar radiculopathy (EQ1) (continued)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Pain (leg)-VAS 100 mm (long-term) (follow-up: range 52 weeks to 5 years) (MID 7 to 11 points)						
2 RCTs	Very serious[b]	Not serious	Not serious	Serious[c]	Short-term improvements in pain persisted over time in both treatment groups. However, between-group differences were minimal by 52w and beyond. Peul et al.[32] (N=283) 52w AMD 0 (95% CI -4.0 to 4.0); 2y AMD 2 (95% CI -2.0 to 6.0), similar at 5y. Osterman et al.[33] (N=56) calculated 52w AMD -7, calculated 2y AMD -13, RM AMD from 0 to 2 years -9.0 (95%CI -20 to 1).	⊕◯◯◯ VERY LOW No difference
Pain (back)-VAS 100 mm (short- and medium-term) (follow-up: range 6 weeks to 26 weeks) (MID 7 to 11 points)						
3 RCTs	Very serious[a]	Not serious	Not serious	Not serious	Pain improved in both treatment groups. Scores decreased by 7 to 24 points more in surgery groups. Peul et al.[32] (N=283) 8w AMD -11.3 (95% CI, -17.4 to -5.6), 26w AMD -2.3 (95% CI, -8.2 to 3.6). Osterman et al.[33] (N=56) calculated 6w, 12w, and 26w AMD -13. Gerszten et al.[41] (N=90) calculated 6w AMD -19 (p=0.0005), calculated 12w AMD -24 (P=0.0001); calculated 26w AMD -21 (P=0.002).	⊕⊕◯◯ LOW Favors surgery
Pain (back)-VAS 100 mm (long-term) (follow-up: range 52 weeks to 5 years) (MID 7 to 11 points)						
2 RCTs	Very serious[b]	Not serious	Not serious	Serious[c]	Short-term improvements in pain persisted over time in both treatment groups. However, between-group differences were minimal by 52w and beyond. Peul et al.[32] (N=283) 52w AMD -2.3 (95% CI -8.2 to 3.6), 2y AMD -1.4 (95% CI, -6.3 to 4.5), 5y AMD 3.1 (95% CI, -4.2 to 10.3). Osterman et al.[33] (N=56) calculated 52w AMD -4, calculated 2y AMD -16, RM AMD from 6w to 2y -7.0 (95%CI -17 to 3).	⊕◯◯◯ VERY LOW No difference
Pain- SF-36 Bodily Pain subscale (short- and medium-term) (follow-up: range 8 weeks to 26 weeks) (MID 3 points)						
4 RCTs	Very serious[d]	Serious[e]	Not serious	Serious[f]	Pain improved in both treatment groups; between-group differences were mixed, some favored surgery while some showed no difference. Weinstein et al. [SPORT][22] (N=501) 12w AMD 2.9 (95% CI, -2.2 to 8.0). Peul et al.[32] (N=283) 8w AMD 8.4 (95% CI 3.2 to 13.5), 26w AMD 3.3 (95% CI, -1.8 to 8.4). Gerszten et al.[41] (N=90) significant between-group difference favoring surgery at 26w (actual values NR, P=0.0039). McMorland et al.[23] (N=40) no difference in RM AMD 6w to 12w (actual values NR, P=0.341).	◯◯◯◯ INSUFFICIENT Mixed findings

(continued)

Table 9. Summary of findings and strength of evidence ratings comparing surgery to nonsurgical interventions for pain in persons with symptomatic lumbar radiculopathy (EQ1) (continued)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Pain- SF-36 Bodily Pain subscale (long-term) (follow-up: range 52 weeks to 8 years) (MID 3 points)						
2 RCTs	Very serious[b]	Not serious	Not serious	Serious[g]	Short-term improvements in pain persisted over time in both treatment groups; no between-group differences were observed in either study. Peul et al.[32] (N=283) 52w AMD 2.7 (95% CI, -2.6 to 7.9), 2y AMD -2.3 (95% CI, -7.3 to 2.7). Weinstein et al. [SPORT][22] (N=501) 52w AMD 2.8 (95% CI, -2.3 to 7.8), 2y AMD 3.2 (95% CI -2.0 to 8.4), 4y AMD 4.5 (95% CI, -1.2 to 10.3), 8y AMD 0.7 (95% CI, -5.2 to 6.6); RM AMD at 12w to 2y, 4y, and 8y were also nonsignificant.	⊕◯◯◯ VERY LOW No difference
Pain-Sciatica index (short- and medium-term) (follow-up: range 8 weeks to 26 weeks) (MID 2.4 points)						
2 RCTs	Very serious[h]	Not serious	Not serious	Not serious	Scores improved in both treatment groups; index scores decreased by 2.1 to 4.0 points more in the surgery group; between-group differences favored surgery in all studies at most follow-up time points. Weinstein et al. [SPORT][22] (N=501) Bothersomeness subscale 12w AMD -2.1 (95% CI, -3.4 to -0.9). Peul et al.[32] (N=283) Frequency subscale: 8w AMD -4.0 (95% CI, -5.3 to -2.7), 26w AMD -0.5 (95% CI, -1.8 to 0.8). Bothersomeness subscale: 8w AMD -3.6 (95% CI, -4.9 to -2.3), 26w AMD -1.2 (95% CI, -1.3 to -0.1).	⊕⊕◯◯ LOW Favors surgery
Pain- Sciatica index (long-term) (follow-up: range 52 weeks to 8 years) (MID 2.4 points)						
2 RCTs	Very serious[h]	Serious[i]	Not serious	Not serious	Short-term improvements in scores persisted over time in both treatment groups; between-group differences varied by study. Weinstein et al. [SPORT][22] 52w AMD -1.6 (95% CI, -2.9 to -0.4), for Bothersomeness subscale and similar significant differences favoring surgery at 2y, 4y, and 8y though AMDs were all less than the MID. Peul et al.[32] (N=283) 52w AMD -0.5 (95% CI, -1.8 to 0.8) for Frequency subscale, AMD -0.4 (95% CI, -1.5 to 0.7) for Bothersomeness subscale.	⊕◯◯◯ VERY LOW No difference[i]

[a] Risk of bias was high in all trials for medium-term outcomes; risk of bias was some concerns in 1 trial for short-term outcomes. Sources of bias: lack of participant and outcome assessor blinding and differential crossover.

[b] Risk of bias was high in both trials. Sources of bias: lack of participant and outcome assessor blinding and differential crossover.

[c] Optimal information size criteria not met in smaller trial[33] sample size can only detect between-group differences of more than 15 points.

[d] Risk of bias was rated as high in 3 trials, and as some concerns in 1 trial.[23] Sources of bias: lack of participant and outcome assessor blinding and extensive and differential crossover.

[e] Studies in the evidence base do not observe a consistent between-group difference; one study that reported a statistical difference between groups did not report the magnitude of effect to assess whether it is meaningful.[41]

[f] Optimal information size criteria not met by some trials: a sample size of 1,398 is required to detect a difference of 3 points, actual values and CI for AMDs not provided by some studies.

[g] Optimal information size criteria not met: a sample size of 1,398 is required to detect a difference of 3 points

[h] Risk of bias was high in all trials. Sources of bias: lack of participant and outcome assessor blinding and extensive and differential crossover, high attrition at longer-term follow-up points in one RCT.[22]

[i] Although statistically significant between-group differences were observed in one trial[22] at all follow-up time points; the AMDs were all less than the MID for this measure; thus we concluded no meaningful clinical difference. Only 1 long-term follow-up result was reported by the other RCT at 52 weeks with a lower magnitude of difference.

Abbreviations: AMD = absolute mean difference; CI = confidence interval; MID = minimally important between-group difference; NR = not reported; RCT = randomized controlled trial; RM = repeated measures; SPORT = Spine Patient Outcomes Research Trial; VAS = Visual Analog Scale; SF-36 = Short Form 36; w = week(s); y = year(s).

VAS Leg Pain

Three RCTs reported leg pain outcomes using VAS 100 mm scales.[32,33,41] On this scale, which ranges from 0 to 100, a higher score indicates worse pain, and a negative AMD favors the surgical intervention over the nonsurgical comparator intervention. All studies reported decreased scores (i.e., improvement) in participants allocated to both the surgical treatment (range 41 to 57 point decrease) and nonsurgical comparator (range 20 to 36.5 point decrease) from baseline through short-term (6- and 8-week follow-up). VAS leg pain scores decreased by 6 to 26 points more among participants allocated to surgery at short- and medium-term follow-up. Of the two RCTs reporting long-term outcomes,[32,33] improvements in leg pain persisted within groups, but between-group differences were minimal. Specific study findings:

- Gerszten et al., which compared plasma disc decompression with epidural steroid injection, reported larger decreases in scores at 6 weeks, 12 weeks, and 26 weeks for participants allocated to plasma disc decompression compared with epidural steroid injection (12-week mean (SD) decrease in VAS 100 mm leg pain: -46 (SD 4) in surgical group, -23 (SD 5) in epidural steroid group, P=0.0001, calculated AMD -23).[41] Between group differences persisted at 26 weeks (calculated AMD -26, P=0.0008).

- Peul et al., which compared microdiscectomy with conservative management, reported differences between groups for short-term and medium-term follow-up that favored microdiscectomy (8-week AMD -17.7 (95% CI, -23.1 to -12.3; 26-week AMD -6.1 (95% CI, -10.0 to -2.2).[32] However, differences between groups were minimal and not statistically different by 52 weeks, 2 years, and 5 years (2-year AMD: 2.0 [95% CI, -2.0 to 6.0]).

- Osterman et al., which compared microdiscectomy with physiotherapy, also reported larger decreases among participants allocated to surgery through short-term and medium-term follow-up but no tests of significance at these time points were reported.[33] The calculated AMDs at 6 weeks, 12 weeks, and 26 weeks were -17, -11, and -13, respectively. Differences between groups were minimal at 52 weeks (calculated AMD -7), but larger decreases were observed again at 2 years among participants allocated to surgery (calculated AMD -13). The study reported a repeated measures AMD between groups from 6 weeks to 2 years of -9 (95% CI, -20 to 1).

VAS Back Pain

The same three RCTs also reported back pain outcomes using VAS 100 mm scale.[32,33,41] In all three RCTs, VAS back pain scores were lower than VAS leg pain scores at baseline, consistent with a study population selected for radicular pain. Back pain scores decreased among participants allocated to both the surgery (range 17 to 32 point decrease) and the nonsurgical comparator groups (range 1 to 19 point decrease) in all RCTs. VAS back pain scores decreased by 7 to 24 points more among participants allocated to surgery in the short-term, and a similar pattern in the medium-term as reported for VAS Leg Pain. Of the two RCTs reporting long-term outcomes,[32,33] improvements in back pain persisted within groups, but between-group differences were not significant. Specific study findings:

- Gerstzen et al. reported similar or increased VAS back pain scores from baseline (range decrease of 0.4 to increase of 7 points) in participants allocated to epidural steroid injection while scores decreased in participants allocated to plasma disc decompression (range of decrease 18 to 21 points).[41] This resulted in between-group differences that were significant at 6 weeks (calculated AMD -19, P=0.002), 12 weeks (calculated AMD -24, P=0.001), and 26 weeks (calculated AMD -21, P=0.0008).

- Peul et al. reported an AMD between microdiscectomy and conservative management of -11.3 (95% CI, -17.4 to -5.6) at 8 weeks, -2.3 (95% CI, -8.2 to 3.6) at 26 and 52 weeks, and -1.4 (95% CI, -6.3 to 4.5) at 2 years.[32]

- Osterman et al. reported larger decreases in participants allocated to microdiscectomy compared to physiotherapy at 6 weeks, 12 weeks, and 26 weeks (calculated AMD -13 at all time points), but no tests of significant differences were reported.[33] The difference between groups was minimal at 52 weeks (calculated AMD -4), but larger decreases among participants allocated to surgery were observed again at 2 years (calculated AMD -16). The repeated measures AMD from 6 weeks to 2 years was not statistically significant (AMD -7 (95% CI -17 to 3).

VAS Pain

One RCT also reported pain outcomes using VAS, but it was not specific to leg or back pain. Erginousakis et al. used a 10-cm VAS outcome to assess the effectiveness of percutaneous disc decompression compared with conservative management.[37] Similar to the other three RCTs reporting VAS Pain scores, scores decreased from baseline to follow-up in participants allocated to both the surgery (baseline mean [SD] 7.4 [1.4]) and nonsurgical intervention (baseline mean [SD] 6.9 [1.9]. At 12 weeks follow-up, the mean (SD) among participants allocated to percutaneous disc decompression was 3.0 (2.4) and was 0.9 (2.0) among participants allocated to conservative management; this difference was not statistically different (P > 0.005 [sic]). Pain scores increased among participants allocated to conservative management and decreased among participants allocated to surgery at 52 weeks and 2 years resulting in significant between-group differences favoring surgery (P=0.005 and P=0.004, respectively). We did not use this study in the strength of evidence ratings for the VAS pain outcome because of the lack of specificity for leg versus back pain.

SF-36 Bodily Pain

Four RCTs reported pain outcomes using the Bodily Pain subscale of the SF-36.[22,23,32,41] On this scale, which ranges from 0 to 100, higher scores represent less pain and a positive between-group AMD favors the surgical intervention relative to the nonsurgical intervention. The studies all reported increases in scores from baseline to short-term follow-up among participants allocated to both the surgical intervention (range 14.1 to 40.9 point increase) and the nonsurgical comparator (range 17.3 to 30.5 point increase). However, between-group differences at follow-up varied by study and follow-up time point.

One of the two trials reporting short-term outcomes observed a significant between-group difference favoring surgery;[32] the other reported no difference between groups.[23] The trial reporting a significant difference used a conservative management comparator of education and pain medication and physiotherapy as needed, while the trial that did not observe a difference used a comparator that involved a program of spinal manipulations that involved a mean of 21 treatment sessions.

Three RCTs reported medium-term outcomes; the one RCT demonstrating a between-group difference favoring surgery at 26 weeks used a comparator group that received epidural steroid injection at baseline, and a second injection (if symptoms persisted) 3 weeks later.[41] This study did not report short-term outcomes and the lack of any additional treatment in the comparator group beyond two steroid injections during the first 6 weeks of enrollment may partially explain the between-group difference favoring surgery at 26 weeks relative to the other two RCTs reporting no differences between groups.[22,32] In these RCTs, a suite of conservative management therapies were provided to participants with individualization of therapy encouraged and not time-limited.

Two RCTs reported outcomes at 52 weeks or longer.[22,32] Both studies observed no significant between-group differences between participants allocated to surgery compared with participants allocated to conservative management.

Specific study findings:

- Gerszten et al. reported larger improvements in participants allocated to plasma disc decompression compared to participants allocated to epidural steroid injection at 26 weeks (actual values NR, P=0.0039).[41]
- McMorland et al., which compared microdiscectomy with spinal manipulation, reported no differences in a repeated measures analysis from 6 weeks to 12 weeks (AMD NR, P=0.341).[23] The calculated AMD at 6 weeks was -3.2 and was 11.5 at 12 weeks.
- Weinstein et al. [SPORT], which compared discectomy or microdiscectomy to conservative management, reported nonsignificant differences ranging from 0.7 to 4.5 at follow-up time points between 12 weeks and 8 years and also reported nonsignificant repeated measures differences from 12 weeks to 2 years (AMD NR, P=0.74), 4 years (AMD NR, P=0.15), and 8 years (AMD NR, P=0.22).[22,99,100]

- Peul et al., which compared microdiscectomy to conservative management, reported a significant difference at 8 weeks favoring surgery (AMD 8.4 [95% CI, 3.2 to 13.5]). However, the outcomes reported at 26 weeks (AMD 3.3 [95% CI, -1.8 to 8.4]), 52 weeks (AMD 2.7 [95% CI, -2.6 to 7.9]), and 2 years (AMD -2.3 [95% CI, -7.3 to 2.7]) showed no significant between-group differences.[32,97]

Sciatica Index

Two RCTs reported outcomes measured with the Sciatica Index.[22,32] With this measure, which ranges from 0 to 24, higher scores represent worse symptoms and negative between-group difference favor the surgical intervention. Like other pain outcomes already reported, scores improved at short-term follow-up in both the surgical (range 9.0 to 10.7 point decrease) and nonsurgical intervention groups (range 6.8 to 6.9 point decrease). Sciatica Index scores decreased more by 2.1 to 4.0 points more among participants allocated to surgical intervention compared with participants allocated to nonsurgical interventions at short- and medium-term follow-up. Specific study findings:

- Weinstein et al. [SPORT] reported an AMD of -2.1 (95% CI, -3.4 to -0.9) at 12 weeks favoring surgery compared to conservative management for the Bothersomeness subscale.[22] Significant differences favoring surgery persisted at all additional follow-up time points (52 weeks, 2 years, 4 years, and 8 years) and in repeated measures analyses from 12 weeks to 2 years and from 12 weeks to 8 years.[22,99,100] However, the magnitude of the between-group differences were less than the MID at all time points for this measure.

- Peul et al. reported outcomes for the Frequency and Bothersomeness subscales of this index.[32] For the Frequency subscale, the AMD between groups favored microdiscectomy compared with conservative management at 8 weeks (AMD -4.0 [95% CI, -5.3 to -2.7]) and 26 weeks (AMD -1.8 [95% CI, -1.9 to -0.7]); this benefit did not persist at 52 weeks (AMD -0.5 (95% CI, -1.8 to 0.8]) and was not reported at 2 years or 5 years of follow-up. For the Bothersomeness subscale, authors reported an AMD of -3.6 (95% CI, -4.9 to -2.3) at 8 weeks and -1.2 (95% CI, -1.3 to -0.1) at 26 weeks. By 52 weeks, the difference between groups was not significant (AMD -0.4 [95% CI, -1.5 to 0.7]).

Other Pain Measures

Three RCTs reported other measures related to pain.[23,26,37] These measures confirmed similar findings to other measures of pain previously reported. We did not use these studies in our strength of evidence ratings for pain outcome because they were only used in one study each. Specific study findings:

- McMorland et al. reported pain using three different subscales of the McGill Pain Questionnaire and the Abderdeen Back Pain Scale at 6 week and 12 weeks follow-up.[23] In this RCT, pain decreased significantly over time in participants allocated to both microdiscectomy and spinal manipulation. Though a numerically larger decrease was observed in the surgical group at both 6 weeks and 12 weeks, the between-group differences in the repeated measures analysis were not statistically significant (AMDs NR). The between-group differences in the Aberdeen Back Pain scale were not tested for

significance at single follow-up timepoints; the AMD was less than the MID at 6 weeks, but was largerd than the MID by 12 weeks. However, participants allocated to microdiscectomy had a significantly larger decrease in pain compared to spinal manipulation in repeated measures analysis (AMD NR, P=0.034).

- Weber et al. reported the frequency and proportion of participants within three categories of radiating pain at 4 years and 8 years (no pain, some pain, considerable pain).[26] The proportion of participants with no pain was 63.2% among those allocated to discectomy and 57.6% among those allocated to conservative management at 4 years follow-up (calculated P=0.86). At 10 years, these proportions were 84.3% and 78.8%, respectively (calculated P=0.41).

- Erginousakis et al. reported the proportion of participants within four categories of pain reduction at 2 years (100% pain relief, 50% pain relief, 0% pain relief, and aggravation of pain).[37] The proportion of participants that reported 100% pain relief was 55% among those allocated to percutaneous disc decompression and 19% among those allocated to conservative therapy (calculated P=0.008).

B. Minimally-invasive surgery compared with microdiscectomy or discectomy

Ten RCTs comparing minimally-invasive surgical interventions (tubular/trocar discectomy,[30,40] percutaneous endoscopic discectomy,[34] endoscopic interlaminar or transforaminal discectomy,[31] microendoscopic discectomy,[25,29] sequestrectomy,[28] percutaneous laser disc decompression,[39] microscopically assisted percutaneous nucleotomy,[36] and video-assisted microdiscectomy[43]) to either microdiscectomy or discectomy and reported at least one pain outcome. Four were rated as high risk of bias,[25,30,31,34] five were rated as having some concerns for bias,[28,29,36,39,43] and one was rated as low risk for bias for outcomes up to 2 years and high risk of bias for outcomes longer than 2 years of follow-up because of high attrition.[40] Pain outcomes reported include the VAS 100 mm or 10 cm for leg pain or back pain, the SF-36 Bodily Pain subscale, the Sciatica Index, and the frequency and proportion of participants reporting reduced pain, no pain, or relief from pain. ***Table 10*** summarizes the findings and strength of evidence related to pain outcomes for this comparison. A detailed description of findings follows this table.

Table 10. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for pain in persons with symptomatic lumbar radiculopathy (EQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Pain (leg) VAS 100mm and 10 cm (short- and medium-term) (follow-up: range 4 weeks to 26 weeks) (MID 7 to 11 points)						
5 RCTs	Serious[a]	Not serious	Not serious	Not serious	Pain improved in both surgical groups.[28,29,31,39,40] Between-group differences were not significant in any study. Arts et al.[40] (N=328) 4w AMD 4.5 (95% CI, -0.3 to 9.3), 8w AMD 4.5 (95% CI, -0.4 to 9.3). Brouwer et al.[39] (N=115) 4w AMD 7.4 (95% CI, -1.9 to 16.8), 8w AMD 5.7 (95% CI, -3.7 to 15.0). Pooled between-group mean difference in scores at 12w to 26w 0.3 (95% CI, -2.2 to 2.9, 4 RCTs, 642 participants, I^2=0%).[b]	⊕⊕⊕◯ MODERATE No difference
Pain (leg) VAS 100mm and 10 cm (long-term) (follow-up: range 52 weeks to 5 years) (MID 7 to 11 points)						
5 RCTs	Serious[a]	Not serious	Not serious	Not serious	Improvements in pain persist in both surgical groups in all studies; however, between-group differences not significant at any single long-term follow-up point.[28,29,31,39,40] Pooled between-group mean difference in scores at 52w to 1.5 y 1.6 (95% CI, -1.5 to 4.6, 4 RCTs, 640 participants, I^2=28.1%) and at 2y -0.1 (95% CI, -2.7 to 2.4, 4 RCTs, 619 participants, I^2=0%).[b] Arts et al.[40] (N=328) RM AMD: 4w to 52w 4.2 (95% CI 0.9 to 7.5), 4w to 2y 3.3 (95% CI 0.2 to 6.2). Brouwer et al.[39] (N=115) RM AMD: 4w to 52w 6.9 (95% CI, 1.3 to 12.6), 4w to 2y 5.0 (95% CI, -0.2 to 10.2). Arts et al.[40,48] 5y AMD 0.2 (95% CI, -5.5 to 6.0).	⊕⊕⊕◯ MODERATE No difference
Pain (back)-VAS 100mm (short-term) (follow-up: range 4 weeks to 8 weeks) (MID 7 to 11 points)						
2 RCTs	Not serious	Not serious	Not serious	Serious[c]	Pain improved in both surgical groups. Between-group differences were not significant in either study. Arts et al.[40] (N=328) 4w AMD 3.1 (95% CI, -1.9 to 8.1), 8w AMD 3.8 (95% CI, -1.3 to 8.8). Brouwer et al.[39] (N=115) 4w AMD -2.0 (95% CI, -11.3 to 7.2), 8w AMD 6.3 (95% CI, -2.9 to 15.5).	⊕⊕⊕◯ MODERATE No difference
Pain (back) VAS 100mm and 10 cm (medium-term) (follow-up: range 12 weeks to 26 weeks) (MID 7 to 11 points)						
5 RCTs	Serious[a]	Serious[d]	Not serious	Not serious	Improvements in pain persist in both surgical groups.[28,29,31,39,40] Between-group differences were significant in only 1 RCT: Brouwer et al.[39](N=115) 26w AMD 9.4 (95% CI, 0.1 to 18.6). Pooled between-group mean difference in scores at 12w to 26w was 1.3 (95% CI, -3.5 to 6.2; 4 RCTs; 642 participants; I^2=61.7%).[b]	⊕⊕◯◯ LOW No difference

(continued)

Table 10. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for pain in persons with symptomatic lumbar radiculopathy (EQ1) (continued)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Pain (back) VAS 100mm and 10 cm (long-term) (follow-up: range 52 weeks to 5 years) (MID 7 to 11 points)						
5 RCTs	Serious[a]	Serious[d]	Not serious	Not serious	Improvements in pain persist in both surgical groups in all studies.[28,29,31,39,40] Between-group differences were not significant at any single follow-up time point. Pooled between-group mean difference in scores at 52w to 1.5y follow-up 1.5 (95% CI, -3.0 to 5.9; 4 RCTs; 640 participants; I^2=57.6%) [b] Pooled mean difference at 2y -0.8 (95% CI, -5.7, to 4.1; 4 RCTs; 619 participants; I^2=59.0%).[b] Arts et al.[40,101] (N=328) RM AMD 4w to 52w 3.5 (95% CI, 0.1 to 6.9), 4w to 2y 3.0 (95% CI, -0.2 to 6.3). Brouwer et al.[39,102] (N=115) RM AMD 4w to 52w 7.6 (95% CI, -1.7 to 16.9), 4w to 2y -1.5 (95% CI, -11.0 to 8.0). Arts et al.[40,48] 5y AMD 0.4 (95% CI -5.9 to 6.7).	⊕⊕◯◯ LOW No difference
Pain-SF-36 Bodily Pain (short-term) (follow-up: range 4 weeks to 8 weeks) (MID 3 points)						
2 RCTs	Not serious	Serious[e]	Not serious	Serious[f]	Increase in scores from baseline in both surgical groups; no significant between-group differences. Arts et al.[40] (N=328) 4w AMD -1.6 (95% CI, -6.7 to 3.6), 8w AMD -5.1 (95% CI, -10.3 to 0.1. Brouwer et al.[39] (N=115) 4w AMD 4.1 (95% CI, -4.8 to 12.9), 8w AMD 0.6 (95% CI, -9.1 to 9.3).	⊕⊕◯◯ LOW No difference
Pain-SF-36 Bodily Pain (medium-term) (follow-up: range 12 weeks to 26 weeks) (MID 3 points)						
3 RCTs	Serious[g]	Serious[d]	Not serious	Serious[h]	Improvements in pain persist in both surgical groups;[28,39,40] 1 of 3 studies reports a significant between-group difference (Brouwer et al.[39] [N=115] 26w AMD -11.3 [95% CI, -20.1 to -2.4]). Pooled between-group mean difference at 12 to 26w -3.0 (95% CI, -12.8 to 6.8), 3 RCTs, 500 participants, I^2=75.4%).	⊕◯◯◯ VERY LOW No difference
Pain-SF-36 Bodily Pain (long-term) (follow-up: range 52 weeks to 2.8 years) (MID 3 points)						
3 RCTs	Serious[i]	Serious[d]	Not Serious	Serious[h]	Increase in scores from baseline to 52w, 2y, and 2.8y in both surgical groups in all studies Arts et al.[40] (N=328), Brouwer et al.[39] (N=115), and Ryang et al.[30] (N=60). Between-group differences ranged from -11 to 2.5 points; none of these between-group differences were statistically significant.	⊕◯◯◯ VERY LOW No difference

(continued)

Table 10. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for pain in persons with symptomatic lumbar radiculopathy (EQ1) (continued)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Sciatica index (short-, medium-, and long-term) (follow-up: range 4 weeks to 2 years) (MID 2.4 points)						
2 RCTs	Not Serious	Not serious	Serious[j]	Not serious	Decreases in scores from baseline to all follow-up time points in both surgical groups for both the Bothersomeness and Frequency subscales; no significant between-group differences at any single time point, or in repeated measures analyses in either study. Arts et al.[40] (N=328) RM AMD 4w to 52w 0.7 (95% CI, -0.1 to 1.5), 0 to 2y 0.5 (95% CI, -0.3 to 1.3) for Bothersomeness subscale; similar findings for Frequency subscale. See Appendix D, Table D-3 for findings related to Brouwer et al.[39] (N=115).	⊕⊕⊕◯ MODERATE No difference

[a] One trial was rated as low risk of bias for outcomes at 2 years or less, and high risk of bias for outcomes longer than 2 years[40], 3 were rated as some concerns for bias,[28,29,43] and 1 was rated as high risk of bias[30]. Sources of bias: lack of participant and outcome assessor blinding (in all but the 1 trial[40]), inadequate randomization and allocation concealment in one trial rated as high risk of bias.[30]

[b] One of the 5 trials did not provide measures of variance required for pooling effect estimate.

[c] Optimal information size criterion not met; the smaller trial[39] is unable to detect differences less than 11 points, and the larger trial[40] is unable to detect differences of less than 7 points.

[d] I^2 statistic indicates moderate heterogeneity; trials are somewhat inconsistent with respect to magnitude and direction of effect and overlap in confidence intervals.

[e] Confidence intervals are modestly wide and spanning thresholds that would favor both interventions, less overlap in confidence intervals for 8 week outcomes.

[f] Optimal information size criterion not met: a sample size of 1,398 required to detect a difference of 3 points; the smaller trial[39] is unable to detect differences less than 11 points, and the larger trial[40] is unable to detect differences of less than 7 points.

[g] Two of 3 trials had some concerns for bias,[28,39] the other trial was low risk for bias for outcomes at 2 years or less.[40] Sources of bias: lack of participant and outcome assessor blinding. Sources of bias: lack of participant and outcome assessor blinding.

[h] Optimal information size criterion not met: a sample size of 1,398 required to detect a difference of 3 points, the pooled estimate for medium-term outcomes is based on a sample size of 500. For long-term outcomes, the largest trial[40] would be unable to detect a difference less than 7 points, and the two other studies would not be able to detect differences smaller than approximately 11 points.

[i] One trial was low risk of bias for outcomes at 2 years or less,[40] 1 trial was some concerns for bias,[39] and 1 trial was high risk of bias.[30] Sources of bias: lack of participant and outcome assessor blinding (all trials) and inadequate randomization and allocation concealment [30].

[j] One of the trials[39] used an adapted approach to score this measure.

Abbreviations: AMD = absolute mean difference; CI = confidence interval; MID = minimally important between-group difference; N = number; NR = not reported; RCT = randomized controlled trial; RM = repeated measures; SF-36 = Short Form 36; VAS = visual analog scale; w = week(s); y = year(s).

VAS Leg and Back Pain

Five RCTs reported leg or back pain outcomes using the VAS 100 mm or 10 cm scale.[28,29,31,39,40] A decrease in score represents improvements in pain, a negative AMD between groups favors the minimally-invasive surgical procedure. Two RCTs[39,40] reported outcomes at 4 weeks and 8 weeks follow-up and all reported medium-term outcomes. The range of decreases in leg pain

scores at the earliest follow-up in each study (4 weeks to 26 weeks) was 42.5 to 69 points among participants allocated to minimally-invasive surgery and 29.8 to 62 points among participants allocated to standard surgery. Between-group differences in short-term outcomes were not significant and with one exception (Brouwer et al.[39]) between-group differences in medium-term outcomes were also not significant. The pooled between-group difference in VAS 100 mm leg pain at 12 to 26 weeks was 0.3 (95% CI, -2.2 to 2.9, 4 RCTs, 642 participants, I^2=0%, ***Appendix G, Figure G-1***). The pooled between-group difference in VAS 100 mm back pain at 12 to 26 weeks was 1.3 (95% CI, -3.5 to 6.2, 4 RCTs, 642 participants, I^2=61.7%, ***Appendix G, Figure G-2***).

All 5 RCTs also reported long-term outcomes at 52 weeks and 2 years and the reductions in leg and back pain observed in the short and medium-term generally persisted. Between-group differences were reported for both single points in follow-up and using repeated measures in some studies. The pooled between-group difference in VAS 100 mm leg pain at 52 weeks to 1.5 years was 1.6 (95% CI, -1.5 to 4.6, 4 RCTs, 640 participants, I^2=28.1%) and at 2 years was -0.1 (95% CI, -2.7 to 2.4; 4 RCTs, 640 participants, I^2=0%), as shown in ***Appendix G, Figure G-1***. The pooled between-group difference in VAS 100 mm back pain at 52 weeks to 1.5 years was 1.5 (95% CI, -3.0 to 5.9, 4 RCTs, 640 participants, I^2=57.6%) and at 2 years was -0.8 (95% CI, -5.7 to 4.1; 4 RCTs, 640 participants, I^2=59.0%) as shown in ***Appendix G, Figure G-2***.

One RCT reported outcomes at 5 years and reported persistence of pain reduction observed at earlier follow-up time points with no significant between-group differences.[40,48]

Specific study findings:

- Arts et al.,[40] which compared tubular discectomy to microdiscectomy, reported between-group differences of 4.5 (95% CI, -0.3 to 9.3) at 4 weeks and 4.5 (95% CI, -0.4 to 9.3) at 8 weeks for VAS 100 mm leg pain scores. The between-group difference at 26 weeks was 2.0 (95% CI, -2.9 to 6.8). Between-group differences at single follow-up time points at 52 weeks or 2 years were also not significant; however, the repeated measures AMD for 4 weeks to 52 weeks was 4.2 (95% CI, 0.9 to 7.5) and for 0 to 2 years was 3.3 (95% CI, 0.2 to 6.2) with both estimates favoring microdiscectomy compared with tubular discectomy.[40,101] However, the between-group differences at 5 years were not significant (AMD 0.2 [95% CI, -5.5 to 6.0]) and the repeated measures AMD from 4 weeks to 5 years was also not significant (calculated AMD 1.8, P=0.13).[48] Study authors reported similar findings for VAS 100 mm back pain scores.

- Brouwer et al.,[39] which compared percutaneous laser disc decompression to microdiscectomy, reported between-group differences of 7.4 (95 % CI, -1.9 to 16.8) at 4 weeks and 5.7 (95% CI, -3.7 to 15.0) at 8 weeks for VAS 100 mm leg pain scores. The between-group difference at 26 weeks was 4.2 (95% CI, -5.2 to 13.6). Between-group differences at single follow-up time points at 52 weeks or 2 years were also not significant; however, the repeated measures AMD for 4 weeks to 52 weeks was 6.9 (95% CI, 1.3 to 12.6) favoring microdiscectomy but was not significant for 4 weeks to 2 years (repeated measures AMD 5.0 [95% CI, -0.2 to 10.2]).[39,102] Nonsignificant between-group findings

were also observed for VAS 100 mm back pain scores at 4 weeks and 8 weeks, 52 weeks and 2 years, but a statistically significant between-group difference of 9.4 (95%, CI, 0.1 to 18.6) was observed at 26 weeks. The repeated measures AMD for VAS back pain scores found no significant between-group differences for 4 to 52 weeks or 4 weeks to 2 years.

- Ruetten et al.,[31] which compared endoscopic discectomy with microdiscectomy, reported a decrease in VAS 100 mm leg pain score of 75 at baseline to 6 at 12 weeks and 9 at 26 weeks among participants allocated to endoscopic discectomy and a decrease from 71 at baseline to 9 at 12 weeks and 7 at 26 weeks; between-group differences were reported as nonsignificant. No significant between-group differences were reported at 52 weeks or 2 years.

- Teli et al.,[29] which compared microendoscopic discectomy with microdiscectomy, reported leg pain outcomes using the VAS 10 cm scale. A decrease from baseline to 26 weeks of 8 (SD 1) to 2 (SD 1) was observed among those allocated to microendoscopic discectomy, and a decrease from 8 (SD 1) to 2 (SD 1) was observed among those allocated to microdiscectomy. Decreases persisted at 52 weeks and 2 years with no significant between-group differences were reported at any time point (p=0.73). A similar pattern was observed for VAS 10 cm back pain scores, though baseline scores started lower than leg pain scores.

- Thome et al.,[28] which compared sequestrectomy with microdiscectomy, also reported leg pain outcomes using the VAS 10 cm scale. Scores decreased from 5.9 (SD 2.6) at baseline to 0.7 (SD 1.7) at 12 to 26 weeks among those allocated to sequestrectomy and from 6.7 (SD 2.3) to 1.3 (SD 2.5) among those allocated to microdiscectomy (calculated AMD 0.2 adjusted for baseline). No significant between-group differences were reported at 52 weeks or 2 years. A similar pattern was observed for VAS 10 cm back pain scores.

VAS Pain

Three RCTs reported on other variants of the VAS pain measure.[30,36,43] These findings were consistent with the previously reported VAS outcomes, but we did not use these studies in our strength of evidence ratings for VAS pain outcome because of the lack of specificity for leg versus back pain or because of unspecified follow-up times. Specific study findings:

- Franke et al.,[36] which compared microscopically-assisted percutaneous nucleotomy with microdiscectomy, reported a repeated measures sum of VAS 10 cm leg and back pain scores from 8 weeks to 52 weeks. Although pain scores decreased over time in participants allocated to microscopically-assisted percutaneous nucleotomy and among participants allocated to microdiscectomy, significant between-group differences favoring the minimally-invasive surgical procedure (actual values NR, P=0.006) were only observed at one of the two clinical centers that enrolled patients. Further, when the leg and pain scores were considered in a post-hoc analysis, between-group differences were only seen for back pain.

- Hermantin et al.,[43] which compared video-assisted arthroscopic microdiscectomy with discectomy, reported outcomes using the VAS 10 cm scale, but did not specify whether it was for leg pain or back pain, and the follow-up time period was also not specified. Pain decreased from 6.8 (SD NR) to 1.2 (SD NR) among participants allocated to video-assisted arthroscopic microdiscectomy and from 6.6 (SD NR) to 1.9 (SD NR) among participants allocated to discectomy with laminotomy. No statistical tests of between-group differences were reported.

- Ryang et al.,[28] which compared trocar microdiscectomy with microdiscectomy, reported VAS 10 cm pain scores (unspecified as to leg or pain) over long-term follow-up. Scores decreased from baseline to both follow-up time points among participants allocated to trocar discectomy and among patients allocated to microdiscectomy, but between-group differences at 1.33 years (P=0.86) and 2.8 years (P value reported as NS) suggested no differences.

SF-36 Bodily Pain

Four RCTs compared minimally-invasive surgical interventions to microdiscectomy and reported pain outcomes using the Bodily Pain subscale of the SF-36.[28,30,39,40] For this outcome, a higher score represents less severe pain and a positive between-group AMD favors the minimally-invasive surgical procedure. In all studies, pain scores improved from baseline to short-, medium-, and long-term follow-up among participants allocated to both surgical groups. Increases in scores at the earliest follow-up in each study (range 4 weeks to 26 weeks) ranged from 6.7 to 46.5 among participants allocated to minimally-invasive surgery and from 5.9 to 51.1 among participants allocated to standard surgery. With one exception, no between-group differences were observed at any follow-up time. The pooled mean difference in SF-36 Bodily Pain scores at 12 to 26 weeks was -3.0 (95 % CI, -12.8 to 6.8, 3 RCTs, 500 participants, I^2=75.4%, ***Appendix G, Figure G-3***). Specific study findings:

- Arts et al. [40] (tubular discectomy) and Brouwer et al. [39] (percutaneous laser disc decompression) reported nonsignificant between-group differences at 4 weeks and 8 weeks that ranged from -5.1 to 0.6 points. At 26 weeks, Arts et al. reported an AMD of -4.9 (95% CI, -10.0 to 0.3) while Brouwer et al. reported a significant between-group difference favoring microdiscectomy compared with percutaneous laser disc decompression (AMD -11.3 [95% CI, -20.1 to -2.4]).

- Thome et al.,[28] which compared sequestrectomy with microdiscectomy, observed a nonsignificant, between-group difference (calculated AMD 3.0, P=0.14) at 12 to 26 weeks.

- Arts et al.,[40] Brouwer et al.,[39] Thome et al.,[28] and Ryang et al.[30](trocar discectomy) reported nonsignificant between-group differences ranging from -11 points to 2.5 points at 52 weeks to 2.8 years follow-up. This measure was not reported in 5 year results reported by Arts et al.[48]

Sciatica Index

Two RCTs also reported outcomes measured with the Sciatica index.[39,40] With this measure, a higher score represents more severe symptoms and a negative between-group difference favors the minimally-invasive surgical approach. Like other pain outcomes already reported, scores improved over time in both surgical groups; scores decreased from 4 to 8.5 points among participants allocated to minimally-invasive surgery and from 3.2 to 8.7 points among participants allocated to standard surgery. Between-group differences were nonsignificant for both the Bothersomeness subscale and the Frequency subscale at all single follow-up time points in both trials. Further, repeated measures analyses from 4 weeks to 52 weeks and 4 weeks to 2 years also found nonsignificant between-group differences. Arts et al. reported a repeated measures AMD from 4 weeks to 52 weeks of 0.7 (95% CI, -0.1 to 1.5) and from 4 weeks to 2 years of 0.5 (95% CI, -0.3 to 1.3). This measure was not reported in 5 year results reported by Arts et al.[48] See ***Appendix D, Table D-3*** for detailed findings from Brouwer et al.,[39] which are not directly comparable because of an adaptation to the standard scoring approach used by this study (personal communication with author, February 8, 2018).

Other Pain Measures

Three RCTs reported on the frequency and proportion of participants with improvement in their pain. These findings are largely consistent with previously reported pain outcomes for this comparison; however, we did not use findings from these studies in the strength of evidence ratings because of the heterogeneity in outcome definitions used. Specific study findings:

- Mayer et al., which compared percutaneous endoscopic discectomy with microdiscectomy, reported the frequency and proportion of patients reporting low back pain decreased from 20 (100%) at baseline to 4 (20%) at 2 years follow-up among participants allocated to percutaneous endoscopic discectomy and from 20 (100%) to 7 (35%) among participants allocated to microdiscectomy (calculated P=0.48).[34] However, the proportion reporting low back pain at 2 years was higher in the participants allocated to percutaneous endoscopic discectomy (47.4% versus 20%, calculated P=0.18).

- Ryang et al.,[30] reported the frequency and proportion of participants with radicular pain was 27 (90%) among those allocated to trocar discectomy and 29 (97%) among those allocated to microdiscectomy, these frequencies and proportions decreased to 1 (3%) and 5 (17%) after a mean follow-up of 1.3 years (P=0.11).

- Sasaoka et al.,[25] reported the proportion of participants with residual low back pain or lumbar discomfort at 52 weeks follow-up. Among participants allocated to microendoscopic discectomy, 36.7% endorsed residual pain, among participants allocated to microdiscectomy, 66.7% endorsed residual pain, the P value for comparison was reported as not significant.

C. Microdiscectomy compared with discectomy

Three trials comparing microdiscectomy to discectomy reported pain outcomes using a VAS at 4 and 6 weeks,[35] at 52 weeks,[27] and at 26 weeks, 52 weeks, and 2 years.[29] All three were rated as some concerns for bias. Pain outcomes reported included VAS 10 cm measures for leg pain and

VAS 10 cm measures for back pain. ***Table 11*** summarizes the findings and strength of evidence related to pain outcomes for this comparison.

In Henriksen et al., actual VAS values were not reported but no differences were reported between groups for both VAS leg pain and VAS back pain at 4 weeks and at 6 weeks follow-up.[35] Tullberg et al. reported a mean baseline VAS 10 cm leg pain score of 7.0 (SD NR) among participants allocated to microdiscectomy and 7.0 (SD NR) among participants allocated to discectomy.[27] The mean scores at 52 weeks were 2.1 (SD NR) and 2.3 (SD NR), respectively (AMDs and P value NR). The reduction in VAS 10 cm back pain scores was also similar in both surgical groups (baseline 3.6 and 3.7, respectively, 52 weeks 1.6 and 1.8, respectively; AMDs and P value NR). Teli et al. reported VAS 10 cm leg pain score of 8 (SD 1) at baseline decreasing to 2 (1 SD) at 26 weeks, 1 (SD 1) at 52 weeks, and 2 (SD 1) at 2 years in both surgical groups (P=0.73 for between-group differences).[29] A similar finding was observed for VAS 10 cm back pain scores (P=0.75 for between-group differences).

Table 11. Summary of findings and strength of evidence ratings comparing microdiscectomy to discectomy for pain in persons with symptomatic lumbar radiculopathy

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Pain (leg and back)-VAS 100 mm (short-term) (follow-up: range 4 weeks to 6 weeks) (MID 7 to 11 points)						
1 RCT	Serious[a]	Not serious[b]	Not serious	Very serious[c]	Single study body of evidence. Pain decreased in both surgical groups, with no between-group differences. Henriksen et al.[35] (N=80) only depicted outcomes on a figure, actual values and variance were NR.	◯◯◯◯ INSUFFICIENT
Pain (leg and back)-VAS 10 cm (medium- and long-term) (follow-up: range 26 weeks to 2 years) (MID 7 to 11 points)						
2 RCTs	Serious[d]	Not serious	Not serious	Very serious[e]	Improvements in pain persist in both surgical groups, however no between-group differences. Tullberg et al.[27] (N=60) calculated 52w AMD 0.2 for leg pain, 0.1 for back pain (P Values NR). Teli et al.[29] (N=142) calculated AMD 0 at 26w, 52w, and 2y for leg pain (P=0.73 for between-group differences). Similar pattern for back pain (P=0.75 for between-group differences).	⊕◯◯◯ VERY LOW No difference

[a] This RCT was rated as some concerns for bias. Sources of bias: inadequate randomization or allocation concealment and lack of participant and outcome assessor blinding.

[b] Not applicable since body of evidence has only 1 study.

[c] Optimal information size criterion net met: the sample size in this study would be unable to detect differences less than approximately 14 points, actual measured values and measures of variance not provided.

[d] Both studies were rated as some concerns for risk of bias. Sources of bias: inadequate randomization or allocation concealment and lack of participant and outcome assessor blinding.

[e] Optimal information size criterion not met: the sample sizes of both studies would be unable to detect differences less than approximately 16 points in one study[27] and less than 10 points in the other.[29]; actual AMDs not reported and unable to calculate confidence intervals based on data provided.

Abbreviations: AMD = absolute mean difference; MID = minimally important between-group difference; N = number; NR = not reported; RCT = randomized controlled trial; VAS = visual analog scale; w = week(s); y = year(s).

Functioning/Disability

A. Surgery compared with nonsurgical interventions

Five RCTs reported various measures of physical, mental, emotional, and social functioning or disability.[22,23,32,33,41] Three were rated as high risk of bias,[22,32,33] one was rated as some concerns for bias,[23] and one was rated as high risk for outcomes at 12 weeks or later or some concerns for outcomes at 6 weeks or later.[41] Functional outcomes reported include the Oswestry Disability Index, the Physical Functioning subscale of the SF-36, the Roland-Morris Disability Questionnaire, the Prolo Scale, and various other subscales of the SF-36. We were unable to conduct any quantitative synthesis for functional outcomes within this comparison because of outcome measure or reporting heterogeneity or because some studies did not report measures of variance needed to conduct a meta-analysis. ***Table 12*** summarizes the findings and strength of evidence related to functioning/disability outcomes for this comparison. A detailed description of findings follows this table.

Table 12. Summary of findings and strength of evidence ratings comparing surgery to nonsurgical interventions for functioning/disability in persons with symptomatic lumbar radiculopathy (EQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Functioning/Disability-Oswestry Disability Index (short- and medium-term) (follow-up: range 6 weeks to 26 weeks) (MID 8 to 11 points)						
3 RCTs	Very serious[a]	Not serious	Not serious	Serious[b]	Improvements in both treatment groups. Scores decreased by 4 to 10 points more in surgery group. Weinstein et al. [SPORT][22] (N=501) reported larger improvement in surgical group compared to conservative management group at 12w (AMD -4.7 [95% CI -9.3 to -0.2]). Osterman et al.[33] (N=56) calculated 6w AMD -6, 12w AMD -6, 26w AMD -4 (P values NR). Gerszten et al.[41] (N=90) calculated AMDs 6w -8, 12w -9, 26w -10 (P=0.002 for comparison at each timepoint).	⊕◯◯◯ VERY LOW Favors surgery
Functioning/Disability-Oswestry Disability Index (long-term) (follow-up: range 52 weeks to 8 years) (MID 8 to 11 points)						
2 RCTs	Very serious[c]	Not serious	Not serious	Serious[d]	Short-term improvements persist over time in both treatment groups; however, no significant between-group differences. Osterman et al.[33] (N=56) RM AMD 6w to 2y AMD -3 (95% CI, -10 to 4). Weinstein et al. [SPORT][22] (N=501) RM AMDs from 12w to 2y (NR, P=0.21), 12w to 4y (AMD NR, P=0.074), and 12w to 8y (AMD NR, P=0.096).	⊕◯◯◯ VERY LOW No difference

(continued)

Table 12. Summary of findings and strength of evidence ratings comparing surgery to nonsurgical interventions for functioning/disability in persons with symptomatic lumbar radiculopathy (EQ1) (continued)

Certainty Assessment					Summary of Findings	CERTAINTY /Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Functioning/Disability-Roland-Morris Disability Questionnaire (short- and medium-term) (follow-up: range 6 weeks to 26 weeks) (MID 2 to 5 points)						
2 RCTs	Very serious[e]	Serious[f]	Not serious	Serious[g]	Improvements in both treatment groups. Mixed findings with respect to between-group differences. Peul et al.[32] (N=283) 8w AMD -3.1 (95% CI, -4.3 to -1.7), 26w AMD -0.8 (95% CI, -2.1 to 0.5). McMorland et al.[23] (N=40) RM AMD 6w to 12w NR, P=0.199.	◯◯◯ INSUFFICIENT Mixed findings
Functioning/Disability-Roland-Morris Disability Questionnaire (long-term) (follow-up: range 52 weeks to 5 years) (MID 2 to 5 points)						
1 RCT	Very serious[h]	Not serious[i]	Not serious	Not serious	Short-term improvements persisted over time in both treatment groups. Between-group differences were not significant. Peul et al.[32] (N=283) 52w AMD -0.4 (95% CI, -1.7 to 0.9), 2y AMD -0.5 (95% CI, -1.8 to 0.8), and 5y AMD 0.1 (95% CI, -1.3 to 1.4). Cumulative score 8w to 52w and 8w to 2y also with no significant difference between groups.	◯◯◯ INSUFFICIENT Single study
Functioning/Disability-SF-36 Physical Functioning subscale (short- and medium-term) (follow-up: range 6 weeks to 26 weeks) (MID 3 points)						
3 RCTs	Very serious[a]	Serious[j]	Not serious	Serious[k]	Improvements in both treatment groups. Mixed between-group differences in short-term. Peul et al.[32] (N=283) AMD 8w 9.3 (95% CI, 4.4 to 14.2), AMD 26w 1.5 [95% CI, -3.4 to 6.4]). McMorland et al.[23] (N=40) calculated AMDs 6w 1.2, 12w 11.5, RM AMD 6w to 12w NR, P=0.720). Weinstein et al. [SPORT][22] (N=501) AMD 12w 2.8 (95% CI, -2.5 to 8.1).	◯◯◯ INSUFFICIENT Mixed findings
Functioning/Disability-SF-36 Physical Functioning subscale (long-term) (follow-up: range 52 weeks to 8 years) (MID 3 points)						
2 RCTs	Very serious[l]	Not serious	Not serious	Serious[m]	Short-term improvements persisted over time in both treatment groups, but no significant difference between groups. Peul et al.[32] (N=283) 52w AMD 2.2 (95% CI, -2.8 to 7.2). Weinstein et al. [SPORT][22] (N=501) RM AMD 12w to 2y (NR, P=0.71), 12w to 4y (NR, P=0.42), and 12w to 8y (NR, P=0.47).	⊕◯◯◯ VERY LOW No difference

[a] Risk of Bias was high in all trials for outcomes at 12 weeks or later, some concerns for bias in Gerszten et al. 6 week outcomes.[41] Sources of bias: lack of participant and outcome assessor blinding and extensive and differential crossovers, high attrition.

[b] Optimal information size criterion not met in smaller trial:[33] sample size able to detect differences larger than approximately 12.5 points.

[c] Risk of Bias was high in both trials, sources of bias: lack of participant and outcome assessor blinding, extensive and differential crossovers.

[d] Optimal information size criterion not met in one trial: sample size able to detect differences larger than approximately 12.5 points in 1 trial.[33] AMDs and confidence intervals NR in one trial.[22]

[e] Risk of Bias was high in 1 trial,[32] and was some concerns in the other trial.[23] Sources of bias: lack of participant and outcome assessor blinding in both trials, extensive crossovers in the trial rated as high risk of bias.

[f] One study reported meaningful difference at 8 weeks,[32] however the other study reported no meaningful difference in repeated measures from 6 weeks to 12 weeks. [23]

[g] Optimal information size criterion not met: the smaller trial did not report AMD or confidence intervals and sample size only able to detect differences larger than approximately 4.5 points.

[h] Risk of Bias was rated as high in this trial. Sources of bias: lack of participant and outcome assessor blinding and extensive crossovers.

[i] Not applicable as only 1 study is in this body of evidence.

[j] One of the three RCTs reported meaningful difference at 8 weeks,[32] however, the other two studies reported no meaningful differences.[22,23]

[k] Optimal information size criterion not met: Sample size of 1,398 required for between-group difference of 3 points; smallest trial[23] unable to detect differences smaller than approximately 20 points, sample sizes of other trials unable to detect differences smaller than 7 to 8 points.

[l] Risk of Bias high in both studies. Sources of bias: lack of participant and outcome assessor blinding and extensive and differential crossovers, and high attrition at long-term follow-up in one study.[22]

[m] Optimal information size criterion not met: Sample size of 1,398 required for between-group difference of 3 points; sample size of trials unable to detect differences smaller than 7 to 8 points. AMDs and confidence intervals NR in larger trial.[22]

Abbreviations: AMD = absolute mean difference; CI = confidence interval; MID = minimally important between-group difference; N = number; NR = not reported; RCT = randomized controlled trial; RM = repeated measures; SF-36 = Short Form 36; SPORT = Spine Patient Outcomes Research Trial; w = week(s); y = year(s).

Oswestry Disability Index

Three RCTs reported outcomes using the Oswestry Disability Index.[22,33,41] On this index, which ranges from 0 to 100, higher scores represent worse functional status and a negative between-group difference favors surgery. Across studies, function improved in both participants allocated to surgery (range 12 to 26 point decrease) and in participants allocated to nonsurgical interventions (range 5 to 21.3 decrease) at short-term follow-up. Scores on this index decreased by 4.7 to 10 points more among participants allocated to surgery at short- and medium-term follow-up. Between-group differences did not persist in the long-term. Specific study findings:

- Weinstein et al. [SPORT], which compared discectomy/microdiscectomy to conservative management reported this index at 12 weeks, 52 weeks, 2 years, 4 years, and 8 years.[22,99,100] Only the AMD at 12 weeks demonstrated a significant difference between treatment groups, favoring discectomy/microdiscectomy (AMD -4.7 [95% CI, -9.3 to -0.2]). All other follow-up time points demonstrated a larger numeric improvement for participants allocated to surgery, but these differences were not significant. Further, repeated measures from 12 weeks to 2 years, 4 years, and 8 years also did not demonstrate a significant difference (AMDs NR, P=0.21, P=0.074, and P=0.096, respectively).

- Osterman et al. reported improvements in the index among participants allocated to microdiscectomy and to participants allocated to physiotherapy at all follow-up time points (6 weeks, 12 weeks, 26 weeks, 52 weeks, and 2 years).[33] Calculated AMD at 6 weeks was -6 at both 6 weeks and 12 weeks and was -4 at 26 weeks (P values NR). The repeated measures AMD from 6 weeks to 2 years follow-up did not demonstrate a difference between treatment groups (AMD -3 [95% CI, -10 to 4]).

- Gerszten et al. reported improvements in the index among participants allocated to plasma disc decompression and among participants allocated to epidural steroid injection. Significant between-group differences favoring surgery were observed at 6 weeks, 12 weeks, and 26 weeks (calculated AMDs ranged from -8 to -10, all comparisons P=0.002).[41]

Roland-Morris Disability Questionnaire

Two RCTs reported outcomes using the Roland-Morris Disability Questionnaire.[23,32] On this scale, which ranges from 1 to 24, higher scores represent worse functional status and a negative mean difference favors surgery. Across studies, function improved in participants allocated to both surgery (range 0.7 to 10.4 point decrease) and nonsurgical intervention (range 2.5 to 7.1 point decrease). Short-term between group differences were inconsistent between RCTs; one suggested a favorable effect for surgery[32] while the other suggested no meaningful difference.[23] Differences in how this outcome was analyzed (one used single time point analysis,[32] while the other used repeated measure analysis[23]) may explain this inconsistency. The study reporting a favorable effect at 8 weeks did not observe a meaningful difference at 26 weeks;[32] the other study did not report any outcomes beyond 12 weeks.[23] Because of these inconsistencies, we concluded the evidence was insufficient to draw a conclusion for this outcome in the short- and medium-term. Because of a single-study body of evidence with very serious concerns in at least one domain, we concluded the evidence was also insufficient to draw a conclusion in the long

- McMorland et al. reported significant improvements in function between baseline and 12 weeks in participants allocated to microdiscectomy and in participants allocated to spinal manipulation.[23] However, the repeated measures analysis from 6 weeks to 12 weeks found no significant differences between treatment groups (AMDs NR, P=0.199).

- Peul et al. reported a larger improvement in participants allocated to microdiscectomy compared with participants allocated to conservative management (AMD -3.1 [95% CI, -4.3 to -1.7]) at 8 weeks.[32] This difference between groups did not persist at 26 weeks, 52 weeks, 2 years, or 5 years and cumulative scores at 52 weeks and 2 years also demonstrated no significant difference between groups.

SF-36 Physical Functioning

Three RCTs reported outcomes using the Physical Functioning subscale of the SF-36.[22,23,32] On this scale, which ranges from 0 to 100, higher scores represent better functional status and a positive between-group difference favors surgery. Across studies, function improved in participants allocated to both surgery (range 8.6 to 37.3 point increase) and nonsurgical intervention (range 7.4 to 27.3 point increase). One of the three RCTs reported a larger increase among participants allocated to surgery compared with participants allocated to nonsurgical interventions in the short-term,[32] but these differences did not persist in the medium- or long-term. The other two RCTs observed no between-group differences in the short-, medium-, or long-term.[22] The specific study findings:

- McMorland et al. reported improvements at 6 weeks and 12 weeks among participants allocated to microdiscectomy and among participants allocated to spinal manipulation.[23]

However, repeated measures analysis found no significant difference between treatment groups from 6 weeks to 12 weeks (AMD NR, P=0.720).

- Peul et al. reported improvements among participants allocated to microdiscectomy and in participants allocated to conservative management; the between-group difference favored the surgical group at 8 weeks (AMD 9.3 [95% CI, 4.4 to 14.2]) but was not significantly different at 26 weeks (AMD 1.5 [95% CI, -3.4 to 6.4]) or 52 weeks (AMD 2.2 [95% CI, -2.8 to 7.2]).

- Weinstein et al. [SPORT] reported improvements among participants allocated to microdiscectomy/discectomy and in participants allocated to conservative management. Between-group differences were slightly larger at most follow-up time points (range of AMD 0 to 2.8 at 12 weeks, 52 weeks, 2 years, 4 years, 8 years) among participants allocated to discectomy/microdiscectomy compared to participants allocated to conservative management, but these differences were not statistically significant at any follow-up time point or in any of the three repeated measures analyses (12 weeks to 2 years [AMD NR, P=0.71], 4 years [AMD NR, P=0.42], and 8 years [AMD NR, P=0.47]).

Other Measures of Function

Two RCTs also reported on other various subscales of the SF-36, including social functioning, mental functioning, role emotional, and role physical.[23,32] Most scores improved over time among participants allocated to surgical interventions and among participants allocated to nonsurgical interventions but no significant differences in improvement between treatment groups were observed. One RCT reported outcomes using the functional and economic subscales of the Prolo scale.[32] Peul observed a significant improvement in the functional score at 26 weeks (AMD 0.5 [95% CI, 0.2 to 0.7]), but not at 8 weeks (AMD 0.8 [95% CI -0.6 to 1.1]) or 52 weeks (AMD -0.04, [95% CI -0.3 to 0.2]).

B. Minimally-invasive surgery compared to microdiscectomy or discectomy

Eight RCTs reported at least one outcome related to functioning or disability. One was rated as low risk of bias for outcomes at 2 years or less and high risk of bias for outcomes longer than 2 years,[40] four were rated as some concerns for bias,[28,29,36,39] and three were rated as high risk of bias.[30,31,42] Functional outcomes reported include the Oswestry Disability Index, the Physical Functioning subscale of the SF-36, the Roland-Morris Disability Questionnaire, the Prolo Scale, and other subscales of the SF-36. ***Table 13*** summarizes the findings and strength of evidence related to functioning/disability outcomes for this comparison. A detailed description of findings follows this table.

Table 13. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for functioning/disability in persons with symptomatic lumbar radiculopathy (EQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Oswestry Disability Index (medium- and long-term) (follow-up: range 12 weeks to 2.8 years) (MID 8 to 11 points)						
4 RCTs	Very serious[a]	Not serious	Not serious	Serious[b]	Function improved in both surgical groups. No significant between-group differences at any follow-up time. Ruetten et al.[31](N=200) calculated AMDs range -5 to -6 at 12w, 26w, 52w, 2y. Franke et al.[36](N=100) RM AMD 8w to 52w NR (P=0.08). Ryang et al.[30](N=60) calculated AMDs range -1.98 to 3.6, P=0.83 for between-group difference at 1.3y, P reported as NS for between-group difference at 2.8y. Teli et al.[29](N=142) calculated AMD range 1 to 2 at 26w, 52w, 2y.	⊕◯◯◯ VERY LOW No difference
Roland-Morris Disability Questionnaire (short-term) (follow-up: range 4 weeks to 8 weeks) (MID 2 to 5 points)						
2 RCTs	Not serious	Serious[c]	Not serious	Serious[b]	Function improved in both surgical groups. Between-group differences were significant at 4w in Brouwer et al.[39] (N=115)(AMD -2.5 (95% CI, -4.7 to -0.2) but not at 8w (AMD 0.1 [95% CI, -2.1 to 2.3). Between-group differences reported by Arts et al.[40] (N=328) were not significant (4w AMD 0.2 [95% CI, -1.1 to 1.4], 8w AMD 0.8 [95% CI, -0.4 to 2.1]).	⊕⊕◯◯ LOW No difference
Roland-Morris Disability Questionnaire (medium- and long-term) (follow-up: range 26 weeks to 5 years) (MID 2 to 5 points)						
3 RCTs	Very serious[d]	Not serious	Not serious	Serious[b]	Improvements in function persist in both surgical groups; with one exception, between-group differences were not significant. Haines et al.[42] (N=34) 26w calculated AMD 0.02 (P=0.74) Arts et al.[40] (N=328) 26w AMD 1.0 (95% CI, -0.2 to 2.3), 52w AMD 1.3 (95% CI, 0.03 to 2.6), 2y AMD 0.8 (95% CI, -0.5 to 2.1), 5y AMD 0.9 (95% CI, -0.6 to 2.2); RM AMD 4w to 52w 0.8 (95 %CI -0.2 to 1.7) and 4w to 2y 0.6 (-0.3 to 1.6), RM AMD 4w to 5y P=0.30 (calculated AMD 0.4) Brouwer et al.[39] (N=115) 26w AMD 2.2 (95% CI, -0.1 to 4.4), 52w AMD 1.1 (-1.1 to 3.4), 2y AMD -0.1 (95% CI, -2.4 to 2.2). RM AMD 4w to 52w 0.1 (95 %CI, -1.2 to 1.6); RM AMD 4w to 2y 0 (95% CI -1.3 to 1.3).	⊕◯◯◯ VERY LOW No difference
SF-36 Physical Functioning (short-term) (follow up: range 4 weeks to 8 weeks) (MID 3 points)						
2 RCTs	Not serious	Very serious[c]	Not serious	Serious[e]	Function improves in both surgical groups. Between-group differences were not significant in Arts et al.[40](N=328) 4w AMD -1.1 (95% CI, -5.6 to 3.3); 8w AMD -3.3 (95% CI, -7.8 to 1.1). Brouwer et al.[39](N=115) between-group differences favored minimally-invasive surgery at 4w (AMD 18.4 [95% CI, 10.0 to 26.8]) but no differences at 8w (AMD 5.6 [95% CI, -2.7 to 13.9]).	◯◯◯◯ INSUFFICIENT[c]

(continued)

Table 13. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for functioning/disability in persons with symptomatic lumbar radiculopathy (EQ1) (continued)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
SF-36 Physical Functioning (medium- term) (follow up: range 12 weeks to 26 weeks) (MID 3 points)						
4 RCTs	Very serious[f]	Not serious	Not serious	Serious[g]	Improvements in function persist over time in both surgical groups. No significant between-group differences. Pooled between-group mean difference in scores at 12w to 26w -2.4 (95% CI, -6.1 to 1.2; 4 RCTs; 527 participants; I^2=0%).[28,39,40,42]	⊕◯◯◯ VERY LOW No difference
SF-36 Physical Functioning (long-term) (follow up: range 52 weeks to 2 years) (MID 3 points)						
4 RCTs	Very serious[h]	Serious[i]	Not serious	Serious[j]	Improvements in function persist over time in both groups, but mixed findings for between-group differences. Ryang et al.[30](N=60) calculated AMD at 1.3y -10.1 (P=0.64) and calculated AMD at 2.8y 6.0 (P=0.436). Thome et al.[28](N=84) calculated AMD at 2y 7.0 (P=0.026), favoring minimally-invasive surgery. Arts et al.[40](N=328) AMD 52w -4.8 (95% CI, -9.3 to -0.2) favoring microdiscectomy, but between-group differences at 2y and in RM 4w to 52w and to 2y were NS. Brouwer et al.[39](N=115) AMDs at 52w and 2y were NS, RM AMD 4w to 52w 4.3 (95% CI, -4.5 to 13.2), but RM AMD 4w to 2y favored minimally-invasive surgery (6.1 [95% CI, 0.5 to 11.7]).	◯◯◯◯ INSUFFICIENT[c]
Prolo Scale (short-term) (follow-up: range 4 weeks to 8 weeks) (MID 0.4 points)						
2 RCTs	Not serious	Serious[c]	Serious[k]	Not serious	Function improved in both surgical groups. Arts et al.[40] (N=328) between-group differences were not significant in either the functional subscale (4w AMD 0 [95% CI, -0.3 to 0.2]; 8w AMD -0.1 [95% CI, -0.3 to 0.2]) or the economic subscale (4w AMD 0.2 [95% CI, -0.1 to 0.5]; 8w AMD 0.1[95% CI, -0.2 to 0.4]). Brouwer et al. significant between-group difference in economic subscale at 4w (AMD 1.1 [95% CI, 0.5 to 1.6]) but not at 8w (AMD 0.2 [95% CI, -0.3 to 0.8)] or in functional subscale (4w AMD 0.2 [95% CI, -0.2 to 0.6]; 8w AMD -0.2 [95% CI 0.6 to 0.3]).	⊕⊕◯◯ LOW No difference

(continued)

Table 13. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for functioning/disability in persons with symptomatic lumbar radiculopathy (EQ1) (continued)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Prolo Scale (medium- and long-term) (follow-up: range 12 weeks to 2 years) (MID 0.4 points)						
3 RCTs	Serious[l]	Not serious	Serious[k]	Not serious	Improvements in function persist in both surgical groups in Arts et al.[40] (N=328) and Brouwer et al.[39] (N=115) with no significant between-group differences at 26w or 52w or in repeated measures from 0 to 52w. No significant between-group differences observed at 12w to 26w in Thome et al.[28] (N=84) in proportion of reporting sum of functional and economic subscale scores > 7 (0 [worse] to 10 [best] on a modified Prolo scale); 92% among participants allocated to sequestrectomy compared to 76% among participants allocated to discectomy (P=0.11).	⊕⊕◯◯ LOW No difference

[a] Two trials were rated as high risk of bias,[31,110] 2 trials were rated as some concerns.[29,36] Sources of bias: lack of blinding intervention and outcome assessors (all trials), and inadequate randomization and allocation concealment.[31,110]

[b] Optimal information size criterion not met: trials did not have a sample size to detect a between-group difference of at least 2 points (for Roland-Morris Disability Questionnaire) or 5 points (for Oswestry Disability Index), which is the low end of the minimally important clinical difference for these measures.

[c] Inconsistent finding at 4w; much larger treatment effect in Brouwer et al.[39] favoring minimally-invasive surgery; this finding is not consistent with the findings at 8w or with the findings at 4w and 8w in the other RCT (Arts et al.[40]). For the SF-36 Physical Functioning measure, the between-group difference was large enough that we considered the inconsistency a very serious concern and led us to conclude mixed effects and we assigned a rating of 'insufficient'.

[d] One trial was low risk of bias,[40] 1 trial was some concerns for bias,[39] and 1 trial was high risk of bias.[42] Sources of bias: Sources of bias: lack of blinding intervention and outcome assessors[39,42] and inadequate randomization and allocation concealment and high attrition.[42]

[e] Optimal information size criterion not met: sample size of 1,398 required to detect a difference of 3 points; sample size in Arts et al.[40] unable to detect differences smaller than 6 points and Brouwer et al[39] unable to detect differences smaller than approximately 11 points.

[f] Risk of bias was low in 1 trial[40], some concerns in 2 trials[28,39] and high in 1 trial.[42] Sources of bias: lack of participant and outcome assessor blinding in all but the low risk of bias trial, and inadequate randomization and high attrition.[42]

[g] Optimal information size criterion not met; a sample size of 1,398 required to detect a difference of 3 points, the pooled sample size was only 527 participants; the largest trial sample size only able to detect differences of approximately 6 points or greater[40], and the smallest trial[42] sample size only able to detect differences of approximately 15 points or greater.

[h] Risk of bias was low in 1 trial[40], some concerns in 2 trials[28,39] and high in 1 trial.[30] Sources of bias: lack of participant and outcome assessor blinding in all but the low risk of bias trial, and inadequate randomization and high attrition.[30]

[i] Some inconsistency in magnitude and direction of effects across the 4 studies.

[j] Optimal information size criterion not met: sample size of 1,398 required to detect a difference of 3 points; largest trial sample size (Arts et al.[40]) unable to detect differences smaller than 6 points; smallest trial sample size (Ryang et al.[30] unable to detect differences smaller than approximately 12.5 points.

[k] This scale is based on an observers' assessment of the patients' functional status; thus, is less direct than other measures based on patient self-report.

[l] The risk of bias was low in 1 trial[40], and some concerns in the other 2 trials.[28,39] Sources of bias lack of participant and outcome assessor blinding.

Abbreviations: AMD = absolute mean difference; CI = confidence interval; N = number; NS = not significant; NR = not reported; RCT = randomized controlled trial; RM = repeated measure; SF-36 = Short Form 36; w = week(s); y = year(s).

Oswestry Disability Index

Four RCTs reported outcomes using the Oswestry Disability Index.[29-31,36] Higher scores on this index represent worse functional status and a negative between-group difference favors minimally-invasive surgery. The range of decreases in scores at the earliest follow-up (12 weeks to 26 weeks) in each RCT was 28 to 53 points among participants allocated to minimally-invasive surgery and 29 to 47 points among participants allocated to standard surgery. No significant between-group differences were observed in the medium- or long-term. Specific study findings:

- Ruetten et al.,[31] which compared endoscopic discectomy with microdiscectomy, reported that scores at 12 weeks decreased from baseline by 53 points among participants allocated to endoscopic discectomy and by 47 points among participants allocated to microdiscectomy; however, between group differences were reported as not significant (calculated AMD -6). These decreases persisted at 26 weeks, 52 weeks, and 2 years (calculated AMDs range -5 to -6, all results reported as not significant).

- Teli et al.,[29] which compared microendoscopic discectomy with microdiscectomy, reported decreases from baseline to 26 weeks of 28 points among participants allocated to microendoscopic discectomy and 29 points among participants allocated to microdiscectomy; between-group differences were not significant (calculated AMD 1). These decreases persisted at 52 weeks (calculated AMD 2) and 2 years (calculated AMD 1), with between-group differences remaining nonsignificant.

- Franke et al.,[36] which compared microscopically-assisted percutaneous nucleotomy with microdiscectomy, reported decreased scores among both groups, but no significant between-group difference (P=0.08) at 52 weeks or 2 years (AMDs NR and unable to be calculated).

- Ryang et al.,[30] which compared trocar microdiscectomy with microdiscectomy, reported decreases of 41 points among participants allocated to trocar microdiscectomy and 44.7 points among participants allocated to microdiscectomy at 1.3 years (P=0.83, calculated AMD 3.6). Although scores increased slightly at 2.8 years among participants allocated to microdiscectomy, the between-group differences remained nonsignificant (calculated AMD -1.98).

Roland Morris Disability Questionnaire

Three RCTs reported outcomes with the Roland-Morris Disability Questionnaire.[39,40,42] Higher scores represent worse functional status and a negative mean difference favors minimally-invasive surgery. The range of decreases in scores at the earliest follow-up in each study (4 weeks to 26 weeks) was 4.9 to 9.7 points among participants allocated to minimally-invasive surgery and 2.3 to 10.6 points among participants allocated to standard surgery. Between-group differences favoring minimally-invasive surgery at 4 weeks was observed by one RCT,[39] this difference did not persist at 8 weeks and with one exception,[40] no other between-group differences were observed at any other time points. Specific study findings:

- Arts et al.,[40] which compared tubular discectomy to microdiscectomy, observed an AMD of 0.2 (95% CI, -1.1 to 1.4) at 4 weeks and 0.8 (95% CI, -0.4 to 2.1) at 8 weeks. Nonsignificant between-group differences persisted at 26 weeks (AMD 1.0 [95% CI -0.2 to 2.3]). For long-term outcomes, a very small, significant between-group difference was observed at 52 weeks (AMD 1.3 [95% CI 0.03 to 2.6]) but not at 2 years (AMD 0.8 [95 % CI -0.5 to 2.1]) or 5 years (AMD 0.9 [95% CI, -0.6 to 2.2]).[40,48,101] Repeated measures between-group differences from 4 weeks to 52 weeks (P=0.11), 4 weeks to 2 years (P=0.17), and 4 weeks to 5 years (P=0.30) were not significant.

- Brouwer et al.,[39] which compared percutaneous laser disc decompression to microdiscectomy, observed a significant between-group difference at 4 weeks (AMD -2.5 [95% CI, -4.7 to -0.2]) but not at 8 weeks (AMD 0.1 [95% CI, -2.1 to 2.3]). Nonsignificant between-group differences persisted at 26 weeks (AMD 2.2 [95% CI, -0.1 to 4.4]), 52 weeks (AMD 1.1 (95% CI, -1.1 to 3.4), and 2 years -0.1 (95% CI, -2.4 to 2.2). In addition, nonsignificant repeated measures between-group differences were observed for 4 weeks to 52 weeks and for 4 weeks to 2 years.

- Haines et al.,[42] which compared automated percutaneous or endoscopic discectomy with discectomy, reported no significant between-group difference (P=0.74) in change from baseline scores at 26 weeks (calculated AMD 0.02).

SF-36 Physical Functioning

Five RCTs reported outcomes with the SF-36 Physical Functioning subscale.[28,30,40,42,39] Higher scores on this index represent better functional status and a positive between-group difference favors minimally-invasive surgery. The range of increases in scores at the earliest follow-up in each study (4 weeks to 26 weeks) was 27.2 to 41.8 points among participants allocated to minimally-invasive surgery and 2.6 to 51.9 points among participants allocated to standard surgery. Similar to the Roland-Morris Disability outcome previously reported for this comparison, between-group differences favoring minimally-invasive surgery at 4 weeks were observed by one study.[39] This difference did not persist at 8 weeks. In the medium-term (12 weeks to 26 weeks), the pooled between-group mean difference was -2.4 (95 % CI, -6.1 to 1.2, 4 RCTs, 527 participants, I^2=0.0%, ***Appendix G, Figure G-4***). Some between-group differences were observed at 52 weeks and 2 years; but the findings were mixed with respect to which group was favored.

- Arts et al.[40] and Brower et al.[39] reported increases in scores from baseline to 4 weeks and 8 weeks among participants allocated to the minimally-invasive surgical interventions and among participants allocated to microdiscectomy. Arts et al. reported no between-group differences at 4 weeks (AMD -1.1 (95% CI, -5.6 to 3.3) or 8 weeks (AMD -3.3 (95% CI -7.8 to 1.1).[40] In contrast, Brouwer et al. observed a significant, and clinically-relevant between group difference favoring percutaneous laser disc decompression compared with microdiscectomy at 4 weeks (AMD 18.4 [95% CI, 10.0 to 26.8]); this difference did not persist at 8 weeks (AMD 5.6 [95% CI, -2.7 to 13.9]).[39] No between group differences in the medium-term were observed by Arts et al. (AMD at 26 weeks -3.9 [95% CI, -8.3 to 0.6]),[40] or Brouwer et al. (AMD at 26 weeks -3.2 [95% CI, -11.6 to 5.1]).[39] In the long-term,

Brouwer et al.[39] observed an AMD of -3.2 [95% CI -11.6 to 5.2] at 52 weeks and 4.3 [95% CI -4.5 to 13.2] at 2 years. Repeated measures AMD for 4 weeks to 52 weeks was 5.3 (95% CI, -0.7 to 11.2) but was significant for 4 weeks to 2 years (RM AMD 6.1 [95% CI, 0.5 to 11.7]). In contrast, Arts et al. reported a significant between-group difference favoring microdiscectomy at 52 weeks (AMD -4.8 [95% CI, -9.3 to -0.2]), but this difference was not present at 2 years (AMD 0.8 [95% CI, -0.5 to 2.1]).[40] Repeated measures between group differences from 4 weeks to 52 weeks (-3.1 [95% CI, -6.8 to 0.7]), 4 weeks to 2 years (-2.8 [95% CI, -6.5 to 0.9])). Arts et al did not report 5 year outcomes for this measure.[48]

- Thome et al.[28] reported no significant between-group differences at 12 to 26 weeks (calculated AMD 2.2, P=0.32 for between-group difference at follow-up). However, between-group differences observed at 2 years were significant and favored minimally-invasive surgery (calculated AMD 7, P=0.026 for between-group difference at follow-up).

- Haines et al.[42] reported no significant between-group differences at 26 weeks (calculated AMD 2.9, P=0.96).

- Ryang et al.[30] reported no significant between-group differences at 1.3 years (calculated AMD -10.1, reported P for between-group comparison at follow-up=0.64) or at 2.8 years (calculated AMD 6.0, reported P for between-group comparison at follow-up=0.436).

Prolo Scale

Three RCTs reported functioning/disability using the Prolo Scale, which has a functional subscale and an economic subscale.[28,39,40] On this measure, higher scores represent better function and a positive AMD favors minimally-invasive surgery. Function improved in both surgical groups; with one exception at 4 weeks in one study, between-group differences were not significant at any time point.

- Brouwer et al.[39] reported no significant between-group differences at 4 weeks or 8 weeks in the functional subscale at 4 or 8 weeks (AMD 0.2 [95% CI, -0.2 to 0.6]; AMD -0.2 [95% CI -0.6 to 0.3], respectively) or in the economic scale at 8 weeks (AMD 0.2 [95% CI, -0.3 to 0.8), but did observe a significant difference in the economic subscale at 4 weeks (AMD 1.1 [95% CI, 0.5 to 1.6]).

- Arts et al.[40] reported no significant between-group differences in either subscale at 4 weeks or at 8 weeks. At 26 weeks and at 52 weeks, no significant between-group differences were observed in either study and repeated measures between-group differences from 4 weeks to 52 weeks were also not significant. Arts et al did not report this measure at 2 years or 5 years of follow-up.

- Thome et al.[28] also reported outcomes at 12 to 26 weeks and at 2 years using the Prolo scale, but used a modified scale that rated each subscale using a 5-point Likert scale (instead of a 4 point) and summed the functional and economic subscales to obtain a total score (range of total score could vary from 2 to 10, higher scores represent more improvement). The proportion of participants with a total score greater than or equal to 7 at

12 to 26 weeks was 92% among participants allocated to sequestrectomy and 76% among participants allocated to microdiscectomy (P=0.11). This study also compared the proportion of participants with scores in the following categories: 1 to 4, 5 to 6, 7 to 8, and 9 to 10. No difference in proportion of participants in these categories was observed at 12 to 26 weeks (P=0.852) or at 2 years (P=0.20).

Other Measures of Function

Two RCTs reported outcomes related to function/disability in the medium- and long-term using various subscales of the SF-36 including role-emotional, role-physical, and social functioning.[28,30] With few exceptions, no between-group differences were observed at any follow-up.

C. Microdiscectomy compared to discectomy

One RCT rated as some concerns for bias compared microdiscectomy to standard discectomy and reported outcomes with the Oswestry Disability Index.[29] ***Table 14*** summarizes the findings and strength of evidence related to functioning/disability outcomes for this comparison. Teli et al. observed scores improve at 26 weeks, 52 weeks, and 2 years among participants allocated to microdiscectomy and in participants allocated to discectomy. Among participants allocated to microdiscectomy, score decreases from baseline (40 [SD 4]) ranged from 25 to 29 points; among participants allocated to discectomy score decreases from baseline (39 [SD 4]) ranged from 24 to 27 points. No significant between-group differences were observed (P=0.81).

Table 14. Summary of findings and strength of evidence ratings comparing microdiscectomy to discectomy for functioning/disability in persons with symptomatic lumbar radiculopathy

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Functioning/disability- Oswestry Disability Index (medium- and long-term) (follow-up: range 26 weeks to 2 years) (MID 8 to 11 points)						
1 RCT	Serious[a]	Not serious[b]	Not serious	Serious[c]	Function improved at 26w, 52w, and 2y in both surgical groups, but no significant between-group differences. Teli et al.[29] (N=142) score decreases from baseline (40 [SD 4]) ranged from 25 to 29 points among participants allocated to microdiscectomy and score decreases from baseline (39 [SD 4]) ranged from 24 to 27 points among participants allocated to discectomy. No significant between-group differences were observed (calculated AMDs 26w 2, 52w 2, 2y 1, P=0.81 across comparisons at all time points).	◯◯◯◯ INSUFFICIENT

[a] This RCT was rated as some concerns for bias. Sources of bias: lack of participant and outcome assessor blinding.

[b] Not applicable since body of evidence has only 1 study.

[c] This study's sample size would be unable to detect differences less than approximately 8 points.

Abbreviations: AMD = absolute mean difference; MID = minimally important between-group difference; RCT = randomized controlled trial; SD = standard deviation; w = week(s); y = year(s).

Quality of life

A. Surgery compared with nonsurgical interventions

Two RCTs reported health-related QOL outcomes. One was rated as some concerns for bias[23] and one was rated as high risk of bias.[33] These studies reported outcomes using the total SF-36 score[23] (sum of all normed subscales, possible range 0 to 800) and the 15D QOL measure (range 0 to 1.0).[33] For both measures, a higher score represents better QOL and a positive between-group difference favors surgery. ***Table 15*** summarizes the findings and strength of evidence related to QOL outcomes for this comparison. A detailed description of findings follows this table.

Table 15. Summary of findings and strength of evidence ratings comparing surgery to nonsurgical interventions for quality of life in persons with symptomatic lumbar radiculopathy (EQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Quality of Life- (short-term) (follow-up: range 6 weeks to 26 weeks)						
2 RCTs	Serious[a]	Not serious	Not serious	Very Serious[b]	QOL improved in both treatment groups. McMorland et al.[23] (N=40) observed no between-group differences in RM AMD for cumulative total SF-36 score at 12w (NR, P=0.382). Osterman et al.[33](N=56) calculated between-group AMDs in 15D QOL measure ranging from 0.01 to 0.05 at 6w, 12w, and 26w.	⊕◯◯◯ VERY LOW No difference
Quality of life- (long-term) (52w to 2 years)						
1 RCT	Very serious[c]	Not serious[d]	Not serious	Serious[e]	QOL improvements in both treatment groups. Osterman et al.[33](N=56) calculated between-group AMDs in 15D QOL measure ranging from 0.01 to 0.05 at the various time points reported. RM AMD from 6w to 2y was 0.03 (95% CI -0.01 to 0.07).	◯◯◯◯ INSUFFICIENT

[a] Risk of bias was some concerns in one study, [23] and high in the other study.[33] Sources of bias: lack of participant and outcome assessor blinding in both studies and extensive and differential crossovers in the study rated as high risk of bias.

[b] A sample size of 518 would be required to detect a minimum between-group difference of 10% of the value of baseline scores (approximately 37 points for total SF-36 score).

[c] Risk of bias was high. Sources of bias: lack of participant and outcome assessor blinding

[d] Consistency unknown as is a single-study body of evidence.

[e] A sample size of 24 is required to detect a minimum between-group difference of 10% of the value of baseline scores (approximately 0.08 for 15D HRQOL measure). A sample size of 75 is required to detect a significant between-group difference of 0.03.

Abbreviations: MID = minimally important between-group difference; AMD = absolute mean difference; CI = confidence interval; QOL = quality of life; HRQOL = health-related quality of life; N = number; NR = not reported; RCT = randomized controlled trial; RM = repeated measures; SF-36 = Short Form 36; w = week(s); y = year(s).

In both studies, QOL improved from baseline to follow-up in the surgery and nonsurgical comparator groups; no significant differences between groups were observed by either study.

- McMorland et al.[23] reported a mean (SD) total SF-36 score of 379.5 (149.8) in participants allocated to microdiscectomy and 381.3 (161.9) among participants allocated to spinal manipulation at baseline. These scores improved to 429.1 (157.3) and 445.6 (142.8) at 6 weeks and 500.3 (179.7) and 484.6 (148.9) at 12 weeks, respectively. The repeated

measures AMD from 6 weeks to 12 weeks found no between-group difference (NR, P=0.382).

- Osterman et al.[33] reported the 15D QOL measure at baseline and at various follow-up times from 6 weeks to 2 years. Baseline QOL was 0.83 (SD 0.07) among participants allocated to microdiscectomy and 0.84 (SD 0.06) among participants allocated to physiotherapy. Calculated AMDs between groups were small ranging from 0.01 to 0.05 at 6 weeks, 12 weeks, 26 weeks, 52 weeks, and 2 years (P values NR). The repeated measures AMD in scores over 6 weeks to 2 years was -0.03 (95% CI, -0.07 to 0.01).

B. Minimally-invasive surgery compared to microdiscectomy or discectomy

Three RCTs reported health-related QOL. Two RCTs were rated as some concerns for bias[28,29] and one was rated as high risk of bias.[30] ***Table 16*** summarizes the findings and strength of evidence related to QOL outcomes for this comparison. A detailed description of findings follows this table.

Table 16. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for quality of life in persons with symptomatic lumbar radiculopathy (EQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Quality of Life -SF-36 Physical Health and Mental Health Component Summary (follow-up: range 12 weeks to 3 years) (MID 2 points for PCS and 3 points for MCS)						
3 RCTs	Very serious[a]	Not serious	Not serious	Serious[b]	QOL improves in both surgical groups. With one exception, no between-group differences observed. Ryang et al.[30] (N=60) SF-36 MCS at 1.3y 51.9 (SD 7.8) vs. 44 (SD 13.2), P=0.03; no difference at 2.8y and no between-group difference in PCS. Teli et al.[29] (N=142) no between-group differences in PCS (P=0.68) and MCS (P=0.78) across time at 26w, 52w, and 2y. Thome et al.[28] (N=84) at 12w to 26w: PCS 43.6 (SD 9.7) vs. 41.5 (SD 10.7), P=0.41; MCS 53.6 (SD 9.8) vs. 50.6 (SD 12.0), P=0.26.	⨁◯◯◯ VERY LOW No difference

[a] 1 trial was rated as high risk of bias[30] and two were rated as having some concerns for bias.[28,29] Source of bias: lack of participant and outcome assessor blinding in all trials; inadequate randomization and allocation concealment in 1 trial.[30]

[b] Optimal information size criterion not met: the sample size of the largest RCT is unable to detect differences less than about 4 points, the others are unable to detect differences of less than about 6 points. Actual AMDs with confidence intervals NR.

Abbreviations: MID = minimally important difference**;** MCS = mental health component summary score; N = number; PCS = physical health component summary score; Short Form 36 = SF-36; QOL = quality of life; RCT = randomized controlled trial; SD = standard deviation; w = week(s); y = year(s).

Three RCTs reported SF-36 physical health (PCS) and mental health (MCS) component summary scores over 12 weeks to 2.8 years.[28-30] In all studies, quality of life as measured by both component scores improved over time in both intervention groups, and with one exception, no statistically significant between-group differences were observed.

- Ryang et al.[30] observed a significant difference in the SF-MCS at the 1.3 year follow-up; participants allocated to microdiscectomy had a higher score (mean 51.9 [SD 7.8]) compared with participants allocated to minimal access trocar microdiscectomy (mean 44.0

[SD 13.2], P=0.03), but it is not clear whether this comparison adjusted for small differences in baseline scores. No significant difference in this score was observed at the 2.8-year follow-up (mean 48.8 [SD 10.5] and 48.4 [SD 9.4), P=0.892).

- Teli et al.[29] reported no significant between groups differences in the PCS (P=0.68) or MCS (P=0.78) across time at 26 weeks, 52 weeks, and 2 years.

- Thome et al.[28] reported no significant difference between groups in the PCS (P=0.41) and the MCS (P=0.26) at 12 to 26 weeks follow-up.

C. Microdiscectomy compared with discectomy

One RCT rated as some concerns for bias compared microdiscectomy to standard discectomy and reported outcomes with the SF-36 physical health component summary (PCS) score and the mental health component summary (MCS) score at 26 weeks, 52 weeks, and 2 years.[29] ***Table 17*** summarizes the findings and strength of evidence related to quality of life for this comparison. For PCS, Teli et al. reported increases from baseline (21 [SD 4]) ranging from 19 to 23 points among participants allocated to microdiscectomy at the various follow-up time points compared with increases from baseline (22 [SD 4]) ranging from 18 to 22 points among participants allocated to discectomy. No significant between-group differences were observed (P=0.68). Similar findings were reported for the MCS (P=0.78 for between-group differences).

Table 17. Summary of findings and strength of evidence ratings comparing microdiscectomy to discectomy for quality of life in persons with symptomatic lumbar radiculopathy (EQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Quality of life- SF-36 physical and mental health component summary (medium- and long-term) (follow-up: range 26 weeks to 2 years) (MID 2 points for PCS and 3 points for MCS)						
1 RCT	Serious[a]	Not serious[b]	Not serious	Serious[c]	QOL improvements at 26w, 52w, and 2y in both surgical groups, but no significant between-group differences. Teli et al.[29] (N=142 analyzed) PCS increases from baseline (21 [SD 4]) ranging from 19 to 24 points among participants allocated to microdiscectomy compared with increases from baseline (22 [SD 4]) ranging from 16 to 22 points among participants allocated to discectomy. No significant between-group differences were observed (calculated AMD 26w 1, 52w 2, 2y 3, P=0.68 for between-group differences across time). Similar findings were reported for the MCS (calculated AMD 26w 2, 52w 0, 2y 2, P=0.78 for between-group differences across time).	⊕⊕◯◯ LOW No difference

[a] This RCT was rated as some concerns for bias. Sources of bias: lack of participant and outcome assessor blinding.

[b] Not applicable since body of evidence has only 1 study.

[c] Optimal information size criterion not met: the sample size in this study is only able to detect differences of about 4 points.

Abbreviations: AMD = absolute mean difference; MCS = mental health component summary score; MID = minimally important difference; QOL = quality of life; Short Form 36 = SF-36; N = number; PCS = physical health component summary score; RCT = randomized controlled trial; SD = standard deviation; w = week(s); y = year(s)

Neurological symptoms

A. Surgery compared with nonsurgical interventions

Two RCTs reported outcomes related to neurological symptoms, specifically sensory or motor deficits.[33,41] One was rated as high risk of bias[33] and was rated as some concerns for bias for 6-week outcomes and high risk of bias for outcomes at 12 weeks and later.[41] ***Table 18*** summarizes the findings and strength of evidence related to neurological symptoms for this comparison.

Gerszten et al. evaluated neurologic symptoms on each side and at each lumbosacral nerve root level (i.e., 8 comparisons for muscle strength and 8 comparisons for tactile sensitivity); no significant differences between groups (plasma disc decompression vs. epidural steroid injection) were observed in all but one of these 16 comparisons at 6 weeks follow-up (actual values NR, P value NR).[41] Osterman et al. reported a similar proportion of participants with muscle weakness among those allocated to microdiscectomy (N=28) compared with those allocated to physiotherapy (N=28), respectively, at 6 weeks (53.8% vs. 46.2%), 12 weeks (42.3% vs. 46.2%), and 52 weeks (28.6% vs. 30%). This study did not report between-group statistical significance tests.[33]

Table 18. Summary of findings and strength of evidence ratings comparing surgery to nonsurgical interventions for neurologic symptoms in persons with symptomatic lumbar radiculopathy (EQ1)

Certainty Assessment					Summary of Findings	CERTAINTY /Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Neurologic symptoms (short- medium- and long-term) (follow-up: range 6 weeks to 52 weeks)						
2	Serious[a]	Not serious	Not serious	Very serious[b]	Sensory and motor deficits improved over time in both treatment groups; no difference between groups. Gerszten et al.[41] (N=90) no difference in proportion of participants with motor or sensory deficits between plasma disc decompression and epidural steroid injection at 6w; Osterman et al.[33] (N=56) no difference in proportion of participants with muscle weakness at 6w (53.8% vs. 46.2%), 12w (42.3% vs. 46.2%), and 52w (28.6% vs. 30%).	⨁◯◯◯ VERY LOW No difference

[a] Risk of bias was high in 1 trial, and there were some concerns in the other trial.

[b] Both trials were underpowered to detect differences in proportions less than 35% (N=56) and 25% (N=90).

Abbreviations: N = number; NR = not reported; RCT = randomized controlled trial; w = week(s).

B. Minimally-invasive surgery compared to microdiscectomy or discectomy

Six RCTs comparing minimally-invasive surgery to microdiscectomy[28,30,31,34,36] and discectomy[43] reported outcomes related to neurological symptoms. Three RCTs were rated as some concerns for bias[28,36,43] and three were rated as high risk of bias.[30,31,34] ***Table 19*** summarizes the findings and strength of evidence related to neurological outcomes for this comparison. A detailed description of findings follows this table.

Table 19. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for neurologic symptoms in persons with symptomatic lumbar radiculopathy (EQ1)

Certainty Assessment					Impact	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Neurological symptoms (medium and long-term) (follow-up: range 12 weeks to 2 years)						
6 RCTs	Very serious[a]	Not serious	Not serious	Very serious[b]	Neurological symptoms improved in both surgical groups with no significant differences between surgical groups in 5 of the RCTs;[28,30,31,34,43] one RCT did not report findings by group.[36]	⊕◯◯◯ VERY LOW No difference

[a] Three RCTS were rated as some concerns for bias[28,36,43] and three were rated as high risk of bias.[30,31,34] Sources of bias: lack of participant and outcome assessor blinding (all trials), inadequate randomization and allocation concealment.[30,31,34]

[b] Optimal information size criterion not met: measures of variance not provided in many studies, unclear what effect size is meaningful and whether most studies had adequate sample sizes to detect a minimally important difference.

Abbreviations: RCT = randomized controlled trial.

Findings were not reported by group in one RCT,[36] the remaining five studies observed no between-group differences. Three RCTs reported no statistical difference in neurological symptoms between intervention groups[28,30,31] We calculated no differences in the other 2 RCTs.[34,43]

- Ruetten et al.[31] reported mean North American Spine Society Neurology Scores ranging from 1.9 to 2.1 over 12 weeks to 2 years follow-up in participants allocated to endoscopic discectomy and mean scores ranging from 1.7 to 2.3 among participants allocated to microdiscectomy; differences between groups were reported as not statistically significant (P value NR).

- Ryang et al.[30] reported no difference in the proportion of participants with sensory deficits (40% vs. 43%) in participants allocated to minimal access trocar discectomy compared with participants allocated to microdiscectomy, respectively, over an average of 1.3 years follow-up (P=0.31). Similar findings were observed for the proportion with motor deficits (27% vs. 23%, P=0.86).

- Thome et al.[28] reported improvements in both sensory and motor deficits from baseline to 2 years in participants allocated to both sequestrectomy and microdiscectomy.[28] However, the study reported no between-group differences in a composite measure of neurologic symptoms at 2 years (actual value NR, P=0.278) or in the proportion of participants with sensory deficits (actual value NR, P=0.52) or motor deficits (actual value NR, P=0.74) at 12 to 26 weeks follow-up.

- Hermantin et al.[43] reported the proportion of participants with postoperative sensory deficits (53.3% vs. 60.0%, calculated P=0.79) and motor weakness (16.7% vs. 33%, calculated P=0.23) in participants allocated to video-assisted arthroscopic microdiscectomy and those allocated to discectomy respectively at an unspecified follow-up time.

- Mayer et al.[34] reported the proportion of participants with sensory deficit (5% vs. 25%, calculated P=0.18) and motor deficit (0% vs. 0%) at 2 years follow-up among those

allocated to percutaneous endoscopic discectomy and those allocated to microdiscectomy, respectively.

- Franke et al.[36] reported that overall 83% of motor deficits and 68% of sensory deficits were resolved completely at 52 weeks; outcomes were not reported by intervention group.

C. Microdiscectomy compared with discectomy

No studies reported outcomes related to neurological symptoms for this comparison.

Return to work

A. Surgery compared with nonsurgical interventions

Five RCTs reported various outcomes related to "return to work."[22,26,33,37,41] Some measures captured actual return to work, whereas others reflected somewhat indirect measures, such as self-reported ability to work, receipt of disability benefits, or pain affecting occupational status. We rated all RCTs as high risk of bias for this outcome. ***Table 20*** summarizes the findings and strength of evidence related to return to work outcomes for this comparison. A detailed description of findings follows this table.

Table 20. Summary of findings and strength of evidence ratings comparing surgery to nonsurgical interventions for outcomes related to return to work in persons with symptomatic lumbar radiculopathy (EQ1)

Certainty Assessment					Summary of Findings	CERTAINTY /Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Return to work						
5 RCTs	Very serious[a]	Not serious	Serious[b]	Serious[c]	Except for one study, no between-group differences in measures relating to return to work. Erginousakis et al.[37] (N=62) pain affecting occupational status (12w 12.9% vs. 9.7%, calculated P=1.0; 52w 12.9% vs. 71%, calculated P< 0.001). Weber et al.[26] (N=126) receiving disablement benefits (4y 5% vs. 12.1%, calculated P=0.21; 10y 11.7% vs. 12.1%, calculated P=1.0). Gerszten et al.[41] (N=90) return to work (69% vs. 70%,). Osterman et al.[33] (N=56) VAS 100 work ability RM 0 to 2y AMD 5 (95% CI, -7 to 18). Weinstein et al. [SPORT][22] (N=501) 2y AMD proportion working full time -2.2% (95% CI, -10.6% to 6.2%).	⊕◯◯◯ VERY LOW No difference[d]

[a] Risk of bias was high in all trials. Sources of bias: lack of participant and outcome assessor blinding and inadequate in all trials and inadequate randomization and allocation concealment[37] or extensive and differential crossovers.[22,26,33,41]

[b] Measures used in some studies were indirect: Erginousakis et al.[37] (i.e. pain affecting occupational status); Osterman et al.[33] work ability based on self-reported VAS 100 mm; Weber et al.[26] reported on proportion receiving permanent disability benefits.

[c] Optimal information size criterion not met: variation in measures used and lack of clarity with respect to minimally important differences, all but one study[22] have sample sizes that are too small to detect small to modest differences in dichotomous outcomes.

[d] Though one study did find a difference, it did not directly measure return to work; it measured 'pain affecting occupational status", thus we did not consider the findings mixed.

Abbreviations: AMD = absolute mean difference; N = number; RCT = randomized controlled trial; RM = repeated measure; SPORT = Spine Patient Outcomes Research Trial; VAS = visual analog scale; w = week(s); y = year(s).

Return to work outcomes were measured at various follow-up times across this body of evidence. With one exception[37]; no between-group differences in return to work outcomes were observed.

- Erginousakis et al.[37] reported minimal difference in the proportion of participants reporting pain that affected their occupational status at 12 weeks follow-up (12.9% percutaneous disc decompression, 9.7% conservative management, calculated P=1.0). This proportion remained the same among participants that received percutaneous disc decompression but increased to 71% at 52 weeks and 2 years among participants that received conservative therapy (calculated P<0.001 at both time points).

- Weber et al.[26] reported the proportion of participants with a permanent incapacitation and receiving disablement benefits. At 4 years follow-up, this proportion was 5% in participants allocated to discectomy compared with 12.1% in participants allocated to conservative management (calculated P=0.21). At 10 years follow-up the proportions were 11.7% and 12.1%, respectively (calculated P=1.0).

- Osterman et al.[33] reported an increase in work ability as measured by a 0 to 100 VAS score over five follow-up time points from 6 weeks to 2 years in participants allocated to microdiscectomy and in participants allocated to physiotherapy. Although participants allocated to microdiscectomy had higher VAS work ability scores at all time points, particularly at 12 weeks and 26 weeks, the between-group AMD in repeated measures analysis over 6 weeks to 2 years was not significant (AMD 5 [95% CI, -7 to 18]).

- Weinstein et al. [SPORT][22] reported in increase in the proportion of participants working full or part time in participants allocated to discectomy and in participants allocated to conservative management from baseline to all follow-up times. However, between-group differences in the proportion working full-time were not significant at any time point (12 weeks AMD -5.6% [95% CI, -14.5% to 3.4%], 52 weeks AMD -0.6% [95% CI, -8.6% to 7.3%], 2 years AMD -2.2% [95% CI, -10.6% to 6.2%], 4 years AMD −3.8 [95% CI, −13.3 to 5.8]).

- Gerszten et al.[41] reported that the proportion of participants working full or part time at 26 weeks was similar (69% to 70%) among participants allocated to plasma disc decompression compared with participants allocated to epidural steroid injection.

B. Minimally-invasive surgery compared to microdiscectomy or discectomy

Six RCTs reported various outcomes related to "return to work", though in some studies this outcome was not reported by group.[28,31,34,36,38,43] Four RCTs were rated as having some concerns for bias[28,36,38,43] and two were rated as having high risk of bias.[31,34] ***Table 21*** summarizes the findings and strength of evidence related to work outcomes for this comparison. A detailed description of findings follows this table.

Table 21. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for return to work outcomes in persons with symptomatic lumbar radiculopathy (EQ1)

Certainty Assessment					Summary of Findings	CERTAINTY /Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Return to Work						
6 RCTs	Very serious[a]	Not serious[b]	Not serious	Very serious[c]	Mean duration of postoperative work disability is lower by a range of 3.4w to 15.2w among participants allocated to minimally-invasive surgery compared to standard surgery. Hermantin et al.[43] (N=60) 3.9w (SD NR) vs. 7w (SD NR), P value NR; Mayer et al.[34] (N=40) 7.7w (range 1 to 26w) vs. 22.9w (range 4 to 52w), P value NR; Ruetten et al.[31] (N=200) 3.57w (SD NR) vs. 7w (SD NR), P< 0.01. Thome et al.[28](N=84) reported no between-group differences using a multi-level categorical outcome of "work impairment". Two RCTs did not report findings by group.[36,38]	⊕◯◯◯ VERY LOW Favors minimally-invasive surgery[b]

[a] Two trials were rated as high risk of bias[31,34] and four were rated as some concerns.[28,36,38,43] Sources of bias: lack of participant and outcome assessor blinding in all trials and inadequate randomization and allocation concealment in the two trials rated as high risk of bias.

[b] Three of the 4 RCTs that reported between-group differences evaluated return to work using a continuous measure of postoperative work disability in week and reported a magnitude of effect that seems clinically important. The RCT that reported no between-group differences in return to work used a 4-level categorical measure of work impairment. The difference in type of measure may explain why no between-group differences were observed in this study and we did not consider this an inconsistent finding. Thus, we assessed the overall direction of effect for this body of evidence as favoring minimally-invasive surgery.

[c] Optimal information size criterion not met: two studies did not report findings by group to allow for estimate of between-group difference; three studies did not report measures of variance, unclear what a minimally important difference is for this outcome and what sample size would be required to detect a small to modest difference.

Abbreviations: N = number; NR = not reported; RCT = randomized controlled trial; SD = standard deviation; w = week(s).

Of the 4 RCTs that reported between-group differences, three RCTs[31,34,43] suggest that participants allocated to minimally-invasive surgery return to work sooner than participants allocated to standard surgery as measured by weeks of postoperative disability. The range of this difference is 3.4 weeks to 15.2 weeks. The remaining RCT[28] reported no significant between-group differences; however; this study used a multi-level categorical measure of work impairment, which may be measuring a related, but different construct compared to the other three RCTs.

- Ruetten et al.[31] reported the mean duration of postoperative work disability was 3.57 (SD NR) weeks in participants allocated to endoscopic discectomy compared with 7 (SD NR) weeks in participants allocated to microdiscectomy (P< 0.01).

- Hermantin et al.[43] reported a mean duration of postoperative disability in time lost from work or until able to resume normal activity of 3.9 weeks among participants allocated to video-assisted arthroscopic microdiscectomy compared with 7 weeks among participants allocated to discectomy (measures of variance and P values NR).

- Mayer et al.[34] reported a mean duration of postoperative disability of 7.7 weeks (range 1 to 26) among participants allocated to percutaneous endoscopic discectomy compared with

22.9 weeks (range 4 to 52) among participants allocated to microdiscectomy (measures of variance and P values NR). The overall N (%) returning to work was 19 (95%) among participants allocated to minimally-invasive surgery and 13 (65%) among participants allocated to microdiscectomy (calculated P = 0.004).

- Thome et al. reported specific categories of impairment of work at 12 to 26 weeks and at 2 years.[28] Thirty-one percent of participants allocated to sequestrectomy reported that their work impairment was "much better" at 12 to 26 weeks compared with 33% of participants allocated to microdiscectomy. At 2 years, the proportions were 37% and 31%, respectively. The proportion of participants endorsing various categories of work impairment were not significantly different between groups (P=0.415 at 12 to 26 weeks, P=0.112 at 2 years).

- Chatterjee et al.[38] reported that 92.5% of participants allocated to microdiscectomy returned to work or their previous level of activity by 12 weeks follow-up. The number of participants returning to work in the group allocated to automated percutaneous lumbar discectomy was not reported.

- Franke et al.[36] reported the mean duration of postoperative inability to work overall was 7 weeks; this duration was not reported by intervention groups (microscopically assisted percutaneous nucleotomy vs. microdiscectomy).

C. Microdiscectomy compared with discectomy

One RCT rated as some concerns for bias compared microdiscectomy to discectomy and reported on outcomes related to "return to work".[27] Tullberg et al. reported a mean duration of postoperative, full-time sick leave of 10.4 (SD NR) weeks in participants allocated to microdiscectomy compared with 10.1 (SD NR) weeks in participants allocated to discectomy (P value NR). The proportion out of work at an unspecified follow-up time point was 16.7% among those allocated to microdiscectomy and 6.7% among those allocated to discectomy (calculated P =0.42). ***Table 22*** summarizes the findings and strength of evidence related to return to work for this comparison.

Table 22 Summary of findings and strength of evidence ratings comparing microdiscectomy to discectomy for return to work outcomes in persons with symptomatic lumbar radiculopathy (EQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Return to work (follow-up varies)						
1 RCT	Serious[a]	Not serious[b]	Not serious	Very serious[c]	No between-group differences in return to work outcomes. Tullberg et al.[27] (N=60) mean duration of postoperative, full-time sick leave 10.4w (SD NR) vs. 10.1w (SD NR) (P value NR).The proportion out of work at an unspecified follow-up time point 16.7% vs. 6.7% (calculated P=0.42).	⊕◯◯◯ VERY LOW No difference

[a] This RCT was rated as some concerns for bias. Source of bias: lack of participant and outcome assessor blinding and inadequate information to evaluate randomization and allocation concealment process.

[b] Not applicable since body of evidence has only 1 study.

[c] Optimal information size criterion not met: no measures of variance provided for continuous measure and a sample size of 322 would have been required to detect a significant between-group difference for the difference in proportion observed by the study.

Abbreviations: N = number; NR = not reported; RCT = randomized controlled trial; SD = standard deviation; w = week(s).

Other Efficacy Outcomes

A. Surgery compared with nonsurgical interventions

Four RCTs reported other efficacy outcomes related to perceived recovery, overall time to recovery, overall result, patient satisfaction with symptoms, and self-reported progress.[22,26,32,33,100] These outcomes were consistent with previously reported efficacy outcomes that suggest more favorable outcomes for participants who are allocated to surgery in the short and medium-term. Results from three of these studies also suggest some favorable outcomes in the long-term. We did not use these outcomes in our strength of evidence ratings because of heterogeneity in outcome definition. Specific study findings:

- Osterman et al.[33] found a higher proportion of participants reporting full recovery at 6 weeks (19.2% vs. 0%, P< 0.05) among participants allocated to microdiscectomy compared with participants allocated to physiotherapy. The difference in proportion was similar at 12 weeks (19.2% vs. 15.4%, calculated P=1.0) and 52 weeks (33.3% vs. 25%, calculated P=0.73).

- Peul et al.[32] reported a significant difference in median time to recovery (4.0 weeks (95% CI, 3.7 to 4.4) vs. 12.1 weeks (95% CI 9.5 to 14.9; AMD NR, P< 0.001) among participants allocated to microdiscectomy compared with participants allocated to conservative management. The relative difference in time to "complete" or "nearly complete" recovery at 52 weeks favored microdiscectomy (hazard ratio 1.97 [95% CI 1.72 to 2.22]. However, the proportion of participants reporting complete or nearly complete recovery (on a 7-point Likert scale of self-perceived recovery) by 2 years was not different between groups (81.3% vs. 83.6%, AMD -2.4% [95% -12.0% to 7.2%]).

- Weber et al.[26] reported the proportion of participants achieving good, fair, poor, or bad results overall. This assessment was based on an outcome assessor's evaluation of the patient's neurological deficits, working capacity, pain, and mobility of the lumbar spine. At 52 weeks, 65.0% of participants allocated to discectomy had achieved "good" results compared with 36.4% of participants allocated to conservative management. The difference in proportion between groups across all four categories of results (good, fair, poor, bad) was significant (P=0.0015) at 52 weeks, but these differences did not persist at 4 years or 10 years (P reported as not significant for both).

- Weinstein et al. [SPORT] [22] reported on proportion of patients endorsing various categories of satisfaction with symptoms. At 12 weeks, a higher proportion (54.3%) of participants allocated to discectomy/microdiscectomy reported being very or somewhat satisfied with symptoms compared with 43.0% of participants allocated to conservative management (AMD 11.3% [95% CI 1.6% to 20.9%]). This proportion remained numerically higher at all subsequent follow-up time points (52 weeks, 2 years, 4 years, 8 years) among participants allocated to discectomy/microdiscectomy but the difference between groups

was not statistically significant at any single follow-up time point. Repeated measures of this proportion over 12 weeks to 2 years and 12 weeks to 4 years also observed no significant difference between groups; however, repeated measures of this proportion over 12 weeks to 8 years was statistically significant (P=0.013) favoring surgery.[100]

B. Minimally-invasive surgery compared to microdiscectomy or discectomy

Ten RCTs reported other efficacy outcomes, related to perceived recovery, overall time to recovery, overall result, and patient satisfaction with symptoms. With few exceptions, most observed no significant differences between groups. We did not use these outcomes in our strength of evidence ratings because of heterogeneity in outcome definition. Specific study findings:

- Arts et al.,[40] which compared tubular discectomy with microdiscectomy reported the frequency and proportion of patients with "complete" or "nearly complete" recovery based on a 7-point Likert scale at 4 weeks, 8 weeks, 26 weeks, 52 weeks, 2 years, 3, years, 4 years, and 5 years. The differences between groups were only significant at 8 weeks (63% vs. 75%, odds ratio (OR) 0.56 [95% CI, 0.35 to 0.92]) and 52 weeks (69% vs. 79%, OR 0.59 [95% CI, 0.35 to 0.99]). No relative difference between groups in time to recovery by 52 weeks (hazard ratio (HR) 0.92 [95% CI, 0.73 to 1.17]) or by 2 years (HR 0.93 [95% CI, 0.93 0.74 to 1.17]) was observed.

- Brouwer et al.,[39] which compared percutaneous laser disc decompression with microdiscectomy, used the same 7-point Likert scale as Arts et al. and reported no significant difference in the frequency and proportion achieving complete or nearly complete recovery (69% vs. 75% at 52 weeks [OR 0.81, 95% CI, 0.4 to 1.9]; 70.8% vs. 60.8% at 2 years [OR 1.6, 95% CI, 0.7 to 3.6]). Unlike Arts et al. which did not find a relative difference in time to recovery between groups, Brouwer et al. observed a significantly slower recovery among participants allocated to percutaneous laser disc decompression (HR 0.64 [95% CI, 0.42 to 0.97] at 52 weeks and similar findings at 2 years.

- Three RCTs reported outcomes using the MacNab criteria, a 4-point Likert scale that rates the overall outcome as excellent, good, fair, or poor. Chatterjee et al.[38] reported a significant difference in participants achieving an excellent or good outcome among participants allocated to automated percutaneous lumbar discectomy (9 [29%])) compared with participants allocated to microdiscectomy (32 [80%], p< 0.001). Haines et al.[42] reported 11 (64.7%) of participants allocated to automated or endoscopic percutaneous discectomy achieved an excellent or good outcome compared with 6 (60%) allocated to discectomy (P=0.81). Lastly, Huang et al.[24] reported 9 (90%) of participants allocated to microendoscopic discectomy reported an excellent or good outcome compared with 11 (91.6%) of participants allocated to discectomy (calculated P=1.0).

- Other studies used various measures of satisfaction with outcome, satisfaction with surgery, or global outcome rating. Mayer et al.[34] reported the frequency and proportion with self-reported success of surgery at 2 years; 9 (47%) of participants allocated to

percutaneous endoscopic discectomy compared with 8 (40%) of participants allocated to microdiscectomy. Hermantin et al.[43] reported no difference in frequency and proportion with satisfactory outcome between participants allocated to video-assisted arthroscopic microdiscectomy (29 [97%]) compared with 28 [93%]) allocated to discectomy (calculated P=1.0). Ryang et al.[30] used a VAS 10 cm to report overall improvement from baseline to 2.8 years follow-up; although scores reflected overall improvement among participants allocated to trocar discectomy and participants allocated to microdiscectomy, the scores at follow-up were similar (4.92 vs. 4.64, P value NR). Sasaoka et al.[25] reported a mean percentage improvement in the Japanese Orthopaedic Association Score at 52 weeks of 84.7% among participants allocated to microendoscopic discectomy compared with 88.6% among participants allocated to microdiscectomy (P value NR). Thome et al.[28] reported no differences in patient satisfaction with surgery scores at 12 to 26 weeks or 2 years.

C. Microdiscectomy compared to discectomy

One RCT rated as some concerns for bias comparing microdiscectomy to discectomy reported the frequency and proportion of participants with a specified opinion on recovery at 52 weeks (total recovery, almost recovered, good, unchanged, or worse).[27] Among participants allocated to microdiscectomy, 11 (37.9%) reported total recovery, and 8 (27.6%) reported almost recovered. Among participants allocated to discectomy these outcomes were 6 (20.7%) and 14 (28.3%), respectively (calculated P=0.25 and 0.18, respectively).

3.2.2 Efficacy Question 2

In adults with symptomatic lumbar radiculopathy, does effectiveness or comparative effectiveness of surgical interventions vary for patients who are not employed because of disability or patients who are undergoing recurrent surgery for relapse?

We did not identify any studies that reported outcomes specifically for patients not employed because of disability. We identified two studies focused on the efficacy[46] or comparative effectiveness[47] of revision surgery for relapse. Both were rated as high risk of bias. Study and population characteristics of these trials are summarized in ***Table 23***.

Table 23. Study and population characteristics of the two randomized controlled trials comparing revision surgical interventions to spinal cord stimulation or an alternative revision surgery for the management of lumbar radiculopathy relapses (EQ2)

Author (Year); Country; Risk of Bias	Population Characteristics Age, mean (SD); Women N (%); Duration of symptoms, mean (SD)	*Surgical Intervention* (SG1); N randomized; N analyzed (% of randomized); N crossovers (% of randomized)	*Comparator(s)* (SG2 or NS1); N randomized; N analyzed (% of randomized); N crossovers (% of randomized)	*Primary Outcome*[a] (effect size detectable with 80% power, α = 5%); • Other outcomes
North (2005)[46] United States; High	Age: 52.0 (13.5) Women: 26 (52%) Mean (SD) number of prior operations: SG1: 2.5 (1.1) NS1: 2.5 (1.1)	*Repeat lumbosacral decompression* N randomized: 26; N analyzed: 26 (100%); N crossovers: 14 (54%)	*Spinal cord stimulation* N randomized: 24; N analyzed: 19 (79.2%); N crossovers: 5 (20.8%)	*Successful treatment (NR);* • Impairment from pain in performing everyday activities • Return to work • N (%) with successful treatment
Ruetten (2009)[47] Germany; High	Age: 39 (range 23 to 59) Female: 44 (44%) Duration of symptoms, mean (range) in weeks: 9.85 (0.14 to 56)	*Revision endoscopic discectomy* N randomized: 50; N analyzed: 45 (90%); N crossovers: 0 (0%)	*Revision microdiscectomy* N randomized: 50; N analyzed: 42 (84%); N crossovers: 0 (0%)	*NR (NR);* • VAS 100 leg pain • VAS 100 back pain • NASS pain score • NASS neurology score • Oswestry Disability Index • Postoperative work disability • N (%) with no leg pain • N (%) satisfied with surgery and would undergo the operation again

[a] As specified and reported by study authors.

Abbreviations: N = number of participants; NASS = North American Spine Society; NR = not reported; NS = nonsurgical group; SD = standard deviation; SG = surgical group; VAS = visual analog scale.

The strength of evidence ratings and summary of findings from these studies are provided in ***Table 24*** and ***Table 25.*** A detailed description of study characteristics and findings follows these tables. ***Appendix D, Tables D-1*** and ***D-2*** provides evidence tables with individual study and population characteristics. ***Appendix D, Tables D-3*** and ***D-4*** provide evidence tables with detailed individual study outcomes related to efficacy.

Table 24. Summary of findings for pain, functioning, neurological symptoms and quality of life for RCTs for repeat lumbosacral decompression surgery compared with spinal cord stimulation for treatment of lumbar radiculopathy relapses (EQ2)

Certainty Assessment					Summary of Findings	Certainty/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Pain						
0 RCTs	--					◯◯◯◯ INSUFFICIENT
Functioning/Disability-impairment in everyday activities due to pain (long-term) (follow-up: range 1.8 years to 5.7 years)						
1 RCT	Very serious[a]	Not serious[b]	Serious[c]	Very serious[d]	North et al.[46] (N=50) reported higher levels of impairment from pain in performing everyday activities among participants allocated to repeat surgery compared to participants allocated to spinal cord stimulation, values were only depicted on a figure and differences between groups were reported as NS.	◯◯◯◯ INSUFFICIENT
Quality of life						
0 RCTs	--					◯◯◯◯ INSUFFICIENT
Neurologic symptoms						
0 RCTs	--					◯◯◯◯ INSUFFICIENT
Return to work (long-term) (follow-up: range 1.8 years to 5.7 years)						
1 RCT	Very serious[a]	Not serious[b]	Serious[c]	Very serious[d]	North et al.[46] (N=50) reported no significant differences in return to work but actual values were NR.	◯◯◯◯ INSUFFICIENT

[a] The trial was rated as high risk of bias. Sources of bias: lack of participant and outcome assessor blinding, extensive deviations from intended interventions and differential attrition.

[b] Not applicable as this body of evidence has only one study.

[c] This measure was not well defined and seems to assess pain as well as functional impairment.

[d] Optimal information size criterion net met: actual values and measures of variance were not reported and study sample size is unlikely to be able to detect anything but a very large difference.

Abbreviations: N= number; NR = not reported; NS = not significant; RCT = randomized controlled trial.

Table 25. Summary of findings for pain, functioning, quality of life, neurologic symptoms and return to work comparing revision endoscopic discectomy to revision microdiscectomy for treatment of lumbar radiculopathy relapses (EQ2)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Pain (leg)-VAS 100 mm (medium- and long-term) (follow-up: range 12 weeks to 2 years) (MID 7 to 11 points)						
1 RCT	Very serious[a]	Not serious[b]	Not serious	Very serious[c]	Ruetten et al.[47] (N=100) reported improvements over time from baseline to 12w, 26w, 52w, and 2y among participants in both surgical groups. Between-group differences were reported as NS.	◯◯◯◯ INSUFFICIENT
Pain (back)-VAS 100 mm (medium- and long-term) (follow-up: range 12 weeks to 2 years) (MID 7 to 11 points)						
1 RCT	Very serious[a]	Not serious[b]	Not serious	Very serious[c]	Ruetten et al.[47] (N=100) reported improvements over time from baseline to 12w, 26w, 52w, and 2y among participants in both surgical groups. Between-group differences were reported as NS.	◯◯◯◯ INSUFFICIENT
Pain-North American Spine Society Pain Score (medium- and long-term) (follow-up: range 12 weeks to 2 years)						
1 RCT	Very serious[a]	Not serious[b]	Not serious	Serious[d]	Ruetten et al.[47] (N=100) reported improvements over time from baseline to 12w, 26w, 52w, and 2y among participants in both surgical groups. Between-group differences were reported as NS.	◯◯◯◯ INSUFFICIENT
Functioning/Disability-impairment in activities due to pain (medium- and long-term) (follow-up: range 12 weeks to 2 years)						
1 RCT	Very serious[a]	Not serious[b]	Not serious	Serious[d]	Ruetten et al.[47] (N=100) reported improvements as measured by the Oswestry Disability Index from baseline to 12w, 26w, 52w, and 2y among participants allocated to both surgical groups. Between-group differences were reported as NS.	◯◯◯◯ INSUFFICIENT
Quality of life						
0 RCTs	--					◯◯◯◯ INSUFFICIENT
Neurological symptoms (medium- and long-term) (follow-up: range 12 weeks to 2 years)						
1 RCT	Very serious[a]	Not serious[b]	Not serious	Serious[d]	Ruetten et al.[47] (N=100) reported improvements in both surgical groups over time in the mean North American Spine Society Neurology scores from baseline to 12w, 26w, 52w, and 2y. Between-group differences were reported as NS.	◯◯◯◯ INSUFFICIENT
Return to work-mean duration of postoperative disability						
1 RCT	Very serious[a]	Not serious[b]	Not serious	Serious[d]	Ruetten et al.[47] (N=100) reported a significant different in mean duration of postoperative disability ($P < 0.01$): 4w (SD NR) among participants allocated to revision endoscopic discectomy and 7.4w (SD NR) among participants allocated to revision microdiscectomy.	◯◯◯◯ INSUFFICIENT

[a] This trial was rated as high risk of bias. Sources of bias: lack of participant and outcome assessor blinding, inadequate randomization and allocation concealment process.

[b] Not applicable as this body of evidence has only one study.

[c] Optimal information size criterion not met: measures of variance not reported and sample size only able to detect differences of about 15 points.

[d] Optimal information size criterion not met: measures of variance were not reported, lack of clarity about minimally important differences for these measures.

Abbreviations: MID = minimally important between-group difference; N= number; NR = not reported; NS = not significant; RCT = randomized controlled trial; SD = standard deviation; VAS = visual analog scale; w = week(s); y = year(s).

3.2.2.1 Study Characteristics

Both included studies were parallel-group RCTs. North et al.[46] was conducted in the United States and compared repeat lumbosacral decompression (N randomized=26) with spinal cord stimulation (N randomized=24) in adults with persistent radicular pain despite one or more prior lumbosacral spine surgeries. The mean number of prior surgeries was 2.5 in both surgical groups, and overall 30% of participants reported receiving workers compensation benefits. Ruetten et al.[47] was conducted in Germany and compared revision endoscopic discectomy (N randomized=50) to revision microdiscectomy (N randomized=50) in adults who had a previous conventional discectomy with acute occurrence of radicular leg symptoms after a pain-free interval in combination with a recurrent disc herniation on MRI. The mean duration of symptoms overall in this study was 9.85 weeks (range 0.14 to 56). The proportion that were disabled or receiving disability benefits was not reported. North et al. was rated high risk for bias because of substantial crossover between groups, differential attrition and Ruetten et al. was rated high risk for bias because randomization and allocation concealment were inadequate. In addition, both studies did not blind participants or outcome assessors.

3.2.2.2 Findings

Pain

North et al.[46] did not report any outcomes related to pain. Ruetten et al.[47] reported improvement in VAS 100 mm leg pain score from baseline to 12 weeks, 26 weeks, 52 weeks, and 2 years among participants allocated to revision endoscopic discectomy and among participants allocated to revision microdiscectomy. Between-group differences were reported as not significant at any follow-up time point (AMDs were NR). A similar pattern was observed for VAS 100 mm back pain scores and North American Spine Society pain scores. At 2 years follow-up, the number and proportion of participants reporting no leg pain was 37 (82%) among those allocated to revision endoscopic discectomy and 32 (76%) among those allocated to revision microdiscectomy (P value for comparison NR). Because no studies reported pain outcomes for one comparison and only a single study body of evidence for the other comparison, we assessed the evidence as insufficient to draw a conclusion for pain outcomes.

Functioning/disability

North et al.[46] reported qualitatively that higher levels of impairment from pain in performing everyday activities among participants allocated to repeat lumbosacral decompression compared with participants allocated to spinal cord stimulation at a mean of 2.9 years follow-up; however, actual numeric values of impairment were NR and differences were reported as nonsignificant. Ruetten et al.[47] reported outcomes using the Oswestry Disability Index. Improvements in disability from baseline to 12 weeks, 26 weeks, 52 weeks, and 2 years were observed among participants allocated to revision endoscopic discectomy and among participants allocated to revision microdiscectomy. The between-group differences were reported as not significant (AMDs and P values for between-group differences NR). Because of single study bodies of evidence for each comparison, we assessed the evidence as insufficient to draw a conclusion for function and disability outcomes.

Quality of life

Neither study reported outcomes related to overall quality of life; thus, we assessed the evidence as insufficient for drawing conclusions about quality of life outcomes.

Neurological symptoms

North et al.[46] did not report any outcomes related to neurologic symptoms. Ruetten et al.[47] reported mean North American Spine Society Neurology scores at 12 weeks, 26 weeks, 52 weeks, and 2 years.[47] Participants allocated to revision endoscopic discectomy had a mean baseline score of 3 and participants allocated to revision microdiscectomy had a mean baseline of 5. Scores in both groups improved over time; at 2 years the scores were 2.1 and 2.3, respectively. The differences between groups were reported as not significant (AMDs and P values for between-group differences NR). Because no studies reported neurological symptom outcomes for one comparison and only a study body of evidence reported these outcomes for the other comparison, we assessed the evidence as insufficient to draw a conclusion for neurological symptom. outcomes.

Return to work

North et al.[46] reported no significant differences in return to work at a mean follow-up of 2.9 years, but actual values were not reported. Ruetten et al.[47] reported a significant difference between groups in the mean duration of postoperative disability. Among participants allocated to revision endoscopic discectomy the mean was 4 weeks (SD NR) and among participants allocated to revision microdiscectomy the mean was 7.4 weeks (SD NR) ($P < 0.01$). Because of single study bodies of evidence for each comparison, we assessed the evidence as insufficient to draw a conclusion for return to work outcomes.

Other efficacy outcomes

North et al.[46] reported on the frequency and proportion of successful treatment over a mean follow-up time of 2.9 years (range 1.8 years to 5.7 years). Success was defined as at least 50% pain relief and patient satisfaction with treatment. A significant difference in treatment success was observed ($P < 0.01$). Among those allocated to repeat lumbosacral decompression, successful treatment was observed in 3 (12%). Among those allocated to spinal cord stimulation, successful treatment was observed in 9 (47%). Ruetten et al.[47] reported on the frequency and proportion of patient satisfaction with surgery and whether participants would undergo the operation again. Among those allocated to revision endoscopic discectomy, 43 (95%) were satisfied; among those allocated to revision microdiscectomy 36 (86%) were satisfied (P value for comparison NR).

3.3 Safety

Safety Question 1

In adults with symptomatic lumbar radiculopathy, what are the adverse events associated with surgical interventions?

All 24 RCTs included for EQ1 and the two RCTs included for EQ2 also provided evidence for safety outcomes. A summary of included studies is provided in ***Tables 4, 5, and 6***. ***Appendix D, Tables D-1*** and ***D-2*** provide detailed individual study and population characteristics. ***Appendix D, Table D-5*** provides detailed individual study outcomes related to safety.

3.3.1 Study Characteristics

The study characteristics for the 24 RCTs included for safety outcomes were previously described in EQ1 and EQ2.

3.3.2 Findings

3.3.2.1 Mortality

A. Surgery compared with nonsurgical interventions

Six RCTS reported on mortality.[22,23,26,32,33,41] All of these studies were rated as low risk of bias for this specific outcome. ***Table 26*** summarizes the findings and strength of evidence related to mortality. A description of findings follows this table. Surgical mortality is not relevant as a comparative outcome given the nonsurgical comparison group. Thus, the strength of evidence for surgical mortality reflects our certainty about the absolute incidence of surgical mortality in the surgical intervention group.

Table 26. Summary of findings and strength of evidence ratings comparing surgery to nonsurgical interventions for mortality in persons with symptomatic lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Surgical mortality						
5 RCTs	Not serious[a]	Not serious	Not serious	Very serious[b]	No studies reported any surgery-related deaths.[22,23,32,33,41]	⊕⊕◯◯ LOW[c] NA[c]
All-cause mortality (medium- and long-term)						
3 RCTs	Not serious[a]	Not serious	Not serious	Very serious[d]	All-cause mortality similar between groups. Gerszten et al.[41] 1 death in each treatment group at 26w, unrelated to surgery. Weinstein et al. [SPORT][22](N=501) 3 deaths (1.91%) vs. 4 deaths (2.63%) by 8y. Weber et al.[26](N=126) 3 deaths (5.0%) vs 0 (0%) by 10y.	⊕⊕◯◯ LOW No difference

[a] Though all of these trials were rated as either having some concerns for bias or high risk of bias for efficacy outcomes and other safety outcomes, the risk of bias for mortality outcomes are low since non-comparative surgical mortality and all-cause mortality outcomes are unlikely to be influenced by lack of participant or outcome assessor blinding or crossovers.

[b] Optimal information size criterion not met: no events occurred.

[c] Because the comparator intervention is nonsurgical; this strength of evidence rating reflects the absolute incidence of surgical mortality, not the relative incidence with respect to a comparator.

[d] Optimal information size criterion not met: very rare events occurred.

Abbreviations: N = number; RCT = randomized controlled trial; SPORT = Spine Patient Outcomes Research Trial; w = week(s); y = year(s).

Of the five RCTs that reported surgical mortality, no studies reported any deaths relating to percutaneous disc decompression,[41] microdiscectomy,[23,33] discectomy,[32] or discectomy/microdiscectomy procedures.[22] We concluded with low certainty that deaths related to surgery are rare.

Three RCTs reported all-cause mortality.[22,26,41] Weinstein et al. [SPORT] reported three deaths (1.91%) among participants allocated to discectomy/microdiscectomy and four deaths (2.63%) among participants allocated to conservative management at 8 years.[22] Weber et al. reported three deaths (5.0%) among participants allocated to discectomy and no deaths among participants allocated to conservative management at 10 years.[26] One death was due to cancer and two due to heart disease. Gerszten et al. reported one death in each intervention group at 26 weeks follow-up because of myocardial infarction and acute pyelonephritis.[41] We concluded with low certainty that no difference in all-cause mortality between surgical interventions and nonsurgical interventions exists.

B. Minimally-invasive surgery compared with microdiscectomy or discectomy

Five RCTs reported mortality outcomes.[24,29,31,34,43] ***Table 27*** summarizes the findings and strength of evidence related to mortality for this comparison. A description of findings follows this table.

Table 27. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for mortality in persons with symptomatic lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Surgical mortality						
5 RCTs	Not serious[a]	Not serious	Not serious	Very serious[b]	No studies reported any surgery-related deaths.[24,29,31,34,43]	⊕⊕◯◯ LOW No difference
All-cause mortality						
2 RCTs	Not serious[a]	Not serious	Not serious	Very serious[c]	Ruetten et al.[31] (N=200) reported one death (0.5% of total) unrelated to treatment but did not specify which surgical group the death occurred. Arts et al.[40,48] (N=328) reported 2 deaths by 5y among participants allocated to tubular discectomy and 3 deaths among participants allocated to microdiscectomy.	⊕⊕◯◯ LOW No difference

[a] Though these trials were rated as some concerns for bias or high risk for bias for efficacy outcomes and other safety outcomes, the risk of bias for mortality outcomes is low as these outcomes are unlikely to be influenced by lack of participant or outcome assessor blinding or crossovers.

[b] Optimal information size criterion not met: no events occurred.

[c] Optimal information size criterion not met: very rare events and one study does not report which treatment group the death occurred in.

Abbreviation: N = number; RCT = randomized controlled trial; y = year(s).

No surgery-related deaths were reported in the RCTs that compared percutaneous endoscopic discectomy,[34] endoscopic discectomy,[31] microendoscopic discectomy,[24,29] or video-assisted arthroscopic microdiscectomy[43] to microdiscectomy or discectomy. We concluded with low certainty that no difference in surgical mortality between minimally-invasive surgery and standard surgery exists.

Two RCTs reported all-cause mortality.[31,40,48] Ruetten et al. reported one death (0.5%) unrelated to surgery; the authors did not specify whether this death occurred among participants allocated to the minimally-invasive surgery or among participants allocated to microdiscectomy. Arts et al. reported 2 deaths by 5 years among participants allocated to tubular discectomy and 3 deaths among participants allocated to microdiscectomy. We concluded with low certainty that no difference in all-cause mortality between minimally-invasive surgery and standard surgery exists.

C. Microdiscectomy compared with discectomy

Only one RCT reported a surgical mortality. Teli et al. reported no surgical deaths in either group.[29] No RCTs reported all-cause mortality. ***Table 28*** summarizes the findings and strength of evidence related to mortality for this comparison. Because only 1 RCT was available, we rated the strength of evidence as insufficient for both surgical mortality and all-cause mortality.

Table 28. Summary of findings and strength of evidence ratings comparing microdiscectomy to discectomy for mortality in persons with symptomatic lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Surgical mortality						
1 RCT	Not serious[a]	Not serious[b]	Not serious	Very serious[c]	Teli et al.[29](N=142) reported no surgery-related deaths in either treatment group.	◯◯◯◯ INSUFFICIENT
All-cause mortality						
0 RCTs	--					◯◯◯◯ INSUFFICIENT

[a] Though this trial was rated as some concerns for bias for efficacy outcomes and other safety outcomes, the risk of bias for mortality outcomes is low as these outcomes are unlikely to be influenced by lack of participant or outcome assessor blinding.

This RCT was rated as some concerns for bias. Sources of bias: lack of participant and outcome assessor blinding.

[b] Not applicable as only 1 study.

[c] Optimal information size criterion not met: no events occurred.

Abbreviations: N = number; RCT = randomized controlled trial.

D. Repeat lumbosacral decompression compared with spinal cord stimulation

North et al.[46] reported no deaths among participants allocated to repeat lumbosacral decompression and one death (2%) unrelated to treatment (sudden cardiac event) among participants allocated to spinal cord stimulation at 26 weeks. ***Table 29*** summarizes the findings and strength of evidence related to mortality. Because of only 1 study, we rated the strength of evidence as insufficient for mortality outcomes.

Table 29. Summary of findings and strength of evidence ratings of repeat lumbosacral decompression compared with spinal cord stimulation for mortality in persons with recurrent lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Surgical mortality						
1 RCT	Not serious[a]	Not serious[b]	Not serious	Very serious[c]	North et al.[46] (N=26 allocated to surgery) reported no surgery-related deaths.	◯◯◯◯ INSUFFICIENT
All-cause mortality						
1 RCT	Not serious[a]	Not serious[b]	Not serious	Very Serious[d]	North et al[46] (N=50 total) reported 1 death among spinal cord stimulation participants; the death was because of a sudden cardiac event.	◯◯◯◯ INSUFFICIENT

[a] Though this trial was rated as high risk of bias for efficacy outcomes and other safety outcomes, the risk of bias for mortality outcomes is low as these outcomes are unlikely to be influenced by lack of participant or outcome assessor blinding.

[b] Single study body of evidence, unable to assess consistency.

[c] Optimal information size criterion not met: no events occurred.

[d] Optimal information size criterion not met: only 1 death occurred.

Abbreviations: N = number; RCT = randomized controlled trial.

E. Revision endoscopic surgery compared with microdiscectomy

Ruetten et al.[47] compared revision endoscopic discectomy with revision microdiscectomy and reported no deaths (surgery-related or all-cause mortality) ***Table 30*** summarizes the findings and strength of evidence related to mortality. Because of only 1 study, we rated the strength of evidence as insufficient for mortality outcomes.

Table 30. Summary of findings and strength of evidence ratings of revision endoscopic surgery compared with revision microdiscectomy for mortality in persons with recurrent lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Surgical mortality						
1 RCT	Serious[a]	Not serious[b]	Not serious	Very serious[c]	Ruetten et al.[47] (N=100) reported no surgery-related deaths.	◯◯◯◯ INSUFFICIENT
All-cause mortality						
1 RCT	Serious[a]	Not serious[b]	Not serious	Very Serious[c]	Ruetten et al.[47] (N=100) reported no deaths in either surgical group.	◯◯◯◯ INSUFFICIENT

[a] Though this trial was rated as high risk of bias for efficacy outcomes and other safety outcomes, the risk of bias for mortality outcomes has only serious concerns as these outcomes are unlikely to be influenced by lack of participant or outcome assessor blinding. However, bias because of inadequate randomization remains.

[b] Single study body of evidence, unable to assess consistency.

[c] Optimal information size criterion not met: no events occurred.

Abbreviations: N = number; RCT = randomized controlled trial.

3.3.2.2 Surgical Morbidity

A. Surgery compared with nonsurgical interventions

All but one[26] RCT reported surgical morbidity outcomes.[22,23,32,33,37,41] ***Table 31*** summarizes the findings and strength of evidence related to the absolute incidence of surgical morbidity in the

surgical intervention group as comparative surgical morbidity outcomes are not relevant with a nonsurgical comparison group. A description of findings follows this table.

Table 31. Summary of findings and strength of evidence ratings for surgical morbidity in persons with symptomatic lumbar radiculopathy who undergo surgical intervention (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY / Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Surgical morbidity						
6 RCTs	Not serious[a]	Not serious	Not serious	Very serious[b]	Surgical complications among participants allocated to surgical groups were generally rare. The largest trial, Weinstein et al. [SPORT][22] (N=245 allocated to surgery), reported 10 (4%) dural tears or spinal leaks, 4 (1.6%) superficial wound infections, 1 (0.4%) vascular injury, 2 (0.8%) other intraoperative complication, and 9 (3.6%) other postoperative complication. Gerszten et al.[41](N=90 total) reported 5 (11%) adverse events among participants who underwent surgery and 7 (18%) among participants who underwent epidural steroid injection (calculated P=0.55). Osterman et al.[33](N=28 allocated to surgery) reported 1 (3.6%) case of urosepsis. Peul et al.[32](N=141 allocated to surgery) reported 2 dural tears and 1 wound hematoma. Erginousakis et al.[37](N=31 allocated to surgery) and McMorland et al.[23](N=20) reported 0 adverse events.	⊕⊕◯◯ LOW[c] NA[c]

[a] Though these trials were rated as either having some concerns for bias or high risk of bias for efficacy and other safety outcomes, the risk of bias for this outcome is low since non-comparative surgical morbidity outcomes are unlikely to be influenced by lack of participant or outcome assessor blinding or crossovers.

[b] Optimal information size criterion not met: events were rare in most studies.

[c] Because the comparator intervention is nonsurgical; this strength of evidence rating reflects the absolute incidence of adverse events related to surgery, not the relative incidence with respect to a comparator.

Abbreviations: N = number; RCT = randomized controlled trial; SPORT = Spine Patient Outcomes Research Trial.

We concluded with low certainty that surgical complications were generally rare among participants who underwent surgical intervention. Specific study findings:

- Weinstein et al. [SPORT][22] reported 10 (4.0%) dural tear or spinal fluid leaks, 4 (1.6%) superficial postoperative wound infection, 1 (0.40%) vascular injury, 2 (0.81%) other intraoperative complications, and 9 (3.6%) other unspecified postoperative complications among participants who underwent microdiscectomy.

- Gerszten et al.[41] reported 5 (11%) procedure-related adverse events among participants who underwent plasma disc decompression participants, compared with 7 (18%) procedure-related adverse events among participants who underwent epidural steroid injection participants. The authors used a broad definition for adverse events that included pain at the injection site, increased radicular pain, increased weakness, increased back pain, light headedness, and muscle tightness or spasms.

- Other reported morbidity outcomes among participants allocated to surgical interventions include one case of urosepsis (3.6%) reported by Osterman et al.[33] and one wound hematoma and two dural tears (combined 1.6%) reported by Peul et al.[32]

- McMorland et al.[23] reported no operative complications among participants who underwent microdiscectomy and Erginousakis et al.[37] reported no operative complications among participants who underwent percutaneous disc decompression.

B. Minimally-invasive surgery compared to microdiscectomy or discectomy

Ten RCTs[24,28-31,34,36,39,40,43] reported surgical morbidity outcomes. Because comparative surgical morbidity outcomes are likely to be influenced by the lack of outcome assessor blinding and studies did not provide any detail for us to assess whether adverse event ascertainment was equal, valid, and reliable, we rated this outcome as high risk of bias for most studies. ***Table 32*** summarizes the findings and strength of evidence related to morbidity for this comparison. A description of findings follows this table.

Table 32. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for surgical morbidity in persons with symptomatic lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of Effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Surgical morbidity						
10 RCTs	Very serious[a]	Not serious	Not serious	Very serious[b]	The most commonly reported complications were dural tears and spinal fluid leaks. Between-group differences were generally similar between groups with one exception—.Ruetten et al.[31] (N=200) found significantly fewer complications (P<0.05) in participants who underwent endoscopic discectomy compared with microdiscectomy.	⨁◯◯◯ VERY LOW No difference

[a] The risk of bias for this outcome is high since surgical morbidity outcomes are likely to be influenced by lack of outcome assessor blinding and studies generally did not report enough detail to assess whether adverse event ascertainment was equal, valid, and reliable in both study arms.

[b] Optimal information size criterion not met: though the total N randomized was 1,151, events were rare, and minimally important differences not clear, and most studies probably do not have sample sizes to detect small to modest differences.

Abbreviations: N = number; RCT = randomized controlled trial.

The most common complications reported were those relating to dural tear and spinal fluid leak. In nine of the 10 RCTs, morbidity incidence was similar between groups, though few reported statistical significance testing. One RCT reported significantly fewer complications among participants who underwent endoscopic discectomy compared to participants who underwent microdiscectomy.[31] We concluded with very low certainty that surgical morbidity was similar between minimally-invasive surgery and standard surgery. Specific study findings:

- Ruetten et al.[31] reported significantly fewer complications (P<0.05) among participants who underwent endoscopic discectomy compared to participants who underwent microdiscectomy participants. Complication included transient postoperative dysesthesia (3.3% vs 5.7%), postoperative bleeding (0% vs 2.3%), delayed wound healing (0% vs 1.1%), and soft tissue infection (0% vs 1.1%).

- In Arts et al.,[40] dural tears were the most common intraoperative complication reported in both groups, though the frequency did not differ between groups (actual values NR, P=0.18) and both intraoperative complications (11% vs. 9%, P=0.27) and postoperative complications (11% vs. 9%, P=0.47) were not different between groups.

- Brouwer et al.[39] reported 3 cases of transient nerve root injury among participants who underwent laser disc decompression and 6 cases of adverse events (3 CSF leaks, 1 transient nerve root injury, and 1 surgery at wrong level) among participant who underwent microdiscectomy (P values NR).[39] In addition, 5 (9%) cases of technical failure were observed among participants who underwent percutaneous laser disc decompression.

- Franke et al.[36] reported 2 dural tears among participants who underwent microscopically assisted percutaneous nucleotomy and 3 among participants who underwent microdiscectomy (calculated P=0.67).

- Ryang et al.[30] also reported dural tears (0 among participants who underwent trocar microdiscectomy and 2 among those who underwent microdiscectomy, calculated P=0.49).

- Hermantin et al.[43] reported 1 spinal fluid leak among participants who underwent discectomy and 0 among participants who underwent video-assisted arthroscopic microdiscectomy; no infections or neurovascular injuries were reported in either group.

- Huang et al.[24] reported 1 nerve root sleeve tear among participants who underwent microendoscopic discectomy and 0 cases among participants who underwent discectomy.

- Teli et al.[29] compared microendoscopic discectomy to microdiscectomy and reported dural tears (6 vs 2), root injury (2 vs 0), spondylodiscitis (1 vs 0), and worsening motor deficit (2 vs 1) and wound infection (0 vs. 4) among participants who underwent microendoscopic discectomy and microdiscectomy, respectively.

- Thome et al.[28] also reported similar frequency of complications among those who underwent sequestrectomy compared with those who underwent microdiscectomy (0 vs. 0 intraoperative complications, 0 vs. 1 wound infection, 1 vs.1 nerve root sheath tear, 1 vs. 0 dural leak, 1 vs. 0 discitis).

- Mayer et al.[34] reported no complications in either surgical group.

C. Microdiscectomy compared with discectomy

Three RCTs reported surgical morbidity,[27,29] but one did not report by group. ***Table 33*** summarizes the findings and strength of evidence related to morbidity for this comparison. In one RCT, the overall frequency of surgical infection was 6.3%.[35] The other two RCTs reported similar frequency of complications between groups, but no statistical testing was performed. Teli et al.[29] reported the frequency of dural tear (2 vs. 2), nerve root injury (0 vs. 0), wound infection (4 vs. 3), and worsening motor deficit (1 vs. 0), among participants who underwent microdiscectomy compared with participants who underwent discectomy, respectively. Tullberg et al.[27] reported the frequency of nerve root sheath (1 vs. 1), dural leak (1 vs. 0), and discitis (1

vs. 0) among participants who underwent microdiscectomy compared with participants who underwent discectomy, respectively. We concluded with very low certainty that surgical morbidity was similar for microdiscectomy compared with discectomy.

Table 33. Summary of findings and strength of evidence ratings comparing microdiscectomy to discectomy for surgical morbidity in persons with symptomatic lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY /Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Surgical morbidity						
3 RCTs	Very serious[a]	Not serious	Not serious	Very serious[b]	No between-group differences in various surgical morbidity outcomes. Teli et al.[29] (N=142) reported 2 dural tears in each group, no differences in nerve root injury, wound infection, spondylodiscitis or worsening motor function (all P values > 0.37 for comparisons). Tullberg et al.[27] (N=60) reported 1 nerve root sheath tear in each group, and 1 dural leak and 1 discitis among participants who underwent microdiscectomy and 0 among those who underwent discectomy (calculated P values were NS). Henriksen et al.[43] (N=80) reported 5 (6.3%) wound infections overall (NR by group).	⊕◯◯◯ VERY LOW No difference

[a] The risk of bias in these studies for this outcome is high since surgical morbidity outcomes are likely to be influenced by lack of outcome assessor blinding and studies generally did not report enough detail to assess whether adverse event ascertainment was equal, valid, and reliable in both study arms.

[b] Optimal information size criterion not met: rare to no events occurred and most study sample sizes are likely to small to detect small to modest differences.

Abbreviations: N = number; NS = not significant; NR = not reported; RCT = randomized controlled trial.

D. Repeat lumbosacral decompression compared with spinal cord stimulation

One RCT that compared repeat lumbosacral decompression to spinal cord stimulation reported surgical morbidity outcomes. ***Table 34*** summarizes the findings and strength of evidence related to surgical morbidity. North et al.[46] reported 0 (0%) site infections among repeat lumbosacral decompression participants and 1 (4.2%) among participants who underwent spinal cord stimulation. No other surgical complications were noted.

Table 34. Summary of findings and strength of evidence ratings of repeat lumbosacral decompression compared with spinal cord stimulation for surgical morbidity in persons with recurrent lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Surgical morbidity						
1 RCT	Very serious[a]	Not serious[b]	Not serious	Serious[c]	North et al.[46] (N=50) reported 1 (4.2%) site infection among participants who underwent spinal cord stimulation compared to 0 among participants who underwent revision microdiscectomy. No other surgical complications were reported.	◯◯◯◯ INSUFFICIENT

[a] The risk of bias in for this outcome is high since surgical morbidity outcomes are likely to be influenced by lack of outcome assessor blinding and did not report enough detail to assess whether adverse event ascertainment was equal, valid, and reliable in both study arms.

[b] Single study body of evidence, unable to assess consistency.

[c] Optimal information size criterion may not have been met: rare events.

Abbreviations: N = number; RCT = randomized controlled trial.

E. Revision endoscopic surgery compared with microdiscectomy

One RCT that compared revision endoscopic discectomy to revision microdiscectomy reported surgical morbidity outcomes. ***Table 35*** summarizes the findings and strength of evidence related to surgical morbidity. Ruetten et al.[47] significantly less serious complications (not further described) among revision endoscopic discectomy participants (6% vs 21%). All other complications, including dural injury (1 vs 3), transient postoperative dysesthesia (2 vs 5), delayed wound healing (0 vs 2), and soft tissue infection (0 vs 1) occurred with less frequency among participants who underwent revision endoscopic discectomy.

Table 35. Summary of findings and strength of evidence ratings of revision endoscopic surgery compared with revision microdiscectomy for surgical morbidity in persons with recurrent lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Surgical morbidity						
1 RCT	Very serious[a]	Not serious[b]	Not serious	Serious[c]	Ruetten et al.[47] (N=100) reported significantly fewer serious complications (not further described) in participants who underwent revision endoscopic discectomy compared to revision microdiscectomy group (6% vs. 21%, P<0.05). Other complications reported: dural injury (1 vs. 3), transient postoperative dysesthesia (2 vs. 5), soft tissue infection (0 vs. 1).	◯◯◯◯ INSUFFICIENT

[a] The risk of bias for this outcome is high since surgical morbidity outcomes are likely to be influenced by lack of outcome assessor blinding and did not report enough detail to assess whether adverse event ascertainment was equal, valid, and reliable in both study arms.

[b] Single study body of evidence, unable to assess consistency.

[c] Optimal information size criterion may not have been met: rare events.

Abbreviations: N = number; NR = not reported; RCT = randomized controlled trial.

3.3.2.3 Reoperations

This section reports on outcomes related to reoperations. This outcome should be interpreted with caution because of the variability in how studies characterized and reported this outcome. For example, some studies reported 'all-cause' reoperations, some studies only report reoperations for technical failures, and some studies did not define or explain how reoperations were defined. Further, studies varied considerably with respect to the timing of reoperations and whether offering a reoperation in the scenario of an unsuccessful outcome was a formalized part of the study protocol, and whether participants who underwent reoperations were included or excluded from analyses of efficacy outcomes.

A. Surgery compared with nonsurgical interventions

Five RCTs reported the incidence of reoperations in participants that were allocated to and underwent the surgical intervention; some studies also reported reoperations among participants who crossed over from the nonsurgical intervention to surgery.[22,23,32,33,37] ***Table 36*** summarizes the findings and strength of evidence related to reoperation for this comparison. Reoperations is not relevant as a comparative outcome given the nonsurgical comparison group. Thus, the strength of evidence for reoperation reflects our certainty about the absolute incidence of reoperations among those who underwent surgery, whether initially allocated to the surgical group or among those who crossed over to surgery at some point during the trial. A description of findings follows this table.

Table 36. Summary of findings and strength of evidence ratings comparing surgery to nonsurgical interventions for reoperations in persons with symptomatic lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Reoperations						
5 RCTs	Very serious[a]	Not serious	Not serious	Very serious[b]	The incidence of reoperations in study groups varied between 0% to 10.1%. Weinstein et al. [SPORT][22] (N=245 plus crossovers) 25 (10.1%) had a reoperation within 2y. Peul et al.[32] (N=142 allocated to surgery) 7 (6%) had reoperation by 2y and 9 (7%) by 5y. Osterman et al.[33] (N=28) allocated to surgery) 2 (7.1%) reoperations and Erginousakis[37] (N=31 allocated to surgery) 1 (3.2%). McMorland et al.[23] (N=20 allocated to surgery) reported no reoperations.	⨁◯◯◯ VERY LOW[c]

[a] This outcome was rated as high risk of bias in all studies because it includes reoperations among those who crossed over.

[b] Optimal information size criterion not met: unclear what a minimally important difference is, all but the largest trial have sample sizes that are unlikely to detect small to modest differences.

[c] Because the comparator intervention is nonsurgical; this strength of evidence rating reflects the absolute incidence of reoperations related to surgery, not the relative incidence with respect to a comparator.

Abbreviations: N = number; NR = not reported; RCT = randomized controlled trial; SPORT = Spine Patient Outcomes Research Trial; y = year(s).

We concluded with very low certainty that the incidence of reoperations varies between 0% and 10%. Specific study findings:

- Peul et al.[32] reported that 7 (6%) participants allocated to microdiscectomy had reoperations for recurrent sciatic within 2 years and 9 (7%) by 5 years. Among participants allocated to conservative management who crossed over to receive surgery, 4 (6%) underwent a reoperation by 2 years and 8 (12%) by 5 years.

- Weinstein et al. [SPORT] [2] reported that 25 (10.1%) of participants who underwent microdiscectomy/discectomy (including crossover) had reoperations for recurrent herniation, complication, or other reasons within 2 years.

- Erginousakis et al.[37] reported 1 reoperation among participants allocated to percutaneous disc decompression; follow-up time period was not reported but other outcomes were reported through 2 years follow-up.

- McMorland et al.[23] reported no reoperations among participants allocated to microdiscectomy; the follow-up time period was not reported but other outcomes were reported through 52 weeks.

- Osterman et al.[33] reported 2 (7.1%) reoperations among participants who underwent microdiscectomy; the follow-up time period was not reported but other outcomes were reported through 2 years.

B. Minimally-invasive surgery compared to microdiscectomy or discectomy

Ten RCTs reported reoperation rates.[28-31,34,36,38-40,43] ***Table 37*** summarizes the findings and strength of evidence related to reoperation for this comparison. A description of findings follows this table.

Table 37. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for reoperations in persons with symptomatic lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Reoperations						
10 RCTs	Very serious[a]	Serious[b]	Not serious	Very Serious[c]	The proportion of participants in each study group who had a reoperation varied from 2% to 64.5%. Between-group differences were not significant in 8 studies, but favored standard surgery in 2 studies.[38,39]. The pooled RR was 1.37 (95% CI 0.74 to 2.52, 10 RCTs, 1,172 participants, I^2=60.8%)) and pooled ARD 7% (95% CI, -2% to 17%, I^2=86.1%). See text for complete details.	◯◯◯◯ INSUFFICIENT

[a] One trial was rated as low risk for bias,[40] six were rated as some concerns for bias[28,29,36,38,39,43]and three were rated as high risk for bias.[30,31,34] Sources of bias: lack of participant and outcome assessor blinding in all but the low risk of bias trial, inadequate randomization and allocation concealment[30,31,34] and deviations from intended intervention.[34]

[b] The magnitude of between-group differences in two of the RCTs[38,39] were much larger than any of the other 8 RCTs in this body of evidence. I^2 statistic confirms statistical heterogeneity in pooled result. Sensitivity analysis of pooled estimate without the larger of the two outliers results in a pooled ARD of 2% (95% CI, -4% to 8 %; 9 RCTs, 1,101 participants; I^2= 60.7%) and pooled RR of 1.17 (95% CI, 0.70 to 1.97; I^2= 44.4%).

[c] Optimal information size criterion not met: reoperations rare in some studies, most study sample sizes unable to detect small to modest differences between groups. The width of the confidence intervals around pooled estimate cannot exclude a clinically meaningful favorable or unfavorable difference.

Abbreviations: N = number; NR = not reported; CI = confidence interval; RCT = randomized controlled trial; RR = relative risk ratio; ARD = absolute risk difference.

The proportion of participants that had reoperations varied extensively across study groups (from 2.5 % to 64.5). Two studies observed significantly higher frequency of reoperations among participants who underwent minimally-invasive surgery compared to standard surgery.[38,39] These findings were inconsistent with findings from the other 8 RCTs, which observed similar incidence of reoperations between surgical groups. The pooled absolute risk difference (ARD) was 7 % (95% CI, -2% to 17%; 10 RCTs, 1,172 participants, I^2 = 86.1%) and pooled relative risk ratio (RR) was 1.37 (95% CI, 0.74 to 2.52; I^2 = 60.8%) (***Appendix G***, ***Figure G-5***) suggesting no difference between treatment groups. Because of heterogeneity in the pooled estimate, we explored the two studies with significantly higher incidences of reoperations among the minimally-invasive surgery group.[38,39]

- Brower et al.[39] reported 24 (44%) reoperations within 52 weeks among those who underwent percutaneous laser decompression compared with 9 (16%) reoperations among those who underwent microdiscectomy (calculated P=0.002).[39] The authors of this study reported technical failures in 9% of the minimally-invasive group resulting from failure to reach the disc space. However, this study also reported higher reoperation rates in the standard surgery group, suggesting possible differences in the enrolled study population (e.g. duration or severity of disease) or differences in pre-, intra-, or post-operative care compared with other studies.

- Chatterjee et al.[38] reported 20 (64.5%) reoperations among participants who underwent automated percutaneous discectomy and 1 (2.5%) among participants who underwent microdiscectomy (calculated P<0.001). This study originally planned to enroll 160 participants with small, contained lumbar disc herniations but was halted after only 71 patients were enrolled as it became clear to the surgeon performing all the procedures that the minimally-invasive surgery had markedly inferior outcomes relative to the standard surgery group. Patients randomized to minimally-invasive surgery who were judged by a blinded observer to have had an unsuccessful outcome were offered a subsequent microdiscectomy and 20 of the 22 eligible participants accepted this offer. Although enrolled patients had confirmation of nerve root compression with MRI, the authors postulate that CT or discography would have better characterized the disc herniations resulting exclusion of participants who would be poor candidates for this minimally-invasive technique. The authors also postulate that their enrolled study population may have had longer duration of disease, and more dehydration of the disc, also contributing to poorer outcomes.

Because of the unique circumstances in the Chatterjee et al [38] study, we excluded it from the pooled estimate in a sensitivity analysis. The pooled ARD without it was 2% (95% CI, -4% to 8%; 9 RCTs, 1,101 participants; I^2= 60.7%) and the pooled RR was 1.17 (95% CI, 0.70 to 1.97; I^2= 44.4%). We believe the residual inconsistency is likely explained by varying definitions and

ascertainment methods (e.g., timing of measurement) for identifying reoperations, and because some studies may have been more or less aggressive in offering participants reoperations for residual symptoms. Excluding Chatterjee et al. reduced the heterogeneity of the pooled estimates somewhat, but the confidence intervals around the point estimate are still quite wide and suggest that the incidence of reoperations could be up to 26% lower to more than two times higher for participants allocated to minimally-invasive surgery compared to standard surgery. As a result, we assessed the evidence as insufficient for drawing a conclusion about the comparative difference in the incidence of reoperations between minimally-invasive and standard surgery.

Specific study findings of the other 8 RCTS:

- Arts et al.[40] reported 23 (15%) reoperations within 2 years among participants who underwent tubular discectomy and 14 (10%) reoperations among participants who underwent microdiscectomy (P=0.22). By 5 years, 30 (18%) and 21 (13%) had undergone a reoperation (P=0.29). Some participants had multiple reoperations; the number of reoperations by 5 years among participants who underwent tubular discectomy was 39 compared with 23 among participants that underwent microdiscectomy (P=0.10). There was no difference between groups in the number of participants who had two reoperations (P=0.45) or three reoperations (P=0.50); however, there was a significant difference in the proportion of patients who underwent instrumented fusion by 5 years (6/166 vs. 0 /159, calculated P= 0.01).

- Franke et al.[36] reported 2 (3.9%) reoperations among those who underwent microscopically-assisted percutaneous nucleotomy (1 was for relapse same level/same side and 1 was for progressive disc degeneration and segmental instability) and 5 (10.4%) reoperations among those who underwent microdiscectomy (4 were for relapse and 1 was for progression). This difference in proportion was not significant (calculated P=0.26).

- Mayer et al.[34] reported 3 (15%) reoperations among those who underwent percutaneous endoscopic discectomy compared with 1 (5%) reoperations among those who underwent microdiscectomy (calculated P=0.61).

- Ruetten et al.[31] reported 7 (7.7%) reoperations among participants who underwent endoscopic discectomy compared with 10 (11.5%) reoperations among those who underwent microdiscectomy (calculated P=0.45).

- Ryang et al.[30] reported 2 (6.6%) reoperations among participants who underwent trocar discectomy; one was during the initial hospital stay and one was within 6 weeks. Among participants who underwent microdiscectomy, 4 (13.3%) reoperations were performed (one during the initial hospital stay and one at 8 weeks, 28 weeks, and 1.2 years. The difference in proportion of participants who underwent reoperations between groups was not significant (calculated P=0.67).

- Teli et al.[29] reported 8 (11.4%) reoperations among participants who underwent microendoscopic discectomy compared with 3 (4.2%) reoperations among participants who underwent microdiscectomy (calculated P=0.13).

- Thome et al.[28] reported 2 (5%) reoperations for recurrent herniation within 1.5 years among participants who underwent sequestrectomy compared with 4 (10%) reoperations among participants who underwent microdiscectomy. This difference was not significant (calculated P=0.68). Reherniations (with or without reoperation) were reported in 12.5% of participants allocated to sequestrectomy compared with 10.5% of participants allocated to microdiscectomy (calculated P=1.0).

- Hermantin et al.[43] reported 1 (3.3%) reoperations among participants who underwent video-assisted arthroscopic microdiscectomy for treatment of mild lateral stenosis that had not been recognized at the time of surgery. Among participants who underwent discectomy, 2 (6.7%) reoperations were performed; one was for repair of a dural sac/spinal fluid leak and the other was for persistent radicular symptoms. The difference in proportion between groups was not significant (P=1.0).

C. Microdiscectomy compared with discectomy

Two RCTs reported on reoperations.[27,29] ***Table 38*** summarizes the findings and strength of evidence related to reoperation for this comparison. Tullberg et al.[27] reported 1 (3.3%) reoperation by 52 weeks in each surgical group (microdiscectomy and discectomy). Teli et al.[29] reported 3 (4.2%) reoperations among participants who underwent microdiscectomy compared with 2 (3%) among participants who underwent discectomy (calculated P=1.0).

Table 38. Summary of findings and strength of evidence ratings comparing microdiscectomy to discectomy for reoperations in persons with symptomatic lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Reoperations						
2 RCTs	Serious[a]	Not serious	Not serious	Very serious[b]	No between-group differences in frequency of reoperations. Tullberg et al.[27] (N=60) reported 1 (3.3%) reoperation in each surgical group by 52w (calculated P=1.0). Teli et al.[29] (N=142) reported 3 (4.2%) reoperations among participants who underwent microdiscectomy compared with 2 (3%) among participants who underwent discectomy (calculated P=1.0).	⊕◯◯◯ VERY LOW No difference

[a] Both studies were rated as some concerns for risk of bias. Sources of bias: lack of participant and outcome assessor blinding.

[b] Optimal information size criterion not met: reoperations were rare events, study sample sizes unable to detect small to modest differences between groups.

Abbreviations: N = number; NR = not reported; RCT = randomized controlled trial; w = week(s).

D. Repeat lumbosacral decompression compared with spinal cord stimulation

One RCT reported reoperations.[46] ***Table 39*** summarizes the findings and strength of evidence related to reoperation for this comparison. North et al.[46] reported 0 (0%) reoperations among

participants who underwent repeat lumbosacral decompression and 3 (12.5%) hardware revisions among participants who underwent spinal cord stimulation. Because of only one study, we assessed the evidence as insufficient to draw a conclusion about the incidence of reoperations.

Table 39. Summary of findings and strength of evidence ratings of repeat lumbosacral decompression compared with spinal cord stimulation for reoperations in persons with recurrent lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Reoperations						
1 RCT	Very serious[a]	Not serious	Not serious	Very serious[b]	North et al.[46] (N=50) reported 0 reoperations among participants who underwent repeat decompression, and 3 (12.5%) hardware revisions among participants who underwent spinal cord stimulation (P value NR).	◯◯◯◯ INSUFFICIENT

[a] Trial was rated as high risk of bias. Sources of bias: lack of participant and outcome assessor blinding and extensive crossovers and differential attrition.[46]

[b] Rare events, optimal information size criteria not met.

Abbreviations: N = number; NR = not reported; RCT = randomized controlled trial.

E. Revision endoscopic surgery compared with microdiscectomy

One RCT reported reoperations. ***Table 40*** summarizes the findings and strength of evidence related to reoperation for this comparison. Ruetten et al.[47] reported 2 reoperations at 2 years among participants who underwent revision endoscopic discectomy compared with 3 reoperations among participants who underwent revision microdiscectomy. Because of only one study, we assessed the evidence as insufficient to draw a conclusion about the incidence of reoperations.

Table 40. Summary of findings and strength of evidence ratings of revision endoscopic surgery compared with revision microdiscectomy for reoperations in persons with recurrent lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Reoperations						
2 RCTs	Very serious[a]	Not serious	Not serious	Very serious[b]	Ruetten et al.[47] (N=100) reported 2 (4%) reoperations among participants who underwent revision endoscopic discectomy and 3 (6%) reoperations among participants who underwent revision microdiscectomy (calculated P=0.67)	◯◯◯◯ INSUFFICIENT

[a] Trial was rated as high risk of bias. Sources of bias: lack of participant and outcome assessor blinding and extensive crossovers and differential attrition, and inadequate randomization and allocation concealment.[47]

[b] Rare events, optimal information size criteria not met.

Abbreviations: N = number; RCT = randomized controlled trial.

3.3.2.4 Persistent Opioid Use

A. Surgery compared with nonsurgical interventions

Only one RCT reported outcomes related to persistent opioid use.[41] ***Table 41*** summarizes the findings and strength of evidence related to persistent opioid use for this comparison. Gerszten et al. reported that reduction in use of narcotics was not significantly different between participants who underwent percutaneous disc decompression and those who underwent conservative management participants at 26 weeks (actual values NR, P value NR). Because of only one study, we assessed the evidence as insufficient for persistent opioid use outcomes.

Table 41. Summary of findings and strength of evidence ratings comparing surgery to nonsurgical interventions for persistent opioid use in persons with symptomatic lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Persistent opioid use						
1 RCT	Very serious[a]	Not serious[b]	Not serious	Very serious[c]	Gerszten et al.[41] (N=90) reported no significant difference in reduction in narcotics between participants who underwent percutaneous disc decompression and compared with conservative management participants at 26w (actual values NR, P value NR).	◯◯◯◯ INSUFFICIENT

[a] Risk of bias is high for 12 week and later outcomes. Sources of bias: lack of participant and outcome assessor blinding, and deviations from intended interventions.

[b] Not applicable as only 1 study.

[c] Optimal information size criterion not met: actual values and measures of variance not reported.

Abbreviations: N = number; NR = not reported; RCT = randomized controlled trial; w = week(s).

B. Minimally-invasive surgery compared to microdiscectomy or discectomy

Only one RCT reported outcomes related to persistent opioid use.[43] ***Table 42*** summarizes the findings and strength of evidence related to persistent opioid use for this comparison. The duration of postoperative narcotic use ranged from 0.43 to 2 weeks (average 1 week) for participants who underwent video-assisted arthroscopic microdiscectomy and 1 to 8 weeks (average 3.65 weeks) for participants who underwent discectomy (P value NR). Because of only one RCT, we assessed the evidence as insufficient for persistent opioid use outcomes.

Table 42. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for persistent opioid use in persons with symptomatic lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Persistent opioid use						
1 RCT	Serious[a]	Not serious[b]	Not serious	Very serious[c]	Hermantin et al.[43] (N=60) reported mean duration of postoperative narcotic use in video-assisted arthroscopic microdiscectomy participants of 1 week (range 0.43w to 2w) compared to 3.65 weeks (range 1w to 8w) in participants who underwent discectomy (P value NR).	◯◯◯◯ INSUFFICIENT

[a] This RCT was rated as some concerns for bias. Sources of bias: lack of participant and outcome assessor blinding.

[b] Not applicable as only 1 study.

[c] Optimal information size criterion not met: measures of variance or confidence intervals not provided.

Abbreviations: N = number; NR = not reported; RCT = randomized controlled trial; w = week(s).

C. Microdiscectomy compared with discectomy

No RCTs reported persistent opioid use outcomes; thus, we assessed the evidence as insufficient for persistent opioid use outcomes.

D. Repeat lumbosacral decompression compared with spinal cord stimulation

One RCTs reported outcomes related to opioid use.[46] ***Table 43*** summarizes the findings and strength of evidence related to persistent opioid use for this comparison. North et al. reported that by 2.9 years, 15 (58%) participants who underwent repeat lumbosacral decompression participants reported stable or decreased opioid use compared to 20 (87%) participants who underwent spinal cord stimulation participants (P=0.025).[46] Because of only one RCT, we assessed the evidence as insufficient for persistent opioid use outcomes.

Table 43. Summary of findings and strength of evidence ratings of repeat lumbosacral decompression compared with spinal cord stimulation for persistent opioid use in persons with recurrent lumbar radiculopathy (SQ1)

Certainty Assessment					Summary of Findings	CERTAINTY/ Direction of effect
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Persistent opioid use (follow-up 1.8 to 5.7 years)						
1 RCTs	Very serious[a]	Not serious[b]	Not serious	Not serious	North et al.[46] (N=50) 15 (58%) participants with stable or decreased opioid use among participants allocated to repeat surgery compared to 20 (80%) participants allocated to spinal cord stimulation (P=0.025).	◯◯◯◯ INSUFFICIENT

[a] Trial was rated as high risk of bias. Sources of bias: lack of participant and outcome assessor blinding and extensive crossovers and differential attrition.[46]

[b] Not applicable as this body of evidence has only one study.

Abbreviations: N = number; RCT = randomized controlled trial.

E. Revision endoscopic surgery compared with microdiscectomy

No RCTs reported persistent opioid use outcomes; thus, we assessed the evidence as insufficient for persistent opioid use outcomes.

3.4 Cost and Cost-Effectiveness

Cost Question 1

In adults with symptomatic lumbar radiculopathy, what is the cost-effectiveness of surgical interventions?

We identified seven eligible studies reporting cost.[29,44,49-53] Five studies reported cost-effectiveness analyses related to RCTs that we also included for efficacy and safety outcomes (Peul et al., Weinstein et al., Arts et al., Brouwer et al., and Chatterjee et al.).[49-53] One study, Teli et al.,[29] was a trial we also included for efficacy and safety outcomes and reported on surgical costs of three alternative surgical interventions. Lastly, Malter et al.[44] reported a cost-effectiveness analysis using cost and effectiveness inputs from a variety of sources. A summary of included studies is provided in ***Table 44***. ***Appendix D, Tables D-1*** and ***D-2*** provide detailed individual study and population characteristics for the cost studies related to included RCTs and ***Table D-6*** provides detailed information related to methods and findings specific to the seven cost studies. Studies reported cost findings using different currency and base years; thus, we converted all figures to 2010 U.S. dollars (see ***Appendix C*** for details on conversion) for this report. ***Appendix F, Table F-6, F-7, and F-8*** provide quality assessments of individual studies specific to the cost-effectiveness analyses.

Table 44. Study characteristics of the seven studies that evaluated cost effectiveness of surgery for lumbar radiculopathy (CQ1)

Author (Year) Related RCT (Year) Country Quality	Surgical Intervention (N randomized)	Comparator (N randomized)	Key Analysis Parameters	Outcomes Reported
Cost-effectiveness of surgery compared with nonsurgical interventions				
Van den Hout (2008) [49] Peul (2007)[32] Netherlands Good	Discectomy (141)	Conservative management (142)	Year/Currency: 2008 € Discount rate: 0% Time Horizon: 52w Costs included: direct and indirect QOL measure: EQ-5D with U.K. norms	• Mean QALY • Mean total costs • Cost/QALY gained • Mean health care costs • Health care costs/QALY gained
Tosteson (2008)[50] Weinstein (2006)[22] [SPORT] U.S. Good	Discectomy /microdiscectomy (245)	Conservative management (256)	Year/ currency: 2004 U.S.$ Discount rate: 3% Time horizon: 2y Costs included: direct and indirect QOL measure: EQ-5D with U.S. norms Other: Based on pooled data from SPORT RCT and observational cohort.	• Mean QALY • Mean total costs • Cost/QALY gained • Mean direct medical costs • Direct medical costs/QALY gained

(continued)

Table 44. Study characteristics of the seven studies that evaluated cost effectiveness of surgery for lumbar radiculopathy (CQ1) (continued)

Author (Year); Related RCT	Surgical Intervention (N randomized)	Comparator (N randomized)	Key Analysis Parameters	Outcomes Reported
Malter (1996)[44] U.S. Cost Data Fair	Discectomy (NA)	Nonsurgical management (NA)	Year/currency: 1993 U.S.$ Discount rate: 5% Time horizon: 10y Costs included: Direct medical costs QOL measure: Author developed time-trade off utility measure Other: Efficacy estimates based on an RCT comparing surgery with nonsurgical treatment (Weber et al. (1983)[26]), and an RCT comparing surgery with chemonucleolysis (Javid et al. (1998)[111]) and a cohort study comparing surgery with nonsurgical treatment (Atlas et al. (1993)[112]).	• QALY • Costs • Costs/QALY gained
Cost-effectiveness of alternative surgical interventions				
Van den Akker (2011)[51] Arts (2009)[40] Good	Tubular discectomy (167)	Microdiscectomy (161)	*Year/Currency:* 2008 U.S.$ *Discount rate:*0% *Time horizon:* 52w *QOL measure:* EQ5-D *Costs included:* direct and indirect	• Mean QALY • Mean total costs • Cost/QALY gained • Mean health care costs • Health care costs/QALY gained
Van den Akker (2017)[52] Brouwer (2015)[39] Good	Percutaneous laser disc decompression (57)	Discectomy, with laminotomy as needed (58)	*Year/ Currency:* 2010 € *Discount rate:* 0% *Time Horizon:* 52w *QOL measure:* EQ-5D with U.S. norms *Costs Included:* direct and indirect	• Mean QALY • Mean total costs • Costs/QALY gained • Health care costs/QALY gained
Stevenson (1995)[53] Chatterjee[38] (1995) Poor	Automated percutaneous lumbar discectomy (31)	Microdiscectomy (40)	*Year/Currency:* 1992 £ *Discount rate:* NR *Time Horizon:* 26w *Costs Included:* direct and indirect *QOL measure:* NR *Other:* effectiveness was assessed on a 4-pt Likert scale by two clinicians (4=excellent, 1=poor). "Successful outcome" was defined as a 3 or 4.	• Mean total cost • Cost per successful outcome • Cost per point gained on 4-pt Likert scale of effectiveness
Teli (2010)[29] NA	Microendoscopic discectomy (70)	Microdiscectomy (72) Open discectomy (70)	*Year/currency:* Euros, Year NR *Discount rate:* NR *Time horizon:* NA, procedure costs only *Costs included:* procedure costs only	• Mean surgical costs

Abbreviations: N = number; NA = not applicable; NR = not reported; QALY = quality-adjusted life year; QOL = quality of life; RCT = randomized controlled trial; SPORT = Spine Patient Outcomes Research Trial; U.K. = United Kingdom; U.S. = United States; w = week(s); y = year(s).

3.4.1 Study Characteristics

One study was a decision analysis based on published cost and effectiveness data.[44] One study provided cost outcomes as part of the main RCT publication.[29] Five studies were separately published cost-effectiveness analyses of RCTs that we included for efficacy and safety

outcomes[29,49,51-53] One of these analyses, Tosteson et al.,[50] used pooled data from the SPORT trial RCT[22] combined with the SPORT observational study (combined N=1,191). Crossovers in the RCT were included in the treatment group to which they crossed over. Two studies were conducted in the United States;[44,50] the rest were conducted in the Netherlands,[49,51,52] Italy,[29] and the United Kingdom.[53] The time horizon used in studies ranged from 26 weeks to 10 years.

Three studies provided evidence for the cost-effectiveness of surgery compared with nonsurgical treatment.[44,49,50] Four studies[29,51-53] provided evidence for the comparative cost-effectiveness of alternative surgical interventions, including percutaneous laser discectomy,[52] tubular discectomy,[51] and automated percutaneous discectomy[53] compared to microdiscectomy and a three-arm study comparing microendoscopic discectomy, microdiscectomy, and discectomy.[29]

Four studies calculated mean quality adjusted life years (QALY) using the EQ-5D measure of health-related quality of life, and one used an author developed time-tradeoff utility measure.[44] These five studies included both direct and indirect costs. The other two studies did not measure QALYs and only reported direct medical or procedure costs[29,53]

We rated one cost-effectiveness analysis as poor quality[53], one as fair quality [44], and four as good quality.[49-52] We did not assess the quality for Teli et al, as it only reported costs and not cost-effectiveness.

3.4.2 Findings

A. Surgery compared with nonsurgical interventions

Three studies reported cost and cost-effectiveness results.[44,49,50] ***Table 45*** summarizes the findings and strength of evidence related to cost-effectiveness for this comparison. A detailed description of findings follows this table. Two reported findings from both a societal perspective (direct and indirect costs)[49,50] and a payor perspective (direct medical costs); whereas one reported only from a payor perspective.[44] A detailed description of findings follows this table.

Table 45. Summary of findings and strength of evidence ratings comparing surgery to nonsurgical interventions for cost and cost-effectiveness in persons with symptomatic lumbar radiculopathy (CQ1)

Certainty Assessment					Summary of Findings[a]	CERTAINTY
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Cost-effectiveness						
3 studies[b]	Very serious[c]	Not serious	Not serious	Very serious[d]	All studies reported higher QALYs but similar or higher costs among surgical interventions compared to conservative management. The mean cost per QALY gained from the payor perspective ranged from $51,156 to $83,322. Tosteson et al.[50] [SPORT](N=1,191) reported total cost per QALY gained at 2y $80,115 (95% CI, $56,167 to $109,662) and direct medical costs per QALY gained $83,322 (95% CI, $65,189 to $106,655). Van den Hout et al.[49] (N=283) did not report an ICER for total costs per QALY gained (calculated ICER $-419/QALY gained interpreted as surgery more effective and costs less), the health care costs per QALY gained at 52w was $63,011 (95% CI, $21,516 to $660,847). Malter et al.[44] reported undiscounted cost per QALY gained $44,064 (95% CI NR, discounted $51,156 per QALY gained).[d]	⊕◯◯◯ VERY LOW

[a] All costs are reported here in 2010 U.S. Dollars. The costs for the year and currency reported in the published studies is in ***Appendix D***, ***Table D-6***.

[b] Two studies were cost-effective analyses were conducted concurrent to RCTs; one study was a cost-effectiveness analysis conducted using effectiveness inputs from published RCTs and cost inputs from a commercial database.

[c] The RCTs related to two of the cost studies were rated as high risk for bias. Sources of bias: lack of participant and outcome assessor blinding and extensive crossovers. The cost analyses associated with these studies were rated as some concerns for bias.

[d] One study did not provide confidence intervals around incremental cost-effectiveness ratio estimate and we were unable to calculate it based on available data[44]; the other 2 studies have confidence intervals around their incremental cost-effectiveness ratio estimate that span the range of cost savings, probably cost-effective, and not at all cost-effective.

Abbreviations: ICER = incremental cost-effectiveness ratio; N = number; NR = not reported; CI = confidence interval; QALY = quality-adjusted life year; RCT = randomized controlled trial; SPORT = Spine Patient Outcomes Research Trial; w = week(s); y = year(s).

All studies reported higher QALYs among participants allocated to surgical interventions, but similar or higher costs. The mean cost per QALY gained from the payor perspective ranged from $51,156 to $83,322; we assessed the certainty of these estimates as very low.

- Tosteson et al.[50] used cost and effectiveness inputs from Weinstein et al. [SPORT],[22] an RCT conducted in the United States, and reported outcomes at 2 years using 2004 U.S. dollars ($). The difference in QALYs was 0.21 (95% CI 0.16 to 25), favoring microdiscectomy/discectomy compared with conservative management. The total cost among participants who received surgery was $31,561 (95% CI, $29,877 to $33,244) and the total cost among participants who received conservative management was $15,162 (95% CI, $12,979 to $17,202); the AMD was not reported but we calculated it to be $16,399 (95% CI, 95% CI, $16,289 to $16,509). The study reported the cost per QALY gained (also known as the incremental cost-effectiveness ratio, ICER) was $80,115 (95% CI, $57,167 to $109,662). When limited to direct medical costs, the cost among participants who received surgery was $23,361 (95% CI, $22,295 to $24,426) and the cost

among participants who receive conservative management was $6,700 (95% CI, $5,355 to $8,045); we calculated this AMD to be $16,661 (95% CI, $16,590 to $16,732). The direct medical costs per QALY gained was $83,322 (95% CI, $65,189 to $106,655), which means that although surgery is more effective, it also costs more. Whether this estimate is cost-effective depends on the threshold of additional costs that payors are willing to pay for an additional QALY.

- Van den Hout et al.[49] used cost and effectiveness inputs from an RCT conducted in The Netherlands[32] and reported outcomes at 52 weeks using 2008 Euros (€). Microdiscectomy resulted in a higher mean QALY (AMD 0.044 [95% CI, 0.005 to 0.083] but no significant differences in total costs (AMD -$18.44 [95% CI, $-6,192 to $6,157) compared with conservative management. The authors concluded that microdiscectomy dominates conservative management (i.e., is more effective and costs less). However, no cost per QALY gained (i.e., ICER) was reported for the societal perspective. We calculated the ICER as $-419 per QALY gained (95% CI unable to be calculated). Using the 95% confidence intervals provided for the difference in QALYs and the difference in costs by the published study, we calculated the possible range of this estimate to be from $-74,602/QALY (best case) to $1,231,400 (worse case). When limited to health care costs only, the mean cost among participants in the surgical group was $8,646 (SD $5,955) and $5,851 (SD $6,512) for an AMD of $2,796 (95% CI, $1,294 to $4,288). This results in an ICER of $63,011 (95% CI, $21,516 to $660,847), which means that surgery was more effective but also costs more.

- Lastly, Malter et al.[44] conducted a cost-effectiveness analysis over a 10-year horizon using effectiveness data inputs from several published RCTs and observational studies of surgery compared with nonsurgical interventions, including conservative management and chemonucleolysis comparators, and an author-developed time-tradeoff utility measure.[26,111,112] Health care cost inputs were obtained from a commercial database of U.S. costs from 1987 to 1989 and were adjusted to 1993 dollars for reporting in the analysis. The difference in undiscounted QALYs was 0.43 (95% CI NR). The health care costs associated with surgery were estimated to be $25,684 (95% CI NR) and the costs associated with nonsurgical management were estimated to be $6,745 (95% CI NR) (AMD $18,938 [95% CI NR]). The undiscounted cost per QALY gained (ICER) was $44,064 (95% CI NR) and the discounted cost per QALY gained was $51,156 (95% CI NR), which means that surgery was more effective but also costs more.

B. Minimally-invasive surgery compared to microdiscectomy or discectomy

Four studies reported cost and cost-effectiveness results.[29,51-53] One reported findings only from a societal perspective,[53] one reported findings only from a payor perspective,[29] and two reported findings from both a societal perspective and a payor perspective.[51,52] ***Table 46*** summarizes the findings and strength of evidence related to cost-effectiveness for this comparison. A detailed description of findings follows this table.

Table 46. Summary of findings and strength of evidence ratings comparing minimally-invasive surgery to standard surgery for cost and cost-effectiveness in persons with symptomatic lumbar radiculopathy (CQ1)

Certainty Assessment					Summary of Findings[a]	CERTAINTY
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Cost and cost-effectiveness						
4 studies[b]	Serious[c]	Serious[d]	Not serious	Very serious[e]	Inconsistent findings across studies. Teli et al.[29] (N=142) calculated AMD for surgical costs $722 (95% CI, $551 to $892) comparing microendoscopic discectomy to microdiscectomy. Stevens et al.[53] (N=70) calculated AMD cost per successful outcome at 26w $3,573 comparing automated percutaneous lumbar discectomy to microdiscectomy. Van den Akker[51] (N=325) no significant differences in QALYs, total costs, or health care costs; costs and health care costs per QALY are NR but point estimates for differences in QALYs and costs suggest minimally-invasive surgery is less effective and costs more. Van den Akker[52] (N=115) no significant differences in QALYs or total costs but some difference in health care costs (AMD $-2,393 (95% CI, $-4,376 to $-409); costs and health care costs per QALY NR, but point estimates for differences in QALYs and costs suggest that microdiscectomy may be more effective but also costs more (calculated cost per QALY $97,424).	◯◯◯◯ INSUFFICIENT

[a] All costs are reported here in 2010 U.S. Dollars. The costs for the year and currency reported in the published studies is in ***Appendix D***, ***Table D-6***.

[b] All studies were conducted concurrent to RCTs.

[c] The risk of bias in one of the related RCTs was low and the risk of bias in the other three related RCTs was some concerns. Sources of bias: lack of participant and outcome assessor blinding. The risk of bias specific to the cost analyses was some concerns.

[d] Three studies show that minimally-invasive surgery is slightly more expensive, one study shows that minimally-invasive surgery is significantly less expensive. Inconsistency in findings related to cost-effectiveness among the two studies that reported QALYs.

[e] Confidence intervals span thresholds of cost effectiveness and cost savings, some studies do not report estimates for the incremental cost-effectiveness ratio or measures of variance.

Abbreviations: ICER = incremental cost-effectiveness ratio; AMD = absolute mean difference; CI = confidence interval; N = number; NR = not reported; QALY = quality-adjusted life year; RCT = randomized controlled trial; w = week(s); y = year(s).

The findings were inconsistent across studies, likely because of differences in study methods. Thus, we assessed the evidence as insufficient to draw a conclusion about the cost-effectiveness of minimally-invasive surgery compared with standard surgery.

- Teli et al.[29] reported surgical costs using Euros (year unspecified) as part of reporting results from an RCT conducted in Italy. The mean surgical cost among participants who underwent microendoscopic discectomy was $3,878 (SD $580) and was $3,156 (SD $438) among participants who underwent microdiscectomy. The AMD was not reported but we

calculated it to be $722 (95% CI, $551 to $892), which suggests that the minimally-invasive approach was more expensive than microdiscectomy.

- Stevenson et al.[53] used cost inputs from an RCT conducted in the United Kingdom[38] and reported outcomes at 26 weeks using 1992 British pounds (£). This study reported costs and cost per successful outcome, where two clinicians measured success using a 4-point Likert scale (1=poor, 4=excellent) (i.e., MacNab criteria). Success was defined as a 3 or 4 on this scale. In the related RCT, 9 (29%) of participants allocated to automated percutaneous lumbar discectomy had a successful outcome compared with 32 (80%) of participants allocated to microdiscectomy (p< 0.001).[38] The mean differences in success scores were not reported in the main RCT or in the cost study. The mean total cost was $6,340 among participants allocated to automated percutaneous lumbar discectomy and $4,288 among participants allocated to microdiscectomy (AMD NR, calculated to be $2,052). Cost estimates for both groups included the cost of additional surgeries for failed initial surgery. The cost per successful outcome among participants allocated to automated percutaneous lumbar discectomy was $8,931 versus $5,358 among participants allocated to microdiscectomy (calculated AMD $3,573). The authors calculated the cost per additional point gained on the 4-pt. Likert scale of success as $3,770 vs. $2,091 for the two surgical groups, respectively (calculated AMD $1,688).

- Van den Akker et al.[51] used cost and effectiveness inputs from an RCT conducted in The Netherlands[40] and reported outcomes at 52 weeks using 2008 U.S.$. The mean difference in QALYs was not significantly different between participants who underwent tubular discectomy compared with participants who underwent microdiscectomy (AMD -0.12 [95% CI, -0.046 to 0.021]). Similarly, no significant difference in mean total costs or health care costs were observed (AMD $1,510 [95% CI, -$1,352 to $4,373] and AMD $466 ([95% CI, -$246 to $1,178], respectively). The total costs per QALY (ICER) was not reported. Point estimates for the difference in effectiveness and cost were not significant but the direction of their differences suggests that microdiscectomy dominates (i.e., is more effective and costs less) minimally-invasive surgery.

- Van den Akker et al.[52] also reported cost and cost-effectiveness from a different RCT conducted in The Netherlands.[39] This analysis reported outcomes at 52 weeks using 2010 Euros. The mean difference in QALYs was -0.033 (95% CI NR, p=0.27) comparing percutaneous laser disc discectomy with microdiscectomy. The difference in total costs was $-3,215 (95% CI, -$10,294 to $3,865) and the difference in health care costs was $-2,393 (95% CI, -$4,376 to $-409). Point estimates for the difference in effectiveness and total costs were not significant but the direction of their differences suggests that microdiscectomy is more effective but also costs more compared to minimally-invasive surgery. Calculated cost per QALY (ICER) was $97,424 for microdiscectomy compared to minimally-invasive surgery.

C. Microdiscectomy compared with discectomy

Teli et al.[29] also reported surgical costs of microdiscectomy to discectomy, but did not report cost-effectiveness. The cost of microdiscectomy was $3,156 (SD $438) and the cost of

discectomy was $2,976 (SD $322). The AMD was not reported but we calculated it to be $65 (95% CI, $52 to $307). ***Table 47*** summarizes the findings and strength of evidence related to cost for this comparison. Because of only one study, we assessed the evidence as insufficient to draw a conclusion about the costs of microdiscectomy compared to discectomy.

Table 47. Summary of findings and strength of evidence ratings comparing microdiscectomy to discectomy for direct surgical costs in persons with symptomatic lumbar radiculopathy (CQ1)

Certainty Assessment					Summary of Findings[a]	CERTAINTY
№ of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision		
Direct surgical costs						
1 study[b]	Serious[c]	Not serious[d]	Not serious	Not serious	Teli et al.[29] (N=142) reported slightly higher costs for microdiscectomy compared to discectomy ($3,156 (SD $438) vs $2,976 (SD $322). Calculated AMD $65 (95% CI, $52 to $307).	◯◯◯◯ INSUFFICIENT

[a] All costs are reported here in 2010 U.S. Dollars. The costs for the year and currency reported in the published studies is in ***Appendix D***, ***Table D-6***.

[b] Cost effectiveness analysis conducted concurrent to an RCT

[c] The related RCT was rated as some concerns for bias Sources of bias: lack of participant and outcome assessor blinding. The cost analysis associated with this RCT was rated as some concerns for bias.

[d] Not applicable as only 1 study.

Abbreviations: AMD = absolute mean difference; N = number; RCT = randomized controlled trial; SD = standard deviation.

3.5 Clinical Practice Guideline Synthesis

We identified 14 relevant clinical practice guidelines (CPGs) or interventional procedures guidance related to the use of surgical interventions for lumbar radiculopathy; these are summarized in ***Table 48.*** We identified CPGs from the American Pain Society (2009), the American Society of Interventional Pain Physicians (2013), the North American Spine Society (2012), the National Institute for Health and Care Excellence (U.K.) (2016), and the American College of Occupational and Environmental Medicine (2016). In addition, we identified nine "interventional procedures guidance" related to minimally-invasive spine procedures from the National Institute for Health and Care Excellence (U.K.). The CPGs vary with respect to how they were developed and the types of studies that were included to inform clinical recommendations. For example, many of the interventional procedure guidance documents were partly based on case series, and most included cohort studies or other nonrandomized trial study designs. The strength of evidence that the guidelines are based on also varied. Only one CPG (National Institute for Health and Care Excellence, "Low Back pain and Sciatica in over 16s") used the GRADE approach; this CPG rated the evidence as low or very low for nearly all comparisons and outcomes considered.[45]

Table 48. Clinical practice guidelines related to lumbar radiculopathy or herniated intervertebral lumbar disc

Organization *Guideline Title (Year)* Guideline Quality[a]	Recommendation[b]	Evidence Base	Rating/Strength of Evidence Narrative Assessment
National Institute for Health and Care Excellence (United Kingdom) *Low back pain and sciatica in over 16s: assessment and management-Invasive treatments (2016)*[45] Quality Rating: 6 out of 7	Consider spinal decompression for sciatica (includes laminectomy, foraminotomy, and/or discectomy) when nonsurgical treatment has not improved pain or function and their radiological findings are consistent with sciatica symptoms.	9 RCTs comparing surgery to nonsurgical treatment including epidural steroids, analgesics and anti-inflammatory medication, physical therapy[c] 4 cohort studies comparing decompression to fusion or conservative treatment	Low or very low for nearly all comparisons and outcomes[d] Sciatic symptoms tend to improve naturally with time without treatment, but earlier symptom resolution with surgical intervention should be an option for people.
North American Spine Society *Clinical Guidelines for Diagnosis and Treatment of Lumbar Disc Herniation with Radiculopathy (2012)*[57] Quality Rating: 5 out of 7	Discectomy is suggested to provide more effective symptom relief than medical/interventional care for patients with lumbar disc herniation with radiculopathy whose symptoms warrant surgical intervention. In patients with less severe symptoms, surgery or medical/interventional care appear to be effective for both short- and long-term relief.	3 RCTs 2 prospective comparative cohort studies	Grade: B[e]
	Surgical intervention prior to 6 months is suggested in patients with symptomatic lumbar disc herniation whose symptoms are severe enough to warrant surgery. Earlier surgery (within 6 months to 1 year) is associated with faster recovery and improved long-term outcomes.	4 studies (unclear study design)	Grade: B[e]
	The performance of surgical decompression is suggested to provide better medium-term (1 to 4 years) symptom relief as compared with medical/interventional management of patients with radiculopathy from lumbar disc herniation whose symptoms are severe enough to warrant surgery.	3 RCTs 1 prospective comparative cohort study	Grade: B[e]
	Surgical decompression provides long-term (greater than four years) symptom relief for patients with radiculopathy from lumbar disc herniation whose symptoms warrant surgery. It should be noted that a substantial portion (23-28%) of patients will have chronic back or leg pain.	1 retrospective comparative cohort study 5 retrospective case series	Level of Evidence: IV[e]

(continued)

Table 48. Clinical practice guidelines related to lumbar radiculopathy or herniated intervertebral lumbar disc (continued)

Organization *Guideline Title (Year)* Guideline Quality[a]	Recommendation[b]	Evidence Base	Rating/Strength of Evidence Narrative Assessment
North American Spine Society (continued)	When surgery is indicated, performance of sequestrectomy or aggressive discectomy is recommended for decompression in patients with lumbar disc herniation with radiculopathy since there is no difference in rates of reherniation.	1 RCT 1 prospective comparative cohort study	Grade: B[e]
	Use of an operative microscope is suggested to obtain comparable outcomes to open discectomy for patients with lumbar disc herniation with radiculopathy whose symptoms warrant surgery.	2 RCTs	Grade: B[e]
	Endoscopic percutaneous discectomy is suggested for carefully selected patients to reduce early postoperative disability and reduce opioid use compared with open discectomy in the treatment of patients with lumbar disc herniation with radiculopathy.	3 RCTs	Grade: B[e]
	Endoscopic percutaneous discectomy may be considered for the treatment of lumbar disc herniation with radiculopathy.	3 RCTs 4 retrospective case series	Grade: C[e]
	Automated percutaneous discectomy may be considered for the treatment of lumbar disc herniation with radiculopathy.	2 RCTs 4 prospective case series	Grade: C[e]
	In a select group of patients automated percutaneous lumbar discectomy (APLD) may achieve equivalent results to open discectomy, however, this equivalence is not felt to be generalizable to all patients with lumbar disc herniation with radiculopathy whose symptoms warrant surgery.	3 RCTs	Level of Evidence: II/III[e]
	There is insufficient evidence to make a recommendation for or against the following: Urgent surgery for patients with motor deficits Use of spinal manipulation as an alternative to discectomy The specific surgical approach for far lateral disc herniation Use of tubular discectomy compared with open discectomy Use of medial facetectomy with discectomy Use of fusion for specific patient populations with lumbar disc herniation and radiculopathy	--	Grade: I[e]

(continued)

Table 48. Clinical practice guidelines related to lumbar radiculopathy or herniated intervertebral lumbar disc (continued)

Organization *Guideline Title (Year)* Guideline Quality[a]	Recommendation[b]	Evidence Base	Rating/Strength of Evidence Narrative Assessment
North American Spine Society (continued)	Use of percutaneous electrothermal disc decompression Use of intradiscal high-pressure saline injection Use of automated percutaneous discectomy compared with open discectomy Use of plasma disc decompression/nucleoplasty Use of plasma disc decompression as compared with transforaminal epidural steroid injections in patients with lumbar disc herniation who have previously failed transforaminal epidural steroid injection therapy		
American Pain Society *Interventional Therapies, Surgery, and Interdisciplinary Rehabilitation for Low Back Pain (2009)*[55] Quality Rating: 5 out of 7	Open discectomy or microdiscectomy for radiculopathy with prolapsed disc.	4 RCTs comparing surgery to conservative management	Level B/Good[f] Moderate net benefit for short-term outcomes (up to 12w) only
	Insufficient evidence for determining superiority of open vs. micro approaches.		
	Insufficient evidence to evaluate alternative surgical methods, including laser- or endoscopic-assisted techniques."		
American Society of Interventional Pain Physicians *An Update of Comprehensive Evidence-Based Guidelines for Interventional Techniques in Chronic Spinal Pain (2013)*[21,56] Quality Rating: 4 out of 7	For lumbar disc prolapse, protrusion, and extrusion: automated percutaneous lumbar decompression (APLD), percutaneous lumbar disc decompression (PLDD), and mechanical decompression with nucleoplasty are recommended in select cases.	19 observational studies for APLD. 15 observational studies for laser-assisted PLDD 1 SR of 3 observational studies PLDD with DeKompressor. 1 RCT and 14 observational studies for nucleoplasty.	The evidence is limited for APLD, PLDD, and percutaneous disc decompression with DeKompressor. The evidence is limited to fair for mechanical lumbar disc decompression with nucleoplasty.

(continued)

Table 48. Clinical practice guidelines related to lumbar radiculopathy or herniated intervertebral lumbar disc (continued)

Organization *Guideline Title (Year)* Guideline Quality[a]	Recommendation[b]	Evidence Base	Rating/Strength of Evidence Narrative Assessment
American College of Occupational and Environmental Medicine *Low back disorders. In occupational medicine practice guidelines: evaluation and management of common health problems and functional recovery in workers (2016)*[54] Quality Rating: Unknown[g]	Patients with evidence of specific nerve root compromise confirmed by appropriate imaging studies may be expected to potentially benefit from surgery. Quality evidence indicates that patient outcomes are not adversely affected by delaying nonemergent surgery for weeks or a few months and continued conservative care is encouraged in patients with stable or improving deficits who desire to avoid surgery. However, patients with either moderate to severe neurological deficits that are not improving or trending to improvement at 4 to 6 weeks may benefit from earlier surgical intervention. Those with progressive neurological deficit(s) are believed to have indications for immediate surgery. Those with severe deficits that do not rapidly improve are also candidates for earlier testing and referrals.	Unknown[g]	Unknown[g]
National Institute for Health and Care Excellence (United Kingdom) *Percutaneous transforaminal endoscopic lumbar discectomy for sciatica: Interventional procedures guidance [IPG 556] (2016)]*[58] Quality Rating: 2 out of 7	Current evidence on the safety and efficacy of percutaneous transforaminal endoscopic lumbar discectomy for sciatica is adequate to support the use of this procedure provided that standard arrangements are in place for clinical governance, consent and audit. Percutaneous transforaminal endoscopic lumbar discectomy for sciatica is a procedure that needs particular experience. Surgeons should acquire the necessary expertise through specific training and mentoring. It should only be done by surgeons who do the procedure regularly.	1 SR of observational studies 1 retrospective comparative cohort study 2 prospective case series 5 retrospective case series	None provided

(continued)

Table 48. Clinical practice guidelines related to lumbar radiculopathy or herniated intervertebral lumbar disc (continued)

Organization *Guideline Title (Year)* Guideline Quality[a]	Recommendation[b]	Evidence Base	Rating/Strength of Evidence Narrative Assessment
National Institute for Health and Care Excellence (United Kingdom) *Percutaneous interlaminar endoscopic lumbar discectomy for sciatica: Interventional procedures guidance[IPG555](2016)*[59] Quality Rating: 2 out of 7	Current evidence on the safety and efficacy of percutaneous interlaminar endoscopic lumbar discectomy for sciatica is adequate to support the use of this procedure provided that standard arrangements are in place for clinical governance, consent and audit. Percutaneous interlaminar endoscopic lumbar discectomy for sciatica is a procedure that needs particular experience. Surgeons should acquire the necessary expertise through specific training and mentoring. It should only be done by surgeons who do the procedure regularly.	2 RCTs 2 retrospective comparative cohort studies 4 retrospective case series	None provided
National Institute for Health and Care Excellence (United Kingdom) *Percutaneous coblation of the intervertebral disc for low back pain and sciatica Interventional procedures guidance[IPG543](2016)*[60] Quality Rating: 2 out of 7	Current evidence on percutaneous coblation of the intervertebral disc for low back pain and sciatica raises no major safety concerns. The evidence on efficacy is adequate and includes large numbers of patients with appropriate follow-up periods. Therefore, this procedure may be used provided that normal arrangements are in place for clinical governance, consent and audit. As part of the consent process, patients should be informed that there is a range of treatment options available to them and that further procedures may be needed.	1 SR 2 RCTs 1 case series	None provided
National Institute for Health and Care Excellence (United Kingdom) Percutaneous electrothermal treatment of the intervertebral disc annulus for low back pain and sciatica Interventional procedures guidance[IPG544](2016)[61] Quality Rating: 2 out of 7	Current evidence on percutaneous electrothermal treatment of the intervertebral disc annulus for low back pain and sciatica raises no major safety concerns. The evidence on efficacy is inconsistent and of poor quality. Therefore, this procedure should only be used with special arrangements for clinical governance, consent and audit or research.	1 SR 1 RCT 1 Cohort study	None provided

(continued)

Table 48. Clinical practice guidelines related to lumbar radiculopathy or herniated intervertebral lumbar disc (continued)

Organization *Guideline Title (Year)* Guideline Quality[a]	Recommendation[b]	Evidence Base	Rating/Strength of Evidence Narrative Assessment
National Institute for Health and Care Excellence (United Kingdom) *Percutaneous intradiscal radiofrequency treatment of the intervertebral disc nucleus for low back pain. Interventional procedures guidance[IPG545] (2016)*[62] Quality Rating: 2 out of 7	Current evidence on percutaneous intradiscal radiofrequency treatment of the intervertebral disc nucleus for low back pain raises no major safety concerns. The evidence on its efficacy is limited in quantity and quality. Therefore, this procedure should only be used with special arrangements for clinical governance, consent and audit or research.	1 RCT 1 nonrandomized CT 2 case series	None provided
National Institute for Health and Care Excellence (United Kingdom) *Epiduroscopic lumbar discectomy through the sacral hiatus for sciatica Interventional procedures guidance[IPG570] (2016)*[63] Quality Rating: 2 out of 7	Current evidence on the safety and efficacy of epiduroscopic lumbar discectomy through the sacral hiatus for sciatica is limited in quantity and quality. Therefore, this procedure should only be used in the context of research.	1 Cohort study	None provided
National Institute for Health and Care Excellence (United Kingdom) *Percutaneous intradiscal laser ablation in the lumbar spine. Interventional procedures guidance[IPG357] (2010)*[64] Quality Rating: 2 out of 7	Current evidence on the safety and efficacy of percutaneous intradiscal laser ablation in the lumbar spine is adequate to support the use of this procedure provided that normal arrangements are in place for clinical governance, consent and audit. Patients selected for the procedure should be limited to those with severe pain refractory to conservative treatment, in whom imaging studies show bulging of an intact disc, and who do not have neurological deficit requiring surgical decompression.	1 RCT 2 Cohort studies 2 Case series	None provided

(continued)

Table 48. Clinical practice guidelines related to lumbar radiculopathy or herniated intervertebral lumbar disc (continued)

Organization *Guideline Title (Year)* Guideline Quality[a]	Recommendation[b]	Evidence Base	Rating/Strength of Evidence Narrative Assessment
National Institute for Health and Care Excellence (United Kingdom) *Automated percutaneous mechanical lumbar discectomy: Interventional procedures guidance[IPG141])(2005)*[65] Quality Rating: 2 out of 7	Current evidence suggests that there are no major safety concerns associated with automated percutaneous mechanical lumbar discectomy. There is limited evidence of efficacy based on uncontrolled case series of heterogeneous groups of patients, but evidence from small randomized controlled trials shows conflicting results. In view of the uncertainties about the efficacy of the procedure, it should not be used without special arrangements for consent and for audit or research. Clinicians wishing to undertake automated percutaneous mechanical lumbar discectomy should take the following actions. Inform the clinical governance leads in their Trusts. Ensure that patients understand the uncertainty about the procedure's efficacy and provide them with clear written information. In addition, use of the Institute's information for the public is recommended. Audit and review clinical outcomes of all patients having automated mechanical percutaneous lumbar discectomy.	3 RCTs 5 case series	None provided
National Institute for Health and Care Excellence (United Kingdom) *Endoscopic laser foraminoplasty. Interventional procedures guidance[IPG31] (2003)*[66] Quality Rating: 2 out of 7	Current evidence of the safety and efficacy of endoscopic laser foraminoplasty does not appear adequate to support the use of this procedure without special arrangements for consent and for audit or research. Clinicians wishing to undertake endoscopic laser foraminoplasty should inform the clinical governance leads in their Trusts. They should ensure that patients offered the procedure understand the uncertainty about its safety and efficacy and should provide them with clear written information. Use of the Institute's information for the public is recommended. Clinicians should ensure that appropriate arrangements are in place for audit or research. Further research into safety and efficacy outcomes will be useful in reducing the current uncertainty. NICE is not undertaking further investigation at present.	3 Cohort studies 2 Case series	None provided

[a] We assessed the quality of guideline using the Appraisal of Guidelines For Research & Evaluation II (AGREE II) Instrument, version 2017.[21] The lowest quality score possible is 1, the highest possible quality score is 7.

[b] Only recommendations from the guideline pertinent to surgical interventions for lumbar radiculopathy are summarized.

[c] One included trial was for treatment of sciatica with spinal stenosis, the rest were for treatment of lumbar radiculopathy

[d] Based on GRADE.

[e] Level 1=high quality RCTs or SRs of RCTs; Level II=lesser quality RCTs, prospective comparative studies, SRs that include Level II studies; Level III=Case control or retrospective cohort studies, SRs of Level III studies, Level 4=case series; Level 5= Expert Opinion, Grade A=Good evidence (Level 1 studies with consistent findings); Grade B=Fair evidence (Level II or III studies with consistent findings), Grade C=Poor evidence (Level IV or V studies); Grade I=insufficient or conflicting evidence not allowing a recommendation

[f] One included trial was for treatment of sciatica with spinal stenosis, the rest were for treatment of lumbar radiculopathy

[g] The complete guideline is not publicly accessible; thus, a full quality appraisal and summary of the evidence base and strength of evidence ratings were not possible.

Abbreviations: RCT = randomized controlled trial; SR = systematic review; CT = controlled trial; w = week(s); y = year(s); APLD = automated percutaneous lumbar decompression; PLDD = percutaneous lumbar disc decompression.

Overall, the guidelines we identified were in general agreement about considering discectomy or microdiscectomy (and related decompressive procedures) as acceptable treatment based on evidence that it improves outcomes in the short- to medium-term. One guideline specifies that this surgery can be considered when symptoms have not improved with conservative therapy.[45] Another guideline suggests that conservative therapy is reasonable for patients with nonprogressive symptoms who wish to delay surgery.[54] The guideline recommendations relating to minimally-invasive spine surgery varied; one did not consider these specific procedures within their scope.[45] Three of the guidelines were developed 5 or more years ago; thus may not include the most recent evidence for these procedures.[55-57]

4. Discussion

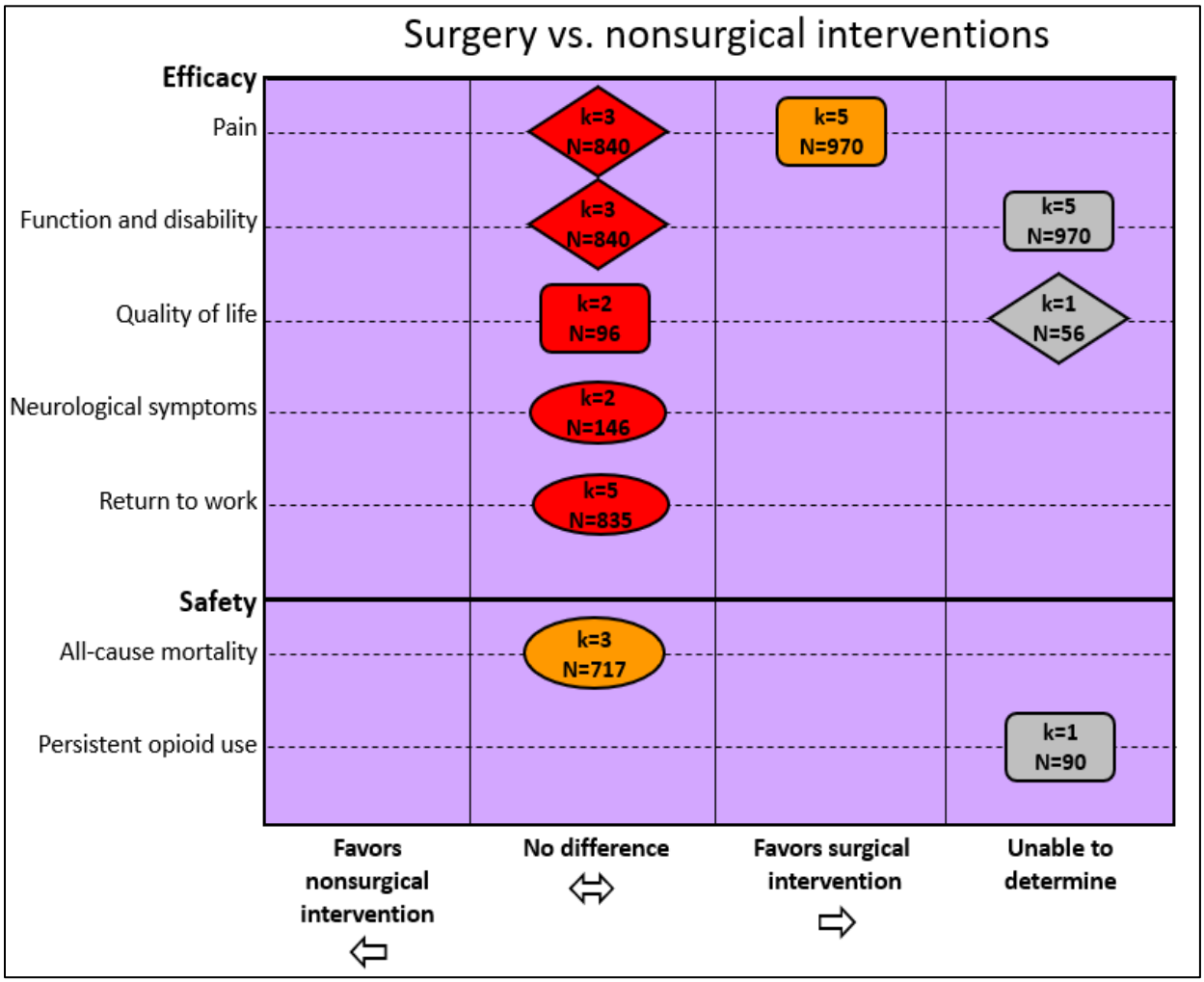
4.1 Summary of the Evidence

Evidence maps summarizing the overall findings and strength of evidence are provided in ***Figures 3, 4, 5, 6,*** and ***7***. On these maps, we place a shape that indicates the number of studies and participants contributing to each of the major efficacy and safety domains (e.g., pain, morbidity). On these maps, we place the shape according to whether the evidence shows a meaningful difference favoring surgery or favoring nonsurgical intervention, or whether no difference was observed. We use different shapes when outcomes vary at different follow-up time points. The strength of evidence for each major outcome domain is conveyed by shading the shape various colors; gray indicates insufficient evidence, red indicates very low certainty, orange indicates low certainty, yellow indicates moderate certainty, and green indicates high certainty. With few exceptions, most findings that we considered sufficient for assessment were based on evidence graded as low to very low certainty, primarily because of some or high concerns for bias among included studies and imprecision in study estimates. Most outcomes assessed as having insufficient evidence were single study bodies of evidence, but some were assessed as insufficient because of mixed findings. See Section 4.2 for further discussion of the limitations of this evidence base.

4.1.1 Surgery compared to nonsurgical interventions

Surgery reduces pain more than nonsurgical interventions in the short and medium-term (up to between 6 weeks and 26 weeks), but this difference does not persist in the long-term (***Figure 3***). Several explanations for this are possible. One explanation for the mitigation of benefits observed is that the impact of participants that crossover between groups accumulates over time. For example, Peul et al.[32] reports that of the 142 participants allocated to conservative management, 55 (39%) underwent surgery during the first year after a median of 14.6 weeks, 62 (44%) underwent surgery by 2 years, and 66 (46%) by 5 years. In an intent-to-treat analysis, any treatment effects that might exist are mitigated by these crossovers. See Section 4.2.1 for further discussion of this issue. Another explanation is that long-term outcomes simply reflect the natural history of radiculopathy, particularly radiculopathy that results from disc herniation. Observational studies have suggested that radicular symptoms improve in nearly three-quarters of patients by 12 weeks.[73] Thus, the majority of individuals with symptoms will improve in the long-term regardless of treatment provided, and one would not expect to see between-group differences in outcomes from treatment over longer periods of follow-up.

Figure 3. Evidence map of surgery compared with nonsurgical interventions for treatment of symptomatic lumbar radiculopathy



Note: Outcomes related to surgical mortality (k=5), surgical morbidity (k=6), and reoperations (k=5) were synthesized for the surgical intervention group only as they are not appropriate for comparative evaluation with a nonsurgical intervention group. See Section 3.3.2.1, 3.3.2.2, and 3.3.2.3 for details.

The evidence was insufficient to assess short- and medium-term impact on function because of inconsistent findings across studies, but long-term impact on function suggests no difference between treatments. The impact on other outcomes including quality of life, neurological symptoms, and return to work also found no meaningful differences between treatment groups in the short-, medium-, or long-term.

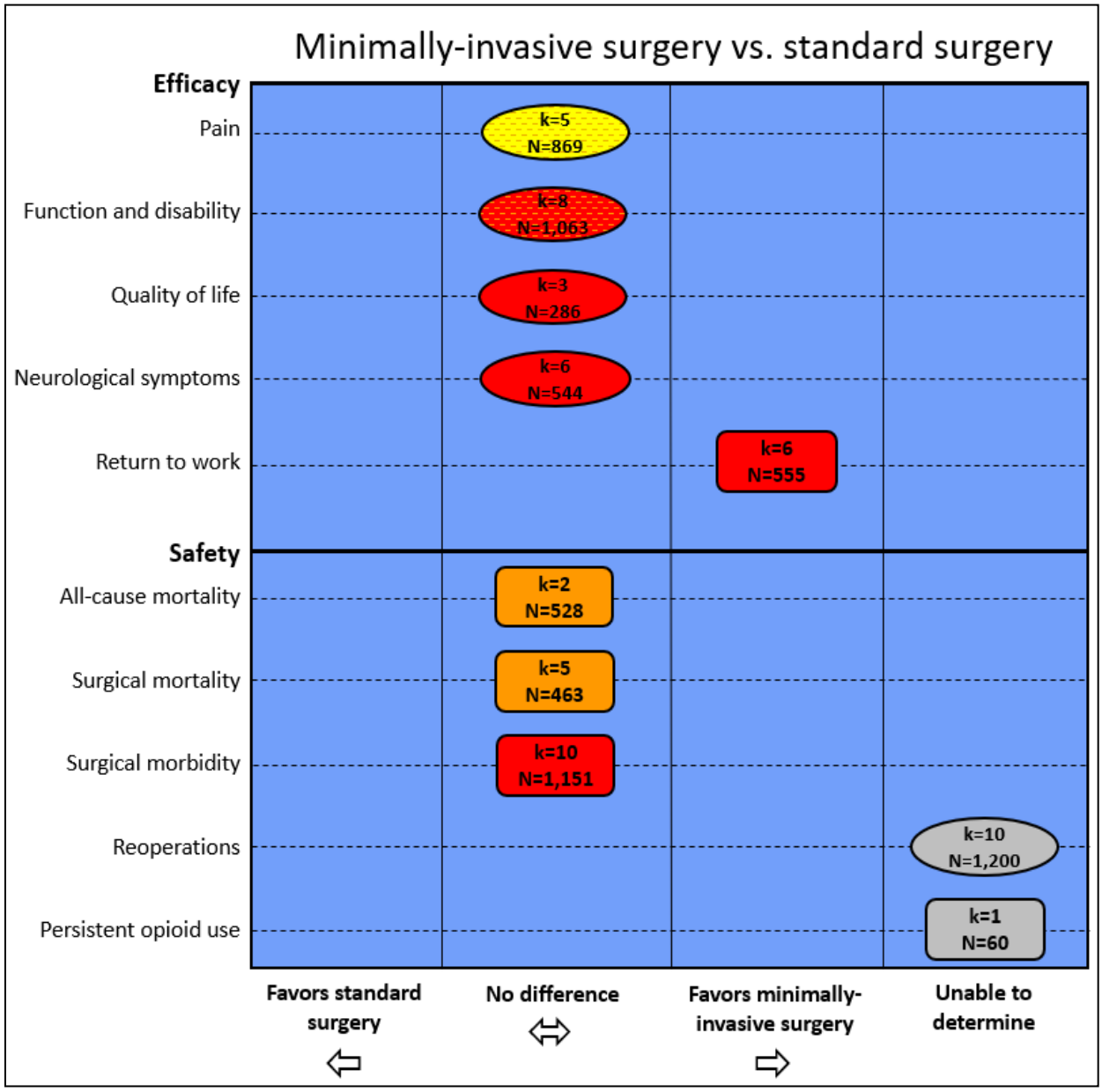
No surgery-related deaths were observed and surgery-related complications were rare, but these findings may not be applicable to community practice where enrolled participants may have more comorbidities than participants enrolled in RCTs. As might be expected, no difference in all-cause mortality was observed. The evidence was insufficient to assess outcomes related to persistent opioid use because of a single-study body of evidence.

We concluded with very low certainty that surgery compared with nonsurgical interventions may be cost-effective depending on a decision-makers willingness to pay threshold. In this HTA, the cost per QALY gained ranged from $51,156 to $83,322 in 2010 U.S. dollars from a healthcare payor perspective. Although no definitive consensus exists, costs per QALY gained of less than $50,000 are generally considered cost-effective, costs between $50,000 and $150,000 are considered of intermediate value, and costs more than $150,000 per QALY gained are considered low value, though we note these thresholds are typically applied to costs from a societal perspective.[113,114] For comparative purposes, the cost per QALY gained (from a payor perspective) of an implantable cardioverter defibrillator (ICD) compared to standard care to prevent sudden cardiac death in patients with left ventricular systolic dysfunction ranges from $37,962 to $78,380 in 2010 U.S. dollars across the six RCTs that have demonstrated mortality reduction from ICDs.[115] And, the cost per QALY gained (from a societal perspective) of biennial screening mammography among women ages 50 to 74 not at increased risk for breast cancer with average/low breast density ranges from $112,380 to $214,348 in 2010 U.S. dollars depending on model used.[116]

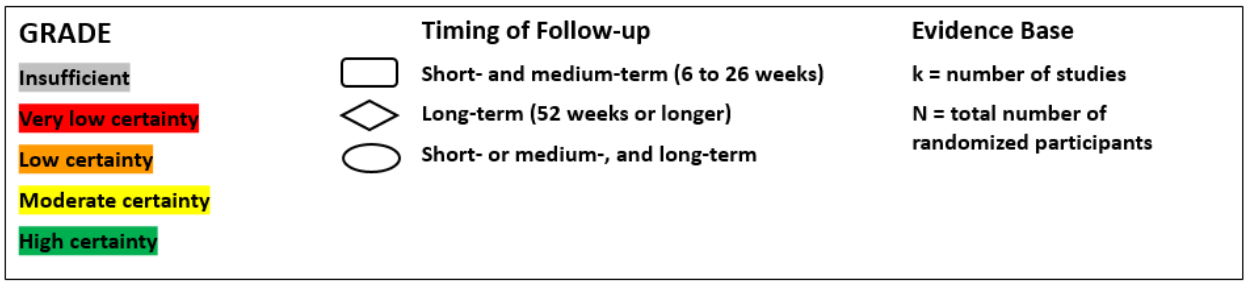
4.1.2 Minimally-invasive surgery compared to standard surgery

For the purposes of this HTA, we synthesized procedures using an endoscopic or percutaneous approach under the broad term “minimally-invasive surgical procedures”, however, these procedures may represent a heterogenous set of interventions. This limitation is described further in Section 4.2.4. With few exceptions, minimally-invasive surgical interventions and standard surgery similarly reduce pain and improve function (***Figure 4***). However, minimally-invasive surgery seems to result in a quicker return to work, though this finding should be interpreted with caution because of the varying definitions of return to work used by studies, differences in work culture between U.S. and European countries, and because the advice given to participants as to when to return to work may be in part based on the procedure they received.

Figure 4. Evidence map of minimally-invasive surgery compared with discectomy or microdiscectomy for treatment of symptomatic lumbar radiculopathy



Note: Pattern-filled outcomes indicate that multiple measures within the outcome were reported but graded as having different levels of certainty.



No surgical deaths were reported and surgical morbidity was similar between both approaches. Although 10 studies reported on the incidence of reoperations, the evidence was insufficient to draw a definitive conclusion because of mixed findings and imprecision in estimates. The evidence for persistent opioid use outcomes was also insufficient because of a single-study body of evidence. The evidence on cost-effectiveness for minimally-invasive surgery compared to standard approaches was also insufficient; further none of the cost analyses were conducted in the U.S.

4.1.3 Microdiscectomy compared to discectomy

Microdiscectomy and discectomy were comparable with respect to pain, surgical morbidity and incidence of reoperations (***Figure 5***). However, the evidence was insufficient to draw conclusions about other outcomes because no studies reported these outcomes (neurological symptoms, persistent opioid use) or these outcomes were reported by only a single-study (function/disability, quality of life, return to work, surgical mortality). The evidence was also insufficient for drawing conclusions about cost or cost-effectiveness.

4.1.4 Repeat surgery for recurrent radiculopathy

Only two RCTs reported on repeat surgery for recurrent radiculopathy, and both used different comparator groups resulting in a single-study body of evidence for each comparison (***Figure 6***, ***Figure 7***). Thus, the evidence was insufficient to draw any conclusions about efficacy, safety, or costs.

Figure 5. Evidence map of microdiscectomy compared with discectomy for treatment of symptomatic lumbar radiculopathy

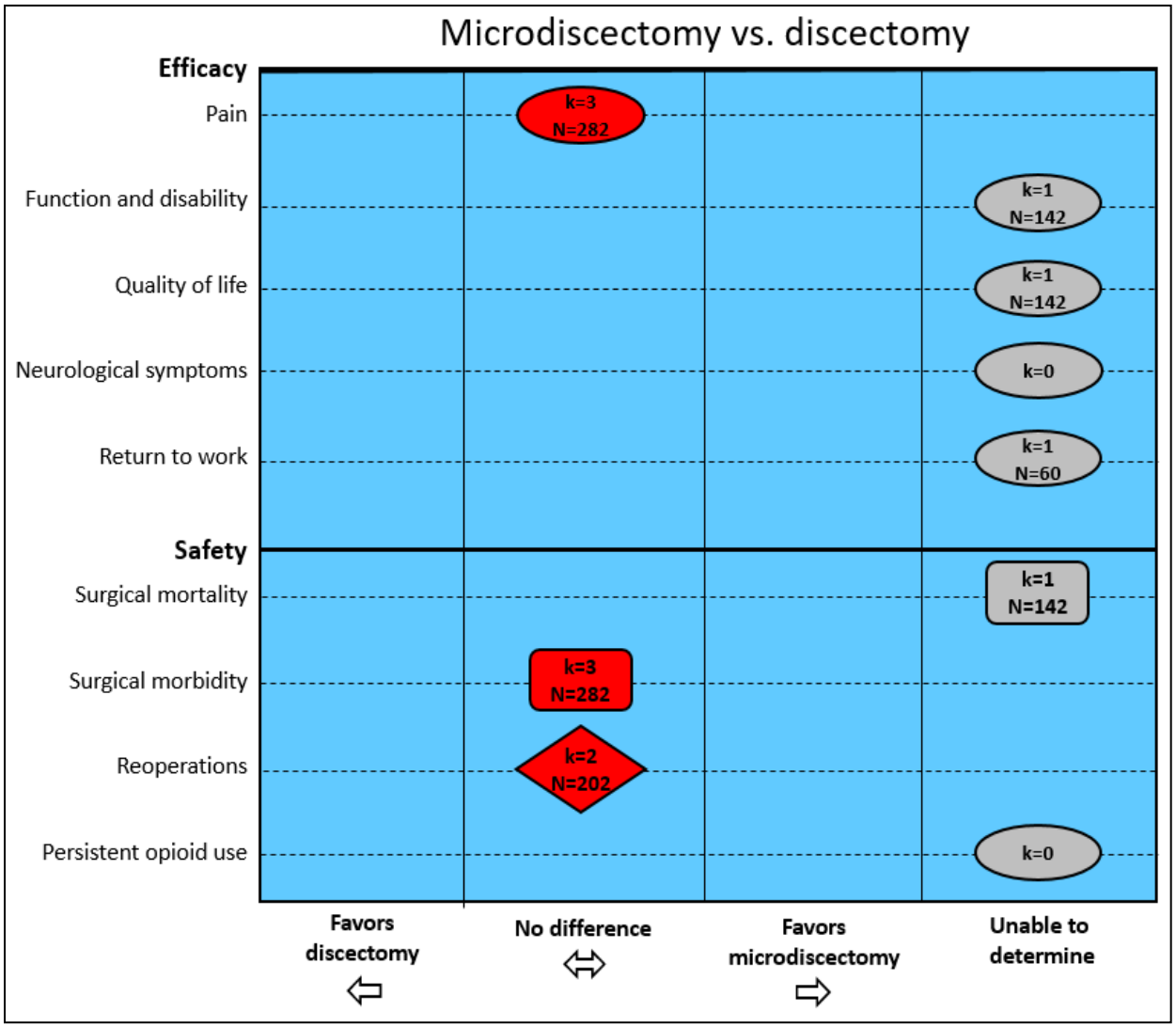


Figure 6. Evidence map of repeat lumbosacral decompression compared with spinal cord stimulation for treatment of recurrent symptomatic lumbar radiculopathy

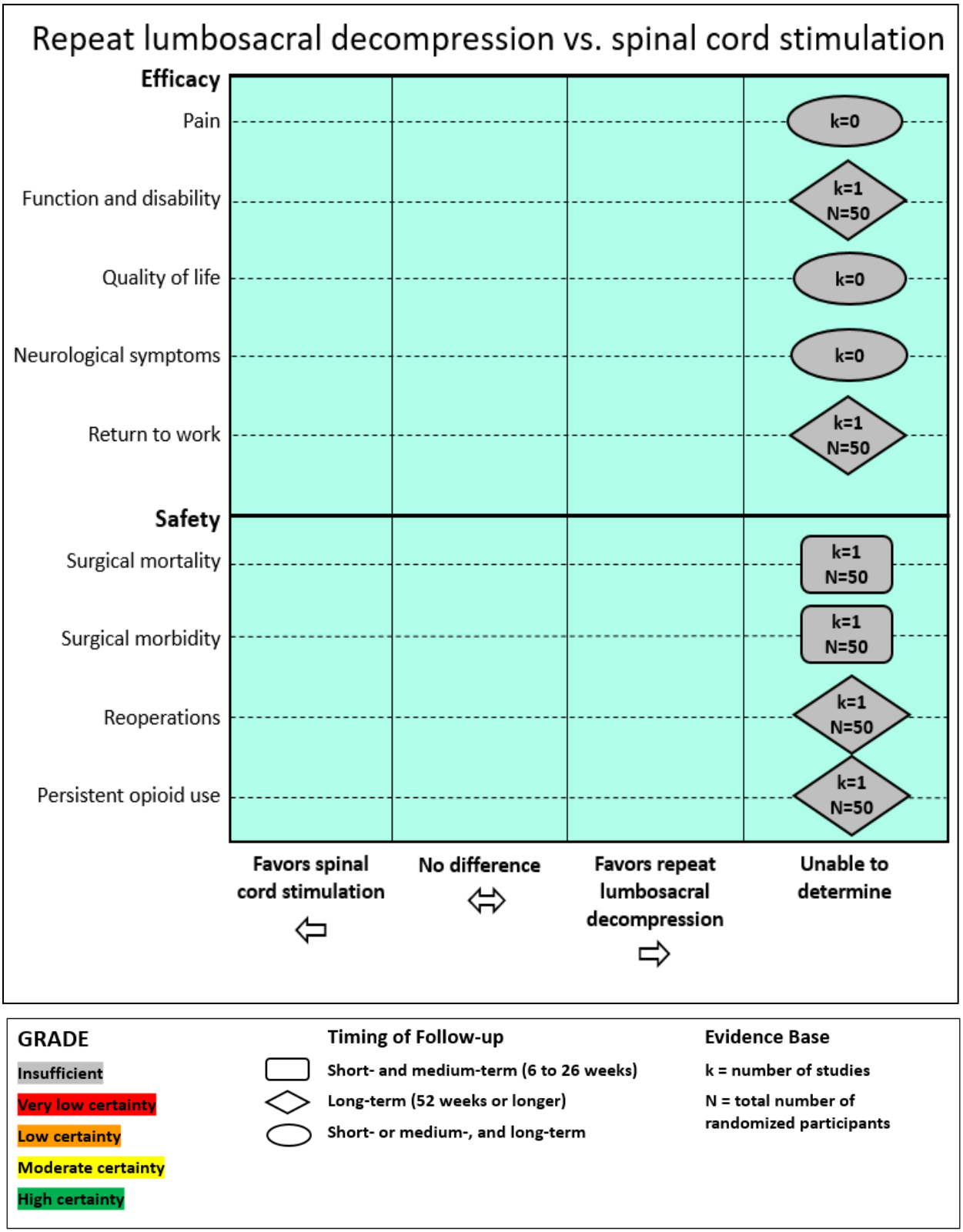
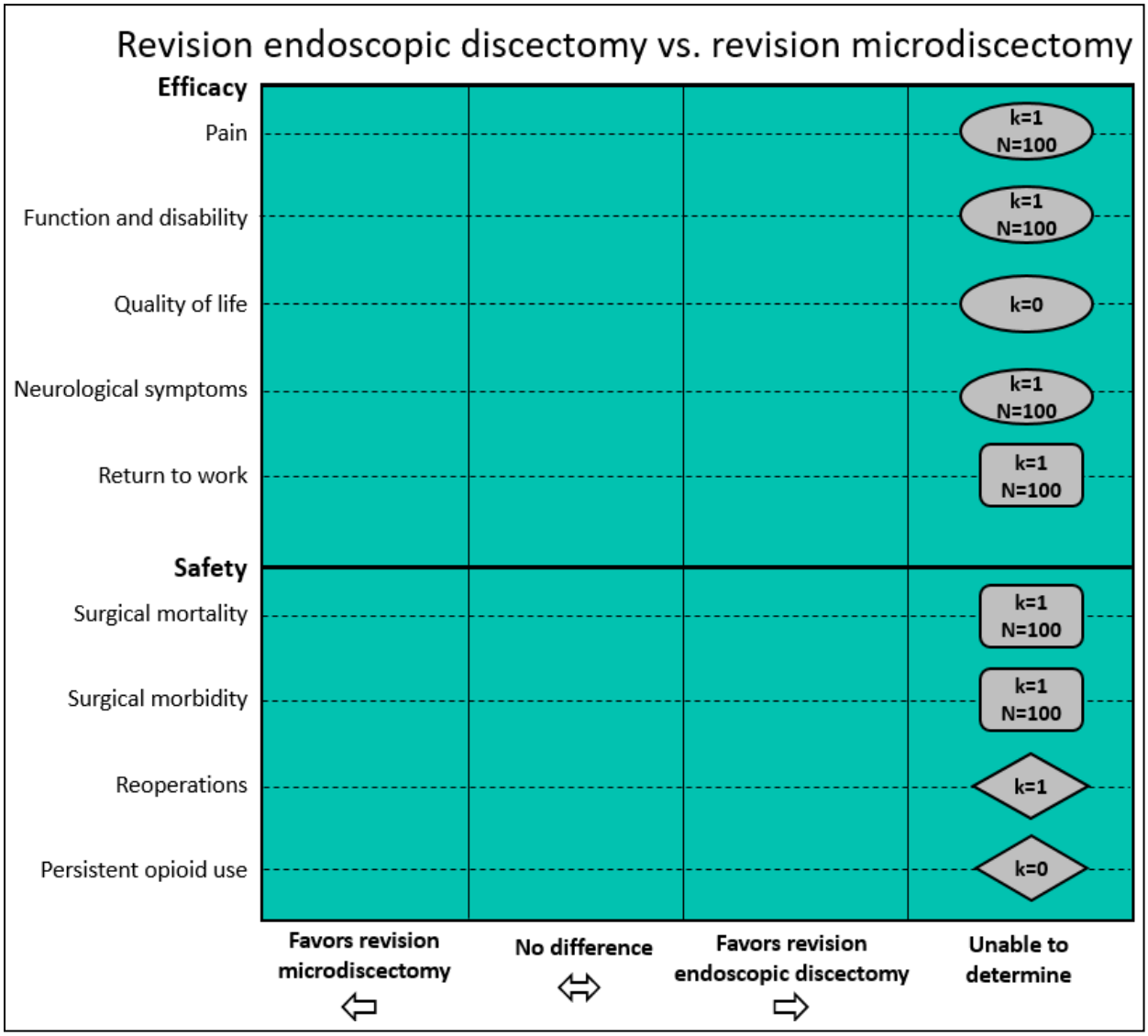
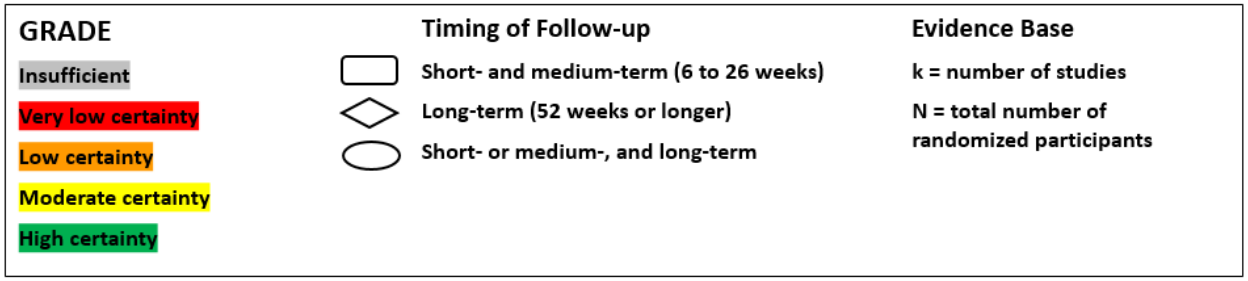


Figure 7. Evidence map of revision endoscopic discectomy compared with revision microdiscectomy for treatment of recurrent symptomatic lumbar radiculopathy





4.2 Limitations of the Evidence Base

The primary research study and clinical practice guideline evidence we identified for inclusion in this HTA has several limitations.

4.2.1 High risk of bias among included studies

We rated nearly half of included studies as having a high risk of bias. Some sources of bias across included studies were common; for example, all but one study did not blind participants, caregivers, or clinicians to treatment allocation and most did not blind outcome assessors. Knowledge of treatment allocation has the potential to influence other decisions about care that may be related to the outcome, and to influence the outcome assessment itself. Because most all studies use patient-reported outcomes, this introduces some concerns for bias. Although blinding treatment allocation can be very challenging to perform in trials of surgery, particularly those comparing surgical interventions to nonsurgical interventions, the risk of bias nonetheless remains and should be acknowledged. The direction of bias from nonblinding largely depends on the beliefs and attitudes of participants, clinicians, and outcome assessors, so cannot always be predicted. The sources of bias contributing to high risk of bias ratings for a few studies was inadequate randomization (e.g., using even/odd sequences) or inadequate allocation concealment.

Deviations from the intended intervention because of crossovers and contamination were the source of bias contributing to high risk of bias ratings for some studies, particularly for studies comparing surgery to non-surgical interventions. In an intent-to treat analysis, which preserves randomization and thus mitigates other sources of bias, the direction of bias from crossovers is predictable and results in bias toward a null effect. Thus, estimates from the intent-to-treat analysis are conservative, and may underestimate the effect when deviations from the intended intervention occur. This is evident with respect to pain and functional outcomes for surgery compared to nonsurgical interventions; the evidence favored surgery in the short- and medium-term, but these differences in the long-term may have been mitigated by cumulative crossovers that occurred over time. Because of extensive crossovers, several included studies also reported 'as treated' analyses. Participants are included in these analyses according to the treatment they received (as opposed to the treatment to which they were randomly allocated). Weinstein et al. [SPORT] reported an as-treated analysis in addition to the intent-to-treat analysis and found favorable effects for discectomy/microdiscectomy compared with conservative management through 2 years of follow-up.[22] The between-group difference at 52 weeks for the SF-36 Bodily Pain subscale was 15.0 (95 % CI, 10.9 to 19.2), the SF-36 Physical Functioning subscale 17.5 (95% CI, 13.6 to 21.5), and the Oswestry Disability Index -15.0 (95% CI, -18.3 to -11.7). Although as-treated analyses may offer some insight into the magnitude of bias toward the null effect, participants generally do not cross over at random but for reasons that are also related to the outcomes (e.g., pain, symptoms); thus, these analyses can introduce other biases on the effect estimate.

Most studies reporting outcomes at time points longer than 2 years reported high attrition at these longer-term follow-ups. For example, Arts et al.[40,48] reported 5 year outcomes for only 63.5% of participants randomized to tubular discectomy and 60.9% of participants randomized to

discectomy. Similarly, Weinstein et al. [SPORT][22,100]reported 8 year outcomes for 64.1% of participants randomized to discectomy/microdiscectomy compared and 59.4% of participants randomized to conservative management. Some studies also had high attrition at short- or medium term follow-up,[41,42] and some studies did not provide any information about missing data or participants lost to follow-up.

4.2.2 Studies generally underpowered for many outcomes of interest

Only 11 of the 24 included RCTs for efficacy and safety designated a primary outcome and described the sample size required to detect an a priori effect size. Few described how this effect size was determined or whether it represented a minimally important clinical difference and whether the analysis was designed for detecting superiority or noninferiority. Eight studies were powered based on pain or function outcomes, three studies were powered based on duration of surgery or hospital stay, and one study was powered based on differences in success. Consequently, study samples were not adequate to detect many of the efficacy outcomes that were reported, or safety outcomes, which occurred at a low frequency. Thus, effect estimates were often imprecise, which resulted in downgrading the strength of evidence ratings from moderate to low, or from low to very low.

4.2.3 Variation in diagnosis and severity of symptoms

Most studies required participants to have a clinical diagnosis of radiculopathy with disc herniation or nerve root compression confirmed by imaging (usually CT or MRI) for enrollment. However, few studies described the criteria for clinical diagnosis. Further, the duration of symptoms and criteria related to provision of conservative therapy prior to enrollment was variable across studies. The duration between enrollment and receiving surgery was also variable, in some cases months.

Based on initial direct from stakeholders at the scoping stage of this HTA, we focused study selection criteria on populations with radiculopathy but without symptoms of neurogenic claudication and spinal stenosis. This resulted in an evidence base primarily focused on populations with disc protrusion or herniation as the etiology of the radiculopathy and populations with radiculopathy in the absence of disc herniation are not represented here. However, when we examine the studies included in the 2016 U.K. National Institute for Clinical Excellence Low Back Pain and Sciatica Evidence Review,[45] which allowed for studies that enrolled participants with sciatica and neurogenic claudication or other symptoms of spinal stenosis to be included, we note that all included RCTs that were identified were focused surgical decompressive procedures among populations with disc herniation, and not neurogenic claudication or other symptoms of central canal spinal stenosis.

4.2.4 Limited number of comparative effectiveness trials for any one procedure

We identified 15 trials comparing minimally-invasive surgery to open surgery. However, most of these interventions were only evaluated by 1 to 3 RCTs and variations in the outcomes reported limited our ability to draw conclusions for any one specific minimally-invasive procedure. Further, variation in nomenclature for these procedures and lack of detail regarding the procedures may have also limited out ability to synthesize findings for specific interventions.

Many studies lacked a full description of the surgical intervention, including the procedure, the skill experience of the surgeon and surgical team, and pre- and postoperative care.

4.2.5 Variation in type, timing, and completeness in reporting outcomes

Some studies reported between-group differences at multiple follow-up time points without a priori specification of a primary time point; others more appropriately used repeated measures analysis, to account for multiple observations over time, and some reported both. Our ability to conduct quantitative syntheses (i.e., meta-analysis) was limited by variation in specific outcomes reported and by incomplete reporting. For efficacy outcomes, studies used a variety of pain and function measures, and measured outcomes at different follow-up time points. Further, some studies only reported adjusted difference-in-difference treatment effects, while other studies reported values only at follow-up, not adjusted for baseline values. Some studies only reported short-term outcomes, others only reported long-term outcomes. Some studies reported results of statistical significance testing for between-group differences, but did not report actual outcome values. Others reported actual outcome values, but no measures of variance. Safety outcomes reported were very heterogenous, particularly with respect to 'reoperations' and 'persistent opioid use'. For example, some studies reported 'all-cause' reoperations, some studies only report reoperations for technical failures, and some studies did not define or explain how reoperations were defined. Combined, these data reporting issues limited the extent to which we could conduct quantitative synthesis.

"Return to work" outcomes were particularly challenging to interpret in this evidence report. First, measures used were varied across studies. Second, the work culture in Europe, where most of the RCTs took place, is quite different from the work culture in the U.S. Finally, return to work outcomes for minimally-invasive surgery compared with standard surgery are particularly challenging to interpret because the advice given to participants as to when to return to work may be in part based on the procedure they received, given that nearly all studies did not blind providers or participants to the procedure received.

4.2.6 Applicability of older studies and RCTs to community practice

Six RCTs were conducted prior to the year 2000.[26,27,34,35,38,43,44] Changes in surgical technique and pre- and post-operative care may limit the applicability of findings from these older studies to current practice. Further, by limiting included studies to trials, we may have underestimated differences in safety outcomes as participants in trials may have fewer comorbidities than individuals within the general population. For example, in a retrospective analysis of discharge data from 1997 to 2007, Martin et al. report ta 14.7% incidence of reoperations at 4 years following lumbar decompression for herniated disc among nonfederal hospitals in Washington state.[117] Similarly, an analysis conducted using South Korean national database of patients who underwent surgery for herniated disc in 2003 reported a cumulative incidence of reoperation at 5 years of 13.7% for open discectomy and 12.4% for endoscopic discectomy.[118]

4.2.7 Limited number of United States cost studies

Whereas efficacy and safety outcomes from studies conducted outside of the United States are likely applicable to U.S. settings, it is not clear cost studies conducted based on RCTs outside of

the United States would apply to U.S. settings. The only RCT conducted in the United States included in this HTA reported direct medical costs in the surgical group that were nearly triple the health care costs in any of the other cost studies that reported health care costs separately.[50] Although the effectiveness inputs from non-U.S. studies used in cost-effectiveness analyses are likely applicable, the extreme differences in how health care services are organized and financed between U.S. and non-U.S. countries probably reduces the applicability of the cost inputs used in non-U.S. studies.

4.2.8 Limitations in the AGREE guideline appraisal instrument

The AGREE guideline appraisal instrument largely focuses on evaluating the processes through which a guideline is developed; it does not assess how well the evidence included in the guideline was evaluated and interpreted correctly, or whether the conclusions of the guideline are consistent with the evidence. Thus, some guidelines may score artificially high and explains why conclusions may differ between guidelines despite having nearly similar quality scores.

4.3 Other related HTAs

The only related HTA that we identified was commissioned by the National Institute for Health Research (U.K.) Health Technology Assessment programme.[67] This HTA included both surgical and nonsurgical interventions for the management of sciatica and used a network meta-analysis to provide a measure of relative therapeutic effect across 18 different treatment categories. The findings suggest that nonopioid medication, epidural corticosteroids injections, and disc surgery are effective for reducing sciatica This HTA also concluded that stepped care approaches to treatment are cost-effective relative to direct referral for surgery.

4.4 Selected payer coverage policies

The Centers for Medicare and Medicaid Services (CMS) does not have a national coverage determination related to open standard or microsurgical decompressive procedures (i.e., discectomy, microdiscectomy, foraminotomy, laminectomy/otomy). With respect to the use of lasers, CMS recognizes their use to alter, revise, or destroy tissues in place of more conventional techniques as part of a surgical procedure. Medicare administrative contractors have been advised to use discretion in determining coverage for procedures performed with a laser when the laser has been FDA-cleared, the procedure is considered reasonable and necessary, and a noncoverage instruction does not exist (effective date May 1, 1997).[68] CMS does have a national coverage determination related to thermal intradiscal procedures; these procedures are not covered (effective date January 1, 2009).[68] Percutaneous disc decompression falls within the category of procedures covered by this determination. ***Table 49*** provides an overview of other payer coverage policies and ***Table 50*** summarizes excerpts from these policies that are relevant to surgery for lumbar radiculopathy and disc herniation.

In general, payers cover decompressive procedures, including discectomy, laminectomy/otomy, foraminectomy/otomy, including microsurgical approaches, for disc herniation with radicular symptoms. Specific criteria vary by payer but often include a failed trial of conservative management for 6 to 12 weeks. Most payers also require imaging confirmation of nerve root

compression that corresponds to symptoms and physical examination findings. The coverage of minimally-invasive procedures varies by payer.

Table 49. Overview of payer coverage policies

Procedure	Medicare	Premera	Regence	Cigna	United	Aetna	Humana	Kaiser
Laminectomy, laminotomy, discectomy, foraminotomy (open technique including microsurgical approaches)	--	✓[a]	--	--	--	✓[a]	✓[a]	--
Automated percutaneous lumbar disc decompression	✕[b]	✕	✕	✕	✕	✓[c]	✕	--
(Percutaneous) endoscopic discectomy	✕	✕	✕	✕	✕	—	No additional reimbursement.	--
(Percutaneous) laser discectomy	✕	✕	✕	✕	✕	No additional reimbursement.	✕	--
Percutaneous nucleoplasty with coblation technology	✕	✕	✕	✕	✕	—	--	--

✓ = covered; ✕ = not covered; — = no policy identified

[a] If specific clinical criteria are met. See Table 50 for details.

[b] All percutaneous disc decompression procedures fall under a Medicare National Coverage Determination related to thermal intradiscal procedures.

[c] Also covers percutaneous manual discectomy, see Table 50 for details.

Table 50. Selected payer coverage for surgery for lumbar radiculopathy

Payer; Effective Date	Policy
Premera (Blue Cross)[119-121] July 1, 2017	Premera may consider lumbar spine decompression surgery (discectomy, foraminotomy, laminotomy) medically necessary for the rapid (48 hours or less) progression of neurologic impairment (e.g., cauda equina syndrome, foot drop, extremity weakness, saddle anesthesia, sudden onset of bladder or bowel dysfunction); or in the absence of rapid progression when all the following criteria are met: • All other sources of low back pain have been ruled out AND • MRI or CT with myelogram within the past 12 months shows nerve root compression that corresponds to symptoms and physical examination findings or there is definitive neurological localization by other means AND • Persistent, debilitating pain radiating from the low back down to the lower extremity is present daily and limits activities of daily living AND • Neurological deficits (e.g., reflex change in the legs, dermatomal sensory loss, motor weakness) or alternative signs of lumbar root irritation (e.g., positive leg raising test) are present on physical examination AND • The patient has failed at least 6 weeks of conservative therapy such as activity modification, oral analgesics/anti-inflammatories, physical therapy, chiropractic manipulation, epidural steroid injections. July 1, 2017 Lumbar laminectomy may also be considered medically necessary for the rapid progression of neurologic impairment or when criteria related to the presence of lumbar spinal stenosis are met. July 1, 2017 Lumbar spine decompression surgery is considered not medically necessary when no clinical indication is documented and there are no confirmatory physical and radiologic findings that meet the relevant criteria listed above. The provider's choice of interventional surgery depends on the specific member's symptoms and imaging findings. July 1, 2017 Automated percutaneous and percutaneous endoscopic discectomy are considered investigational as techniques for intervertebral disc decompression in patients with back pain and/or radiculopathy related to disc herniation in the lumbar, thoracic, or cervical spine. April 1, 2017 Decompression of the intervertebral disc using laser energy (laser discectomy) or radiofrequency coblation (nucleoplasty) are considered investigational as techniques of disc decompression and treatment of associated pain.
Regence (Blue Shield)[122,123] August 1, 2017	Automated percutaneous and percutaneous endoscopic discectomy are considered investigational as techniques for intervertebral disc decompression in patients with back pain and/or radiculopathy related to disc herniation in the lumbar, thoracic, or cervical spine. Decompression of intervertebral discs using laser energy (laser discectomy) or radiofrequency energy (nucleoplasty) are considered investigational for all indications, including but not limited to disc decompression and treatment of associated pain.

(continued)

Table 50. Selected payer coverage for surgery for lumbar radiculopathy (continued)

Payer; Effective Date	Policy
Aetna[124,125] January 17, 2018	Aetna considers lumbar decompression with or without discectomy medically necessary for rapid progression of neurological impairment (e.g., foot drop, extremity weakness, numbness or decreased sensation, saddle anesthesia, bladder dysfunction or bowel dysfunction) confirmed by imaging studies (e.g., CT or MRI). Aetna considers lumbar laminectomy medically necessary for individuals with a herniated disc when *all* the following criteria are met: All other reasonable sources of pain have been ruled out; *and* Central/lateral recess or foraminal stenosis graded as moderate, moderate to severe or severe (not mild or mild to moderate); *and* Imaging studies (e.g., CT or MRI) indicate nerve root compression, that corresponds to the clinical findings of the specific affected nerve root; *and* Member has failed at least 6 weeks of conservative therapy; *and* Member's activities of daily living are limited by persistent pain radiating from the back down to the lower extremity; *and* Presence of neurological abnormalities (e.g., reflex change, positive straight leg raising, sensory loss, weakness) persist on examination and correspond to the specific affected nerve root. Aetna considers percutaneous lumbar discectomy, manual or automated, medically necessary for treatment of herniated lumbar discs when all the following are met: Member is otherwise a candidate for open laminectomy; and Member has failed 6 months of conservative treatment; and Diagnostic studies show that the nuclear bulge of the disc is contained within the annulus (i.e., the herniated disc is contained); and Member has no previous surgery or chemonucleolysis of the disc to be treated; and Member must have typical clinical symptoms of radicular pain corresponding to the level of disc involvement. Aetna considers the following procedures experimental and investigational: Endoscopic disc decompression, ablation, or annular modulation using the DiscFX System; Endoscopic laser foraminoplasty, endoscopic foraminotomy, laminotomy, and rhizotomy (endoscopic radiofrequency ablation) Endoscopic transforaminal discectomy Far lateral microendoscopic discectomy (FLMED) for extra-foraminal lumbar disc herniations or other indications; Far lateral microendoscopic discectomy (FLMED) for extra-foraminal lumbar disc herniations or other indications; Far lateral microendoscopic discectomy (FLMED) for extra-foraminal lumbar disc herniations or other indications; Further Reimbursement Notes: Laser: Clinical studies have not established a clinically significant benefit of use of a laser over a scalpel in spinal surgery. No additional benefit will be provided for the use of a laser in spinal surgery. Microscope and endoscope: Use of a microscope or endoscope is considered an integral part of the spinal surgery and not separately reimbursable.

(continued)

Table 50. Selected payer coverage for surgery for lumbar radiculopathy (continued)

Payer; Effective Date	Policy
United[126] August 1, 2017	United considers percutaneous discectomy and decompression procedures as unproven and not medically necessary for treating discogenic pain, including, but are not limited to, the following procedures: Nucleoplasty [percutaneous disc decompression or percutaneous plasma discectomy] Laser discectomy [laser disc decompression; laser-assisted disc decompression (LADD); or percutaneous endoscopic discectomy, with or without laser] Yeung endoscopic spinal surgery [arthroscopic microdiscectomy or percutaneous endoscopic discectomy] Transforaminal and/or interlaminar [transforaminal and interlaminar approach]
Cigna[127] June 15, 2017	Cigna considers percutaneous, endoscopic laminectomy and disc decompression procedures of the lumbar spine experimental, investigational, and unproven when used to report: Automated percutaneous lumbar discectomy/automated percutaneous nucleotomy Endoscopic anterior spinal surgery/Yeung endoscopic spinal system/percutaneous endoscopic discectomy/arthroscopic microdiscectomy, selective endoscopic discectomy Endoscopic disc decompression, ablation, or annular modulation using the DiscFX™ System Percutaneous laminotomy/laminectomy, percutaneous spinal decompression (e.g., mild® Procedure) Percutaneous laser discectomy /decompression, laser-assisted disc decompression Thermal intradiscal procedures are also considered experimental, investigational or unproven when used to report: Intervertebral disc biacuplasty Intradiscal electrothermal annuloplasty Percutaneous intradiscal radiofrequency thermocoagulation, intradiscal radiofrequency thermomodulation, percutaneous radiofrequency thermomodulation Coblation® Nucleoplasty™, disc nucleoplasty, decompression nucleoplasty plasma disc decompression Intraosseous radiofrequency nerve ablation of basivertebral nerve

(continued)

Table 50. Selected payer coverage for surgery for lumbar radiculopathy (continued)

Payer; Effective Date	Policy
Humana[128] September 28, 2017	Humana members may be eligible for discectomy (including microdiscectomy) for the following indications: Evidence of myelopathy, confirmed by CT or MRI, with both corresponding clinical symptoms and corresponding objective neurological signs Herniated disc, confirmed by imaging studies, when accompanied by radicular pain that has persisted despite 12 consecutive weeks of appropriate conservative treatment (e.g., rest, medications, physical therapy) Rapidly progressive neurological signs/symptoms of lumbar spine compression confirmed by imaging studies. Spinal fractures, infections, and tumors. Humana members may be eligible for a lumbar laminectomy, laminotomy, foraminectomy, foraminotomy or foraminolaminectomy for the following indications: Cauda equina syndrome (bowel or bladder dysfunction, bilateral lower extremity weakness/numbness/decreased sensation, saddle anesthesia) confirmed by imaging studies Herniated disc, foraminal stenosis or spinal stenosis at the level corresponding with clinical findings confirmed by imaging studies when accompanied by both radicular pain that has persisted despite 12 consecutive weeks of appropriate conservative treatment and physical and/or neurological abnormalities suggestive of nerve root or spinal cord compression Rapidly progressive neurologic signs/symptoms of lumbar spine compression confirmed by imaging studies Spinal fractures, infection, injury, tumor. Spondylolisthesis A minimally-invasive approach (e.g., endoscopic) which allows direct visualization of the surgical field and anatomy, is integral to the procedure and is not separately reimbursable. This applies to microendoscopic discectomy, tubular microdiscectomy, and other systems designed for minimally-invasive procedures. Humana members may NOT be eligible for other types of discectomy procedure including, but not limited to, the following procedures considered experimental and investigational: Accurascope DND Automated percutaneous lumbar discectomy, including but not limited to the Stryker Dekompressor lumbar discectomy probe. Laser discectomy, regardless of the approach, including percutaneous laser discectomy, laser-assisted discectomy, laser disc decompression, laser-assisted disc decompression or percutaneous laser disc decompression Percutaneous discectomy techniques not previously listed including, but not limited to, the HydroCision/HydroDiscectomy Humana members may NOT be eligible for other types of laminectomy, laminotomy, foraminectomy, foraminotomy, foraminolaminectomy, laminoplasty, corpectomy or decompression procedure including, but not limited to, the following procedures considered experimental and investigational: ANY percutaneous laminectomy, laminotomy, foraminectomy, foraminotomy, foraminolaminectomy, laminoplasty or corpectomy Laser laminectomy Percutaneous image-guided lumbar decompression including, but not be limited to, the MILD procedure and the Totalis Direct Decompression system

(continued)

Table 50. Selected payer coverage for surgery for lumbar radiculopathy (continued)

Payer; Effective Date	Policy
Kaiser Permanente	No policies specific to the coverage of surgeries for lumbar radiculopathy.
Medicaid[129]	Policies vary by state; example state Medicaid policy for North Carolina: Medicaid and North Carolina Health Check shall cover lumbar decompression surgery (discectomy, microdiscectomy, corpectomy, hemicorpectomy, foraminectomy, foraminoplasty, foraminotomy, laminectomy, hemilaminectomy, laminotomy, laminoplasty, and osteophytectomy) when all other reasonable sources of pain have been ruled out and the beneficiary meets the one or more of following specific criteria: Rapidly progressive neurological findings of nerve root or spinal cord compression, with imaging evidence of pathology that correlates with clinical findings (with or without gait or sphincter disturbance); Elective surgery needed as indicated by all the following when the beneficiary has failed at least six (6) consecutive weeks of conservative medical management (unless imaging indicates the need for urgent intervention): [a] Herniated disc with all the following: i. Nerve or spinal cord impingement seen on imaging studies; ii. Clinical findings consistent with impingement; and iii. All major psychosocial and substance use issues have been addressed. [b] Persistent pain and symptoms or findings that have not improved after at least six (6) consecutive weeks of conservative medical management, consisting of one or more of the following: i. Severe disabling radiculopathy; or ii. Clinical findings of nerve root compromise; 3) Spinal stenosis, spondylolisthesis, spinal fracture, or cauda equina syndrome.

Abbreviations: CT = computed tomography; DND = discectomy and neural decompression procedure; MRI = magnetic resonance imaging.

4.5 Limitations of this HTA

This HTA has several limitations related to the scoping and process and analyses we used to conduct the HTA.

4.5.1 Limitations in scope

This HTA was limited to studies and other information published or publicly available in English. We only included efficacy outcomes reported at 4 weeks or later; thus, immediate and very short-term benefits are not reflected in our synthesis. For example, outcomes related to operative time, intraoperative blood loss, length of inpatient hospital stay, short-term post-operative mobility, etc.

Our HTA excluded observational study designs, which may provide additional information for both efficacy and safety outcomes that could be more generalizable than data from participants in trials, who generally have fewer comorbidities than the general population as discussed in Section 4.2.6. Our HTA also excluded 'as treated' or 'per protocol' analyses, which could offer additional evidence on the efficacy and safety of the surgical interventions of interest in this HTA.

4.5.2 Limitations in process

The electronic search was limited to only three databases. For efficiency, we relied on hand searches of existing systematic reviews to identify eligible studies published prior to 2007. Although this approach may have resulted in missed studies, we think this is unlikely since we hand searched more than 40 systematic reviews. We used a single reviewer to screen titles and abstract; however, we mitigated this risk through reviewer training, quantitative assessment of interrater reliability during initial dual-review of 50 titles/abstracts, and using a low threshold for reviewers to request a second screening by another team member.

4.5.3 Limitations in analysis

Our grouping of minimally-invasive surgical procedures combines procedures that in fact may be heterogenous. Although the surgical approach used may be slightly different (e.g., direct vs. indirect visualization, different ablative techniques), the objective of the procedure (disc removal and decompression) is similar. For outcomes where quantitative synthesis was possible, we did not consistently observe heterogeneity in treatment effects, which suggests that factors other than the specific type of minimally-invasive intervention may explain the heterogeneity of treatment effect where it was observed. For example, we observed no heterogeneity of treatment effect in the pooled estimate of VAS leg pain outcomes for minimally-invasive surgery compared to standard surgery, but modest heterogeneity was observed for VAS back pain outcomes from the same set of studies. Because leg pain is the predominant symptom in sciatica relative to back pain, this suggests that differences in the underlying patient populations (e.g., coexisting morbidities, criteria for diagnosis/enrollment, etc.) may explain the heterogeneity in findings as opposed to differences in the minimally-invasive approach used. We note that our approach to synthesizing these interventions as a class is consistent with the approach used in several other systematic reviews on this topic.[130,131]

4.6 Ongoing Research and Future Research Needs

We did not identify any ongoing trials of surgical interventions specifically for lumbar radiculopathy through our search of the U.S. clinical trials registry (clinicaltrials.gov). Several trials are ongoing related injections of biologics (e.g., condoliase into nucleus pulposus)[69] or pharmacologics (e.g., epidural clonidine)[70] or use of adjunctive treatments (e.g., epidural steroid injections, stem cell injections, annular repair technologies) during or after discectomy to improve outcomes. The challenges faced in conducting methodologically rigorous randomized trials of surgical interventions are well-documented.[71] However, additional trials on treatment of lumbar radiculopathy with the same methodologic flaws will be unlikely to change the certainty of findings. Additional research on patient preferences and values related to timing of treatment or surgery, and establishment of minimally important clinical differences in outcomes that are specific to sciatica would also advance research in this area. Finally, advanced analytic and statistical techniques could be used within trials to quantify and mitigate the impact of crossovers on treatment effects and could be used within observational studies to mitigate biases introduced by nonrandomized study designs, potentially broadening the evidence base available to address important research questions.

5. Conclusion

Most findings in this HTA are based on a body of RCT evidence graded as low to very low certainty.

Surgery (discectomy or microdiscectomy) for symptomatic lumbar radiculopathy reduces pain more in the short and medium-term (up to 26 weeks) compared to nonsurgical interventions, but these findings do not persist at one year or longer follow-up. The evidence is insufficient to assess short- and medium-term impact on function because of inconsistent findings, but long-term impact on function suggests no difference between treatments. Surgery compared with nonsurgical interventions result in similar improvements in neurologic symptoms, quality of life, and return to work. No surgery-related deaths were observed and surgery-related complications were rare. The evidence is insufficient to assess outcomes related to persistent opioid use. Surgery compared with nonsurgical interventions may be cost-effective depending on a decision-makers willingness to pay threshold.

Minimally-invasive surgery is comparable to microdiscectomy or discectomy for reducing pain and improving function, quality of life, and neurological symptoms. No surgery-related deaths were observed and surgical morbidity is similar. The evidence is insufficient for drawing conclusions about differences in incidence of reoperations, persistent opioid use, and cost-effectiveness.

Microdiscectomy compared with discectomy are similar with respect to pain reduction, surgical morbidity, and incidence of reoperations, but the evidence is insufficient for drawing conclusions about differences in other efficacy, safety, and cost outcomes.

The evidence is insufficient for drawing conclusions about repeat surgery among individuals with recurrent radiculopathy.

6. References

1. Tarulli AW, Raynor EM. Lumbosacral radiculopathy. *Neurol Clin.* 2007;25(2):387-405. PMID: 17445735. doi: 10.1016/j.ncl.2007.01.008
2. Hsu PS, Armon C, Levin K. *Acute lumosacral radiculopathy: Pathophysiology, clinical features, and diagnosis.* Waltham, MA: UpToDate Inc.; 2017. http://www.uptodate.com
3. Genevay S, Atlas SJ. Lumbar spinal stenosis. *Best Pract Res Clin Rheumatol.* 2010;24(2):253-265. PMID: 20227646. doi: 10.1016/j.berh.2009.11.001
4. Konstantinou K, Dunn KM. Sciatica: review of epidemiological studies and prevalence estimates. *Spine (Phila Pa 1976).* 2008;33(22):2464-2472. PMID: 18923325. doi: 10.1097/BRS.0b013e318183a4a2
5. St. Francis Medical Technologies, Inc. Draft Package Insert: S STOP Interspinous Process Decompression (IPD) System. Published 2005. Updated November 9. Accessed November 8, 2017.
6. Paradigm Spine. Instructions for Use: coflex@ Interlaminar Technology. https://www.accessdata.fda.gov/cdrh_docs/pdf11/P110008C.pdf. Published n.d. Accessed November 8, 2017.
7. X-Sten Corp. Premarket Notification (51 0(k)] Summary. https://www.accessdata.fda.gov/cdrh_docs/pdf6/K062038.pdf Published 2006. Accessed November 8, 2017.
8. Vavken P, Ganal-Antonio AK, Quidde J, Shen FH, Chapman JR, Samartzis D. Fundamentals of Clinical Outcomes Assessment for Spinal Disorders: Clinical Outcome Instruments and Applications. *Global Spine J.* 2015;5(4):329-338. PMID: 26225283. doi: 10.1055/s-0034-1396046
9. Department of Health and Human Services. Patient-Reported Outcomes Measurement Information System (PROMIS). http://www.healthmeasures.net/explore-measurement-systems/promis. Published 2017. Accessed November 8, 2017.
10. United Nations Development Programme. Human Development Index. Website. http://hdr.undp.org/en/composite/HDI. Published 2016. Accessed October 15, 2017.
11. Washington State Health Care Authority. *Lumbar Fusion for Patients with Degenerative Disc Disease Uncomplicated by Comorbid Spinal Conditions - Re-Review.* Authority WSHC; 2015.
12. Higgins JPT, Sterne JAC, Savović J, et al. A revised tool for assessing risk of bias in randomized trials. In: Chandler J, McKenzie J, Boutron I, Welch V, eds. *Cochrane Methods.* Vol 102016. https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool
13. Chiou CF, Hay JW, Wallace JF, et al. Development and validation of a grading system for the quality of cost-effectiveness studies. *Med Care.* 2003;41(1):32-44. PMID: 12544542. doi: 10.1097/01.Mlr.0000039824.73620.E5
14. Bureau of the Fiscal Service. Treasury Reporting Rates of Exchange. https://www.fiscal.treasury.gov/fsreports/rpt/treasRptRateExch/historicalRates.htm. Published 2017. Updated October 12. Accessed October 19, 2017.
15. United States Department of Labor, Bureau of Labor Statistics. Consumer Price Index. https://www.bls.gov/cpi/. Published n.d. Accessed October 19, 2017.
16. Viechtbauer W. Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software.* 2010;36(3):1-48. PMID: WOS:000281593200001.

17. Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions. www.handbook.cochrane.org. Published 2011. Accessed 2 February, 2018.
18. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924-926. PMID: 18436948. doi: 10.1136/bmj.39489.470347.AD
19. Berkman ND, Lohr KN, Ansari MT, et al. Grading the strength of a body of evidence when assessing health care interventions: an EPC update. *Journal of Clinical Epidemiology*. 2015;68(11):1312-1324. PMID: WOS:000364612900008. doi: 10.1016/j.jclinepi.2014.11.023
20. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *CMAJ*. 2010;182(18):E839-842. PMID: 20603348. doi: 10.1503/cmaj.090449
21. AGREE Next Steps Consortium. The AGREE II Instrument [Electronic Version]. http://www.agreetrust.org. Published 2017. Accessed January 15, 2018.
22. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT): a randomized trial. *Jama*. 2006;296(20):2441-2450. PMID: 17119140. doi: 10.1001/jama.296.20.2441
23. McMorland G, Suter E, Casha S, du Plessis SJ, Hurlbert RJ. Manipulation or microdiskectomy for sciatica? A prospective randomized clinical study. *J Manipulative Physiol Ther*. 2010;33(8):576-584. PMID: 21036279. doi: 10.1016/j.jmpt.2010.08.013
24. Huang TJ, Hsu RW, Li YY, Cheng CC. Less systemic cytokine response in patients following microendoscopic versus open lumbar discectomy. *J Orthop Res*. 2005;23(2):406-411. PMID: 15734255. doi: 10.1016/j.orthres.2004.08.010
25. Sasaoka R, Nakamura H, Konishi S, et al. Objective assessment of reduced invasiveness in MED. Compared with conventional one-level laminotomy. *Eur Spine J*. 2006;15(5):577-582. PMID: 15926058. doi: 10.1007/s00586-005-0912-8
26. Weber H. Lumbar disc herniation. A controlled, prospective study with ten years of observation. *Spine (Phila Pa 1976)*. 1983;8(2):131-140. PMID: 6857385.
27. Tullberg T, Isacson J, Weidenhielm L. Does microscopic removal of lumbar disc herniation lead to better results than the standard procedure? Results of a one-year randomized study. *Spine (Phila Pa 1976)*. 1993;18(1):24-27. PMID: 8434321.
28. Thome C, Barth M, Scharf J, Schmiedek P. Outcome after lumbar sequestrectomy compared with microdiscectomy: a prospective randomized study. *J Neurosurg Spine*. 2005;2(3):271-278. PMID: 15796351. doi: 10.3171/spi.2005.2.3.0271
29. Teli M, Lovi A, Brayda-Bruno M, et al. Higher risk of dural tears and recurrent herniation with lumbar micro-endoscopic discectomy. *Eur Spine J*. 2010;19(3):443-450. PMID: 20127495. doi: 10.1007/s00586-010-1290-4
30. Ryang YM, Oertel MF, Mayfrank L, Gilsbach JM, Rohde V. Standard open microdiscectomy versus minimal access trocar microdiscectomy: results of a prospective randomized study. *Neurosurgery*. 2008;62(1):174-181; discussion 181-172. PMID: 18300905. doi: 10.1227/01.neu.0000311075.56486.c5
31. Ruetten S, Komp M, Merk H, Godolias G. Full-endoscopic interlaminar and transforaminal lumbar discectomy versus conventional microsurgical technique: a prospective, randomized, controlled study. *Spine (Phila Pa 1976)*. 2008;33(9):931-939. PMID: 18427312. doi: 10.1097/BRS.0b013e31816c8af7

32. Peul WC, van Houwelingen HC, van den Hout WB, et al. Surgery versus prolonged conservative treatment for sciatica. *N Engl J Med.* 2007;356(22):2245-2256. PMID: 17538084. doi: 10.1056/NEJMoa064039
33. Osterman H, Seitsalo S, Malmivaara A, Karppinen J. Surgery for disc herniation. A randomized controlled trial with 2-year follow-up. Paper presented at: Proceedings of the International Society for Study of the Lumbar Spine. 2003.
34. Mayer HM, Brock M. Percutaneous endoscopic discectomy: surgical technique and preliminary results compared to microsurgical discectomy. *J Neurosurg.* 1993;78(2):216-225. PMID: 8267686. doi: 10.3171/jns.1993.78.2.0216
35. Henriksen L, Schmidt K, Eskesen V, Jantzen E. A controlled study of microsurgical versus standard lumbar discectomy. *Br J Neurosurg.* 1996;10(3):289-293. PMID: 8799541.
36. Franke J, Greiner-Perth R, Boehm H, et al. Comparison of a minimally invasive procedure versus standard microscopic discotomy: a prospective randomised controlled clinical trial. *Eur Spine J.* 2009;18(7):992-1000. PMID: 19360440. doi: 10.1007/s00586-009-0964-2
37. Erginousakis D, Filippiadis DK, Malagari A, et al. Comparative prospective randomized study comparing conservative treatment and percutaneous disk decompression for treatment of intervertebral disk herniation. *Radiology.* 2011;260(2):487-493. PMID: 21613439. doi: 10.1148/radiol.11101094
38. Chatterjee S, Foy PM, Findlay GF. Report of a controlled clinical trial comparing automated percutaneous lumbar discectomy and microdiscectomy in the treatment of contained lumbar disc herniation. *Spine (Phila Pa 1976).* 1995;20(6):734-738. PMID: 7604351.
39. Brouwer PA, Brand R, van den Akker-van Marle ME, et al. Percutaneous laser disc decompression versus conventional microdiscectomy in sciatica: a randomized controlled trial. *Spine J.* 2015;15(5):857-865. PMID: 25614151. doi: 10.1016/j.spinee.2015.01.020
40. Arts MP, Brand R, van den Akker ME, Koes BW, Bartels RH, Peul WC. Tubular diskectomy vs conventional microdiskectomy for sciatica: a randomized controlled trial. *JAMA.* 2009;302(2):149-158. PMID: 19584344. doi: 10.1001/jama.2009.972
41. Gerszten PC, Smuck M, Rathmell JP, et al. Plasma disc decompression compared with fluoroscopy-guided transforaminal epidural steroid injections for symptomatic contained lumbar disc herniation: a prospective, randomized, controlled trial. *J Neurosurg Spine.* 2010;12(4):357-371. PMID: 20201654. doi: 10.3171/2009.10.spine09208
42. Haines SJ, Jordan N, Boen JR, Nyman JA, Oldridge NB, Lindgren BR. Discectomy strategies for lumbar disc herniation: results of the LAPDOG trial. *J Clin Neurosci.* 2002;9(4):411-417. PMID: 12217670.
43. Hermantin FU, Peters T, Quartararo L, Kambin P. A prospective, randomized study comparing the results of open discectomy with those of video-assisted arthroscopic microdiscectomy. *J Bone Joint Surg Am.* 1999;81(7):958-965. PMID: 10428127.
44. Malter AD, Larson EB, Urban N, Deyo RA. Cost-effectiveness of lumbar discectomy for the treatment of herniated intervertebral disc. *Spine (Phila Pa 1976).* 1996;21(9):1048-1054; discussion 1055. PMID: 8724089.
45. National Guideline Centre. National Institute for Health and Care Excellence: Clinical Guidelines. *Low Back Pain and Sciatica in Over 16s: Assessment and Management.* London: National Institute for Health and Care Excellence; 2016.

46. North RB, Kidd DH, Farrokhi F, Piantadosi SA. Spinal cord stimulation versus repeated lumbosacral spine surgery for chronic pain: a randomized, controlled trial. *Neurosurgery.* 2005;56(1):98-106; discussion 106-107. PMID: 15617591.
47. Ruetten S, Komp M, Merk H, Godolias G. Recurrent lumbar disc herniation after conventional discectomy: a prospective, randomized study comparing full-endoscopic interlaminar and transforaminal versus microsurgical revision. *J Spinal Disord Tech.* 2009;22(2):122-129. PMID: 19342934. doi: 10.1097/BSD.0b013e318175ddb4
48. Overdevest GM, Peul WC, Brand R, et al. Tubular discectomy versus conventional microdiscectomy for the treatment of lumbar disc herniation: long-term results of a randomised controlled trial. *J Neurol Neurosurg Psychiatry.* 2017;88(12):1008-1016. PMID: 28550071. doi: 10.1136/jnnp-2016-315306
49. van den Hout WB, Peul WC, Koes BW, Brand R, Kievit J, Thomeer RT. Prolonged conservative care versus early surgery in patients with sciatica from lumbar disc herniation: cost utility analysis alongside a randomised controlled trial. *Bmj.* 2008;336(7657):1351-1354. PMID: 18502912. doi: 10.1136/bmj.39583.709074.BE
50. Tosteson AN, Skinner JS, Tosteson TD, et al. The cost effectiveness of surgical versus nonoperative treatment for lumbar disc herniation over two years: evidence from the Spine Patient Outcomes Research Trial (SPORT). *Spine (Phila Pa 1976).* 2008;33(19):2108-2115. PMID: 18777603.
51. van den Akker ME, Arts MP, van den Hout WB, Brand R, Koes BW, Peul WC. Tubular diskectomy vs conventional microdiskectomy for the treatment of lumbar disk-related sciatica: cost utility analysis alongside a double-blind randomized controlled trial. *Neurosurgery.* 2011;69(4):829-835; discussion 835-826. PMID: 21623248. doi: 10.1227/NEU.0b013e31822578f6
52. van den Akker-van Marle ME, Brouwer PA, Brand R, et al. Percutaneous laser disc decompression versus microdiscectomy for sciatica: cost utility analysis alongside a randomized controlled trial. *Interv Neuroradiol.* 2017;23(5):538-545. PMID: 28679342. doi: 10.1177/1591019917710297
53. Stevenson RC, McCabe CJ, Findlay AM. An economic evaluation of a clinical trial to compare automated percutaneous lumbar discectomy with microdiscectomy in the treatment of contained lumbar disc herniation. *Spine (Phila Pa 1976).* 1995;20(6):739-742. PMID: 7604352.
54. National Guideline Clearinghouse. *Low back disorders: American College of Occupational and Environmental Medicine.* Rockville MD: Agency for Healthcare Research and Quality (AHRQ); 2016. https://www.guideline.gov/summaries/summary/50687/low-back-disorders?q=lumbar
55. Chou R, Loeser JD, Owens DK, et al. Interventional therapies, surgery, and interdisciplinary rehabilitation for low back pain: an evidence-based clinical practice guideline from the American Pain Society. *Spine (Phila Pa 1976).* 2009;34(10):1066-1077. PMID: 19363457. doi: 10.1097/BRS.0b013e3181a1390d
56. Manchikanti L, Abdi S, Atluri S, et al. An update of comprehensive evidence-based guidelines for interventional techniques in chronic spinal pain. Part II: guidance and recommendations. *Pain Physician.* 2013;16(2 Suppl):S49-283. PMID: 23615883.
57. North American Spine Society. *Clinical Guidelines for Diagnosis and Treatment of Lumbar Disc Herniation with Radiculopathy.* Burr Ridge, IL: Society NAS; 2012.

58. National Institute for Health and Care Excellence. Percutaneous transforaminal endoscopic lumbar discectomy for sciatica. https://www.nice.org.uk/guidance/ipg556/resources/percutaneous-transforaminal-endoscopic-lumbar-discectomy-for-sciatica-pdf-1899871990618309. Published 2016. Accessed January 15, 2018.
59. National Institute for Health and Care Excellence. Percutaneous interlaminar endoscopic lumbar discectomy for sciatica. https://www.nice.org.uk/guidance/ipg555. Published 2016. Updated April 27. Accessed November 8, 2017.
60. National Institute for Health and Care Excellence. Percutaneous coblation of the intervertebral disc for low back pain and sciatica[IPG543]. *Interventional procedures guidance* https://www.nice.org.uk/guidance/ipg543. Published 2016. Updated January 2016. Accessed January 15, 2018.
61. National Institute for Health and Care Excellence. Percutaneous electrothermal treatment of the intervertebral disc annulus for low back pain and sciatica[IPG544]. *Interventional procedures guidance* https://www.nice.org.uk/guidance/ipg544. Published 2016. Updated January 2016. Accessed January 15, 2018, 2018.
62. National Institute for Health and Care Excellence. Percutaneous intradiscal radiofrequency treatment of the intervertebral disc nucleus for low back pain[IPG545]. *Interventional Procedures Guidance* https://www.nice.org.uk/guidance/ipg545. Published 2016. Updated January 2016. Accessed January 15, 2018.
63. National Institute for Health and Care Excellence. Epiduroscopic lumbar discectomy through the sacral hiatus for sciatica[IPG570]. *Interventional procedures guidance* https://www.nice.org.uk/guidance/ipg570. Published 2016. Accessed January 15, 2018.
64. National Institute for Health and Care Excellence. Percutaneous intradiscal laser ablation in the lumbar spine[IPG357]. *Interventional procedures guidance* https://www.nice.org.uk/guidance/ipg357. Published 2010. Updated September 2010. Accessed January 15, 2018.
65. National Institute for Health and Care Excellence. Automated percutaneous mechanical lumbar discectomy. https://www.nice.org.uk/guidance/IPG141. Published 2005. Updated November. Accessed November 8, 2017.
66. National Institute for Health and Care Excellence. Endoscopic laser foraminoplasty[IPG31]. *Interventional procedures guidance* https://www.nice.org.uk/guidance/ipg31. Published 2003. Accessed January 15, 2018.
67. Lewis R, Williams N, Matar HE, et al. The clinical effectiveness and cost-effectiveness of management strategies for sciatica: systematic review and economic model. *Health Technol Assess.* 2011;15(39):1-578. PMID: 22078311. doi: 10.3310/hta15390
68. Centers for Medicare & Medicaid Services. National Coverage Determination (NCD) for Thermal Intradiscal Procedures (TIPs) (150.11). https://www.cms.gov/medicare-coverage-database/details/ncd-details.aspx?NCDId=324&ncdver=1&bc=AAAAgAAAAAAA&. Published n.d. Accessed February 5, 2018.
69. Seikagaku Corporation. A Study of SI-6603 in Patients With Lumbar Disc Herniation. *ClinicalTrials.gov.* Vol NCT01941563. Bethesda, MD: National Library of Medicine; 2000-.
70. Mayo Clinic. Epidural Clonidine for Lumbosacral Radiculopathy. *ClinicalTrials.gov*. Bethesda, MD: National Library of Medicine; 2000-.

71. Cook JA. The challenges faced in the design, conduct and analysis of surgical randomised controlled trials. *Trials.* 2009;10:9-9. PMID: 19200379. doi: 10.1186/1745-6215-10-9
72. Levin K, S. HP, C. A. *Acute lumbosacral radiculopathy: Treatment and prognosis.* Waltham, MA: UpToDate Inc.; 2017. http://www.uptodate.com
73. Vroomen PC, de Krom MC, Knottnerus JA. Predicting the outcome of sciatica at short-term follow-up. *Br J Gen Pract.* 2002;52(475):119-123. PMID: 11887877.
74. Wilson CA, Roffey DM, Chow D, Alkherayf F, Wai EK. A systematic review of preoperative predictors for postoperative clinical outcomes following lumbar discectomy. *Spine J.* 2016;16(11):1413-1422. PMID: 27497886. doi: 10.1016/j.spinee.2016.08.003
75. Casey E. Natural history of radiculopathy. *Phys Med Rehabil Clin N Am.* 2011;22(1):1-5. PMID: 21292142. doi: 10.1016/j.pmr.2010.10.001
76. Rhee JM, Schaufele M, Abdu WA. Radiculopathy and the herniated lumbar disk: controversies regarding pathophysiology and management. *Instr Course Lect.* 2007;56:287-299. PMID: 17472314.
77. Heliövaara M, MÄkelÄ M, Knekt P, Impivaara O, Aromaa A. Determinants of sciatica and low-back pain. *Spine.* 1991;16(6):608-614. PMID: 1830689.
78. Varlotta G, Brown M, Kelsey J, Golden A. Familial predisposition for herniation of a lumbar disc in patients who are less than twenty-one years old. *JBJS.* 1991;73(1):124-128. PMID: 1824705.
79. U.S. Food and Drug Administration. 510(k) Summary for Stryker Dekompressor™ Percutaneous Discectomy Probe. https://www.accessdata.fda.gov/cdrh_docs/pdf3/K032473.pdf. Published 2003. Accessed February 5, 2018.
80. U.S. Food and Drug Administration. Section 510(k) Premarket Notification for Herniatome Percutaneous Discectomy Device. https://www.accessdata.fda.gov/cdrh_docs/pdf14/k141557.pdf. Published 2014. Accessed February 5, 2018.
81. U.S. Food and Drug Administration. 510(k) Summary for Nucleotome®. https://www.accessdata.fda.gov/cdrh_docs/pdf4/K040919.pdf. Published 2004. Accessed February 5, 2018.
82. U.S. Food and Drug Administration. 510(k) Summary for Laurimed Percutaneous Discectomy Sustem. https://www.accessdata.fda.gov/cdrh_docs/pdf8/k082194.pdf. Published 2008. Accessed February 5, 2018.
83. U.S. Food and Drug Administration. Section 5: 510(k) Summary for ENSPIRETM Debrider System. https://www.accessdata.fda.gov/cdrh_docs/pdf9/k090278.pdf. Published 2009. Accessed February 5, 2018.
84. U.S. Food and Drug Administration. 510(k) Summary for enSpire™ Discectomy System. https://www.accessdata.fda.gov/cdrh_docs/pdf11/k110992.pdf. Published 2011. Accessed February 5, 2018.
85. U.S. Food and Drug Administration. 510(k) Summary of Safety and Effectivness for Disc-FX™ System. https://www.accessdata.fda.gov/cdrh_docs/pdf5/K052241.pdf. Published 2006. Accessed February 5, 2018.
86. U.S. Food and Drug Administration. 510(k) Summary for ORATEC(R) Spine CATH™ Intradiscal Catheter. https://www.accessdata.fda.gov/cdrh_docs/pdf/K993967.pdf. Published 1999. Accessed February 5, 2018.

87. U.S. Food and Drug Administration. 510(k) Summary of Safety and Effectiveness for SpineCATH Intradiscal Catheter (Nucleotomy Catheter). https://www.accessdata.fda.gov/cdrh_docs/pdf/K013622.pdf. Published 2002. Accessed February 5, 2018.
88. U.S. Food and Drug Administration. 510(k) Summary for ArthroCare® Corporation, ArthroCare® Coblator IQTM Perc-D® SpineWands®. https://www.accessdata.fda.gov/cdrh_docs/pdf13/K132099.pdf. Published 2013. Accessed February 5, 2018.
89. U.S. Food and Drug Administration. 510(k) Premarket Notification for METRX SYSTEM from Richard W. Treharne, Medtronic Sofamore Danek, Inc. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpmn/pmn.cfm?ID=K002931. Published 2018. Updated January 29. Accessed February 5, 2018.
90. U.S. Food and Drug Administration. 510(k) Summary of Safety and Effectivess for Yeung Endoscopic Spine System. https://www.accessdata.fda.gov/cdrh_docs/pdf/K973405.pdf. Published 1998. Accessed February 5, 2018.
91. North American Spine Society. Minimally Invasive Procedures for Back and Neck Pain. https://northamericanspine.com/. Published n.d. Accessed February 5, 2018.
92. U.S. Food and Drug Administration. 510(k) Summary for Ceralas Diode 980nm Laser System (Models D15, D25). https://www.accessdata.fda.gov/cdrh_docs/pdf7/K072779.pdf Published 2008. Accessed February 5, 2018.
93. Washington State Health Care Authority. Surgery for Symptoms of Lumbar Radiculopathy: Draft key questions: public comment and response. https://www.hca.wa.gov/about-hca/health-technology-assessment/surgery-for-symptomatic-lumbar-radiculopathy Published 2017. Updated November 17. Accessed February 3, 2018.
94. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol.* 2012;8(1):23-34. PMID: 22833776.
95. Gartlehner G, West SL, Mansfield AJ, et al. Clinical heterogeneity in systematic reviews and health technology assessments: synthesis of guidance documents and the literature. *Int J Technol Assess Health Care.* 2012;28(1):36-43. PMID: 22217016. doi: 10.1017/s0266462311000687
96. West SL, Gartlehner G, Mansfield AJ, et al. AHRQ Methods for Effective Health Care. *Comparative Effectiveness Review Methods: Clinical Heterogeneity*. Rockville (MD): Agency for Healthcare Research and Quality (US); 2010.
97. Peul WC, van den Hout WB, Brand R, Thomeer RT, Koes BW. Prolonged conservative care versus early surgery in patients with sciatica caused by lumbar disc herniation: two year results of a randomised controlled trial. *BMJ.* 2008;336(7657):1355-1358. PMID: 18502911. doi: 10.1136/bmj.a143
98. Lequin MB, Verbaan D, Jacobs WC, et al. Surgery versus prolonged conservative treatment for sciatica: 5-year results of a randomised controlled trial. *BMJ Open.* 2013;3(5). PMID: 23793663. doi: 10.1136/bmjopen-2012-002534
99. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical versus nonoperative treatment for lumbar disc herniation: four-year results for the Spine Patient Outcomes Research Trial

(SPORT). *Spine (Phila Pa 1976).* 2008;33(25):2789-2800. PMID: 19018250. doi: 10.1097/BRS.0b013e31818ed8f4

100. Lurie JD, Tosteson TD, Tosteson AN, et al. Surgical versus nonoperative treatment for lumbar disc herniation: eight-year results for the spine patient outcomes research trial. *Spine (Phila Pa 1976).* 2014;39(1):3-16. PMID: 24153171. doi: 10.1097/BRS.0000000000000088
101. Arts MP, Brand R, van den Akker ME, et al. Tubular diskectomy vs conventional microdiskectomy for the treatment of lumbar disk herniation: 2-year results of a double-blind randomized controlled trial. *Neurosurgery.* 2011;69(1):135-144; discussion 144. PMID: 21792119. doi: 10.1227/NEU.0b013e318214a98c
102. Brouwer PA, Brand R, van den Akker-van Marle ME, et al. Percutaneous laser disc decompression versus conventional microdiscectomy for patients with sciatica: Two-year results of a randomised controlled trial. *Interv Neuroradiol.* 2017;23(3):313-324. PMID: 28454511. doi: 10.1177/1591019917699981
103. Gempt J, Jonek M, Ringel F, Preuss A, Wolf P, Ryang Y. Long-term follow-up of standard microdiscectomy versus minimal access surgery for lumbar disc herniations. *Acta Neurochir (Wien).* 2013;155(12):2333-2338. PMID: 24136677. doi: 10.1007/s00701-013-1901-z
104. Barth M, Weiss C, Thome C. Two-year outcome after lumbar microdiscectomy versus microscopic sequestrectomy: part 1: evaluation of clinical outcome. *Spine (Phila Pa 1976).* 2008;33(3):265-272. PMID: 18303458. doi: 10.1097/BRS.0b013e318162018c
105. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res (Hoboken).* 2011;63(S11):S240-252. PMID: 22588748. doi: 10.1002/acr.20543
106. Maruish ME. *User's manual for the SF-36v2 Health Survey.* Lincoln, RI: QualityMetric Inc; 2011.
107. Lauridsen HH, Hartvigsen J, Manniche C, Korsholm L, Grunnet-Nilsson N. Responsiveness and minimal clinically important difference for pain and disability instruments in low back pain patients. *BMC Musculoskeletal Disorders.* 2006;7:82-82. PMID: PMC1635558. doi: 10.1186/1471-2474-7-82
108. Schwind J, Learman K, O'Halloran B, Showalter C, Cook C. Different minimally important clinical difference (MCID) scores lead to different clinical prediction rules for the Oswestry disability index for the same sample of patients. *J Man Manip Ther.* 2013;21(2):71-78. PMID: 24421616. doi: 10.1179/2042618613y.0000000028
109. Fairbank JC, Pynsent PB. The Oswestry Disability Index. *Spine (Phila Pa 1976).* 2000;25(22):2940-2952; discussion 2952. PMID: 11074683.
110. Fu YS, Zeng BF, Xu JG. Long-term outcomes of two different decompressive techniques for lumbar spinal stenosis. *Spine (Phila Pa 1976).* 2008;33(5):514-518. PMID: 18317196. doi: 10.1097/BRS.0b013e3181657dde
111. Javid MJ, Nordby EJ, Ford LT, et al. Safety and efficacy of chymopapain (Chymodiactin) in herniated nucleus pulposus with sciatica. Results of a randomized, double-blind study. *JAMA.* 1983;249(18):2489-2494. PMID: 6341632.

112. Atlas SJ, Singer DE, Keller R, Convery K, Mooney N, Deyo R. The efficacy of lumbar disc surgery in patients with sciatica - preliminary-results from the Maine Lumbar Spine Study. *Clinical Research.* 1993;41(2):A200-A200. PMID: WOS:A1993KW76100484.
113. Anderson JL, Heidenreich PA, Barnett PG, et al. ACC/AHA statement on cost/value methodology in clinical practice guidelines and performance measures: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures and Task Force on Practice Guidelines. *Circulation.* 2014;129(22):2329-2345. PMID: 24677315. doi: 10.1161/cir.0000000000000042
114. Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness--the curious resilience of the $50,000-per-QALY threshold. *N Engl J Med.* 2014;371(9):796-797. PMID: 25162885. doi: 10.1056/NEJMp1405158
115. Sanders GD, Hlatky MA, Owens DK. Cost-effectiveness of implantable cardioverter-defibrillators. *N Engl J Med.* 2005;353(14):1471-1480. PMID: 16207849. doi: 10.1056/NEJMsa051989
116. Trentham-Dietz A, Kerlikowske K, Stout NK, et al. Tailoring Breast Cancer Screening Intervals by Breast Density and Risk for Women Aged 50 Years or Older: Collaborative Modeling of Screening Outcomes. *Ann Intern Med.* 2016;165(10):700-712. PMID: 27548583. doi: 10.7326/m16-0476
117. Martin BI, Mirza SK, Flum DR, et al. Repeat surgery after lumbar decompression for herniated disc: the quality implications of hospital and surgeon variation. *Spine J.* 2012;12(2):89-97. PMID: 22193055. doi: 10.1016/j.spinee.2011.11.010
118. Kim CH, Chung CK, Park CS, Choi B, Kim MJ, Park BJ. Reoperation rate after surgery for lumbar herniated intervertebral disc disease: nationwide cohort study. *Spine (Phila Pa 1976).* 2013;38(7):581-590. PMID: 23023591. doi: 10.1097/BRS.0b013e318274f9a7
119. Premera Blue Cross. Medical Policy - 7.01.93. Decompression of the Intervertebral Disc Using Laser Energy (Laser Discectomy) or Radiofrequency Coblation (Nucleoplasty). https://www.premera.com/medicalpolicies/7.01.93.pdf. Published 2017. Accessed February 5, 2018.
120. Premera Blue Cross. Medical Policy - 7.01.551. Lumbar Spine Decompression Surgery: Discectomy, Foraminotomy, Laminotomy, Laminectomy. https://www.premera.com/medicalpolicies/7.01.551.pdf. Published 2017. Accessed February 5, 2018.
121. Premera Blue Cross. Medical Policy - 7.01.18. Automated Percutaneous and Percutaneous Endoscopic Discectomy. https://www.premera.com/medicalpolicies/7.01.18.pdf Published 2017. Accessed February 5, 2018.
122. Regence BlueCross, BlueShield of Oregon. Surgery, Policy No. 145. Automated Percutaneous and Percutaneous Endoscopic Discectomy. http://blue.regence.com/trgmedpol/surgery/sur145.pdf. Published 2017. Accessed February 5, 2018.
123. Regence BlueCross, BlueShield of Oregon. Surgery, Policy No. 131. Decompression of Intervertebral Discs Using Laser Energy (Laser Discectomy) or Radiofrequency Energy (Nucleoplasty). http://blue.regence.com/trgmedpol/surgery/sur131.pdf. Published 2017. Accessed February 5, 2018.

124. Aetna. Number: 0743. Spinal Surgery: Laminectomy and Fusion. http://www.aetna.com/cpb/medical/data/700_799/0743.html. Published 2007. Accessed February 5, 2018.
125. Aetna. Number: 0016. Back Pain - Invasive Procedures. http://www.aetna.com/cpb/medical/data/1_99/0016.html Published 2007. Accessed February 5, 2018.
126. UnitedHealthcare. Discogenic Pain Treatment. https://www.uhcprovider.com/content/dam/provider/docs/public/policies/comm-medical-drug/discogenic-pain-treatment.pdf Published 2017. Accessed February 5, 2018.
127. Cigna. Minimally Invasive Intradiscal/ Annular Procedures and Trigger Point Injections. https://cignaforhcp.cigna.com/public/content/pdf/coveragePolicies/medical/mm_0139_coveragepositioncriteria_invasive_treatment_for_back_pain.pdf. Published 2017. Accessed February 5, 2018.
128. Humana. Spinal Decompression Surgery. http://apps.humana.com/tad/tad_new/Search.aspx?criteria=laminectomy&searchtype=freetext&policyType=both. Published 2017. Accessed February 5, 2018.
129. NC Division of Medical Assistance. Medicaid and Health Choice. Spinal Surgeries, Clinical Coverage Policy No: 1A-30. https://files.nc.gov/ncdma/documents/files/1A-30.pdf. Published 2017. Accessed February 5, 2018.
130. Rasouli MR, Rahimi-Movaghar V, Shokraneh F, Moradi-Lakeh M, Chou R. Minimally invasive discectomy versus microdiscectomy/open discectomy for symptomatic lumbar disc herniation. *Cochrane Database of Systematic Reviews.* 2014(9). http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD010328.pub2/abstract.
131. Dasenbrock HH, Juraschek SP, Schultz LR, et al. The efficacy of minimally invasive discectomy compared with open discectomy: a meta-analysis of prospective randomized controlled trials. *J Neurosurg Spine.* 2012;16(5):452-462. PMID: 22404142. doi: 10.3171/2012.1.spine11404

Appendix A. State of Washington Health Care Authority Utilization Data

The State of Washington Health Care Authority provided this data and analysis for inclusion in this Health Technology Assessment (HTA).

Populations

This analysis includes member utilization and cost data from the following agencies: PEBB/UMP (Public Employees Benefit Board/Uniform Medical Plan); PEBB Medicare, the Department of Labor and Industries (LNI) Workers' Compensation Plan; and the HCA (Health Care Authority) Medicaid (Fee-for-Service) and the Managed Care (MCO) Medicaid program.

Population Criteria:

- >17 years old at time of service (PEBB) OR >18 years old for MCO and HCA Medicaid AND
- Experiencing at least one of the codes from ***Table A-1***.

Table A-1. Procedure (CPT/HCPCS) Descriptions

CPT/ HCPCS Codes	Description
62380	Endoscopic decompression of spinal cord, nerve root(s), including laminotomy, partial facetectomy, foraminotomy, discectomy and/or excision of herniated intervertebral disc, 1 interspace, lumbar
63030	Laminotomy (hemilaminectomy), with decompression of nerve root(s), including partial facetectomy, foraminotomy and/or excision of herniated intervertebral disc; 1 interspace, lumbar
63042	Laminotomy (hemilaminectomy), with decompression of nerve root(s), including partial facetectomy, foraminotomy and/or excision of herniated intervertebral disc, reexploration, single interspace; lumbar
63044	Laminotomy (hemilaminectomy), with decompression of nerve root(s), including partial facetectomy, foraminotomy and/or excision of herniated intervertebral disc, reexploration, single interspace; each additional lumbar interspace (List separately in addition to code for primary procedure)
63047	Laminectomy, facetectomy and foraminotomy (unilateral or bilateral with decompression of spinal cord, cauda equina and/or nerve root[s], [e.g., spinal or lateral recess stenosis]), single vertebral segment; lumbar
63048	Laminectomy, facetectomy and foraminotomy (unilateral or bilateral with decompression of spinal cord, cauda equina and/or nerve root[s], [e.g., spinal or lateral recess stenosis]), single vertebral segment; each additional segment, cervical, thoracic, or lumbar (List separately in addition to code for primary procedure)
63185	Laminectomy with rhizotomy; 1 or 2 segments
63190	Laminectomy with rhizotomy; more than 2 segments
63191	Laminectomy with section of spinal accessory nerve
63200	Laminectomy, with release of tethered spinal cord, lumbar
0275T	Percutaneous laminotomy/laminectomy (interlaminar approach) for decompression of neural elements, (with or without ligamentous resection, discectomy, facetectomy and/or foraminotomy), any method, under indirect image guidance (e.g., fluoroscopic, CT), single or multiple levels, unilateral or bilateral; lumbar

Abbreviations: CPT = current procedural terminology; HCPCS = Healthcare Common Procedure Coding System.

Methods

Count of surgical treatments were based on an individual experiencing a paid provider-patient face-to-face, on a specific date and including at least one of the Current Procedural Terminology (CPT) codes from ***Table A-1***. Additional analysis included reviewing utilization data based on a claim having or not having a specific primary diagnosis as listed in ***Table A-2***. Data evaluation

included examining utilization by member; by treatment modality (***Table A-1***), and by total claims' cost incurred by a member on the date of their surgery (Total Claims).

Analyzing total claims for the date of service provided an enhanced view of the overall costs for a surgical intervention of symptomatic lumbar radiculopathy (e.g., facility costs, labs, etc.). Unless otherwise noted, "dollars" refers to paid dollars. Denied claims were excluded from the analysis.

Table A-2. Selected Diagnosis Codes and Descriptions

ICD-10 Diagnosis Code	ICD-9 Diagnosis Code	Description
M47.819	721.9	Other spondylosis with radiculopathy
M47.817	721.3	Other spondylosis with radiculopathy, lumbar region
M47.817	721.3	Other spondylosis with radiculopathy, lumbosacral region
M47.817	721.3	Other spondylosis with radiculopathy, sacral and sacrococcygeal region
M51.16	722.93	Intervertebral disc disorders with radiculopathy, lumbar region
M51.17	722.93	Intervertebral disc disorders with radiculopathy, lumbosacral region
M54.1	729.2	Radiculopathy
M54.16	729.2	Radiculopathy, lumbar region
M54.17	724.4	Radiculopathy, lumbosacral region
M54.18	729.2	Radiculopathy, sacral and sacrococcygeal region
M48.061	724.02	Spinal stenosis, lumbar region without neurogenic claudication

Abbreviations: ICD-9, ICD-10 = International Statistical Classification of Diseases and Related Health Problems. The numbers indicate the version of the code.

Demographics

The following figures depict 4 years of population fluctuations for combined Medicaid (HCA and MCO) (***Figure A-1***), and for PEBB/UMP and UMP Medicare (***Figure A-2***). ***Figure A-3*** depicts the distribution of the population by age.

Figure A-1. Medicaid population growth, 2014-2017

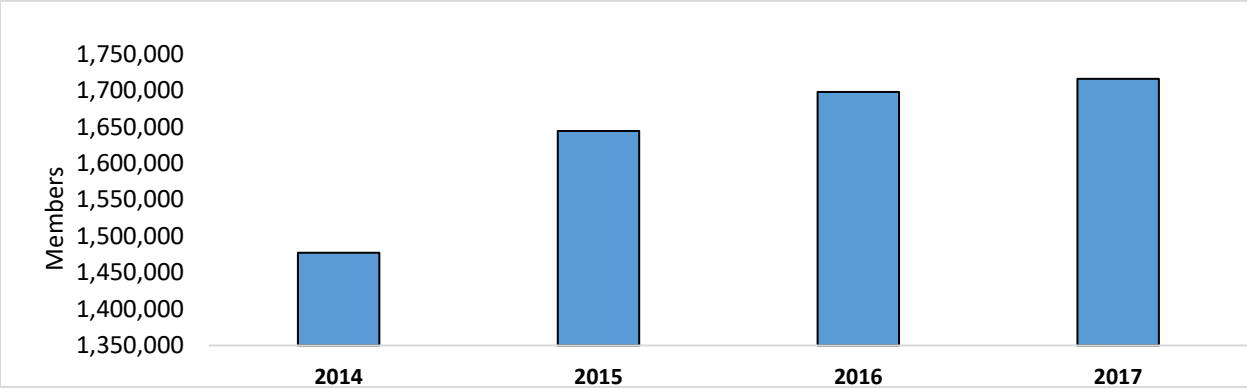


Figure A-2. Public Employees Benefit Board/Uniform Medical Plan Population Growth, 2014-2017

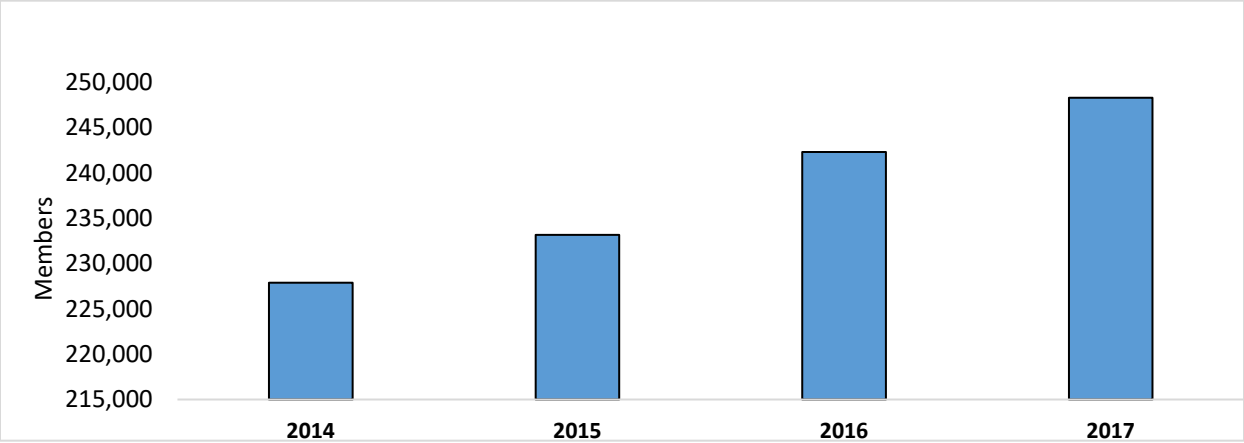
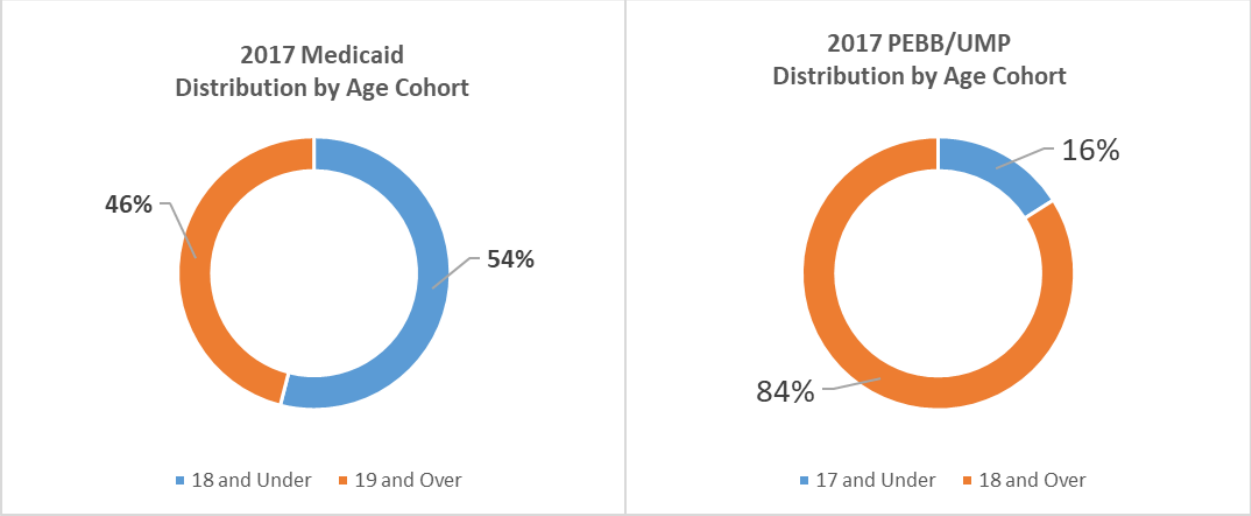


Figure A-3. Distribution of the population by age and cohort, 2017



Abbreviations: PEBB/UMP = Public Employees Benefit Board/ Uniform Medical Plan.

Cost and Utilization Data

Table A-3 provides utilization data for the surgical treatments listed in ***Table A-1*** stratified by presence or absence of the diagnoses indicated in ***Table A-2*** and by cohort. We note that the data for 2017 does not include 90 days of claims run-out.

Table A-3. Utilization data for surgical treatments for lumbar radiculopathy by diagnosis and cohort, 2015-2017

Cohort and Parameter[a]	2015	2016	2017[b]
Medicaid Managed Care Organization			
Unique Patients	320	352	256
Total Treatments with Diagnosis	351	394	242
Treatments without Diagnosis	476	501	378
Dollars Paid by Total Treatments with Diagnosis	$1,835,396	$1,779,602	$1,311,784
Average Paid Dollars/Patient	$5,754	$6,425	$3,780
Medicaid Health Care Authority (Fee for Service)			
Unique Patients	25	29	3
Total Treatments with Diagnosis	25	28	3
Treatments without Diagnosis	42	37	6
Dollars Paid by Total Treatments with Diagnosis	$110,476	$88,391	$24,329
Average Paid Dollars/Patient	$4,419	$3,048	$8,110
Labor and Industries Workers' Compensation Plan[c]			
Unique Patients	223	231	213
Total Treatments with Diagnosis	229	240	216
Treatments without Diagnosis	1	1	1
Dollars Allowed by Total Treatments with Diagnosis	$2,657,263	$3,333,749	$3,243,177
Average Dollars Allowed/Patient	$11,916	$14,431.81	$15,226.18
Public Employee Benefit Board/Uniform Medical Plan			
Unique Patients	76	91	79
Total Treatments with Diagnosis	79	96	83
Treatments without Diagnosis	192	185	142
Dollars Paid by Total Treatments with Diagnosis	$675,955	$943,363	$785,274
Average Paid Dollars/Patient	$8,894	$10,367	$9,940
Public Employee Benefit Board /Medicare			
Unique Patients	73	39	39
Total Treatments with Diagnosis	82	42	41
Treatments without Diagnosis	144	111	101
Dollars Paid by Total Treatments with Diagnosis	$52,834	$35,125	$33,866
Average Paid Dollars/Patient	$724	$901	$868

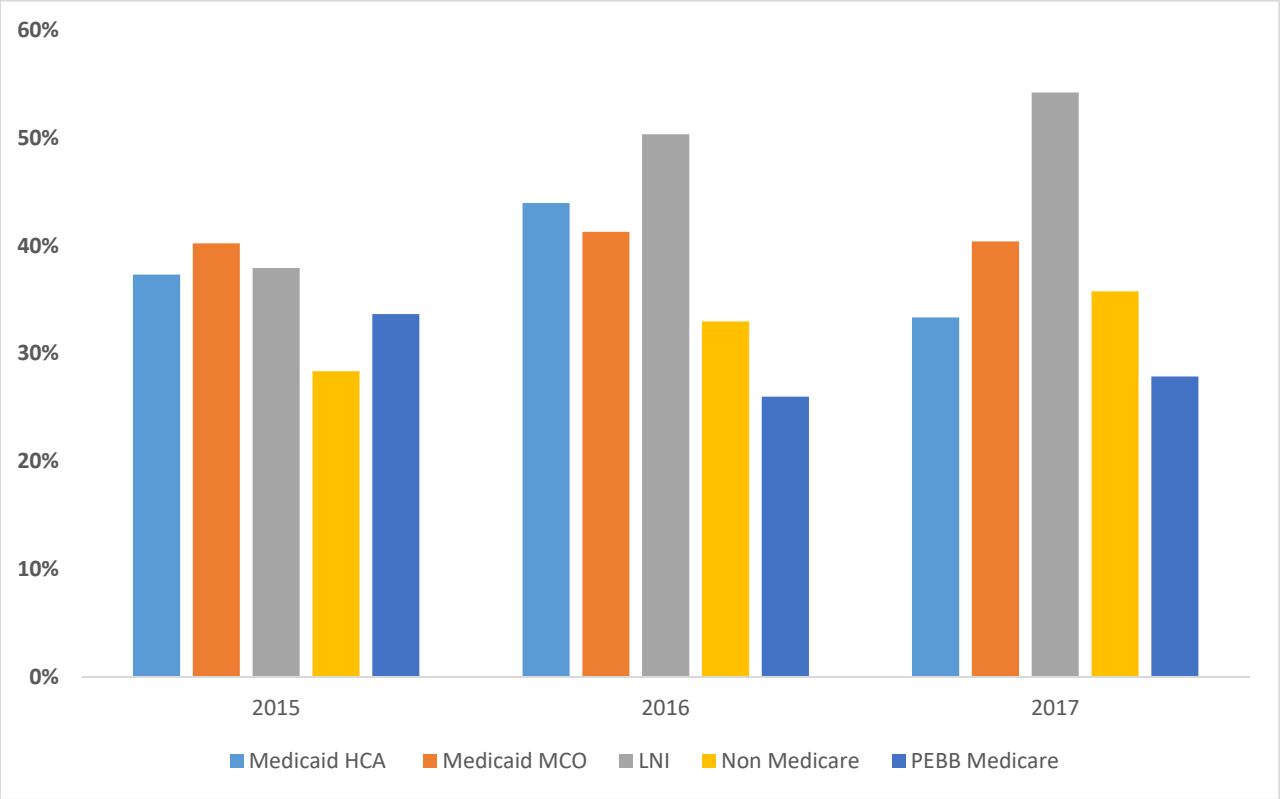
[a] Parameter definition:
Unique Patients = non-duplicated patient by year, reported by agency;
Total treatments with diagnosis = treatment (from Table A-1) of a patient by provider face-to-face on a specific date with a diagnosis from Table A-2;
Total treatments without diagnosis = treatment (from Table A-1) of a patient by a provider face-to-face on a specific date without a diagnosis from Table A-2;
Dollars paid by total treatments with diagnosis = annual dollars paid for treatments (from Table A-1) with diagnosis from Table A-2;
Average paid dollars/patient with diagnosis = average dollars per unique patient

[b] Data does not include full 90 days of claims-run out.

[c] Uses allowed dollars.

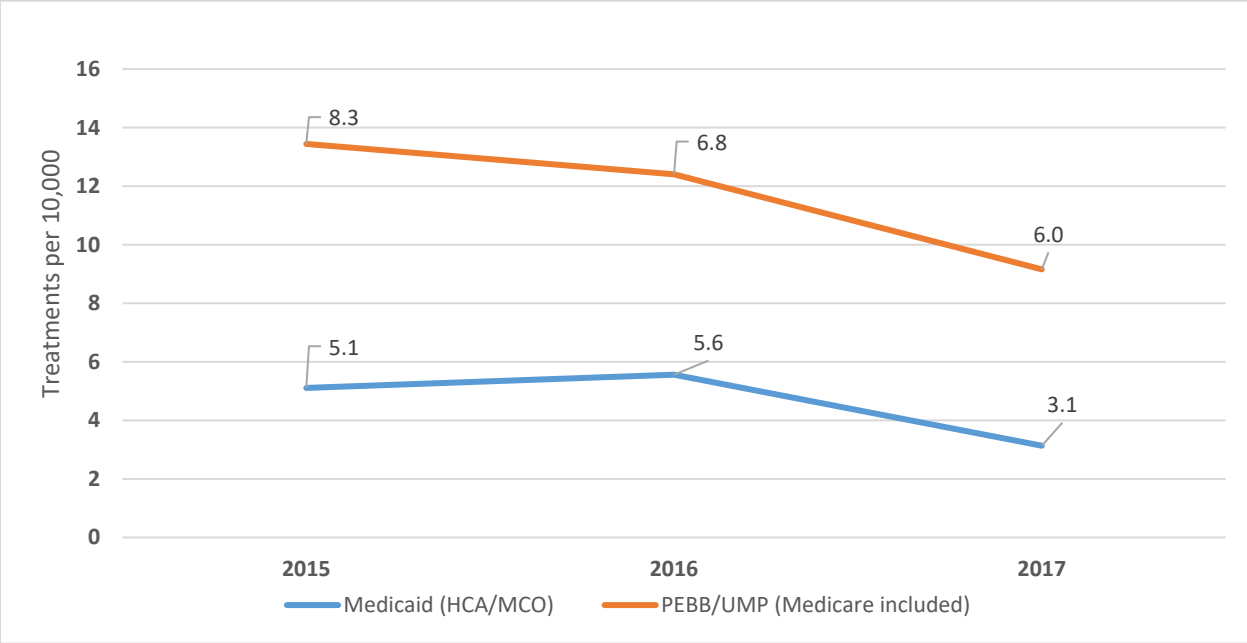
Figure A-4 provides the percent of surgical treatments from ***Table A-1*** that are billed with a primary diagnosis code from ***Table A-2*** and ***Figure A-5*** provides the rate of utilization of the surgical treatments from ***Table A-1*** among individuals with diagnoses in ***Table A-2*** per 10,000 beneficiaries. Data is presented separately for the Medicaid (HCA and MCO) cohorts, the Labor and Industries cohort, and the PEBB/UMP (including Medicare) cohorts.

Figure A-4. Percent of surgery procedures that are billed with a designated primary diagnosis code, 2015-2017



Abbreviations: HCA = Health Care Authority; MCO = Managed Care; LNI = Department of Labor and Industries; PEBB = Public Employees Benefit Board.

Figure A-5. Surgery utilization with diagnosis per 10,000 adult[a] beneficiaries, 2015-2017



[a] Beneficiaries >18 years old for Medicaid; Beneficiaries > 17 years old for Public Employee Benefit Board/Uniform Medical Plan

Abbreviations: HCA = Health Care Authority; MCO = Managed Care; PEBB/UMP = Public Employees Benefit Board/ Uniform Medical Plan.

Appendix B. Search Strategy

PubMed searched from 1/1/2007-11/09/2017

((((((((((((((((((((((((((((((((((((laser disc decompression[Title/Abstract]) OR laser thermo-discoplasty[Title/Abstract]) OR laser discectomy[Title/Abstract]) OR laser disc thermoplasty[Title/Abstract]) OR laser disc thermoplasty[Title/Abstract]) OR ((laser thermoplasty[Title/Abstract] AND disc[Title/Abstract]))) OR laser thermoplasty[Title/Abstract]) OR laser nucleotomy[Title/Abstract]) OR MILD[Title/Abstract]) OR minimally invasive lumbar decompression[Title/Abstract]) OR image-guided lumbar decompression[Title/Abstract]) OR thermal discoplasty[Title/Abstract]) OR ((endoscopic discectomy[Title/Abstract]) OR endoscopic nucleotomy[Title/Abstract])) OR ((percutaneous nucleotomy[Title/Abstract]) OR percutaneous discectomy)) OR thermal nucleoplasty[Title/Abstract]) OR laser discoplasty[Title/Abstract])))) AND Lumbar[Title/Abstract])) OR ((((((((((((((((((((((Laminectomy[MeSH Terms]) OR Laminotomy[MeSH Terms]) OR Diskectomy[MeSH Terms]) OR Diskectomy, percutaneous[MeSH Terms]) OR Foraminotomy[MeSH Terms]) OR Intervertebral Disc/surgery[MeSH Terms]) OR Intervertebral Disc/therapy[MeSH Terms]) OR Intervertebral Disc Displacement/surgery[MeSH Terms]) OR Intervertebral Disc Displacement/therapy[MeSH Terms]) OR Intervertebral Disc Degeneration/surgery[MeSH Terms]) OR Intervertebral Disc Degeneration/therapy[MeSH Terms]) OR Sciatica/surgery[MeSH Terms]) OR Sciatica/therapy[MeSH Terms]) OR Spinal osteophytosis/surgery[MeSH Terms]) OR Spinal Nerve Roots/surgery[MeSH Terms]) OR Spinal osteophytosis/therapy[MeSH Terms]) OR Lumbar vertebrae/surgery[MeSH Terms]) OR Lumbar vertebrae/therapy[MeSH Terms])))))))) NOT (Goats[MeSH Terms] OR Mouse[MeSH Terms] OR Rat[MeSH Terms] OR Dogs[MeSH]))) NOT (Cervical vertebrae[MeSH Terms] OR Cervical vertebrae[Title/Abstract]))) NOT (Cancer[MeSH Terms] OR Fractures[MeSH Terms] OR Cancer Pain[MeSH Terms] OR Tuberculosis[MeSH Terms] OR Spinal Cord Injuries[MeSH Terms] OR Scoliosis[MeSH Terms]))) NOT (Case Reports[Publication Type] OR Comment[Publication Type] OR Letter[Publication Type] OR Patient Education Handout[Publication Type] OR Editorial[Publication Type]))) NOT (children[MeSH Terms] OR infants[MeSH Terms] OR adolescents[MeSH Terms]))) NOT Spinal Fusion[MeSH Major Topic]) NOT Spondylolisthesis/surgery[MeSH Major Topic])) AND (Comparative Study[Publication Type] OR Review[Publication Type] OR Meta-analysis[Publication Type] OR Clinical Study[Publication Type] OR Randomized Controlled Trial[Publication Type] OR Clinical Trial[Publication Type] OR Clinical Trial, Phase I[Publication Type] OR Clinical Trial, Phase II[Publication Type] OR Clinical Trial, Phase III[Publication Type] OR Clinical Trial, Phase IV[Publication Type] OR Observational Study[Publication Type] OR Pragmatic Clinical Trial[Publication Type] OR Controlled Clinical Trial[Publication Type]) Sort by: PublicationDate Filters: Publication date from 2007/01/01 to 2017/11/09; English

Yield: 1,443

Cochrane Library Search from inception to 11/10/2017

Terms: Sciatica, lumbar radiculopathy, lumbar disc disease, minimally-invasive spine

Total Yield: 93

ClinicalTrials.Gov Search from inception to 11/10/2017

Terms: Sciatica, lumbar radiculopathy, lumbar disc disease, minimally-invasive spine; limits: Adult 18-65, Adult 66+, Interventions

Total Yield: 253

Other Data

The following websites were searched using the terms radiculopathy, laminectomy, discectomy, practice guidelines, spine surgery, nerve root compression

United States (U.S.) Food and Drug Administration

Centers for Medicare and Medicaid Services

Aetna

UnitedHealth

Humana

BlueCross BlueShield (Premera and Regence)

Kaiser Permanente

National Institute for Health and Care Excellence (U.K.)

U.S. Agency for Healthcare Research and Quality

North American Spine Society

American Society of Interventional Pain Physicians

American Academy of Orthopaedic Surgeons

American Academy of Neurological Surgeons

American Pain Society

American College of Occupational and Environmental Medicine

Appendix C. Additional Methods

The exchanges rates listed in ***Table C-1*** were used to convert foreign costs to United States (U.S.) dollars and the chain-weighted, average year consumer price indices used to adjust reported costs to 2010 dollars are reported in ***Table C-2***.

Table C-1. Exchange rates used to convert foreign costs to U.S. dollars

	U.S. $	British Pound	Euro €
Year 1992	1	0.568	-
Year 2008	1	-	0.659
Year 2009	1	-	0.789
Year 2010	1	-	0.740

Source: U.S. Department of Treasury. Treasury Reporting Rates of Exchange. Historical Rates for March 31st, 1992; March 31st, 2008; March 31st, 2009; and March 31st, 2010. Available at: https://www.fiscal.treasury.gov/fsreports/rpt/treasRptRateExch/historicalRates.htm Accessed January 21, 2018.

Abbreviations: U.S. = United States.

Table C-2. Chain-weighted, average year consumer price indices

Year	Annual Average CPI
1992	140.3
1993	144.5
2004	188.9
2005	195.3
2008	215.30
2009	214.54
2010	218.06
2014	236.74

Source: U.S. Department of Labor, Bureau of Labor Statistics. CPI Databases. All Urban Consumers (Chained CPI). Average Annual Indices. Available at: https://www.bls.gov/cpi/data.htm . Accessed January 21, 2018.

Abbreviations: CPI = consumer price index.

Appendix D. Evidence Tables

Table D-1. Study characteristics of included studies

Table D-1. Study characteristics of included studies

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48]	Parallel-group RCT; The Netherlands; Low	Government; The Dutch Health Care Insurance Board; The Sciatica Micro-Endoscopic Discectomy Randomized Controlled Trial	Tubular discectomy **N randomized**: 167; **N analyzed**: 166 (99.4%) in main study's primary analyses. 52w: 156 (93.4%) 2y: 154 (92.2%) 3y: 117 (70.1%) 4y: 117 (70.1%) 5y: 106 (63.5%); **N crossovers**: 2 (1.2%); Surgery scheduled within 4w of first visit. A 25-30 mm midline incision used, skin retracted laterally and the guidewire and sequential dilators (METRx, Medtronic, Minneapolis, Minnesota) were placed at the inferior aspect of the lamina under fluoroscopic control. A 14- to 18-mm working channel was introduced over the final dilator and attached to the table. The herniated disk was removed through the tubular retractor with microscopic magnification. Bony lamina removal was minimal, if necessary.	Microdiscectomy **N randomized**: 161; **N analyzed**: 159 (98.8%) in main study's primary analyses. 52w: 151 (93.8%) 2y: 144 (89.4%) 3y: 106 (65.8%) 4y: 102 (63.4%) 5y: 98 (60.9%); **N crossovers**: 0 (0%); Surgery was scheduled within 4w of first visit. A 25-30mm midline skin incision used followed by ipsilateral paravertebral muscle retraction. The herniated disk was removed by the unilateral transflaval approach with the aid of a headlight loupe or microscope magnification, depending on the surgeon's preference. Bony lamina removal was minimal, if necessary.
Brouwer (2015)[39] Brouwer (2017)[102]	Parallel-group RCT; The Netherlands; Some concerns	Government; Healthcare Insurance Board of the Netherlands	Percutaneous laser disc decompression **N randomized**: 57; **N analyzed**: 55 (96.5%); **N crossovers**: Unclear; CT–guided treatment was performed with the patient in prone position under local anesthesia. An 18-G needle was placed centrally in the nucleus pulposus and parallel to the end plates by means of a posterolateral approach. Through the needle, a glass fiber of 600 micron was advanced into the disc, enabling the application of laser energy (diode laser; Biolitec Inc, East Longmeadow, MA, USA; 980 nm, 7 W, 0.6-second pulses, and an interval of 1 second) for a total energy delivered of 1,500 J (2,000 J for L4–L5 level).	Microdiscectomy **N randomized**: 58; **N analyzed**: 57 (98.3%); **N crossovers**: 0 (0%); A discectomy performed under general or spinal anesthesia using loupe magnification or microscope depending on the surgeon's preference. The aim of the surgery was to remove the herniated disc fragment, without any attempt to remove the disc itself, using a unilateral transflaval approach.

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Chatterjee (1995)[38]	Parallel-group RCT; United Kingdom; Some concerns	Government; The Department of Health, London, UK	Automated Percutaneous Lumbar Discectomy **N randomized**: 31; **N analyzed**: 31 (100%); **N crossovers**: 0 (0%); Procedure was performed with a 2-mm nonflexible automated suction nucleotome (Surgical Dynamics, San Leandro, California) under local anesthesia and with biplanar radiologic control. It was necessary to achieve a position that was either exactly central within the disc or slightly posterior to center before disc aspiration was commenced. Disc aspiration was continued until no more nuclear material could be obtained.	Microdiscectomy **N randomized**: 40; **N analyzed**: 40 (100%); **N crossovers**: 0 (0%); Microdiscectomy was performed by standard technique via a 2-cm incision and a transligamentous approach with the removal of not only the herniated portion of the disc but also with clearance of all loose intradiscal material.
Erginousakis (2011)[37]	Parallel-group RCT; Greece; High	Reported as NOT industry supported.	Percutaneous disc decompression **N randomized**: 31; **N analyzed**: 31 (100%); **N crossovers**: 0 (0%) in main analysis; 20 received subsequent microdiscectomy Intervertebral disc decompression using a 17-gauge Dekompressor (Stryker, Kalamazoo, MI) was performed with fluoroscopic guidance. Approximately 1–3 grams of disc material and 1 milliliter of tissue has been removed once the tissue becomes visible at the collection chamber entrance.	Conservative Management **N randomized**: 31; **N analyzed**: 31 (100%); **N crossovers**: 0 (0%); A 6-week course of monitored and registered conservative therapy during which participants received analgesics, anti-inflammatory drugs, muscle relaxants, and physiotherapy. It also included education and counseling.

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Franke (2009)[36]	Parallel-group RCT; Germany; Some concerns	NR	Microscopically assisted percutaneous nucleotomy **N randomized**: 52; **N analyzed**: 52 (100%); **N crossovers**: 0 (0%); The level localization with a spinal needle was done on the opposite side. The pinpoint was directed at the open interlaminar window. The skin incision of 15 mm at the side of the pathology was performed in height of the needle entry point approximately 2 cm paramedian. Both the thoracolumbar fascia and the paraspinal muscles were dilated till the working channel could be brought in; surgery performed under direct vision via a microscope	Microdiscectomy **N randomized**: 48; **N analyzed**: 48 (100%); **N crossovers**: 0 (0%); Procedure not described.
Gerszten (2010)[41]	Parallel-group RCT; United States; Some concerns (6w outcomes) High (12w and later outcomes)	Commercial; ArthroCare Corp.	Plasma disc decompression with coblation technology (PDD) **N randomized**: 46; 45 ITT sample; **N analyzed**: 29 (64% of ITT sample) at 26w; **N crossovers**: 12 were unresolved and received a second, unspecified procedure; Procedure performed on an outpatient basis using the Coblation DLR or DLG SpineWand surgical device (ArthroCare Corp.). Procedure conducted under fluoroscopic guidance. A 17G spinal cannula was introduced into disc using a posterolateral extrapedicular approach and position at the junction of the annulus and nucleus. The SpineWand was introduced through the cannula and positioned within the nucleus then placed in ablation mode and advanced and retracted to create a total of 6 channels.	Epidural steroid injection (ESI) **N randomized**: 44; 40 ITT sample; **N analyzed**: 28 (70% of ITT sample) at 26w; **N crossovers**: 8 were unresolved and received a second, unspecified procedure; Procedure performed under fluoroscopic guidance. The location was determined by the treating physician with the goal of delivering steroids to the site of the disc protrusion and nerve irritation. A transforaminal approach was used. Medication type and dose were left to the discretion of the treating clinician. A second ESI procedure was allowed by the study protocol.

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Haines (2002)[42]	Parallel-group RCT; United States; High	Government; National Institute of Neurological Disorders and Stroke and the Agency for Health Care Research and Quality; LAPDOG	Automated percutaneous discectomy, endoscopic percutaneous discectomy (APD/EPD) **N randomized**: 21; **N analyzed**: 17 (81.0%) at 26w; **N crossovers**: 0 (0%); When this study was conceived there was a single manufacturer of the only device specifically designed for APD (Nucleotome1, Surgical Dynamics, Alameda, CA, USA). Surgical Dynamics had produced an explicit protocol specifying inclusion and exclusion criteria as well as surgical technique. This was adhered to in the APD group. With technology advancement and diffusion, equipment for endoscopic approach became available from several manufacturers and indications for the procedure became less specific. Thus, the intervention protocol was modified after enrollment of the first 26 participants to incorporate epidural endoscopic technique.	Discectomy **N randomized**: 13; **N analyzed**: 10 (76.9%) at 26w; **N crossovers**: 1 (7.69%); Discectomies were done according to the surgeon's usual technique.
Henriksen (1996)[35]	Parallel-group RCT; Denmark; Some concerns	NR	Microdiscectomy **N randomized**: 40; **N analyzed**: 39 (97.5%); **N crossovers**: 0 (0%); The patients were operated in the prone position. A midline skin incision approximately 7 cm long was used over the appropriate disc, and a 7 cm long fascial incision was made. Additional details provided but not clear whether they pertain to SG1 or SG2.	Discectomy **N randomized**: 40; **N analyzed**: 40 (100%); **N crossovers**: 0 (0%); The patients were operated in the prone position. A midline skin incision approximately 7 cm long was used over the appropriate disc, and a 3 cm long fascial incision was made. Additional details provided but not clear whether they pertain to SG1 or SG2. For example, 1/3 of the facet joint was removed. An editorial note associated with the study raised concerns about whether the procedure performed in SG2 was what most would consider a microsurgical discectomy.

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Hermantin (1999)[43]	Parallel-group RCT; United States; Some concerns	NR	Video-assisted arthroscopic microdiscectomy **N randomized** 30; **N analyzed** 30 (100%); **N crossovers** 0 (0%); Performed with the use of an oval 5X8mm cannula introduced through a universal cannula by an 18G needle under fluoroscopy, positioned in the triangular working zone bordered anterolaterally by the exiting nerve root, medially by the traversing nerve root and dura, and caudally by the vertebral plate of the caudad lumbar segment. Herniated disc fragments are pulled back into the intervertebral disc space and then are withdrawn.	Discectomy, with laminotomy **N randomized**: 30; **N analyzed**: 30 (100%); **N crossovers**: 0 (0%); The open laminotomy and discectomy was performed in a standard fashion. A four-centimeter posterior midline incision was made, and a small laminotomy and discectomy was performed at the specified level.
Huang (2005)[24]	Parallel-group RCT; Taiwan; Some concerns	Other; Research Committee of Chang Gung Memorial Hospital, Taiwan	Microendoscopic discectomy **N randomized**: 10; **N analyzed**: 10 (100%); **N crossovers**: 0 (0%); The Vertebroscope System (Zeppelin, Pullach, Germany) was used to perform the endoscopic discectomy procedure.	Discectomy **N randomized**: 12; **N analyzed**: 12 (100%); **N crossovers**: 0 (0%); Intervention not described.
Malter (1996)[44]	Cost-effectiveness Analysis; United States; Fair (See Appendix F, Table F-6)	Government; Agency for Health Care Policy and Research, Seattle VA Medical Center, National Research Service Award	Discectomy **N randomized:** NA	Conservative management **N randomized:** NA

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Mayer (1993)[34]	Parallel-group RCT; Germany; High	NR	Percutaneous endoscopic discectomy **N randomized**: 20; **N analyzed**: 20 (100%); **N crossovers**: 3 (15%); The procedure is performed with the patient under local anesthesia, using approach is as described by Day and Nazarian. Under fluoroscopic control, the tip of a 18G cannula is advanced to the center of the disc. Discography is performed to confirm the indication for percutaneous endoscopic discectomy. A guidewire is advanced through the cannula until the tip reaches the center of the disc; the cannula is then removed. With the wire as a guide, a blunt tapered trocar is advanced to the posterolateral border of the annulus fibrosus through a stab incision. The trocar, in turn, serves as a guide for the introduction of the working cannula (outer diameter 5 mm). Following introduction of the working cannula, the trocar is removed and the disc is entered by cutting a circular window in the annulus fibrosus with the aid of a trephine. Rigid forceps are introduced to remove a small amount of nucleus pulposus from the center of the disc to create a cavity before introducing the endoscope. Endoscopy of the disc is performed with a rigid endoscope. The herniated part of the nucleus pulposus can be removed using reverse-opening forceps as well as flexible forceps. A bilateral approach is used for continuous endoscopy during removal of disc herniations located in the midline.	Microdiscectomy **N randomized**: 20; **N analyzed**: 20 (100%); **N crossovers**: 0 (0%); Intervention not described.

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
McMorland (2010)[23]	Parallel-group RCT; Canada; Some concerns	Other; Supported by a grant from the Foundation for Chiropractic Education and Research.	Microdiscectomy **N randomized**: 20; **N analyzed**: 20 (100%) (outcomes reported only to 12w for ITT analysis) 24w 20 (100%) 52w 15 (75%); **N crossovers**: 3 (15%) enrolled in spinal manipulation 26-34w after surgery so received both interventions. Note, this crossover happened AFTER all ITT outcomes are reported (12w); Surgical microdiscectomies were performed with patients in a prone position supported by bolsters through a standard midline lumbar incision. All procedures were undertaken using microsurgical techniques with the aid of an operating microscope. Laminotomies were created as required at the level of the lumbar disc herniation. Both sequestrectomy and intra-annular discectomy were performed to ensure adequate nerve root decompression.	Spinal manipulation **N randomized**: 20; **N analyzed**: 20 (100%) (outcomes reported only to 12w for ITT analysis) 24w 20 (100%) 52w 17 (85%); **N crossovers**: 8 (40%) underwent microdiscectomy after 12w of spinal manipulation care. Note, this crossover happened AFTER all ITT outcomes are reported (12w); All spinal manipulative therapies were provided by a single chiropractic doctor. Spinal manipulative therapy consisted of side posture, high-velocity, low-amplitude, short lever technique. The decision to administer manual spinal manipulation on each visit was based on that patient's ability to tolerate the position. Cryotherapy or thermotherapy (ice or heat) were used on an "as-needed" basis. All patients were provided with an information/education package and were introduced to rehabilitative exercises. The patients also participated in a supervised rehabilitative (core stability) exercise regimen. Treatments typically required 2 to 3 visits per week for the first 4 weeks reducing to 1 to 2 visits per week for the next 3 to 4 weeks. At the 8-week mark, follow-up visits were scheduled based on the patient's symptoms until the patient's symptoms were deemed stable (i.e., no deterioration or flare up) with a 2-month treatment holiday. Mean number of treatment session was 21 plus an additional 6 supervised rehabilitation sessions over 52 weeks.

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
North (2005)[46]	Parallel-group RCT; United States; High	Commercial; Medtronic, Inc.	Repeat lumbosacral decompression **N randomized**: 26; **N analyzed**: 26 (100%); **N crossovers**: 14 (54%); Laminectomy (N=23) and/or foraminotomy (N=21) and/or discectomy(N=6) with or without fusion (N=3), with or without instrumentation (N=6). Patients randomized to reoperation could cross over to spinal cord stimulation after a 6-month postoperative period.	Spinal cord stimulation **N randomized**: 24; **N analyzed**: 19 (79.2%); **N crossovers**: 5 (20.8%); Percutaneous placement of a temporary electrode (3487A Pisces-Quad; Medtronic, Inc., Minneapolis, Minnesota) for a therapeutic trial lasting at least 3 days. The spinal cord stimulation patients could receive a permanent implant (3487A-56 or 3587A Resume electrode, X-trel or Itrel pulse generator; Medtronic, Inc.) if they reported at least 50% estimated relief of pain by standard pain rating methods and demonstrated stable or improved analgesic medication intake, with improved physical activity commensurate with neurological status and age. Patients randomized to spinal cord stimulation who did not meet these criteria could immediately cross over to reoperation.

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Osterman (2003)[33]	Parallel-group RCT; Finland; High	Government; Finnish Office for Health Technology Assessment at National Research and Development Centre for Welfare and Health, Jorvi Hospital, Helsinki and Uusimaa Hospital District, Espoo, Finland.	Microdiscectomy **N randomized**: 28; **N analyzed** 6w 26 (93%) 12w 26 (93%) 26w 26 (93%) 52w 21 (75%) 2y 26 (93%); **N crossovers**: 0 (0%); A microdiscectomy by a spinal orthopedic surgeon was performed within 2 weeks of randomization. The operation was carried out under general anesthesia in a genupectoral position with fluoroscopic control of the spinal level before draping. The patients were usually discharged from the hospital on the second or third postoperative day. Sick leave and analgesia were prescribed according to individual requirements. Surgical patients were advised to continue with isometric exercises while waiting for the operation and after discharge from the hospital. At follow-up visits, this group received active physiotherapeutic instructions, including stretching, bending, and muscle strengthening exercises. Passive forms of treatment were not recommended.	Physiotherapy **N randomized**: 28; **N analyzed** 6w 26 (93%) 12w 26 (93%) 26w 22 (78.6%) 52w 20 (71.4%) 2y 24 (86%); **N crossovers**" 11 (39.3%) Note: 3 were < 6w, 4 were between 6w and 12 w, 3 were between 12w and 26w, and 1 between 26w and 52w; The control group received physiotherapeutic instructions initially and continued with isometric exercises after randomization. At follow-up visits, as with the surgical group, activity was encouraged. Patients in the control group were informed of symptoms meriting operation, and they were advised to contact the treating physician if the symptoms should get worse.

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98]	Parallel-group RCT; The Netherlands; High	Government; Supported by a grant from the Netherlands Organisation for Health Research and Development (ZonMW) and the Hoelen Foundation, The Hague; Sciatica Trial	Microdiscectomy **N randomized**: 141; **N analyzed** 52w: 140 (99.3%) 2y: 130 (92.2%) 5y: 115 (81.6%); **N crossovers** 52w: 16 (11.3%) 2y: 16 (11.3%) 5y: 16 (11.3%); Surgery was scheduled within 2 weeks after assignment. A minimal unilateral transflaval approach with magnification. The goal was to decompress the nerve root and reduce the risk of recurrent disk herniation by performing an annular fenestration, curettage, and removal of loose degenerated disk material from the disk space with the use of a rongeur, without attempting to perform a subtotal discectomy.	Conservative management **N randomized**: 142; **N analyzed** 52w: 141 (99.3%) 2y: 130 (91.5%) 5y: 116 (81.7%); **N crossovers** 52w: 55 (38.7%) 2y: 62 (43.7%) 5y: 66 (46.5%); General practitioners provided prolonged conservative treatment. Invitation to website for education on natural course of their illness and the expectation of successful recovery, irrespective of the initial intensity of their pain. Treatment aimed at enabling resumption of daily activities. Pain medication as needed. Patients who were fearful of moving were referred to a physiotherapist. If sciatica persisted for 6 months after the patient underwent randomization, microdiscectomy was offered. Patients who had increasing leg pain not responsive to medication or progressive neurologic deficits were offered surgery earlier than 6 months after randomization.

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Ruetten (2008)[31]	Parallel-group RCT; Germany; High	Other; Study reports that no funds or benefits were received or will be received in support of this work.	Endoscopic (interlaminar or transforaminal) discectomy **N randomized**: 100; **N analyzed**: 91 (91%); **N crossovers**: 0 (0%); All the operating instruments and optics were products supplied by WOLF (Richard Wolf GmbH, Knittlingen, Germany). The full-endoscopic transforaminal procedure was used for extraforaminal and intraforaminal herniations. The full-endoscopic interlaminar operation was performed for herniations mainly in the spinal canal. Sequestrotomy alone was performed in small or covered annular defects when the sequestered disc material exceeded the level of the intervertebral space toward cranial or caudal. This occurred in 39 participants.	Microdiscectomy **N randomized**: 100; **N analyzed**: 87 (87%); **N crossovers**: 0 (0%); The conventional discectomy was performed with paramedian or lateral access in known standardized technique using a microscope. Sequestrotomy alone was performed in small or covered annular defects when the sequestered disc material exceeded the level of the intervertebral space toward cranial or caudal. This occurred in 43 participants.

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Ruetten (2009)[47]	Parallel-group RCT; Germany; High	NR	Revision endoscopic discectomy **N randomized**: 50; **N analyzed**: 45 (90%); **N crossovers**: 0 (0%); Used either the transforaminal (TF) or interlaminar (IL) approach. Osseous resection was required in 6% of cases. The TF procedure was performed with access as lateral as possible. A spinal cannula is inserted via the 6mm skin incision. After insertion of a lead wire, the cannulated dilator is pushed in and a surgical sheath is placed. Decompression is performed. If the anatomic osseous diameter of the intervertebral foramen does not permit direct entry into the spinal canal, the opening is expanded. An extraforaminal approach is made at the caudal pedicle in cases where the position of the exiting nerve is not clear. The IL operation was performed using a dilator inserted bluntly to the lateral edge of the interlaminar window and an operation sheath directed toward the ligamentum flavum. The medial edge of the descending facet is located and prepared directly on the bone toward ventral until the medial edge of the ascending facet is visible. Blunt penetration to the floor of the spinal canal and preparation of the ventral epidural space. Bone resection to expand the interlaminar window to enable penetration into the spinal canal with the endoscope is usually not necessary owing to the resection during the primary operation. Sequestrotomy alone was performed in small or covered annular defects when the sequestered disc material has exceeded the level of the intervertebral space. All the operating instruments and optics were products supplied by WOLF (Richard Wolf GmbH, Knittlingen, Germany).	Revision microdiscectomy **N randomized**: 50; **N analyzed**: 42 (84%); **N crossovers**: 0 (0%); The conventional microsurgical operations were performed with paramedian access in known standardized technique using a microscope. Sequestrotomy alone was performed in small or covered annular defects when the sequestered disc material has exceeded the level of the intervertebral space toward cranial or caudal. Osseous resection was required in 94% of cases.

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Ryang (2008)[30] Gempt (2013)[103]	Parallel-group RCT; Germany; High	NR	Trocar microdiscectomy **N randomized**: 30; **N analyzed**: unclear; **N crossovers**: 0 (0%); The skin was incised horizontally over a length of 4 to 5 cm on the affected side after localization of the interlaminar space with lateral x-ray fluoroscopy. The lumbodorsal fascia was incised vertically over a distance of 4 to 5 cm, 0.5 cm paramedially. The paraspinal musculature was partially detached from the hemilamina in a subperiostal fashion, the interlaminar space was visualized, and the retractor placed into position. A second fluoroscopy was obtained to confirm the correct level. The operating microscope (Carl Zeiss Co., Oberkochen, Germany) was put into position and the remaining operation performed in the standard microsurgical fashion with bayoneted microsurgical instruments. Partial hemilaminectomy of the superior and inferior lamina and medial facetectomy, with partial flavectomy were carried out to visualize the compromised nerve root. The herniated sequester was removed. A partial nucleotomy was performed in some cases.	Microdiscectomy **N randomized**: 30; **N analyzed**: unclear; **N crossovers**: 0 (0%); After localization of the interlaminar space, a skin incision measuring 1.6 cm in length was performed 1.5 cm paramedially. The lumbodorsal fascia was bluntly dissected and the trocar, together with the enclosed mandrin, was gently screwed into the paraspinal muscles until the interlaminar window was reached, with the tip of the mandrin pointing medially. In this way, the paraspinal muscle attachments to the laminae and spinous processes could be well preserved in their full integrity. The mandrin was removed and the handle attached to the trocar. After a second fluoroscopy, surgery was performed with the aid of an operative microscope. After exposure of the interlaminar space, a minimal interlaminar fenestration was performed by use of drill of different size Kerrison punches, but only if necessary. Minimal partial flavectomy and bony resection. The nerve root was retracted medically and herniated disc material removed, and if necessary a partial discectomy performed.
Sasaoka (2006)[25]	Parallel-group RCT; Japan; High	NR	Microendoscopic discectomy **N randomized**: 15; **N analyzed**: unclear; **N crossovers**: 0 (0%); Intervention not described.	Microdiscectomy **N randomized**: 11; **N analyzed**: unclear; **N crossovers**: 0 (0%); Intervention not described.

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Teli (2010)[29]	Parallel-group RCT; Italy; Some concerns	NR	Microendoscopic discectomy **N randomized**: NR; **N analyzed**: 70; **N crossovers**: 0 (0%); The Metr'X system (Medtronic Sofamor Danek, Memphis, USA) with a 16- or 18-mm tubular retractor was used. Laminotomy, medial facetectomy when needed and nerve root retraction followed by discectomy were performed.	Microdiscectomy **N randomized**: NR; **N analyzed**: 72; **N crossovers**: 0 (0%); Microdiscectomy with use of a surgical microscope. Laminotomy, medial facetectomy when needed and nerve root retraction followed by discectomy were performed. Discectomy **N randomized**: NR; **N analyzed**: 70; **N crossovers**: 0 (0%); Open discectomy with use of a magnifying loop. Laminotomy, medial facetectomy when needed and nerve root retraction followed by discectomy were performed.

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Thome (2005)[28] Barth (2008)[104]	Parallel-group RCT; Germany; Some concerns	NR	Sequestrectomy **N randomized**: 42; **N analyzed**: 42 (100%); **N crossovers**: 0 (0%); The spinal canal harboring the sequestrated disc material was exposed by performing a minimal interlaminar fenestration in cases of nondislocated or caudally herniated discs. In cases of cranially positioned herniated discs, a translaminar approach was undertaken, if possible. Thus, minimal removal of bone and articular structures was achieved by individualization of the procedure according to the preoperative MR imaging–depicted anatomy. In the sequestrectomy-treated group, only the herniated material was removed and the intervertebral space was not entered.	Microdiscectomy **N randomized**: 42; **N analyzed**: 42 (100%); **N crossovers**: 0 (0%); The spinal canal harboring the sequestrated disc material was exposed by performing a minimal interlaminar fenestration in cases of nondislocated or caudally herniated discs. In cases of cranially positioned herniated discs, a translaminar approach was undertaken, if possible. Thus, minimal removal of bone and articular structures was achieved by individualization of the procedure according to the preoperative MR imaging–depicted anatomy. In the microdiscectomy-treated group, the removal of the herniated material was followed by scalpel incision of the annulus fibrosus and resection of discal tissue from the intervertebral space—particularly the (degenerated) nucleus—with rongeurs.
Tullberg (1993)[27]	Parallel-group RCT; Sweden; Some concerns	NR	Microdiscectomy **N randomized**: 30; **N analyzed**: 29 (97%); **N crossovers**: 0 (0%); A similar dissection and disc herniation removal technique was used in both groups, the only difference being the operating microscope was used in the microdiscectomy group. Dissection involved the space between two adjacent vertebrae. The average length of the skin incision was 3.5 cm (range 2.5 to 4.5 cm). The area of exposed ligamentum flavum and laminae were the same in methods. The disc was opened by sharp dissection and all the material was removed with rongeurs. Finally, the exposed nerve root was covered by a fat graft intended to prevent scar formation.	Discectomy **N randomized**: 30; **N analyzed**: 29 (97%); **N crossovers**: 0 (0%); A similar dissection and disc herniation removal technique was used in both groups, the only difference being the operating microscope was NOT used in the discectomy (standard procedure group).

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
Weber (1983)[26]	Parallel-group RCT; Norway; High	Commercial Norsk Hydro A/S	Discectomy **N randomized**: 60; **N analyzed**: 60 (100%); **N crossovers**: 1 (1.7%); After removal of the ligamentum flavum and in most cases a small resection of the edge of the vertebral arch above and below the exposed intervertebral space, the herniated mass of cartilage was removed extradurally. Excochleation of the disc was then performed.	Conservative management **N randomized**: 66; **N analyzed**: 66 (100%); **N crossovers**: 17 (25.8%); Bed rest, physiotherapy and medication for an average of 6w at a rehabilitation hospital.
Weinstein (2006)[22] Weinstein (2008)[99] Lurie (2014)[100] SPORT	Parallel-group RCT; United States; High	Government; National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS) and the Office of Research on Women's Health, National Institutes of Health; and by the National Institute of Occupational Safety and Health, U.S. Centers for Disease Control and Prevention; Spine Patient Outcomes	Discectomy/microdiscectomy **N randomized**: 245; **N analyzed** 232 in main study's primary analyses. 52w: 202 (82.4%) 2y: 187 (76.3%) 3y: 180 (73.5%) 4y: 157 (64.1%) 8y: 157 (64.1%); **N crossovers** Cumulative crossovers over time: 6w: 171 (69.8%) 12w: 130 (53.0%) 26w: 113 (46.1%) 52w: 107 (43.7%) 2y: 105 (42.9%) 3y: 103 (42.0%) 4y: 101 (41.2%) 8y: 97 (39.6%);	Conservative management **N randomized**: 256; **N analyzed** 240 included in main study's primary analyses. 52w: 213 (83.2%) 2y: 191 (74.6%) 3y: 170 (66.4%) 4y: 159 (62.1%) 8y: 152 (59.4%); **N crossovers** Cumulative crossovers over time: 6w: 44 (17.2%) 12w: 71 (30%) 26w: 93 36.3%) 52w: 103 (40.2%) 2y: 107 (41.8% 3y: 111 (43.4%) 4y: 115 (44.9%) 8y: 122 (47.7%);

(continued)

Table D-1. Study characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Study Design; Country; Risk of Bias	Study Sponsor; Study Sponsor Name; Trial Name (if applicable)	Surgical Intervention; N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description	Comparator(s); N randomized; N analyzed (% of randomized); N crossovers (% of randomized); Intervention Description
		Research Trial (SPORT)	Standard open discectomy with examination of the involved nerve root. The procedure agreed on by all participating centers was performed under general or local anesthesia, with patients in the prone or knee-chest position. Surgeons were encouraged to use loupe magnification or a microscope. Using a midline incision reflecting the paraspinous muscles, the interlaminar space was entered as described by Delamarter and McCullough. In some cases, the medial border of the superior facet was removed to provide a clear view of the involved nerve root. Using a small annular incision, the fragment of disk was removed as described by Spengler. The canal was inspected and the foramen probed for residual disk or bony pathology. The nerve root was decompressed, leaving it freely mobile.	The nonoperative treatment group received usual care, with the study protocol recommending that the minimum nonsurgical treatment include at least active physical therapy, education/counseling with home exercise instruction, and nonsteroidal anti-inflammatory drugs, if tolerated. Other nonoperative treatments were listed, and physicians were encouraged to individualize treatment to the patient; all nonoperative treatments were tracked prospectively.

Abbreviations: cm = centimeter; CT = computed tomography; ITT = intention-to-treat; mm = millimeter; N = number; NA = not applicable; NR = not reported; NS = not significant; RCT = randomized controlled trial; SG = surgical group; NS = nonsurgical group; w = week(s); y = year(s).

Table D-2. Population characteristics of included studies

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48]	356; 328	Inclusion: Age 18 to 70 years with sciatica due to lumbar disk herniation, which lasted more than 6 to 8w and refractory to conservative treatment; nerve root compression was confirmed with MRI. Exclusion: <1/3 of spinal canal diameter disc herniation with doubtful nerve root compression, cauda equina syndrome, previous spinal surgery at the same disc level, spondylolisthesis, central canal stenosis.	**Age** SG1: 41.6 (9.8) SG2: 41.3 (11.7) **Female** SG1: 82 (49%) SG2: 71 (45%) **Nonwhite** SG1: NR SG2: NR	SG1: 110 (66%) SG2: 103 (65%) Defined as: Sick leave from work; NR	**Duration of symptoms, mean (SD) in weeks** SG1: 29.2 (47.4) SG2: 27.8 (23.3)
Brouwer (2015)[39] Brouwer (2017)[102]	NR; 115	Inclusion: Patients between 18 and 70 years with sciatica that was refractory to conservative management for more than 6 to 8 weeks; MRI confirmation of disc herniation at the corresponding level and the herniated fragment was smaller than one-third of the spinal canal. Exclusion: Patients with cauda equina syndrome, previous spinal surgery at the same disc level, lytic or degenerative spondylolisthesis, sequestered disc herniation, disc height less than 7 mm or central canal stenosis.	**Age** SG1: 43.2 (11.8) SG2: 43.7 (9.7) **Female** SG1: 19 (35%) SG2: 24 (42%) **Nonwhite** NR	SG1: 26 (49%) SG2: 31 (55%) Defined as: Sick leave from work; NR	**Duration of sciatica, median (range) in weeks** SG1: 30.0 (9 to 182) SG2: 26.0 (8 to 260)
Chatterjee (1995)[38]	NR; 71 Note, the study originally planned to enroll 160 participants but was halted early because of inferiority of one treatment arm.	Inclusion: Radicular pain as dominant symptom, conventional conservative therapy for a minimum of 6 weeks, MRI-confirmed contained lumbar disc herniation at a single level, the height of which was less than 30% of the sagittal canal size. Exclusion: Dominant symptom of low back pain, MRI-confirmed disc extrusions, sequestrations, subarticular or foraminal stenosis or multiple levels of herniation.	**Age** SG1: 38.9 (range 20 to 56) SG2: 41.3 (range 21 to 67) **Female** SG1: NR (51%) SG2: NR (40%) **Nonwhite** NR	NR; NR	**Duration of low back pain, mean (range) in weeks** SG1: 78.2 (8.7 to 191.2) SG2: 143.4 (8.7 to 260.7) **Duration of current episode of radicular pain, mean (range) in weeks** SG1: 13 (6 to 30) SG2: 20 (6 to 38)

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
Erginousakis (2011)[37]	NR; 62	Inclusion: Adults with sciatica and small- to medium-sized intervertebral disk herniation (occupying less than one-third of the canal diameter at magnetic resonance [MR] imaging) that was symptomatic (leg pain with or without back pain; leg pain greater than back pain when these two coexisted; lancinating, burning, stabbing, or electrical sensation of pain; straight leg raise limited to less than 30 degrees), with the symptoms consistent with the segmental level where herniation was seen at MR imaging, conservative therapy unsuccessful. Exclusion: No neurologic deficit. Response to a 6-week course of rigorous conservative treatment; untreatable coagulopathy; active, systemic, or local infections; herniation occupying more than one-third of the spinal canal diameter and noncorrelating pain. Degenerative disease of the intervertebral disc with a disc height reduction of more than 50%–60%.	**Age** SG1: 38 (4.2) NS1: 36 (5.8) **Female** SG1: 12 (38.7%) NS1: 14 (45.2%) **Nonwhite** NR	NR; NR	NR
Franke (2009)[36]	NR; 100	Inclusion: Disc dislocation grades 3 to 5 according to Kramer et al. Exclusion: lateral disc hernia, protrusions, cauda equina syndrome, coexisting severe lumbar canal stenosis, olisthesis, scoliosis greater than 10 degrees, kyphosis greater than 15 degrees, prior lumbar spine surgeries, malignant or inflammatory disease.	**Age** 44 (11.7) **Female** 40 (40%) **Nonwhite** NR	NR; NR	NR

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
Gerszten (2010)[41]	NR; 90	Inclusion: Adults between 18 and 75 years old, BMI less than 40, radicular pain score of 50 or greater as measured using a 0- to 100mm VAS, had received an epidural corticosteroid injection for the same symptoms between 3 weeks and 6 months previously with no or only partial relief and residual symptoms. Normal neurological function required and imaging evidence of a focal lumbar disc protrusion and disc height of more than 50% of that of the normal adjacent discs. In addition, the level and site of the disc protrusion had to correlate with pattern of pain. Exclusion: Extruded or sequestered disc herniation. Sciatica originating from more than one disc level, more severe axial (back) pain than radicular (leg) pain, cauda equina syndrome, progressive neurological deficit, radiological evidence of spondylolisthesis or moderate or severe stenosis at the level to be treated. History of previous spinal surgery at or adjacent to the level to be treated, spinal fracture, tumor or infection.	**Age** SG1: 46 (12) NS1: 42 (11) **Female** SG1: 24 (53%) NS1: 19 (48%) **Nonwhite** NR	NR; NR	**Duration of symptoms, mean (range)** SG1: 52w (4w to 16y) NS1: 2y (10w to 13y) P = 0.04 for SG1 vs NS1

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
Haines (2002)[42]	95; 34	Inclusion: Age 18 to 65 years with predominantly unilateral leg pain or paresthesiae, at least two of four objective signs (dermatomal sensory loss, myotomal weakness, appropriate reflex loss, appropriate nerve stretch test) and an imaging study confirming disc herniation Exclusion: No previous treatment for lumbar spinal disease, moderate or advanced lumbar spondylosis or central or lateral spinal stenosis, spondylolisthesis, progressive neurologic deficit or technical contraindications to the percutaneous procedure	**Age** SG1: 42.2 (12.0) SG2: 35.4 (10.1) **Female** SG1: 10 (47.6%) SG2: 5 (38.4%) **Nonwhite** SG1: 2 (9.5%) SG2: 1 (7.7%)	NR; NR	NR
Henriksen (1996)[35]	99; 80	Inclusion: Age 20 to 60 years who had conservative management including bed rest, analgesics, muscle relaxers and physical therapy without sufficient improvement, diagnostic studies included positive myelograms, and/or CT scans Exclusion: Obesity, prior back surgery, or symptoms from more than one nerve root	**Age** Median (IQR) SG1: 39.7 (30 to 46) SG2: 42.8 (36 to 48) **Female** SG1: 15 (38.5%) SG2: 14 (35%) **Nonwhite** NR	NR; NR	NR

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
Hermantin (1999)[43]	NR; 60	Inclusion: Single intracanalicular lumbosacral (other than L1/L2) disc herniation with associated radiculopathy; a herniation not exceeding one-half of the anteroposterior diameter of the spinal canal; an absence of central or lateral osseous or ligamentous stenosis; accessibility of the disc for both arthroscopic microdiscectomy and laminotomy; failure to respond to nonoperative measures for 14 weeks; more pain in the lower extremities than in the back; the presence of positive tension signs with or without an accompanying neurological deficit; a dermatomal distribution of pain in the lower extremities matching that seen on imaging studies and specific nerve-root involvement. Exclusion: previous operation on the low back, litigation or Workers' Compensation claim involving the disc herniation, central or lateral stenosis of the spinal canal, severe degenerative narrowing of the intervertebral disc space at the index level, evidence on imaging of global bulging of the intervertebral disc associated with central or lateral stenosis; a sequestered herniation that had migrated, a large central or extraligamentous herniation between L5/S1.	**Age** SG1: Mean 40 (range 18-67) SG2: Mean 39 (range 15-66) **Female** SG1: 13 (43.4%) SG2: 8 (26.7%) **Nonwhite** NR	NR; NR (this is presumably 0 (0%) as patients with any litigation or Workers' Compensation claim involving the disc herniation are excluded)	Minimum duration of nonoperative treatment prior to randomization was 14w in both groups.

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
Huang (2005)[24]	NR; 21 Note: 22 patients were enrolled, but only 21 underwent randomization; one patient insisted on assignment to SG2.	Inclusion: Age criteria NR, symptomatic herniated intervertebral discs who were scheduled to undergo elective lumbar discectomy; 16 patients failed to respond to conservative treatment after three months, and six patients with acute attack of intractable back and leg pain that demonstrated no improvement after 1-2 weeks of absolute bed rest. Exclusion: Recurrent lumbar disc herniation, significant motor deficit or sphincter disturbance	**Age** SG1: 39.2 (10.8) SG2: 39.8 (11.0) **Female** SG1: 4 (40%) SG2: 3 (25%) **Nonwhite** NR	NR; NR	NR
Malter (1996)[44]	2,175; NA	Patients younger than 65 years old identified from the MEDSTAT commercially-available database of non-governmental insurers from all 50 states. Inclusion: Diagnosis of a herniated intervertebral disc confirmed with imaging. Exclusion: NR	**Age** SG1: 46 (NR) NS1: 46 (NR) **Female** SG1: 208 (56) NS1: 829 (46) **Nonwhite** NR	NR; NR	NR

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
Mayer (1993)[34]	NR; 40	Inclusion: Clinical symptoms due to discogenic lumbar nerve root compression with radicular symptoms such as positive straight-leg raising test, sciatica, sensory disturbances, mild motor weakness, and/or reflex differences, failed conservative therapy, "contained" (when the outer border of the annulus fibrosus was still intact.) or small non-contained (extrusion of nucleus pulposus under the posterior longitudinal ligament but still at the level of the disc space and occupying not more than one-third of the sagittal diameter of the spinal canal), imaging confirmation. Exclusion: Severe motor deficits, conus or cauda equina syndrome, or rapidly progressing neurological symptoms, patients with signs of segmental instability or previous surgery at the same site, worker's compensation claims, malformations, tumors, or posttraumatic root compression, large "non-contained" disc herniations extending cranially or caudally to the level of the disc space, spinal stenosis, or spondylolisthesis.	**Age** SG1: 39.8 (10.4) SG2: 42.7 (10) **Female** SG1: 8 (40%) SG2: 6 (30%) **Nonwhite** NR	NR; NR (this is presumably 0 (0%) as patients with worker's compensation claims are excluded)	**Duration of symptoms, mean (SD) in weeks** SG1: 27.6 (NR) SG2: 29.2 (NR)

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
McMorland (2010)[23]	60; 40	Inclusion: Unilateral radiculopathy secondary to lumbar disc herniation at L3/4, L4/5, or L5/S1, with leg-dominant symptoms with objective signs of nerve root tethering ± neurologic deficit correlated with evidence of appropriate root compression on magnetic resonance imaging, failed at least 3 months of nonoperative management including treatment with analgesics, lifestyle modification, physiotherapy, massage therapy, and/or acupuncture. Exclusion: major neurological deficits (Cauda equina syndrome, rapidly progressing neurological symptoms (e.g. foot drop)), previous surgery at symptomatic level, concurrent treatment involving spinal manipulation at time of enrollment, prolonged use of systemic corticosteroids, osteopenia/osteoporosis, spondylolisthesis grade III or IV.	**Age** SG1 male: 42.85 (NR) SG1 female: 40.1 (NR) NS1 male: 36.4 (NR) NS1 female: 48.33 (NR) **Female** SG1: 7 (35%) NS1: 9 (45%) **Nonwhite** NR	SG1: 9 (45%) NS1: 11 (55%) Defined as: "Medical leave" for work status; SG1: 1 (5%) NS1: 1 (5%) **N (%) receiving 3rd party disability insurance** SG1: 0 (0%) NS1: 0 (0%)	**N (%) with duration of complaint 12-26w** SG1: 3 (15%) NS1: 6 (30%) **N (%) with duration of complaint 26-52w** SG1: 5 (25%) NS1: 6 (30%) **N (%) with duration of complaint >52w** SG1: 12 (60%) NS1: 8 (40%)

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
North (2005)[46]	99; 60 (randomized) 50 (treated) Of the 10 randomized but not treated, 9 did not receive authorization from Workers Compensation for study participation and 1 had a stroke prior to treatment). Patients with Worker's Compensation consented to randomization as often as other patients.	Inclusion: Surgically remediable nerve root compression and concordant complaints of persistent or recurrent radicular pain, with or without low back pain, after one or more lumbosacral spine surgeries. Pain refractory to conservative care, with concordant neurological, tension, and/or mechanical signs and imaging findings of neural compression. Exclusion: A disabling neurological deficit (e.g., foot drop, neurogenic bladder) in the distribution of a nerve root or roots caused by surgically remediable compression; radiographically demonstrated (by myelographic block or its magnetic resonance imaging equivalent) critical cauda equina compression; radiographic evidence of gross instability (spondylolisthesis or abnormal subluxation); unresolved issues of secondary gain; a chief complaint of axial (low back) pain exceeding radicular (hip, buttock, and leg) pain.	**Age** 52.0 (13.5) **Female** 26 (52%) **Nonwhite** NR	NR; Overall 15 (30%) receiving workers compensation	**Mean (SD) number of prior operations** SG1: 2.5 (1.1) NS1: 2.5 (1.1)

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
Osterman (2003)[33]	NR; 56	Inclusion: Age 20 to 50 years with 1) below knee radicular pain of 6 to 12 weeks' duration, 2) a CT finding of intervertebral disc extrusion or sequester, 3) at least one specific physical finding (a positive straight leg raising test, muscle weakness, altered reflexes, dermatomal sensory change). Exclusion: 1) previous back surgery, 2) spondylolisthesis, 3) symptomatic spinal stenosis, 4) over 3 months' continuous sick leave because of low back pain or leg pain, 5) a condition confounding evaluation of treatment outcomes (vascular claudication, symptomatic osteoarthritis, previous major trauma, diabetic polyneuropathy), or 6) a contraindication to conservative treatment (cauda equina syndrome, progressive neurologic deficit, or intolerable pain).	**Age** SG1: 37 (7) NS1: 38 (7) **Female** SG1: 13 (46.4%) NS1: 9 (32.1%) **Nonwhite** NR	NR; NR	**Duration of leg pain, mean (SD) in weeks** SG1: 11.0 (4.6) NS1: 8.6 (3.0) **Duration of back pain, mean (SD) in weeks** SG1: 13.4 (6.7) NS1: 10.4 (4.6)
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98]	599; 283	Inclusion: Eligible patients were 18 to 65 years of age, had a radiologically confirmed disk herniation, and had received a diagnosis from an attending neurologist of an incapacitating lumbosacral radicular syndrome that had lasted for 6 to 12 weeks with correlation of MRI findings with symptoms. Exclusion: Cauda equina syndrome, muscle paralysis, insufficient strength to move against gravity, occurrence of another episode of symptoms like those of the current episode during the previous 12 months, previous spine surgery, bony stenosis, and spondylolisthesis.	**Age** SG1: 41.7 (9.9) NS1: 43.4 (9.6) **Female** SG1: 52 (37%) NS1: 45 (32%) **Nonwhite** NR	SG1: 107 (76%) NS1: 116 (82%) Defined as: Sick leave from work; NR	**Duration of symptoms, mean (SD) in weeks** SG1: 9.43 (2.37) NS1: 9.48 (2.11)

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
Ruetten (2008)[31]	NR; 200	Inclusion: Clinically symptomatic disc herniation with radicular pain and neurologic deficits. Exclusion: NR	**Age** 43 (range 20 to 68) **Female** 116 (58%) **Nonwhite** NR	NR; NR	**Duration of symptoms, mean (range) in weeks** 11.71 (0.14 to 68) 162 of the 200 patients had received a mean of 9w of conservative treatment.
Ruetten (2009)[47]	NR; 100	Inclusion: Previous conventional discectomy with acute occurrence of radicular leg symptoms on the same side after a pain-free interval and who showed a recurrent disc herniation in the same level with MRI. Inclusion criteria specific for the full endoscopic transforaminal access were (1) sequestering of material located cranially below the lower edge of the cranial pedicle or caudally not over the middle of the caudal pedicle and (2) lateral radiologic evidence that the foramen was not overlaid by the pelvis beyond the middle of the cranial pedicle. Exclusion: NR	**Age** 39 (range 23 to 59) **Female** 44 (44%) **Nonwhite** NR	NR; NR	**Duration of symptoms, mean (range) in weeks** 9.85 (0.14 to 56)

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
Ryang (2008)[30] Gempt (2013)[103]	NR; 60	Inclusion: 1) single level virgin lumbar disc herniation; 2) typical monoradicular symptoms attributable to the involved lumbar segment with predominant sciatica compared to less severe lower back pain; and 3) failure of 8 to 12 w of conservative treatment, intolerable sciatica, or rapidly progressive neurological deficits including motor deficits, bladder dysfunction, and cauda equina syndrome. Exclusion: (1) history of previous lumbar back surgery or conservatively treated lumbar disc herniation at adjacent levels; (2) signs of spinal instability or other spinal abnormalities such as bone disease, spinal infection, malignancy, or signs of spinal canal stenosis on computed tomography or magnetic resonance imaging and neurogenic claudication; (3) intra- and extraforaminal far lateral disc herniation; (4) chronic pain syndrome and opioid abuse; (5) pending worker's compensation.	**Age** SG1: 38.2 (9.3) SG2: 39.1 (11.3) **Female** SG1: 17 (56.7%) SG2: 11 (36.7%) **Nonwhite** NR	NR; NR	NR
Sasaoka (2006)[25]	NR; 33	NR	**Age** 42.4 (range 20 to 72) SG1: 36.5 (range 25 to 60) SG2: 37.7 (range 20 to 58) **Female** 14 (42.4%) SG1: 9 (60.0%) SG2: 3 (27.3%) **Nonwhite** NR	NR; NR	NR

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
Teli (2010)[29]	NR; 240	Inclusion: Aged 18 to 65 years, symptomatic, single level posterior lumbar disc herniation with diagnosis made by spine specialists (orthopaedic and neurosurgeons) with pain and/or neurological signs in concordant distribution lasting at least 6 weeks despite appropriate conservative treatment consisting of systemic drugs for pain relief and/or epidural steroid administration, imaging confirmation with MRI or a CT scan of the lumbar spine, supplemented with plain X-rays of the lumbar spine including the thoracolumbar tract to exclude or confirm the presence of a segmentation anomaly. Exclusion: cauda equina symptoms, foraminal or extra-foraminal herniations, cervical or lumbar spine stenosis of any etiology, malignancy, previous spine surgery, spinal deformity including spondylolisthesis, concurrent infection and rheumatic disease.	**Age** 39.3 (range 27 to 61) **Female** 73 (34.4% of N analyzed) **Nonwhite** NR	NR; NR	**Duration of pain, mean (SD) in weeks** SG1: 11 (5) SG2: 12 (6) SG3: 11 (5)
Thome (2005)[28] Barth (2008)[104]	221; 84	Inclusion: Age between 18 and 60 years, with MRI- documented intraspinal (not extraforaminal) disc fragment that had perforated the annulus fibrosus. Exclusion: Previous lumbar spine surgery, emergency indication for surgery, MRI-documented lumbar spinal stenosis or spondylolisthesis.	**Age** SG1:42 (9) SG2: 40 (10) **Female** SG1: 18 (42.9%) SG2: 19 (45.2%) **Nonwhite** NR	NR; NR	**Duration of symptoms, mean (SD) in weeks** SG1: 11 (12) SG2: 8 (10) P=0.27 for SG1 vs SG2

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
Tullberg (1993)[27]	NR; 60	Inclusion: Sciatica, failed 2 months of conservative therapy and had CT-confirmed diagnosis of single lumbar disc herniation within the spinal canal Exclusion: Recurrent disc herniation or previous back surgery	**Age** SG1: 40 (range 17 to 59) SG2: 38 (range 18 to 64) **Female** SG1: 12 (40%) SG2: 9 (30%) **Nonwhite** NR	**N (%) with specified duration of disablement:** SG1 none: 6 (NR) <4w: 5 (NR) 4w to 12w: 8 (NR) 16w to 26w: 5 (NR) 27w to 52w: 5(NR) >1y: 1 (NR) SG2 <4w: 0 (NR) 4w to 12w: 6 (NR) 16w to 26w: 7 (NR) 27w to 52w: 13 (NR) >1y: 4 (NR) Defined as: Sick leave from work; NR	**N (%) with specified duration of symptoms**: SG1 <4w:1 (NR) 4w to 12w: 7 (NR) 16w to 26 w: 2 (NR) 27w to 52w: 14 (NR) >52w: 6 (NR) SG2 <4w: 0 (NR) 4w to 12w: 6 (NR) 16w to 26w: 7 (NR) 27w to 52w: 13 (NR) >1y: 4 (NR)
Weber (1983)[26]	NR; 126	Inclusion: Patients admitted to hospital with sciatica, with continued radicular pain provoked by moderate exercise, sitting position, or increased abdominal pressure after an initial 14d period of observation, radiographic confirmation of definite or possible disc herniation based on radiculography with water soluble contrast. Exclusion: Definite indications for surgery (intolerable pain, suddenly occurring or progressive muscle weakness, bladder or rectum paresis), severe or immobile scoliosis; patients with moderate symptoms but who showed signs of continuous improvement because of bed rest, physiotherapy, or medication during the 14d observation period were also excluded.	**Age** SG1: 40.0 (NR) NS1: 41.7 (NR) **Female** SG1: 28 (46.7%) NS1: 30 (45.5%) **Nonwhite** NR	NR; NR	NR

(continued)

Table D-2. Population characteristics of included studies (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	N eligible; N randomized	Population Eligibility	Age, Mean (SD); Women, N (%); Nonwhite, N (%)	Disabled, N (%); Disability Benefits, N (%)	Duration of Symptoms
Weinstein (2006) [22] Weinstein (2008)[99] Lurie(2014)[100] SPORT	1991; 501	Inclusion:18 years and older and diagnosed as having intervertebral disk herniation (confirmed by MRI or CT) and persistent symptoms (radicular pain and evidence of nerve-root irritation with a positive nerve-root tension sign or corresponding neurologic deficit) despite nonoperative treatment for >6 weeks. Exclusion: Prior lumbar surgery, cauda equina syndromes, scoliosis greater than 15 degrees, segmental instability, vertebral fractures, spine infection of tumor, inflammatory spondyloarthropathy.	**Age** Overall: 42.3 (11.6) SG1: 41.7 (11.8) NS1: 43.0 (11.3) **Female** Overall: 194 (41.1%) SG1: 101 (44%) NS1: 93 (39%) **Nonwhite** Overall: 73 (15.5%) SG1: 35 (15.1%) NS1: 38 (15.8%)	Overall: 58 (12.3%) SG1: 27 (12%) NS1: 31 (13%) Defined as: Employment status is "Disabled"; Overall: 76 (16.1%) SG1: 36 (16%) NS1: 40 (17%)	NR

Abbreviations: CI = confidence interval; CT = computed tomography; d = days; mm = millimeter; MRI = magnetic resonance imaging; N = number; NR = not reported; NS = not significant; SD = standard deviation; SE = standard error; SG = surgical group; NS = nonsurgical group; SPORT = Spine Patient Outcomes Research Trial; VAS = visual analog scale; w = week(s); y = year(s).

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48] The Sciatica Micro-Endoscopic Discectomy Randomized Controlled Trial Low for follow-up through 2 years; Some concerns for follow-up longer than 2 years	**SG1:** Tubular discectomy **N randomized:** 167 **N analyzed:** 166 (99.4%) in main study's primary analyses. 52w: 156 (93.4%) 2y: 154 (92.2%) 3y: 117 (70.1%) 4y: 117 (70.1%) 5y: 106 (63.5%) **SG2:** Microdiscectomy **N randomized:** 161 **N analyzed:** 159 (98.8%) in main study's primary analyses. 52w: 151 (93.8%) 2y: 144 (89.4%) 3y: 106 (65.8%) 4y: 102 (63.4%) 5y: 98 (60.9%)	**VAS 100mm leg pain, mean (SD)** Baseline: SG1 62.6 (21.1); SG2 61.7 (24.0) AMD (95% CI); [Negative AMD favors SG1] 4w: 4.5 (−0.3 to 9.3) 8w: 4.5 (−0.4 to 9.3) 26w: 2.0 (−2.9 to 6.8) 52w: 4.4 (−0.5 to 9.4) 2y: 1.3 (-3.6 to 6.2) 3y: −0.6 (−6.0 to 4.7) 4y: −0.4 (−5.9 to 5.2) 5y: 0.2 (−5.5 to 6.0) RM AMD 4w to 52w: 4.2 (0.9 to 7.5), P=0.01 main treatment effect, P=0.12 treatment X time interaction RM AMD 4w to 2y: 3.3 (0.2 to 6.2), P=0.04 main treatment effect, P=0.08 treatment X time interaction RM AMD 4w to 5y: Calculated 1.8, P=0.13 **VAS 100mm back pain, mean (SD)** Baseline: SG1 40.2 (27.0); SG2 38.3 (27.8) AMD (95% CI); [Negative AMD favors SG1] 4w: 3.1 (−1.9 to 8.1) 8w: 3.8 (−1.3 to 8.8) 26w: 3.5 (−1.5 to 8.6) 52w: 4.9 (−0.2 to 10.1) 2y: 4.1 (−1.2 to 9.4) 3y: −1.5 (−7.3 to 4.4) 4y: −0.7 (−6.7 to 5.3) 5y: 0.4 (−5.9 to 6.7) RM AMD 4w to 52w: 3.5 (0.1 to 6.9), P=0.04 main treatment effect, P=0.37 treatment X time interaction RM AMD 4w to 2y: 3.0 (-0.2 to 6.3), P=0.07 main treatment effect, P=0.05 treatment X time interaction RM AMD 4w to 5y: Calculated 2.0, P=0.14	NR	NR

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48] (continued)		**SF-36 Bodily Pain, mean (SD)** Baseline: SG1 27.8 (18.2); SG2 25.2 (17.7) AMD (95% CI); [Positive AMD favors SG1] 4w: -1.6 (−6.7 to 3.6) 8w: -5.1 (−10.3 to 0.1) 26w: -4.9 (−10.0 to 0.3) 52w: -3.8 (−9.0 to 1.5) 2y: -3.2 (-8.6 to 2.3) 3y: NR 4y: NR 5y: NR		
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48] The Sciatica Micro-Endoscopic Discectomy Randomized Controlled Trial Low for follow-up through 2 years; Some concerns for follow-up longer than 2 years (continued)		RM 4w to 52w: -3.3 (-7.3 to 0.7), P=0.10 main treatment effect, P=0.28 treatment X time interaction RM 4w to 2y: -2.8 (-6.7 to 1.0), P=0.14 main treatment effect, P=0.22 treatment X time interaction **Sciatica index, Bothersomeness Score, mean (SD)** Baseline: SG1 14.1 (4.8); 14.2 (5.0) AMD (95% CI); [Negative AMD favors SG1] 4w: 0.3 (−0.7 to 1.4) 8w: 0.8 (−0.3 to 1.8) 26w: 1.1 (0 to 2.1) 52w: 0.9 (−0.1 to 2.0) 2y: 0.2 (-0.8 to 1.3) 3y: NR 4y: NR 5y: NR RM 4w to 52w: 0.7 (-0.1 to 1.5), P=0.10 main treatment effect, P=0.37 treatment X time interaction RM 4w to 2y: 0.5 (-0.3 to 1.3), P=0.26 main treatment effect, P=0.40 treatment X time interaction		

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48] The Sciatica Micro-Endoscopic Discectomy Randomized Controlled Trial Low for follow-up through 2 years; Some concerns for follow-up longer than 2 years (continued)		**Sciatica index, Frequency Score, mean (SD)** Baseline: SG1 16.0 (4.4); SG2 15.5 (4.3) AMD (95% CI); [Negative AMD favors SG1] 4w: 0.3 (-0.8 to 1.4) 8w: 0.8 (-0.4 to 1.9) 26w: 1.0 (-0.1 to 2.1) 52w: 1.0 (-0.1 to 2.2) 2y: 0.3 (-0.9 to 1.5) 3y: NR 4y: NR 5y: NR RM 4w to 52w: 0.7 (-0.2 to 1.7), P=0.14 main treatment effect, P=0.41 treatment X time interaction RM 4w to 2y: 0.5 (-0.5 to 1.4), P=0.32 main treatment effect, P=0.45 treatment X time interaction		

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Brouwer (2015)[39] Brouwer (2017)[102] Some concerns	**SG1:** Percutaneous laser disc decompression **N randomized:** 57 **N analyzed:** 55 (96.5%) **SG2:** Microdiscectomy, **N randomized:** 58 **N analyzed:** 57 (98.3%)	**VAS 100mm leg pain, mean (SD)** Baseline: SG1 56.9 (20.4); SG2 60.7 (19.9) AMD (95% CI); [Negative AMD favors SG1] 4w: 7.4 (-1.9 to 16.8) 8w: 5.7 (-3.7 to 15.0) 26w: 4.2 (-5.2 to 13.6) 52w: 5.7 (-3.8 to 15.2) 2y: -2.7 (-12.3 to 6.8) RM 4w to 52w: 6.9 (1.3 to 12.6) RM 4w to 2y: 5.0 (-0.2 to 10.2), P=0.06 main treatment effect, P=0.42 treatment X time interaction **VAS 100mm back pain, mean (SD)** Baseline: SG1 44.7 (27.6); SG2 45.8 (26.7) AMD (95% CI); [Negative AMD favors SG1] 4w: -2.0 (-11.3 to 7.2) 8w: 6.3 (-2.9 to 15.5) 26w: 9.4 (0.1 to 18.6) 52w: 7.6 (-1.7 to 16.9) 2y: -1.5 (-11.0 to 8.0) RM 4w to 52w: 4.6 (-1.1 to 10.4) RM 4w to 2y: 3.0 (-2.2 to 8.1), P=0.26 main treatment effect, P=0.58 treatment x time interaction **SF-36 Bodily Pain, mean (SD)** Baseline: SG 1 32.8 (20.5); SG2 30.0 (16.1) AMD (95% CI); [Positive AMD favors SG1] 4w: 4.1 (-4.8 to 12.9) 8w: 0.6 (-8.1 to 9.3) 26w: -11.3 (-20.1 to -2.4) 52w: -2.5 (-11.3 to 6.4) 2y: 2.2 (-7.0 to 11.4) RM 4w to 52w: -1.6 (-7.3 to 4.2)	NR	NR

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Brouwer (2015)[39] Brouwer (2017)[102] Some concerns (continued)		RM 4w to 2y: -0.3 (-5.6 to 5.0), P=0.91 main treatment effect, P=0.14 treatment x time interaction Note: the scoring of the following measures was adapted by the study authors, instead of summing items to report a score range of 0 to 24, the study authors reported the average per item, which can range from 1 to 4. **Sciatica index, Bothersomeness, mean (SD)** Baseline: SG1 3.3 (1.2); SG2 3.1 (1.3) AMD (95% CI); [Negative AMD favors SG1] 4w: 0.1 (-0.4 to 0.6) 8w: 0.2 (-0.2 to 0.7) 26w: 0.3 (-0.2 to 0.7) 52w: 0.2 (-0.2 to 0.7) 2y: -0.1 (-0.6 to 0.4) RM 4w to 52w: 0.2 (-0.1 to 0.5) RM 4w to 2y: 0.1 (-0.2 to 0.4), P=0.56 main treatment effect, P=0.89 treatment x time interaction **Sciatica index, Frequency, mean (SD)** Baseline: SG1 3.6 (1.1); SG2 3.8 (1.2) AMD (95% CI); [Negative AMD favors SG1] 4w: 0.1 (-0.4 to 0.6) 8w: 0.1 (-0.4 to 0.5) 26w: 0.1 (-0.4 to 0.6) 52w: 0.1 (-0.3 to 0.6) 2y: 0 (-0.4 to 0.3) RM 4w to 52w: 0.1 (-0.2 to 0.4) RM 4w to 2y: -0.2 (-0.7 to 0.3), P=0.92 main treatment effect, P=0.92 treatment x time interaction		

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Chatterjee (1995)[38] Some concerns	**SG1:** Automated Percutaneous Lumbar Discectomy (APLD) **N randomized:** 31 **N analyzed:** 31 (100%) **SG2:** Microdiscectomy **N randomized:** 40 **N analyzed:** 40 (100%)	NR	NR	NR
Erginousakis (2011)[37] High	**SG1:** Percutaneous Disc Decompression **N randomized:** 31 **N analyzed:** 31 (100%) **NS1:** Conservative management **N randomized:** 31 **N analyzed:** 31 (31%)	**VAS 10 cm pain, mean (SD)** Baseline: SG1 7.4 (1.4); NS1: 6.9 (1.9), P NR 12w: SG1 3.0 (2.4); NS1 0.9 (2.0); P>0.005 (described in text as NS) Calculated AMD 1.6 (adj. for baseline differences) 52w: SG1 1.7 (2.4); NS1 4.0 (3.4); P=0.005 Calculated AMD -2.8 (adj. for baseline differences) 2y: SG1 1.6 (2.5); NS1 4.1 (3.4); P=0.004 Calculated AMD -3.0 (adj. for baseline differences) **Mean % pain reduction at 2y** SG1: 86% NS1: 36% **N (%) with category of pain reduction at 2y** 100% pain relief: SG1 17 (55%); NS1 6 (19%); Calculated P=0.008 50% pain relief: 4 (13%); NS1 2 (6%) 0% pain relief: 2 (6%); NS1 3 (10%) Aggravation of pain: SG1 0 (0%); NS1 2 (6%)	NR	NR

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Franke (2009)[36] Some concerns	**SG1:** Microscopically assisted percutaneous nucleotomy **N randomized:** 52 **N analyzed:** 52 (100%) **SG2:** Microdiscectomy **N randomized:** 48 **N analyzed:** 48 (100%)	**VAS (sum of leg and back) pain (RM at 52w)** Significant within group reduction in pain over time in both groups, P < 0.001 Conflicting between group differences depending on which of the two centers the procedure was performed, P=0.006 at one center, P=0.7 at other center Post hoc analysis at one of the two centers VAS back pain 8w: Larger decrease in SG1, P=0.002 26w: Larger decrease in SG1, P=0.003 52w: No difference, P=0.467 No difference was found for the VAS leg pain at any time points No difference in VAS back or leg pain at the other center at any time point.	**Overall at 52w (NR by group):** 83% of motor deficits resolved completely 68% of sensory deficits resolved completely	NR

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Gerszten (2010)[41] Some concerns (6w outcomes) High (12w and later outcomes)	**SG1:** Plasma disc decompression using coblation technology (PDD) **N randomized:** 46 **N analyzed:** 29 (64% of ITT sample) at 26w **NS1:** Transforaminal epidural steroid injection (TFESI) **N randomized:** 44 **N analyzed**: 28 (70% of ITT sample) at 26w	**VAS 100 mm leg pain, mean (SD)** Baseline: SG1: 72 (13); NS1 75 (14) Change in score, mean (SE); [larger negative change favors SG1] 6w: SG1 -42 (5); NS1 -21 (4); P=0.002, calculated AMD -21 12w: SG1 -46 (4); NS1 -23 (5); P=0.0001, calculated AMD -23 26w: SG1 -47 (6); NS1 -21 (5); P=0.0008, calculated AMD -26 **VAS 100 mm back pain, mean (SD)** Baseline: SG1 44 (24); NS1 53 (23) Change in score, mean (SE); [larger negative change favors SG1] 6w: SG1 -18 (4); NS1 1 (3); P=0.0005, calculated AMD -19 12w: SG1 -17 (5); NS1 7 (4); P=0.0001, calculated AMD -24 26w: SG1 -21 (5); NS1 -0.4 (4); P=0.002, calculated AMD -21 **SF-36 Bodily Pain at 26w** Larger improvement in SG1 compared with NS1, P=0.0039	**N (%) with full muscle strength at 6w** No significant difference between groups on left or right side at L3, L4, L5 or S1 (8 comparisons) **N (%) with normal tactile sensitivity at 6w** Right side S1 (p=0.01) SG1: 39 (98%) NS1: 25 (78%) All other 7 comparisons NS.	NR

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Haines (2002)[42] High	**SG1:** Automated percutaneous discectomy, endoscopic percutaneous discectomy (APD/EPD) **N randomized:** 21 **N analyzed:** 17 (81.0%) at 26w **SG2:** Discectomy **N randomized:** 13 **N analyzed:** 10 (76.9%) at 26w	NR	NR	NR
Henriksen (1996)[35] Some concerns	**SG1:** Microdiscectomy **N randomized:** 40 **N analyzed:** 39 (97.5%) **SG2:** Standard discectomy **N randomized:** 40 **N analyzed:** 40 (100%)	**VAS 100 mm leg pain at 4w and 6w** Actual values only depicted in a figure and variance NR, no differences between the groups reported **VAS 100 mm back pain at 4w and 6w** Actual values only depicted in a figure and variance NR, no differences between the groups reported	NR	NR

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Hermantin (1999)[43] Some concerns	**SG1:** Video-assisted arthroscopic microdiscectomy **N randomized:** 30 **N analyzed:** 30 (100%) **SG2:** Discectomy, with laminotomy **N randomized:** 30 **N analyzed:** 30 (100%)	**VAS 10 cm pain, mean (SD)** Baseline: SG1 6.8 (NR); SG2 6.6 (NR) Unspecified follow-up time: SG1: 1.2 (NR); SG2: 1.9 (NR), calculated AMD -0.9 (adj. for baseline) Note: Mean duration of follow-up was 2.6y (range 1.6y to 3.5y)	Follow-up time for these measures were NR. Note: Mean duration of follow-up was 2.6y (range 1.6y to 3.5y) **N (%) with postoperative reflex abnormalities** SG1: 7 (20.3%) SG2: 6 (20.0%) Calculated P=1.0 **N (%) with sensory deficits** SG1: 16 (53.3%) SG2: 18 (60.0%) Calculated P=0.79 **N (%) with motor weakness** SG1: 5 (16.7%) SG2: 10 (33.3%) Calculated P=0.23	NR
Huang (2005)[24] Some concerns	**SG1:** Microendoscopic discectomy **N randomized:** 10 **N analyzed:** 10 (100%) **SG2:** Discectomy **N randomized:** 12 **N analyzed:** 12 (100%)	NR	NR	NR

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Mayer (1993)[34] High	**SG1:** Percutaneous endoscopic discectomy **N randomized:** 20 **N analyzed:** 20 (100%) **SG2:** Microdiscectomy **N randomized:** 20 **N analyzed:** 20 (100%)	**N (%) of patients with low back pain** Baseline: SG1 19 (95%); SG2 20 (100%) 2y: SG1 9 (47.4%); SG2 4 (20%) Calculated AMD: 32% (adj. for baseline) Calculated P=0.18 **N (%) of patients with sciatica** Baseline: SG1 20 (100%); SG2 20 (100%) 2y: SG1 4 (20.0%); SG2 7 (35%) Calculated AMD: -15% Calculated P=0.48	**N (%) of patients with sensory deficit** Baseline: SG1 13 (65%); SG2 16 (80%) 2y: SG1 1 (5%); SG2 5 (25%); Calculated P=0.18 **N (%) of patients with motor deficit** Baseline: SG1 1 (5%); SG2 4 (20%) 2y: SG1 0 (0%); SG2 0 (0%); Calculated P=1.0 **N (%) of patients with reflex differences** Baseline: SG1 10 (50%); SG2 7 (35%) 2y: SG1 2 (10%); SG2 2 (10%); Calculated P=1.0	NR
McMorland (2010)[23] Some concerns	**SG1:** Microdiscectomy **N randomized:** 20 **N analyzed:** 12w 20 (100%) (outcomes reported only to 12w for ITT analysis) 24w 20 (100%) 52w 15 (75%) **NS1:** Spinal manipulation **N randomized:** 20 **N analyzed:** 12w 20 (100%) (outcomes reported only to 12w for ITT analysis) 24w 20 (100%) 52w 17 (85%)	**SF-36 Bodily Pain, mean (SD)** Baseline: SG1 27.3 (19.8); NS1 28.5 (21.8) 6w: SG1 41.4 (24.1); NS1 45.8 (21.3) Calculated AMD -3.2 (adj. for baseline differences) 12w: SG1 57.4 (22.3); NS1 47.1 (18.4) Calculated AMD 11.5 (adj for baseline differences) RM 6w to 12w: AMD NR, P=0.031 for time effect, P=0.341 for main treatment effect, P=0.367 for treatment X time interaction **McGill Pain Questionnaire, Pain Rating Intensity (rank value), mean (SD) (score range 0 [no pain] to 78 [worst pain])** Baseline: SG1 32.5 (12.9); NS1 28.7 (17.4) 6w: SG1 18.4 (16.3); NS1 21.7 (13.7) Calculated AMD -7.1 (adj. for baseline differences) 12w: SG1 13.0 (16.3); NS1 19.4 (14.3)	NR	**SF-36 Total Score, mean (SD)** Baseline: SG1 379.5 (149.8); NS1: 381.3 (161.9) 6w: SG1 429.1 (157.3); NS1 445.6 (142.8) 12w: SG1 500.3 (179.7); NS1: 484.6 (148.9) RM 6w to 12w AMD (95%CI) [positive AMD favors SG1] AMD NR, P=0.016 for time effect, P=0.382 for main treatment effect, P=0.683 for treatment X time interaction

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
McMorland (2010)[23] Some concerns (continued)		Calculated AMD -10.2 (adj. for baseline differences) RM 6w to 12w: AMD NR, P=0.013 for time effect, P=0.103 for main treatment effect, P=0.754 for treatment X time interaction **McGill Pain Questionnaire, Number of words chosen, mean (SD) (score range 0 [no pain] to 20 [worst pain])** Baseline: SG1 13.2 (5.0); NS1 12.0 (5.5) 6w: SG1 8.8 (6.4); NS1 10.8 (6.1) Calculated AMD -3.2 (adj. for baseline differences) 12w: SG1 5.7 (5.1); NS1 9.6 (6.3) Calculated AMD -5.1 (adj. for baseline differences) RM 6w to 12w: AMD NR, P=0.029 for time effect, P=0.080 for main treatment effect, P=0.574 for treatment X time interaction **McGill Pain Questionnaire, Present pain intensity, mean (SD) (score range 1 [mild pain] to 5 [excruciating pain])** Baseline: SG1 2.7 (1.0); NS1 2.4 (0.8) 6w: SG1 1.6 (1.3); NS1 1.8 (0.7) Calculated AMD -0.5 (adj. for baseline differences) 12w: SG1 1.5 (1.3); NS1 1.6 (0.9) Calculated AMD -0.4 (adj. for baseline differences) RM 6w to 12w: AMD NR, P=0.010 for time effect, P=0.094 for main treatment effect, P=0.736 for treatment X time interaction		

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
McMorland (2010)[23] Some concerns (continued)		**Aberdeen back pain scale, mean (SD) (score range 0 [no pain] to 100 [worst pain])** Baseline: SG1 45.1 (17.8); NS1 44.7 (12.9) 6w: SG1 32.3 (22.2); NS1 34.8 (18.6) Calculated AMD -2.9 (adj. for baseline differences) 12w: SG1 25.8 (23.7); NS1 35.6 (18.9) Calculated AMD -10.2 (adj. for baseline differences) RM 6w to 12w: AMD NR, P= 0.017 for time effect, P=0.034 for main treatment effect (favors SG1), P=0.836 for treatment X time interaction		
North (2005)[46] High	**SG1:** Repeat lumbosacral decompression **N randomized:** 26 **N analyzed:** 26 (100%) **NS1:** Spinal cord stimulation **N randomized:** 24 **N analyzed:** 19 (79.2%)	NR	NR	NR

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Osterman (2003)[33] High	**SG1:** Microdiscectomy **N randomized:** 28 **N analyzed:** 6w 26 (93%) 12w 26 (93%) 26w 26 (93%) 52w 21 (75%) 2y 26 (93%) **NS1:** Physiotherapy **N randomized:** 28 **N analyzed:** 6w 26 (93%) 12w 26 (93%) 26w 22 (78.6%) 52w 20 (71.4%) 2y 24 (86%)	**VAS 100 mm leg pain, mean (SD)** Baseline: SG1 61 (20); NS1 57 (21) [Negative AMD favors SG1] 6w: SG1 12 (20); NS1 25 (27) Calculated AMD -17 (adj. for baseline) 12w: SG1 9 (16); NS1 16 (25) Calculated AMD -11 (adj. for baseline) 26w: SG1 9 (20); NS1 18 (29) Calculated AMD -13 (adj. for baseline) 52w: SG1 6 (11); NS1 9 (19) Calculated AMD -7 (adj. for baseline) 2y: SG1 6 (11); NS1 15 (24) Calculated AMD -13 (adj. for baseline) RM 6w to 2y: AMD -9 (95% CI, -20 to 1) **VAS 100 mm back pain, mean (SD)** Baseline: SG1 53 (25); NS1 47 (28) [Negative AMD favors SG1] 6w: SG1 21 (25); NS1 28 (24) Calculated AMD -13 (adj. for baseline) 12w: SG1 15 (20); NS1 22 (23) Calculated AMD -13 (adj. for baseline) 26w: SG1 13 (22); NS1 20 (28) Calculated AMD -13 (adj. for baseline) 52w: SG1 19 (25); NS1 17 (23) Calculated AMD -4 (adj. for baseline) 2y: SG1 11 (18); NS1 21 (27) Calculated AMD -16 (adj. for baseline) RM 6w to 2y: AMD -7 (95 % CI, -17 to 3)	**N (%) with muscle weakness** 6w: SG1 14 (53.8%); NS1 12 (46.2%) 12w: SG1 11 (42.3%); NS1 12 (46.2%) 26w: NR 52w: SG1 6 (28.6%); NS1 6 (30%) 2y: NR	**15D Health-related quality of life, mean (SD)** Baseline: SG1 0.83 (0.07); NS1 0.84 (0.06) 6w: SG1 0.92 (0.07); NS1 0.89 (0.09) 12w: SG1 0.94 (0.06); NS1 0.91 (0.09) 26w: SG1 0.95 (0.06): NS1 0.90 (0.13) 52w: SG1 0.95 (0.05); NS1 0.94 (0.07) 2y: SG1 0.95 (0.08); NS1 0.93 (0.12) RM 6w to 2y, AMD (95% CI) [positive AMD favors SG1] -0.03 (-0.07 to 0.01)

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98] High	**SG1:** Microdiscectomy **N randomized:** 141 **N analyzed:** 1y: 140 (99.3%) 2y: 130 (92.2%) 5y: 115 (81.6%) **NS1:** Conservative management **N randomized:** 142 **N analyzed:** 1y: 141 (99.3%) 2y: 130 (91.5%) 5y: 116 (81.7%)	**VAS 100 mm leg pain, mean (SD)** Baseline: SG1 67.2 (27.7); NS1 64.4 (21.2) AMD (95% CI); [negative AMD favors SG1] 8w: -17.7 (-23.1 to -12.3) 26w: -6.1 (-10.0 to -2.2) 52w: 0 (-4.0 to 4.0) 2y: 2 (-2.0 to 6.0) 5y: 2.7 (-2.9 to 8.4) **Cumulative score on VAS 100 for leg pain 8w to 52w** SG1: AUC 635.3 (SE 58.6) NS1: AUC 977.0 (SE 68.3) AMD: -341.7 (95% CI, -519.6 to 163.8) **Cumulative score on VAS 100 for leg pain 8w to 2y** SG1: AUC 1,110.2 (SE 133.3) NS1: AUC 1,487.1 (SE 137.7) AMD: -376.8 (95% CI, -754.6 to 0.9) **Cumulative score on VAS 100 for leg pain 8w to 5y** SG1: NR NS1: NR Reported as no significant difference **VAS 100 mm back pain, mean (SD)** Baseline: SG1 33.8 (29.6); NS1 30.8 (27.7) AMD (95% CI); [negative AMD favors SG1] 8w: -11.3 (-17.4 to -5.6) 26w: -2.3 (-8.2 to 3.6) 52w: -2.3 (-8.2 to 3.6) 2y: -1.4 (-6.3 to 4.5) 5y: 3.1 (-4.2 to 10.3)	NR	NR

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98] High (continued)		**Cumulative score on VAS 100 for back pain 0 to 52w** SG1: AUC 884.2 (SE 77.6) NS1: AUC 1047.9 (SE 77.6) AMD: -163.7 (95% CI, -379.9 to 52.5) **Cumulative score on VAS 100 for back pain 0 to 2y** SG1: AUC 1526.3 (SE 169.7) NS1: AUC 1,734.1 (SE 182.3) AMD: -207.8 (95% CI, -702.0 to 286.4) **Cumulative score on VAS 100 for back pain 0 to 5 years:** SG1: NR NS1: NR Reported as no significant difference **SF-36 Bodily Pain, mean (SD)** Baseline: SG1 21.9 (16.6); NS1 23.9 (18.1) AMD (95% CI); [positive AMD favors SG1] 8w: 8.4 (3.2 to 13.5) 26w: 3.3 (-1.8 to 8.4) 52w: 2.7 (-2.6 to 7.9) 2y: -2.3 (-7.3 to 2.7) 5y: NR **Sciatica index, Bothersomeness Score, mean (SD)** Baseline: SG1 14.6 (5.1); NS1 14.5 (4.1) AMD (95% CI); [negative AMD favors SG1] 8w: -3.6 (-4.9 to -2.3) 26w: -1.2 (-1.3 to -0.1) 52w: -0.4 (-1.5 to 0.7) 2y and 5y: NR		

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98] High (continued)		**Sciatica index, Frequency Score, mean (SD)** Baseline: SG1 16.0 (4.6); NS1 16.2 (4.2) AMD (95% CI); [negative AMD favors SG1] 8w: -4.0 (-5.3 to -2.7) 26w: -1.8 (-1.9 to -0.7) 52w: -0.5 (-1.8 to 0.8) 2y and 5y: NR		
Ruetten (2008)[31] High	**SG1:** Endoscopic (interlaminar or transforaminal) discectomy **N randomized:** 100 **N analyzed:** 91 (91%) **SG2:** Microdiscectomy **N randomized:** 100 **N analyzed:** 87 (87%)	**VAS 100 mm leg pain, mean (SD)** Baseline: SG1 75 (NR); SG2 71 (NR) 12w: SG1 6 (NR); SG2 9 (NR) Calculated AMD (adjusted for baseline): -7 26w: SG1 9 (NR); SG2 7 (NR) Calculated AMD (adjusted for baseline): -2 52w: SG1 9 (NR); SG2 11 (NR) Calculated AMD (adjusted for baseline): -6 2y: SG1: 8 (NR): SG2 9 (NR) Calculated AMD (adjusted for baseline): -5 Between-group differences reported as NS. **VAS 100 mm for back pain, mean (SD)** Baseline: SG1 19 (NR); SG2 15 (NR) 12w: SG1 15 (NR); SG2 20 (NR) Calculated AMD (adjusted for baseline): -9 26w: SG1 16 (NR); SG2 22 (NR) Calculated AMD (adjusted for baseline): -10 52w: SG1 17 (NR); SG2 19 (NR) Calculated AMD (adjusted for baseline): -6 2y: SG1 11 (NR); SG2 18 (NR) Calculated AMD (adjusted for baseline): -11 Between-group differences reported as NS.	**NASS Neurology Score, mean (SD)** (range of scores 1 (best) to 6 (worst); negative AMD favors SG1) Baseline: SG1 3.1 (NR); SG2 2.9 (NR) 12w: SG1 2 (NR); SG2 2 (NR) 26w: SG1 2.1 (NR); SG2 2.3 (NR) 52w: SG1 1.9 (NR); SG2 1.7 (NR) 2y: SG1 2.1 (NR); SG2 1.9 (NR) Results reported as NS.	NR

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Ruetten (2008)[31] High (continued)		**North American Spine Society Pain Score, mean (SD)** (range of scores 1 (best) to 6 (worst); negative AMD favors SG1) Baseline SG1 4.6 (NR); SG2 4.2(NR) 12w: SG1 2 (NR); SG2 2.4 (NR) Calculated AMD (adjusted for baseline): -0.8 26w: SG1 2.2 (NR); SG2 2.6 (NR) Calculated AMD (adjusted for baseline): -0.8 52w: SG1 2.2 (NR); SG2 2.4 (NR) Calculated AMD (adjusted for baseline): -0.6 2y: SG1 2.1 (NR); SG2 2.3 (NR) Calculated AMD (adjusted for baseline): -0.6 Between-group differences reported as NS. **N (%) with no leg pain at 2y** SG1: 77 (85%) SG2: 69 (79%) Calculated AMD: -6% Results reported as NS. **N (%) with leg pain occasionally or pain was greatly reduced at 2y** SG1: 12 (13%) SG2: 13 (15%) Calculated AMD: 2% Results reported as NS. **N (%) with no improvement in leg pain at 2y** SG1: 2 (2%) SG2: 5 (6%) Calculated AMD: 4% Results reported as NS.		

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Ruetten (2008)[31] High (continued)		**N (%) with progradient back pain at unspecified time point** SG1: 2 (2.2%) SG2: 10 (11.5%) Calculated AMD: 9.3% P < 0.001		
Ruetten (2009)[47] High	**SG1:** Revision endoscopic discectomy **N randomized:** 50 **N analyzed:** 45 (90%) **SG2:** Revision microdiscectomy **N randomized:** 50 **N analyzed:** 42 (84%)	**VAS 100 mm leg pain, mean (SD)** Baseline: SG1 85 (NR); SG2 79 (NR) 12w: SG1 8 (NR); SG2 12 (NR) 26w: SG1 10 (NR); SG2 12 (NR) 52w: SG1 12 (NR); SG2 9 (NR) 2y: SG1 8 (NR); SG2 10 (NR) Differences between groups reported as NS. **VAS 100 mm back pain, mean (SD)** Baseline: SG1 14 (NR); SG2 15 (NR) 12w: SG1 14 (NR); SG2 13 (NR) 26w: SG1 12 (NR); SG2 12 (NR) 52w: SG1 16 (NR); SG2 15 (NR) 2y: SG1 15 (NR); SG2 14 (NR) Differences between groups reported as NS. **North American Spine Society Pain score, mean (SD)** Baseline: SG1 4 (NR); SG2 4 (NR) 12w: SG1 1.9 (NR); SG2 2.1 (NR) 26w: SG1 1.9 (NR); SG2 2.0 (NR) 52w: SG1 2.1 (NR); SG2 2.2 (NR) 2y: SG1 2.1 (NR); SG2 2.1 (NR) Differences between groups reported as NS. **N (%) with no leg pain at 2y** SG1: 37 (82%) SG2: 32 (76%) Differences between groups reported as NS.	**North American Spine Society Neurology Score, mean (SD)** Baseline: SG1 3 (NR); SG2 5 (NR) 12w: SG1 2.2 (NR); SG2 2.1 (NR) 26w: SG: 2.0 (NR); SG2 2.1 (NR) 52w: SG1 2.2 (NR); SG2 2.3 (NR) 2y: SG1 2.1 (NR); SG2 2.3 (NR) Differences between groups reported as NS.	

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Ryang (2008)[30] Gempt (2013)[103] High	**SG1:** Trocar microdiscectomy **N randomized:** 30 **N analyzed:** unclear **SG2:** Microdiscectomy **N randomized:** 30 **N analyzed:** unclear	Outcomes measured over average follow-up of 1.33y (range 26w to 2.17y) Additional long-term outcomes reported for 38 participants over average follow-up of 2.8y (range 52w to 4.5y) **VAS 10 cm Pain Score, mean (SD)** Baseline: SG1 6.9 (2.4); SG2 7.3 (2.3), P=0.60 Follow-up: SG1 2.1 (2.4); SG2 2.1 (2.4), P=0.86 Calculated AMD: 0.4 (adj. for baseline) Long-term follow-up: SG1 median 0.5 (range 0 to 7); SG2 1.65 (0 to 7.5), P reported as NS **N (%) with radicular pain** Baseline: SG1 27 (90%), SG2 29 (97%), P=0.31 Follow-up: SG1 1 (3%); SG2 5 (17%), P=0.11 Calculated AMD: -7% **SF-36 Bodily Pain, mean(SD)** Baseline: SG1 22.4 (22.8); SG2 19.1 (17.3), P=0.84 Follow-up: SG1 68.9 (31.9); SG2 70.2 (25.1), P=0.95 Calculated AMD: -4.6 (adj. for baseline) Long-term follow-up: SG1 58.0 (27.1); 69.4 (26.6) Calculated AMD: -14.7 P=0.198	Outcomes measured over average follow-up of 1.33y (range 26w to 2.17y) **N (%) with sensory deficits** Baseline: SG1 27 (90%); SG2 22 (73%), P=0.10 Follow-up: SG1 12 (40%); SG2 13 (43%), P=0.31 **N (%) with motor deficits** Baseline: SG1 16 (53%); SG2 15 (50%), P=0.61 Follow-up: SG1 8 (27%); SG2 7 (23%), P=0.86	Outcomes measured over average follow-up of 1.33y (range 26w to 2.17y) Additional long-term outcomes reported for 38 participants over average of 2.8y (range 52w to 4.5y) **SF-36 Physical Component Summary, mean (SD)** Baseline: SG1 29.3 (7.9); SG2 27.3 (5.9), P=0.44 Follow-up: SG1 47.6 (10.7), SG2 47.5 (9.4), P=0.79 Long-term follow-up: SG1 42.6 (10.8); SG2 48.5 (8.7), P=0.081 **SF-36 Mental Component Summary, mean (SD)** Baseline: SG1 39.5 (12.4); SG2 42.3 (14.8), P=0.51 Follow-up: SG1 44.0 (13.2), SG2 51.9 (7.8), P=0.03 Long-term follow-up: SG1 48.4 (9.4); SG2 48.8 (10.5), P=0.892

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Sasaoka (2006)[25] High	**SG1:** Microendoscopic discectomy **N randomized:** 15 **N analyzed:** unclear **SG2:** Microdiscectomy **N randomized:** 11 **N analyzed:** unclear	**N (%) with residual low back pain or lumbar discomfort at 52w** SG1: NR (36.7%) SG2: NR (66.7%) Calculated AMD: 30% P NR but difference reported as significant.	NR	NR
Teli (2010)[29] Some concerns	**SG1:** Microendoscopic discectomy **N randomized:** NR **N analyzed:** 70 **SG2:** Microdiscectomy **N randomized:** NR **N analyzed:** 72 **SG3:** Discectomy **N randomized:** NR **N analyzed:** 70	**VAS 10 cm leg pain, mean (SD)** Baseline: SG1 8 (1); SG2 8 (1); SG3 8 (1) 26w: SG1 2 (1); SG2 2 (1); SG3 2 (1) Calculated AMD: 0 (adj. for baseline) 52w: SG1 1 (1); SG2 1 (1); SG3 1 (1) Calculated AMD: 0 (adj. for baseline) 2y: SG1 2 (1); SG2 2 (1); SG3 2 (1) Calculated AMD: 0 (adj. for baseline) No difference among the three groups at any time point, P=0.73 **VAS 10 cm back pain, mean (SD)** Baseline: SG1 3 (1); SG2 4 (1); SG3 3 (1) 26w: SG1 2 (1); SG2 2 (1); SG3 1 (1) Calculated AMD: 1 (adj. for baseline) 52w: SG1 1 (1); SG2 1 (1); SG3 1 (1) Calculated AMD: 1 (adj. for baseline) 2y: SG1 2 (1); SG2 2 (1); SG3 1 (1) Calculated AMD: 1 (adj. for baseline) No difference among the three groups at any time point, P=0.75	NR	**SF-36 Physical Health Component Summary, mean (SD)** Baseline: SG1 20 (4); SG2 21 (4); SG3 22 (4) 26w: SG1 42 (4); SG2 42 (4); SG3 42 (4) Calculated AMD 1 52w: SG1 44 (4); SG2 45 (4); SG3 44 (4) Calculated AMD 2 2y: SG1 39 (6); SG2 40 (6); SG3 38 (6) Calculated AMD 3 No difference among the three groups at any time point, P=0.68

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Teli (2010)[29] Some concerns (continued)				**SF-36 Mental Health Component Summary** Baseline: SG1 22 (3); SG2 21 (2); SG3 23 (2) 26w: SG1 38 (4); SG2 40 (4); SG3 40 (3) Calculated AMD 2 52w: SG1 40 (4); SG2 40 (4); SG3 42 (3) Calculated AMD 0 2y: SG1 38 (5); SG2 39 (6); SG3 39 (3) Calculated AMD 2 No difference among the three groups at any time point, P=0.78
Thome (2005)[28] Barth (2008)[104] Some concerns	**SG1:** Sequestrectomy **N randomized:** 42 **N analyzed:** 42 (100%) **SG2:** Microdiscectomy **N randomized:** 42 **N analyzed:** 42 (100%)	**VAS 10 cm leg pain, mean (SD)** Baseline: SG1 5.9 (2.6); SG2 6.7 (2.3) 12w to 26w: SG1 0.7 (1.7); SG2 1.3 (2.5) Calculated AMD: 0.2 (adj. for baseline) 52w to 1.5y: SG1 0.6 (1.4); SG2 0.8 (1.7) Calculated AMD: 0.4 (adj. for baseline) 2y: SG1 1.2 (1.8); SG2 1.6 (2.4) P>0.05 for difference between groups over time **VAS 10 cm back pain, mean (SD)** Baseline: SG1 5.2 (2.6); SG2 5.9 (2.5) 12w to 26w: SG1 0.9 (1.4); SG2 1.6 (2.5) Calculated AMD: 0 (adj. for baseline) 52w to 1.5y: SG1 1.0 (1.7); SG2 1.6 (2.1) Calculated AMD: 0.9 (adj. for baseline) 2y: SG1 1.8 (1.9); SG2 2.9 (2.6) P>0.05 for difference between groups over time	**N (%) of patients with improvement in sensory deficit, 12w to 26w** SG1: Exact value NR SG2: Exact value NR P=0.52 **N (%) of patients with improvement in motor deficit, 12w to 26w** SG1: Exact value NR SG2: Exact value NR P=0.74 Change in sensory index, motor grade, straight leg raise test, and reflex index showed no difference between groups between baseline, 12w to 26w and 2y, P for trend for all parameters reported as > 0.278.	**SF-36 Mental Component Summary Score, mean (SD)** Baseline: SG1 46.5 (11.8); SG2 47.0 (12.5), P=0.87 12w to 26w: SG1 53.6 (9.8); SG2: 50.6 (12.0), P=0.26 **SF-36 Physical Component Summary Score, mean (SD)** Baseline: SG1 28.8 (6.6); SG2 28.5 (8.1), P=0.87 12w to 26w: SG1 43.6 (9.7); SG2 41.5 (10.7), P=0.41

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Thome (2005)[28] Barth (2008)[104] Some concerns (continued)		**SF-36 Bodily Pain, mean (SD)** Baseline: SG1 24.4 (13.1); SG2 19.0 (15.0), P=0.09 12w to 26w: SG1 68.6 (19.4); SG2 60.2 (27.6), P=0.14 Calculated AMD: 3 (adj. for baseline) 2y: SG1 71.6 (25.4); SG2 59.8 **N (%) with VAS 10 cm leg pain >3** Baseline: SG1 NR (79), SG2 (NR) 88, P=0.38 12w to 26w: SG1 NR (5%); SG2 NR (18%), P=0.15 Calculated AMD: -4% 52w to 1.5y: SG1 NR (5%); SG2 NR (11%), P=0.43 Calculated AMD: **N (%) with VAS 10 cm back pain >3** Baseline: SG1 NR (74%), SG2 NR (80%), P=0.60 12w to 26w: SG1 NR (5%); SG2 NR (16%), P=0.26 52w to 1.5y: SG1 NR (13%); SG2 NR (19%), P=0.54 (27.7), P=0.064	**N (%) of patients reporting specified changes in sensory deficit at 12w to 26w** Much better: SG1 NR (73%); SG2 NR (71%) Better: SG1 NR (14%); SG2 NR (18%) Equal: SG1 NR (13%); SG2 NR (11%) Worse: SG1 NR (0%); SG2 NR (0%) Much Worse: SG1 NR (0%); SG2 NR (0%) P=0.969 **N (%) of patients reporting specified changes in sensory deficit at 2y** Much better: SG1 NR (68%); SG2 NR (54%) Better: SG1 NR (16%); SG2 NR (20%) Equal: SG1 NR (16%); SG2 NR (11%) Worse: SG1 NR (0%); SG2 NR (6%) Much Worse: SG1 NR (0%); SG2 NR (9%) P=0.061 P for trend over time=0.034 **N (%) of patients reporting specified changes in motor deficit at 12w to 26w** Much better: SG1 NR (67%); SG2 NR (79%) Better: SG1 NR (22%); SG2 NR (15%) Equal: SG1 NR (11%); SG2 NR (3%) Worse: SG1 NR (0%); SG2 NR (3%) Much Worse: SG1 NR (0%); SG2 NR (0%) P=0.390	

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Thome (2005)[28] Barth (2008)[104] Some concerns (continued)			**N (%) of patients reporting specified changes in motor deficit at 2y** Much better: SG1 NR (84%); SG2 NR (63%) Better: SG1 NR (8%); SG2 NR (23%) Equal: SG1 NR (8%); SG2 NR (6%) Worse: SG1 NR (0%); SG2 NR (0%) Much Worse: SG1 NR (0%); SG2 NR (8%) P=0.041 P for trend over time=0.004	
Tullberg (1993)[27] Some concerns	**SG1:** Microdiscectomy **N randomized:** 30 **N analyzed:** 29 (97%) **SG2:** Discectomy **N randomized:** 30 **N analyzed:** 29 (97%)	**VAS 10 cm leg pain, mean (SD)** Baseline: SG1 7.0 (NR); SG2 7.0 (NR) 52w: SG1 2.1 (NR); SG2 2.3 (NR) **VAS 10 cm back pain, mean (SD)** Baseline: SG1 3.6 (NR); SG2 3.7 (NR) 52w: SG1 1.6 (NR); SG2 1.8 (NR)	NR	NR
Weber (1983)[26] High	**SG1:** Discectomy **N randomized**: 60 **N analyzed:** 60 (100%) **NS1:** Conservative management **N randomized:** 66 **N analyzed**: 66 (100%)	**N (%) in specified category of radiating pain at 4y** No pain: SG1 36 (63.2%); NS1 38 (57.6%); Calculated P=0.86 Some pain: SG1 15 (26.3%); NS1 21 (31.8%) Considerable pain: SG1 6 (10.5%); NS1 7 (10.6%) **N (%) in specified category of radiating pain at 10y** No pain: SG1 43 (84.3%); NS1 52 (78.8%); Calculated P=0.41 Some pain: SG1 8 (14.0%); NS1 14 (21.2%) Considerable pain: SG1 0 (0%); NS1 0 (0%)	NR	NR

(continued)

Table D-3. Individual study findings related to efficacy outcomes Part 1 (pain, neurological symptoms, quality of life) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Pain	Neurologic Symptoms	Quality of Life
Weinstein (2006) [22] Weinstein (2008)[99] Lurie (2014)[100] SPORT High	**SG1:** Discectomy/ microdiscectomy **N randomized**: 245 **N analyzed:** 232 in main study's primary analyses. 52w: 202 (82.47%) 2y: 186 (75.9%) 3y: 180 (73.5%) 4y: 149 (60.8%) 8y: 157 (64.1%) **NS1:** Conservative management **N randomized**: 256 **N analyzed**: 240 included in main study's primary analyses. 52w: 213 (83.2%) 2y: 187 (73.0%) 3y: 170 (66.4%) 4y: 150 (58.6%) 8y: 152 (59.4%)	**SF-36 Bodily Pain, mean (SD)** Baseline: SG1 27.1 (18.5); NS1 26.7 (17.4) AMD (95% CI) [positive AMD favors SG1] 12w: 2.9 (-2.2 to 8.0) 52w: 2.8 (-2.3 to 7.8) 2y: 3.2 (-2.0 to 8.4) 4y: 4.5 (-1.2 to 10.3) 8y: 0.7 (-5.2 to 6.6) RM 12w to 2y: AMD NR, P=0.74 RM 12w to 4y: AMD NR, P=0.15 RM 12w to 8y: AMD NR, P=0.22 **Sciatica index, Bothersomeness Score, mean (SD)** Baseline: SG1 15.4 (5.1); NS1 15.0 (5.3) AMD (95% CI) [negative AMD favors SG1] 12w: -2.1 (-3.4 to -0.9) 52w: -1.6 (-2.9 to -0.4) 2y: -1.6 (-2.9 to -0.3) 4y: -1.8 (-3.2 to -0.4) 8y: -1.5 (-2.9 to -0.2) RM 12w to 2y: AMD NR, P=0.003 favoring SG1 RM 12w to 4y: AMD NR, P NR for the ITT analysis RM 12w to 8y: AMD NR, P=0.005 favoring SG1	NR	NR

Abbreviations: AMD = absolute mean difference; AUC = area under the curve; CI = confidence interval; cm = centimeter; ITT = intention-to-treat; mm = millimeter; N = number; NR = not reported; NS = not significant; RM = repeated measure; SD = standard deviation; SE = standard error; SF-36 = Short Form 36; SG = surgical group; NS = nonsurgical group; SPORT = Spine Patient Outcomes Research Trial; VAS = visual analog scale; w = week(s); y = year(s).

Note: For continuous outcome measures, studies either reported 1) the difference in mean scores at a follow-up time point (e.g. mean score in SG1 minus mean score in NS1 at 6 weeks) or 2) mean change from baseline scores at a follow-up time point (e.g., mean change in score in SG1 minus mean change in score for NS1 at 6 weeks). The absolute mean difference (AMD) between groups reported or calculated in this table reflects the mean difference between groups with respect to the change in score.

- For outcomes where a higher score represents fewer symptoms, a positive AMD means the intervention group (SG1) improves symptoms more than the comparator group (NS1 or SG2 or SG3).

- For outcomes where a lower score represents fewer symptoms, a negative AMD means the inter intervention group (SG1) improves symptoms more than the comparator group (NS1 or SG2 or SG3).

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48] The Sciatica Micro-Endoscopic Discectomy Randomized Controlled Trial Low for follow-up through 2 years; Some concerns for follow-up longer than 2 years	**SG1:** Tubular discectomy **N randomized:** 167 **N analyzed:** 159 (98.8%) in main study's primary analyses. 52w: 156 (93.4%) 2y: 154 (92.2%) 3y: 117 (70.0%) 4y: 117 (70.0%) 5y: 106 (63.5%) **SG2:** Microdiscectomy **N randomized:** 161 **N analyzed:** 159 (98.8%) in main study's primary analyses. 52w: 151 (93.8%) 2y: 144 (89.4%) 3y: 106 (65.8%) 4y: 102 (63.4%) 5y: 98 (60.9%)	**Roland-Morris Disability Questionnaire, mean (SD)** Baseline: SG1 16.0 (4.4); SG2 16.3 (4.3) AMD (95% CI); [Negative AMD favors SG1] 4w: 0.2 (−1.1 to 1.4) 8w: 0.8 (−0.4 to 2.1) 26w: 1.0 (−0.2 to 2.3) 52w: 1.3 (0.03 to 2.6) 2y: 0.8 (-0.5 to 2.1) 3y: 0.2 (−1.2 to 1.5) 4y: −0.1 (−1.5 to 1.2) 5y: 0.9 (−0.6 to 2.2) RM AMD 4w to 52w: 0.8 (−0.2 to 1.7), P=0.11 main treatment effect, P=0.50 treatment X time interaction RM AMD 4w to 2y: 0.6 (-0.3 to 1.6), P=0.17 main treatment effect, P=0.15 treatment X time interaction RM AMD 4w to 5y: Calculated 0.4 (NR), P=0.30 **SF-36 Physical Functioning, mean (SD)** Baseline: SG1 36.7 (20.6); SG2 34.9 (20.7) AMD (95% CI); [Positive AMD favors SG1] 4w: -1.1 (−5.6 to 3.3) 8w: -3.3 (−7.8 to 1.1) 26w: -3.9 (−8.3 to 0.6) 52w: -4.8 (-9.3 to -0.2) 2y: -3.4 (-8.2 to 1.4) 3y: NR 4y: NR 5y: NR	NR	**N (%) of patients with complete or nearly complete recovery, based on 7-pt Likert scale** [OR >1 favors SG1] 4w: SG1 NR (62%); SG2 NR (66%); OR 0.84 (95% CI 0.53 to 1.30) 8w: SG1 NR (63%); SG2 NR (75%); OR 0.56 (95% CI 0.35 to 0.92) 26w: SG1 NR (67%); SG2 NR (77%); OR 0.62 (95% CI 0.38 to 1.0) 52w: SG1 NR (69%); SG2 NR (79%), OR 0.59 (95% CI 0.35 to 0.99) 2y: SG1 NR (71%); SG2 NR (77%); OR 0.76 (95% CI 0.45 to 1.28) 3y: SG1 NR (78%); SG2 NR (78%); calculated OR 1.00 (95% CI, 0.51 to 1.83) 4y: SG1 NR (80%); SG2 NR (75%); calculated OR 1.33 (95% CI 0.70 to 2.52) 5y: SG1 NR (74%); SG2 NR (77%); calculated OR 0.85 (95% CI 0.45 to 1.61) **Relative difference in rate of complete or nearly complete recovery, based on 7-pt Likert scale, unadjusted HR (95% CI)** [HR >1 favors SG1] 4w to 52w: 0.92 (95% CI, 0.73 to 1.17) 4w to 2y: 0.93 (95% CI, 0.74 to 1.17)

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48] (continued)		RM AMD 4w to 52w: -3.1 (−6.8 to 0.7), P= 0.11 main treatment effect, P=0.27 treatment X time interaction RM AMD 4w to 2y: -2.8 (-6.5 to 0.9), P=0.14 main treatment effect, P=0.33 treatment X time interaction **Prolo Scale, Functional Score, mean (SD)** Baseline:SG1 0.8 (0.5); SG2 0.7 (0.5) AMD (95% CI); [Positive AMD favors SG1] 4w: 0 (-0.3 to 0.2) 8w: -0.1 (-0.3 to 0.2) 26w: -0.2 (-0.5 to 0) 52w: -0.2 (-0.5 to 0) 2y: NR 3y: NR 4y: NR 5y: NR RM 4w to 52w: -0.1 (-0.3 to 0), P=0.16 main treatment effect, P=0.43 treatment X time interaction **Prolo Scale, Economic Score, mean (SD)** Baseline: SG1 1.5 (1.6); SG2 1.3 (1.6) AMD (95% CI); [Positive AMD favors SG1] 4w: 0.2 (−0.1 to 0.5) 8w: 0.1 (−0.2 to 0.4) 26w: 0.1 (−0.2 to 0.4) 52w: 0 (−0.3 to 0.3) 2y: NR 3y: NR 4y: NR		

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48] (continued)		5y: NR RM 4w to 52w: 0.1 (-0.1 to 0.3); P=0.47 main treatment effect, P=0.76 treatment X time interaction		
Brouwer (2015)[39] Brouwer (2017)[102] Some concerns	**SG1:** Percutaneous laser disc decompression **N randomized:** 57 **N analyzed:** 55 (96.5%) **SG2:** Microdiscectomy **N randomized:** 58 **N analyzed:** 57 (98.3%)	**Roland Morris Disability Questionnaire, mean (SD)** Baseline: SG1 15.7 (4.9);SG2 15.5 (4.7) AMD (95% CI); [Negative AMD favors SG1] 4w: -2.5 (-4.7 To -0.2) 8w: 0.1 (-2.1 to 2.3)[Primary Outcome] 26w: 2.2 (-0.1 to 4.4) 52w: 1.1(-1.1 to 3.4)[Primary Outcome] 2y: -0.1 (-2.4 to 2.2) RM 4w to 52w: 0.2 (-1.2 to 1.6)[Primary Outcome] RM 4w to 2y: 0.0 (-1.3 to 1.3), P=1.00 main treatment effect, P=0.06 treatment X time interaction **SF-36 Physical Functioning, mean (SD)** Baseline: SG1 41.0 (22.6); SG2 38.6 (20.9) AMD (95% CI); [Positive AMD favors SG1] 4w: 18.4 (10.0 to 26.8) 8w: 5.6 (-2.7 to 13.9) 26w: -3.2 (-11.6 to 5.1) 52w: -3.2 (-11.6 to 5.2) 2y: 4.3 (-4.5 to 13.2) RM 4w to 52w: 5.3 (-0.7 to 11.2) RM 4w to 2y: 6.1(0.5 to 11.7), P=0.03 main treatment effect, P=0.001 treatment X time interaction	NR	**N (%) with complete or nearly complete recovery, based on 7-pt Likert scale** [OR > 1 favors SG1] 52w: SG1 NR (69%); SG2 NR (75%); OR 0.81 (95% CI 0.4 to 1.9) 2y: SG1 NR (70.8%); SG2 NR (60.8%); OR 1.6 (95% CI 0.7 to 3.6) **Relative difference in time to complete or nearly complete recovery, HR (95% CI)** [HR > 1 favors SG1] 52w: 0.64 (0.42 to 0.97) 2y: 0.64 (0.43 to 0.96)

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Brouwer (2015)[39] Brouwer (2017)[102] Some concerns (continued)		**Prolo Scale, Functioning Score, mean (SD)** Baseline: SG1 1.1 (0.6); SG2 0.9 (0.5) AMD (95% CI); [Positive AMD favors SG1] 4w: 0.2 (-0.2 to 0.6) 8w: -0.2 (-0.6 to 0.3) 26w: 0.1 (-0.3 to 0.5) 52w: -0.2 (-0.5 to 0.2) RM 4w to 52w: 0.0 (-0.3 to 0.3) **Prolo Scale, Economic Score, mean (SD)** Baseline: SG1 1.7 (1.7); SG2 2.1 (1.7) AMD (95% CI); [Positive AMD favors SG1] 4w: 1.1 (0.5 to 1.6) 8w: 0.2 (-0.3 to 0.8) 26w: -0.7 (-1.3 to 0.2) 52w: 0.1 (-0.4 to 0.7) RM 4w to 52w: 0.3 (-0.1 to 0.7)		
Chatterjee (1995)[38] Some concerns	**SG1:** Automated percutaneous lumbar discectomy (APLD) **N randomized:** 31 **N analyzed:** 31 (100%) **SG2:** Microdiscectomy **N randomized:** 40 **N analyzed:** 40 (100%)	NR	**Returned to work or previous level of activity by 12w** SG1: NR SG2: 37 (92.5%)	**N (%) with excellent/good outcome based on MacNab criteria** SG1: 9 (29%) SG2: 32 (80%) P < 0.001

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Erginousakis (2011)[37] High	**SG1:** Percutaneous disc decompression **N randomized:** 31 **N analyzed:** 31 (100%) **NS1:** Conservative management **N randomized:** 31 **N analyzed:** 31 (100%)	NR	**N (%) reporting that pain affected their occupational status** 12w: SG1 4 (12.9%); NS1 3 (9.7%) Calculated P=1.0 52w: SG1 4 (12.9%); NS1 22 (71%) Calculated P<0.001 2y: SG1 4 (12.9%); NS1 22 (71%) Calculated P<0.001	NR
Franke (2009)[36] Some concerns	**SG1:** Microscopically assisted percutaneous nucleotomy **N randomized:** 52 **N analyzed:** 52 (100%) **SG2:** Microdiscectomy **N randomized:** 48 **N analyzed:** 48 (100%)	**Oswestry Disability Index (RM at 52w)** Significant within group improvement over time in both groups (P< 0.001) No between group difference, P=0.08	**Duration of postoperative inability to work, mean (SD) in weeks** Overall: 7 (NR), NR by group	NR

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Gerszten (2010)[41] Some concerns (6w outcomes) High (12w and 26w outcomes)	**SG1:** Plasma disc decompression with coblation technology (PDD) **N randomized:** 46 **N analyzed:** 29 (64% of ITT sample) at 26w **NS1:** Epidural steroid injection (TFESI) **N randomized:** 44 **N analyzed:** 28 (70% of ITT sample) at 26w	**Oswestry Disability Index, mean (SD)** Baseline: SG1: 42 (14); NS1 43 (17) Change in score, mean (SE); [larger negative change favors SG1] 6w: SG1 -13 (3); NS1 -5 (2); P=0.002, calculated AMD -8 12w: SG1 -11 (3); NS1 -2 (2); P= 0.002, calculated AMD -9 26w: SG1 -14 (4); NS1 -4 (2); P=0.002, calculated AMD -10 **SF-36 Physical Functioning at 26w** Larger improvements in SG1 compared to NS1, P=0.0016 **SF-36 Social Functioning at 26w** Larger improvements in SG1 compared to NS1, P=0.0312 **SF-36 Physical Component Summary at 26w** Larger improvements in SG1 compared to NS1, P=0.0040 No significant difference between groups for the following SF-36 domains: role physical, general health, vitality, role emotional, mental health, or mental components summary	**N (%) of participants working full or part time at 26w** Reported as similar in both groups (69% to 70%)	NR

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Haines (2002)[42] High	**SG1:** Automated percutaneous discectomy, endoscopic percutaneous discectomy (APD/EPD) **N randomized:** 21 **N analyzed:** 17 (81.0%) at 26w **SG2:** Discectomy **N randomized:** 13 **N analyzed:** 10 (76.9%) at 26w	**Modified Roland Disability Score at 26w, mean (SD)** Baseline: SG1 16.9 (4.9) SG2 17.3 (4.1) 26w: SG1 6.12 (7.2) SG2 6.5 (6.1) (P=0.74) Calculated AMD (adjusted for baseline) 0.02 **SF-36 Physical Functioning at 26w, mean (SD)** Baseline: SG1 36.0 (27.1); SG2 37.2 (15.8) 26w: SG1 74.7 (27.6) SG2 73.0 (15.7) (P=0.96) Calculated AMD (adjusted for baseline) 2.9	NR	**N (%) with success at 26w** SG1: 7 (41.2%) SG2: 4 (40%) P=0.95 Success defined as either excellent or good on author defined outcome assessment matrix incorporating 4 dimensions (pain frequency and severity, ability to participate in work activities, ability to participate in leisure activities, and analgesic use **N (%) with success at 26w, based on MacNab Criteria** SG1: 11 (64.7%) SG2: 6 (60%) P=0.81
Henriksen (1996)[35] Some concerns	**SG1:** Microdiscectomy **N randomized:** 40 **N analyzed:** 39 (97.5%) **SG2:** Standard discectomy **N randomized:** 40 **N analyzed:** 40 (100%)	NR	NR	NR

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Hermantin (1999)[43] Some concerns	**SG1:** Video-assisted arthroscopic microdiscectomy **N randomized:** 30 **N analyzed:** 30 (100%) **SG2:** Discectomy, with laminotomy **N randomized:** 30 **N analyzed:** 30 (100%)	NR	**Duration of postoperative disability in time lost from work or until able to resume normal activity, mean (SD) in weeks at unspecified follow-up time** SG1: 3.9 (NR) SG2: 7 (NR)	Follow-up time for these measures NR **N (%) very satisfied with operative result based on self-report** SG1: 22 (73%) SG2: 20 (67%) Calculated P=0.78 **N (%) with satisfactory outcome** SG1: 29 (97%) SG2: 28 (93%) Calculated P=1.0 Satisfactory outcome defined as either an excellent (radicular symptoms had ceased, the tension signs had become negative, the patient had returned to his or her previous occupation or to normal activity, and the patient expressed satisfaction with the results of the operative procedure) or good (if the above criteria were met but the patient had residual back pain and had had to modify his or her occupation) outcome)
Huang (2005)[24] Some concerns	**SG1:** Microendoscopic discectomy **N randomized:** 10 **N analyzed:** 10 (100%) **SG2:** Discectomy **N randomized:** 12 **N analyzed:** 12 (100%)	NR	NR	**N (%) with excellent/good outcome, based on MacNab criteria, follow-up time point unspecified** SG1: 9 (90%) SG2: 11 (91.6%) Calculated P=1.0

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Mayer (1993)[34] High	**SG1:** Percutaneous endoscopic discectomy **N randomized:** 20 **N analyzed:** 20 (100%) **SG2:** Microdiscectomy **N randomized:** 20 **N analyzed:** 20 (100%)	NR	**Duration of postoperative disability, mean (range) in weeks** SG1: 7.7 (1 to 26) SG2: 22.9 (4 to 52) **N (%) returning to work** SG1: 19 (95%) SG2: 13 (65%) Calculated P=0.044	**Clinical score, mean (SD)** Baseline: SG1 4.55 (0.99); SG2 4.2 (0.98) 2y: SG1 8.23 (1.3); SG2 7.67 (1.9) P <0.005, favoring SG1 Clinical scoring system (modified from the system of Suezawa and Schreiber, based on pain, sensory, motor, and reflexes. A total score of 9 to 10 indicates an excellent condition, 7 to 8 good, 6 to 7 moderate, 5 or less poor. **N (%) with specified clinical score at 2y** Excellent: SG1 13 (68.4%); SG2 7 (35%) Good: SG1 4 (21%); SG2 6 (30%) Moderate: SG1 3 (16%); SG2 4 (20%) Bad: SG1 0 (0%); SG2 3 (15%) **N (%) with specified self-reported success of surgery at 2y** Excellent: SG1 9 (47%); SG2 8 (40%) Good: SG1 5 (26%); SG2 3 (15%) Satisfied: SG1 6 (32%); SG2 6 (30%) Bad: SG1 0 (0%); SG2 0 (0%)

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
McMorland (2010)[23] Some concerns	**SG1:** Microdiscectomy **N randomized:** 20 **N analyzed:** 12w 20 (100%) (outcomes reported only to 12w for ITT analysis) 24w 20 (100%) 52w 15 (75%) **NS1:** Spinal manipulation **N randomized:** 20 **N analyzed:** 12w 20 (100%) (outcomes reported only to 12w for ITT analysis) 24w 20 (100%) 52w 17 (85%)	**Roland Score, mean (SD)** Baseline: SG1 10.1 (5.7); NS1 12.0 (5.4) 6w: SG1 9.4 (6.4); NS1 9.5 (6.0) Calculated AMD 1.8 (adj. for baseline) 12w: SG1 7.2 (6.9); NS1 9.0 (6.2) Calculated AMD 0.1 (adj. for baseline) RM 6w to 12w: AMD NR, P=0.033 for time effect, P=0.199 for main treatment effect, P=0.760 for treatment X time interaction **SF-36 Physical Functioning, mean (SD)** Baseline: SG1 42.7 (22.7); NS1 47.4 (24.8) 6w: SG1 51.3 (28.2); NS1 54.8 (24.4) Calculated AMD 1.2 (adj. for baseline) 12w: SG1 65.8 (27.6); NS1 59.0 (25.4) Calculated AMD 11.5 (adj. for baseline) RM 6w to 12w: AMD NR, P=0.034 for time effect, P=0.720 for main treatment effect, P=0.448 for treatment X time interaction **SF-36 Role Physical, mean (SD)** Baseline: SG1: 17.5 (32.5); NS1 18.8 (26.7) 6w: SG1 15.0 (33.8); NS1 26.3 (37.6) 12w: SG1 28.8 (37.4); NS1 32.5 (38.1) RM 6w to 12w: AMD NR, P=0.126 for time effect, P=0.719 for main treatment effect, P=0.038 for treatment X time interaction	NR	NR

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
McMorland (2010)[23] Some concerns (continued)		**SF-36 Role Emotional, mean (SD)** Baseline: SG1 60.8 (41.0); NS1 53.4 (50.0) 6w: SG1 63.3 (47.0); NS1 66.7 (45.7) 12w: SG1 65.0 (43.9); NS1 74.5 (36.4) RM 6w to 12w: AMD NR, P=0.034 for time effect, P=0.715 for main treatment effect, P=0.410 for treatment X time interaction**SF-36 Social Functioning, mean (SD)** Baseline: SG1 50.2 (29.0); NS1 52.9 (33.0) 6w: SG1 54.1 (30.9); NS1 62.7 (20.0) 12w: SG1 67.3 (34.7); NS1 73.6 (19.7) RM 6w to 12w: AMD NR, P=0.138 for time effect, P=0.938 for main treatment effect, P=0.596 for treatment X time interaction **SF-36 Mental Health, mean (SD)** Baseline: SG1 69.2 (13.0); NS1 69.0 (21.3) 6w: SG1 77.7 (16.8); NS1 78.6 (11.6) 12w: SG1 83.2 (10.6); NS1 82.8 (8.7) RM 6w to 12w: AMD NR, P=0.001 for time effect, P=0.905 for main treatment effect, P=0.990 for treatment X time interaction		
North (2005)[46] High	**SG1:** Repeat lumbosacral decompression **N randomized:** 26 **N analyzed:** 26 (100%) **NS1:** Spinal cord stimulation **N randomized:** 24 **N analyzed:** 19 (79.2%)	**Impairment from pain in performing everyday activities at mean 2.9 years follow-up** Reported as higher in SG1 compared with NS1 but actual values are NR and differences reported as NS Everyday activities defined as work, walk, climb stairs, sleep, engage in sex, drive a car, sit at a table to eat	Reported as no significant treatment differences, but actual values NR	**N (%) with successful treatment** SG1: 3 (12%) NS1: 9 (47%) P < 0.01 Success defined as at least 50% pain relief and patient satisfaction with treatment. Long-term follow-up occurred at a mean of 2.9y (range 1.8y to 5.7y)

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Osterman (2003)[33] High	**SG1:** Microdiscectomy **N randomized:** 28 **N analyzed:** 6w 26 (93%) 12w 26 (93%) 26w 26 (93%) 52w 21 (75%) 2y 26 (93%) **NS1:** Physiotherapy **N randomized:** 28 **N analyzed:** 6w 26 (93%) 12w 26 (93%) 26w 22 (78.6%) 52w 20 (71.4%) 2y 24 (86%)	**Oswestry Disability Index, mean (SD)** Baseline: SG1 39 (15); NS1 39 (14) 6w: SG1 16 (16); NS1 22 (16) Calculated AMD -6 (adj. for baseline) 12w: SG1 8 (11); NS1 14 (14) Calculated AMD -6 (adj. for baseline) 26w: SG1 8 (12); NS1 12 (15) Calculated AMD -4 (adj. for baseline) 52w: SG1 10 (13); NS1 11 (14) Calculated AMD -1 (adj. for baseline) 2y: SG1 6 (9); NS1 11 (16) Calculated AMD -5 (adj. for baseline) RM 6w to 2y, AMD (95% CI) [negative AMD favors SG1] -3 (-10 to 4)	**VAS100 work ability, mean (SD)** 6w: SG1 68 (27); NS1 63 (32) Calculated AMD: 5 (95% CI, -11.5 to 21.5) 12w: SG1 84 (14); NS1 70 (31) Calculated AMD: 14 (95% CI, 0.60 to 27.4) 26w: SG1 87 (18); NS1 75 (30) Calculated AMD: 12 (95% CI, -2.1 to 26.1) 52w: SG1 82 (26); NS1 81 (27) Calculated AMD: 1 (95% CI, -15.7 to 17.7) 2y: SG1 89 (16); NS1 79 (28) Calculated AMD: 10 (95% CI, -2.8 to 22.8) RM 6w to 2y, AMD (95% CI) [positive AMD favors SG1] 5 (-7 to 18)	**N (%) reporting full recovery** 6w: SG1 5 (19.2%); NS1 0 (0%), P < 0.05 12w: SG1 5 (19.2%); NS1 4 (15.4%), Calculated P=1.0 26w: NR 52w: SG1 7 (33.3%); NS1 5 (25%), Calculated P=0.73 2y: NR
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98] High	**SG1:** Microdiscectomy **N randomized:** 141 **N analyzed:** 52w: 140 (99.3%) 2y: 130 (92.2%) 5y: 115 (81.6%) **NS1:** Conservative management **N randomized:** 142	**Roland Disability Questionnaire, mean (SD)** Baseline: SG1 16.5 (4.4); NS1 16.3 (3.9) AMD (95% CI) [negative AMD favors SG1] 8w: -3.1 (-4.3 to -1.7) 26w: -0.8 (-2.1 to 0.5) 52w: -0.4 (-1.7 to 0.9) 2y: -0.5 (-1.8 to 0.8) 5y: 0.1 (-1.3 to 1.4)	NR	**Median time to recovery, weeks** SG1: 4.0 (95% CI, 3.7 to 4.4) NS1: 12.1 (95% CI, 9.5 to 14.9) AMD NR, P<0.001 **Relative difference in time to complete or nearly complete recovery at 52w, HR (95%CI)**; [HR >1 favors SG1] 1.97 (95% CI, 1.72 to 2.22)

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98] High (continued)	**N analyzed:** 52w: 141 (99.3%) 2y: 130 (91.5%) 5y: 116 (81.7%)	**Cumulative score on Roland 0 to 52w** SG1: AUC 273.9 (SE 20.7) NS1: AUC 316.3 (SE 18.8) AMD: -42.5 (95% CI, -97.4 to 12.4) **Cumulative score on Roland 0 to 2 y** SG1: AUC 397.2 (SE 39.0) NS1: AUC 458.2 (SE 35.3) AMD: -61.0 (95% CI, -164.5 to 42.5) **SF-36 Physical Functioning, mean (SD)** Baseline: SG1 33.9 (19.6); NS1 34.6 (19.0) AMD (95% CI); [positive AMD favors SG1] 8w: 9.3 (4.4 to 14.2) 26w: 1.5 (-3.4 to 6.4) 52w: 2.2 (-2.8 to 7.2) **Prolo Scale, Functional Score, mean (SD)** Baseline: NR AMD (95% CI); [positive AMD favors SG1] 8w: 0.8 (-0.6 to 1.1) 26w: 0.5 (0.2 to 0.7) 52w: -0.04 (-0.3 to 0.2) 2y: NR 5y: NR **Prolo Scale, Economic Score, mean (SD)** Baseline: NR AMD (95% CI); [positive AMD favors SG1] 8w: -0.5 (-0.8 to -0.1) 26w: 0.1 (-0.3 to 0.5) 52w: -0.2 (-0.6 to 0.2) 2y: NR 5y: NR		**Likert scale (7-pt) global perception of recovery, mean score (SE)** [negative AMD favors SG1] 8w: SG1 2.2 (0.1); NS1 3.1 (0.1); AMD -0.9 (95% CI, -1.2 to -0.6) 26w: SG1 2.1 (0.1); NS1 2.3 (0.1); AMD -0.2 (95% CI, -0.5 to 0.1) 52w: SG1 1.9 (0.1); NS1 2.1 (0.1); AMD -0.2 (95% CI, -0.4 to 0.1) **N (%) reporting complete or nearly complete recovery, based on 7-pt Likert scale at 2y** SG1: NR (81.3%) NS1: NR (83.6%) AMD: -2.4% (95% CI, -12.0% to 7.2)

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98] High (continued)		**SF-36 Social Functioning, mean (SD)** Baseline: SG1 44.6 (30.1); NS1 43.4 (27.1) AMD (95% CI); [positive AMD favors SG1] 8w: 2.3 (-3.7 to 8.3) 26w: 4.5 (-1.4 to 10.6) 52w: 1.3 (-4.7 to 7.3) **SF-36 Role Emotional, mean (SD)** Baseline: SG1 51.0 (46.0): NS1 52.4 (46.0) AMD (95% CI); [positive AMD favors SG1] 8w: 3.1 (-3.0 to 9.3) 26w: 3.9 (-2.3 to 10.1) 52w: -1.4 (-7.6 to 4.8)		
Ruetten (2008)[31] High	**SG1:** Endoscopic (interlaminar or transforaminal) discectomy **N randomized:** 100 **N analyzed:** 91 (91%) **SG2:** Microdiscectomy **N randomized:** 100 **N analyzed:** 87 (87%)	**Oswestry Disability Index, mean (SD)** Baseline: SG1 75 (NR); SG2 73 (NR) 12w: SG1 22 (NR); SG2 26 (NR) Calculated AMD (adjusted for baseline): -6 26w: SG1 21 (NR); SG2 24 (NR) Calculated AMD (adjusted for baseline): -5 52w: SG1 19 (NR); SG2 23 (NR) Calculated AMD (adjusted for baseline): -6 2y: SG1 20 (NR); SG2 24 (NR) Calculated AMD (adjusted for baseline): -6 Between-group differences reported as NS.	**Duration of postoperative work disability, mean (SD) in weeks** SG1: 3.57 (NR) SG2: 7 (NR) P < 0.01	NR

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Ruetten (2009)[47] High	**SG1:** Revision endoscopic discectomy **N randomized:** 50 **N analyzed:** 45 (90%) **SG2:** Revision microdiscectomy **N randomized:** 50 **N analyzed:** 42 (84%)	**Oswestry Disability Index, mean (SD)** Baseline: SG1 80 (NR); SG2 84 (NR) 12w: SG1 22 (NR); SG2 18 (NR) 26w: SG1 24 (NR); SG2 19 (NR) 52w: SG1 18 (NR); SG2 23 (NR) 2y: SG1 20 (NR); SG2 21 (NR) Differences between groups reported as NS.	**Postoperative work disability, mean (SD) in weeks** SG1: 4 (NR) SG2: 7.4 (NR) P < 0.01	**N (%) satisfied with surgery and would undergo the operation again** SG1: 43 (95%) SG2: 36 (86%) Calculated P=0.15
Ryang (2008)[30] Gempt (2013)[103] High	**SG1:** Trocar microdiscectomy **N randomized:** 30 **N analyzed:** unclear **SG2:** Microdiscectomy **N randomized:** 30 **N analyzed:** unclear	Outcomes measured over average follow-up of 1.33y (range 26w to 2.17y) Additional long-term outcomes reported for 38 participants over average of 2.8y (range 52w to 4.5y) **Oswestry Disability Index, mean (SD)** Baseline: SG1 53.1 (19.2); SG2 56.7 (23.1), P=0.48 Follow-up: SG1 12.0 (14.0); SG2 12 (18.8), P=0.83 Calculated AMD (adjusted for baseline): 3.6 Long-term follow-up: SG1 12.95 (11.2); SG2 18.53 (15.37), P reported as NS Calculated AMD (adjusted for baseline): -1.98 **SF-36 Physical Functioning, mean (SD)** Baseline: SG1 33 (23.6); SG2 28.5 (26.1), P=0.45 Follow-up: SG1 74.8 (23.3); SG2 80.4 (19.6), P=0.64, calculated AMD (adjusted for baseline) -10.1 Long-term follow-up: SG1 72.4 (25.0); SG2 78.2 (20), P=0.436, calculated AMD (adjusted for baseline) 6	NR	Long-term outcomes reported for 38 participants over average of 2.8y (range 52w to 4.5y) **VAS10 for improvement from baseline to long-term follow-up (95%CI)** SG1: 4.92 (3.23; 6.61); P<0.001 SG2: 4.64 (3.03; 6.26); P<0.001 Between group P value NR.

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Ryang (2008)[30] Gempt (2013)[103] High (continued)		**SF-36 Role physical, mean (SD)** Baseline: SG1 11 (20.5); SG2 14.5 (28.7), P=0.87 Follow-up: SG1 66.3 (41); SG2 73.2 (37.3), P=0.55 Long-term follow-up: SG1 58.3 (42.0), SG2 81.9 (36.2), P= 0.080 **SF-36 Social Functioning, mean (SD)** Baseline: SG1 44 (32.5); SG2 46.5 (30), P=0.67 Follow-up: SG1 78.3 (27.2); SG2: 88.4 (18), P=0.18 Long-term follow-up: SG1 80.9 (26.1); SG2 88.2 (20.6), P=0.350 **SF-36 Role emotional, mean (SD)** Baseline: SG1 38.7 (45.8); SG2 43 (48.1), P=0.8 Follow-up: SG1 60.6 (46.7); SG2 85.2 (31.1), P=0.03 Long-term follow-up: SG1 77.8 (37.9); SG2 77.8 (42.8), P=1.00		
Sasaoka (2006)[25] High	**SG1:** Microendoscopic discectomy **N randomized:** 15 **N analyzed:** unclear **SG2:** Microdiscectomy **N randomized:** 11 **N analyzed:** unclear	NR	NR	**Mean (SD) percentage improvement of the Japanese Orthopaedic Association (JOA) score at 52w** SG1: 84.7% (NR) SG2: 88.6% (NR) Difference reported as not significant The JOA score comprises of back pain symptoms, leg pain and/or tingling, gait, clinical signs (straight leg-raising test, sensory disturbance, motor disturbance), restriction of activities of daily living, and urinary bladder function

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Teli (2010)[29] Some concerns	**SG1:** Microendoscopic discectomy **N randomized:** NR **N analyzed:** 70 **SG2:** Microdiscectomy **N randomized:** NR **N analyzed:** 72 **SG3:** Discectomy **N randomized:** NR **N analyzed:** 70	**Oswestry Disability Index, mean (SD)** Baseline: SG1 40 (4); SG2 41 (4); SG3 39(4) 26w: SG1 12 (4); SG2 12 (4); SG3 12 (4) Calculated AMD (adjusted for baseline) SG1/SG2 1 Calculated AMD (adjusted for baseline) SG2/SG3 2 52w: SG1 14 (4); SG2 13 (4); SG3 13 (4) Calculated AMD (adjusted for baseline) SG1/SG2 2 Calculated AMD (adjusted for baseline) SG2/SG3 2 2y: SG1 14 (6); SG2 16 (5); SG3 15 (3) Calculated AMD (adjusted for baseline) SG1/SG2 1 Calculated AMD (adjusted for baseline) SG2/SG3 1 No difference among the three groups at any time point, P=0.81	NR	NR
Thome (2005)[28] Barth (2008)[104] Some concerns	**SG1:** Sequestrectomy **N randomized:** 42 **N analyzed:** 42 (100%) **SG2:** Microdiscectomy **N randomized:** 42 **N analyzed:** 42 (100%)	For all SF-36 scores, higher values represent more favorable outcomes. **SF-36 Physical Functioning, mean (SD)** Baseline: SG1 37.9 (21.8); SG2 34.4 (26.9), P=0.53 12w to 26w: SG1 75.1 (24.0); SG2 69.4 (25.6), P=0.32 Calculated AMD 2.2 2y: SG1 82.4 (20.9); SG2 71.9 (23.2), P=0.026 Calculated AMD 7	**N (%) reporting specified categories of impairment of work at 12w to 26w** Much better: SG1 NR (31%); SG2 NR (33%) Better: SG1 NR (33%); SG2 NR (19%) Equal: SG1 NR (25%); SG2 NR (27%) Worse: SG1 NR (8%); SG2 NR (15%)	**N (%) with specified patient satisfaction index scores at 12w to 26w** [Lower scores represent more favorable outcome] 1: SG1 NR (67%); SG2 NR (67%) 2: SG1 NR (30%); SG2 NR (24%) 3: SG1 NR (0%); SG2 NR (6%) 4: SG1 NR (3%); SG2 NR (3%) P=0.693

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Thome (2005)[28] Barth (2008)[104] Some concerns (continued)		**SF-36 role physical, mean (SD)** Baseline: SG1 14.7 (29.0); SG2 16.4 (29.0), P=0.79 12w to 26w: SG1 53.21 (42.2); SG2 49.5 (43.8), P=0.72 2y: SG1 67.1(44.7); SG2 55.0 (45.7), P=0.249 **SF-36 social functioning, mean (SD)** Baseline: SG1 54.0(26.6); SG2 56.8 (26.6), P=0.64 12w to 26w: SG1 83.6 (23.8); SG2 85.1 (20.2), P=0.77 2y: SG1 89.8 (18.6); SG2 79.6 (22.5), P=0.013 **SF-36 role emotional, mean (SD)** Baseline: SG1 58.8 (46.8); SG2 57.1 (48.2), P=0.88 12w to 26w: SG1 82.4 (36.1); SG2 79.8 (36.0), P=0.76 2y: SG1 85.9 (32.5); SG2 73.56 (44.8), P=0.311 Note: the Prolo Scale used in this study was modified from original version, it used a 0 to 5 Likert scale for each subscale, and summed to obtain a total score that ranged from 2 to 10. **N (%) with total Prolo score >=7, 12w to 26w** [higher proportion reflects more favorable outcome] SG1: NR (92%) SG2: NR (76%) P= 0.11	Much worse: SG1 NR (3%); SG2 NR (7%) P=0.415 **N (%) reporting specified categories of impairment of work at 2y** Much better: SG1 NR (37%); SG2 NR (31%) Better: SG1 NR (37%); SG2 NR (11%) Equal: SG1 NR (12%); SG2 NR (41%) Worse: SG1 NR (11%); SG2 NR (11%) Much worse: SG1 NR (3%); SG2 NR (6%) P=0.112	**N (%) with specified patient satisfaction index scores at 2y** [Lower scores represent more favorable outcome] 1: SG1 NR (68%); SG2 NR (56%) 2: SG1 NR (29%); SG2 NR (29%) 3: SG1 NR (3%); SG2 NR (9%) 4: SG1 NR (0%); SG2 NR (6%) P=0.087 **N (%) unsatisfied with surgery as measured by patient satisfaction index score of 3 or 4, 12w to 26w** SG1: NR (3%) SG2: NR (18%) P = 0.06

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Thome (2005)[28] Barth (2008)[104] Some concerns (continued)		**N (%) with specified total Prolo scores at 12w to 26w** 1-4: SG1 NR (3%); SG2 NR (9%) 5-6: SG1 NR (6%); SG2 NR (9%) 7-8: SG1 NR (31%); SG2 NR (12%) 9-10: SG1 NR (60%); SG2 NR (70%) P=0.852 **N (%) with specified Prolo scores at 2y** 1-4: SG1 NR (0%); SG2 NR (3%) 5-6: SG1 NR (3%); SG2 NR (12%) 7-8: SG1 NR (23%); SG2 NR (19%) 9-10: SG1 NR (74%); SG2 NR (66%) P=0.20		
Tullberg (1993)[27] Some concerns	**SG1:** Microdiscectomy **N randomized:** 30 **N analyzed:** 29 (97%) **SG2:** Discectomy **N randomized:** 30 **N analyzed:** 29 (97%)	NR	**Postoperative sick leave, mean (SD) in weeks** Full time sick leave SG1: 10.4 (NR) SG2: 10.1 (NR) Half time sick leave SG1: 2.6 (NR) SG2: 2.9 (NR) **N (%) of patients out of work at unspecified time point during follow-up** SG1: 5 (16.7%) SG2: 2 (6.7%) Calculated P=0.42	**N (%) with specified opinion on recovery at 52w** SG1 Totally recovered: SG1: 11 (37.9%); SG2: 6 (20.7%); Calculated P=0.25 Almost recovered: SG1: 8 (27.6%); SG2: 14 (48.3%); Calculated P=0.18 Good: SG1: 6 (20.7%); SG2: 6 (20.7%); Calculated P=1.0 Unchanged: SG1: 4 (13.8%); SG2: 2 (6.9%); Calculated P=0.67 Worse: SG1: 0 (0%); SG2: 1 (3.4%); Calculated P=1.0

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Weber (1983)[26] High	**SG1:** Discectomy **N randomized**: 60 **N analyzed:** 60 (100%) **NS1:** Conservative management **N randomized:** 66 **N analyzed**: 66 (100%)	NR	**Cumulative N (%) with permanent incapacitation and receiving disablement benefit** 0-4y: SG1 3 (5%); NS1 8 (12.1%); Calculated P=0.21 5-10y: SG1 7 (11.7%); NS1 8 (12.1%); Calculated P=1.0	Result (good, fair, poor, or bad) based on author's evaluation of neurological deficits, working capacity, pain, and mobility of lumbar spine. **N (%) with specified result at 52w** Good: SG1 39 (65.0%); NS1 24 (36.4%) Fair: SG1 16 (26.7%); NS1 28 (42.4%) Poor: SG1 5 (8.3%); NS1 13 (19.7%) Bad: SG1 0 (0%); NS1 1 (1.5%) P= 0.0015 **N (%) with specified result at 4y** Good: SG1 40 (70.2%); NS1 34 (51.5%) Fair: SG1 9 (15.8%); NS1 24 (36.4%) Poor: SG1 8 (14.0%); NS1 5 (7.6%) Bad: SG1 0 (0%); NS1 3 (4.5%) Not examined: SG1 2 (3.3); NS1 0 (0%) P reported as NS **N (%) with specified result at 10y** Good: SG1 35 (63.6%); NS1 37 (56.0%) Fair: SG1 16 (29.1%); NS1 25 (37.9%) Poor: SG1 4 (7.3%); NS1 4 (6.1%) Bad: SG1 0 (0%); NS1 0 (0%) Not examined: SG1 2 (3.3); NS1 0 (0%) P reported as NS **Relapses** 0 to 4y: SG1 8 (13.3%); NS1 14 (6.1%) 5 to 10y: SG1 13 (21.7%); NS1 11 (16.7%) (5 patients in this period also had a relapse in the prior period in both groups)

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Weinstein (2006)[22] Weinstein (2008)[99] Lurie(2014)[100] SPORT High	**SG1:** Discectomy /microdiscectomy **N randomized**: 245 **N analyzed:** 232 in main study's primary analyses. 52w: 202 (82.47%) 2y: 186 (75.9%) 3y: 180 (73.5%) 4y: 149 (60.8%) 8y: 157 (64.1%) **NS1:** Conservative management **N randomized**: 256 **N analyzed**: 240 included in main study's primary analyses. 52w: 213 (83.2%) 2y: 187 (73.0%) 3y: 170 (66.4%) 4y: 150 (58.6%) 8y: 152 (59.4%)	**SF-36 Physical Functioning, mean (SD)** Baseline: SG1 39.7 (24.9); NS1 39.2 (25.7) AMD (95% CI) [positive AMD favors SG1] 12w: 2.8 (-2.5 to 8.1) 52w: 1.2 (-4.1 to 6.5) 2y: 0 (-5.4 to 5.5) 4y: 2.2 (-3.7 to 8) 8y: 1.7 (-4.0 to 7.4) RM 12w to 2y: AMD NR, P=0.71 RM 12w to 4y: AMD NR, P=0.42 RM 12w to 8y: AMD NR, P=0.47 **Oswestry Disability Index, mean (SD)** Baseline: SG1 47.5 (21.4); NS1 46.3 (20.6) AMD (95% CI) [negative AMD favors SG1] 12w: -4.7 (-9.3 to -0.2) 52w: -3.2 (-7.8 to 1.3) 2y: -2.7 (-7.4 to 1.9) 4y: -3.6 (-8.6 to 1.4) 8y: -4.2 (-9.0 to 0.7) RM 12w to 2y: AMD NR, P=0.21 RM 12w to 4y: AMD NR, P=0.074 RM 12w to 8y: AMD NR, P=0.096	**N (%, (SE %) working full or part-time** Baseline: SG1 142 (61.2% (NR)); NS1 148 (61.7% (NR)) 12w: SG1 NR, 63.8% (3.3%); NS1 NR, 69.4% (3.1%); AMD (95% CI): -5.6% (-14.5% to 3.4%) 52w: SG1 NR, 76.4% (2.9%); NS1 NR 77.0% (2.8%); AMD (95% CI): -0.6% (-8.6% to 7.3%) 2y: SG1 NR, 74.2% (3.1%); NS1 NR, 76.4% (3.0%); AMD (95% CI): -2.2% (-10.6% to 6.2%) 4y: SG1 71.4% (NR); NS1 75.1% (NR); AMD (95% CI): −3.8 (−13.3, 5.8)	**Very/somewhat satisfied with symptoms, mean % (SE)**; [positive AMD favors SG1] 12w: SG1 54.3% (3.5%); NS1 43.0% (3.4%) AMD:11.3% (95% CI, 1.6% to 20.9%) 52w: SG1 64.7% (3.4%); NS1 58.5% (3.4%) AMD: 6.1% (95% CI, -3.5% to 15.5%) 2y: SG1 68.3% (3.4%); NS1 64.4% (3.5%) AMD: 4.0% (95% CI, -5.6% to 13.5%) 4y: SG1 64.7% (NR); NS1 61.3% (NR) AMD: 3.4% (95% CI, -7.7% to 14.6%) 8y: SG1 74.3% (NR); NS1 67.4% (NR) AMD: 6.8% (95% CI, -3.4% to 17%) RM 12w to 2y: AMD NR, P=0.17 RM 12w to 4y: AMD NR, P reported as not significant RM 12w to 8y: AMD NR, P=0.013 favoring SG1 **% (SE %) Self-rated progress: major improvement**; [positive AMD favors SG1] 12w: SG1 66.3% (3.3%); NS1 62.1% (3.4%)

(continued)

Table D-4. Individual study findings related to efficacy outcomes Part 2 (functioning/disability, return to work, other efficacy outcomes) (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N randomized N analyzed	Physical, Social, or Psychological Functioning and Disability	Return to Work	Other
Weinstein (2006)[22] Weinstein (2008)[99] Lurie (2014)[100] SPORT High (continued)				AMD: 4.2% (95% CI, -5.1% to 13.5%) 52w: SG1 75.7% (3.0%); NS1 66.7% (3.2%) AMD: 9.0% (95% CI, 0.3% to 17.6%) 2y: SG1 76.3% (3.1%); NS1 69.3% (3.3%) AMD: 7.0% (95% CI, -1.9% to 15.9%) 4y: SG1 72.5% (NR); NS1: 65% (NR) AMD: 7.5% (95% CI, -3.2% to 18.1%) 8Y: SG1 62.3%; NS1: 58.2% AMD: 4.1% (95% CI, -7% to 15.3%) RM 12w to 2y: AMD NR, P=0.04 favoring SG1 RM 12w to 4y: AMD NR, P reported as NS RM 12w to 8y: AMD NR, P= 0.013 favoring SG1

Abbreviations: AMD = absolute mean difference; AUC = area under the curve; CI = confidence interval; N = number; NR = not reported; NS = not significant; RM = repeated measure; HR = hazard ratio; SD = standard deviation; SE = standard error; SF-36 = Short Form 36; SG = surgical group; NS = nonsurgical group; SPORT = Spine Patient Outcomes Research Trial; VAS = visual analog scale; w = week(s); y = year(s).

Note: For continuous outcome measures, studies either reported 1) the difference in mean scores at a follow-up time point (e.g. mean score in SG1 minus mean score in NS1 at 6 weeks) or 2) mean change from baseline scores at a follow-up time point (e.g., mean change in score in SG1 minus mean change in score for NS1 at 6 weeks). The absolute mean difference (AMD) between groups reported or calculated in this table reflects the mean difference between groups with respect to the change in score.

- For outcomes where a higher score represents fewer symptoms, a positive AMD means the intervention group (SG1) improves symptoms more than the comparator group (NS1 or SG2 or SG3).
- For outcomes where a lower score represents fewer symptoms, a negative AMD means the inter intervention group (SG1) improves symptoms more than the comparator group (NS1 or SG2 or SG3).

Table D-5. Individual study findings related to safety outcomes

Main Study Author (Year); Follow-up Studies Author (Year); Trial Name; Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48] The Sciatica Micro-Endoscopic Discectomy Randomized Controlled Trial Low for follow-up through 2 years; Some concerns for follow-up longer than 2 years	**SG1:** Tubular discectomy **N randomized:** 167 **N analyzed:** 166 (99.4%) in main study's primary analyses. 52w: 156 (93.4%) 2y: 154 (92.2%) 3y: 117 (70.1%) 4y: 117 (70.1%) 5y: 106 (63.5%) **SG2:** Microdiscectomy **N randomized:** 161 **N analyzed:** 159 (98.8%) in main study's primary analyses. 52w: 151 (93.8%) 2y: 144 (89.4%) 3y: 106 (65.8%) 4y: 102 (63.4%) 5y: 98 (60.9%)	**All-cause mortality** 1y: SG1 0 (0%); SG2 0 (0%) 2y: SG1 0 (0%); SG2 1 (0.6%) 3y: SG1 0 (0%); SG2 3 (1.9%) 4y: SG1 0 (0%); SG2 3 (1.9%) 5y: SG1 2 (1.2%); SG2 3 (1.9%) **Surgical mortality** NR	**N (%) of intraoperative complications (a patient could have more than one complication)** SG1: 20 (12%) SG2: 13 (8%) P=0.27 **N (%) with dural tears** SG1: 14 SG2: 7 P=0.18 **N (%) with nerve root injuries** SG1: 3 SG2: 3 Calculated P=1.00 **N (%) with exploration started at wrong level** SG1: 1 SG2: 5 Calculated P=0.11 **N (%) with other intraoperative complications** SG1: 2 SG2: 0 Calculated P=0.50 **N (%) of postoperative complications (a patient could have more than one complication)** SG1: 19 (11%) SG2: 14 (9%) P=0.47	**N (%) with reoperations at 52w** SG1: 17 (10) SG2: 11 (7) P=0.33 Reason for reoperations was recurrent disc herniation in 20 (71%) cases. **N (%) with reoperations at 2y** SG1: 23 (15) SG2: 14 (10) P=0.22 Reason for repeated surgery was recurrent disc herniation (same level) in 25 (68%) cases. **N reoperations at 5y** (patients may have had multiple reoperations) SG1: 39 SG2: 23 P=0.10 **N (%) with any reoperation at 5y** SG1: 30 (18.1) SG2: 21 (13.2) P=0.29 **N (%) with one reoperation at 5y** SG1: 23 (13.8) SG2: 19 (11.9) P=0.62	NR

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48] (continued)			**N (%) with wound hematoma** SG1: 2 SG2: 1 Calculated P=1.00 **N (%) with wound infections** SG1: 0 SG2: 0 Calculated P=1.00 **N (%) with urinary tract infections** SG1: 0 SG2: 1 Calculated P=0.49 **N (%) with cerebrospinal fluid leaks** SG1: 1 SG2: 2 Calculated P=0.62 **N (%) with micturition disturbances requiring a catheter** SG1: 3 SG2: 2 Calculated P=1.00 **N (%) with deep venous thromboses in the leg** SG1: 0 SG2: 0 Calculated P=1.00 **N (%) with increase in sensory deficit** SG1: 5 SG2: 6 Calculated P=0.77	Reasons for reoperation was recurrent disc herniation (same level) in 36 (85.7%) of cases, disc herniation (other level) in 3 (7.1%) of cases, and instrumented fusion in 3 (7.1%) of cases. **N (%) with two reoperations at 5y** SG1: 5 (3.0) SG2: 2 (1.3) P=0.45 Reasons for reoperation was recurrent disc herniation (same level) in 12 (85.7%) of cases, disc herniation (other level) in 0 (0%) of cases, and instrumented fusion in 2 (14.3%) of cases. **N (%) with three reoperations at 5y** SG1: 2 (1.2) SG2: 0 (0) P=0.50 Reasons for reoperation was recurrent disc herniation (same level) in 4 (66.7%) of cases, disc herniation (other level) in 1 (16.7%) of cases, and instrumented fusion in 1 (16.7%) of cases. **N (%) with instrumented fusion at 5y** SG1: 6 (3.6%) SG2: 0 (0%) Calculated P=0.01	

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48] (continued)			**N (%) with increase in motor deficit** SG1: 0 SG2: 3 Calculated P=0.12 **N (%) with other postoperative complication (includes allergic reaction, micturition disturbances not requiring a catheter, deep venous thrombosis of arm, sensory deficit of arm, sensory cerebrovascular accident, fever without focus and psychiatric dysfunction)** SG1: 11 SG2: 1 Calculated P=0.006		
Brouwer (2015)[39] Brouwer (2017)[102] Some concerns	**SG1:** Percutaneous laser disc decompression **N randomized:** 57 **N analyzed:** 55 (96.5%) **SG2:** Microdiscectomy **N randomized:** 58 **N analyzed:** 57 (98.3%)	**All-cause mortality** NR **Surgical mortality** NR	**N (%) with surgical complications** SG1: 3 (5%) (all transient nerve root injury) SG2: 6 (11%) (3 CSF leak, 1 micturition problem requiring catheter, 1 transient nerve root injury, 1 surgery at wrong level) Calculated P=0.49 **N (%) with technical failure** SG1: 5 (9%) SG2: NA	**N (%) with reoperations at 52w** SG1: 24 (44%) SG2: 9 (16%) Calculated P=0.002	NR

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Chatterjee (1995)[38] Some concerns	**SG1:** Automated percutaneous lumbar discectomy **N randomized:** 31 **N analyzed:** 31 (100%) **SG2:** Microdiscectomy **N randomized:** 40 **N analyzed:** 40 (100%)	**All-cause mortality** NR **Surgical mortality** NR	NR	**N (%) with reoperations at unspecified follow-up time point** SG1: 20 (64.5%) (offered microdiscectomy) SG2: 1 (2.5%) (for recurrent disc protrusion) Calculated P<0.001	NR
Erginousakis (2011)[37] High	**SG1:** Percutaneous disc decompression **N randomized:** 31 **N analyzed:** 31 (100%) **NS1:** Conservative management **N randomized:** 31 **N analyzed:** 31 (100%)	**All-cause mortality** NR **Surgical mortality** NR	**N (%) with intraoperative or postoperative complications up to 2y** SG1: 0 (0%) NS1: NA	**N (%) with reoperations at unspecified follow-up time** SG1: 1 (3.2%) NS1: NA	NR

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Franke (2009)[36] Some concerns	**SG1:** Microscopically-assisted percutaneous nucleotomy **N randomized:** 52 **N analyzed:** 52 (100%) **SG2:** Microdiscectomy **N randomized:** 48 **N analyzed:** 48 (100%)	**All-cause mortality** NR **Surgical mortality** NR	**N (%) with intraoperative dural tear** SG1: 2 (3.9%) SG2: 3 (6.3%) Calculated P=0.67	**N (%) with reoperations** SG1: 2 (3.9%) SG2: 5 (10.4%) Calculated P=0.26 **N (%) with reoperations due to relapse (same level, same side)** Overall: 5 (5%) SG1: 1 (1%) SG2: 4 (4%) Calculated P=0.19 **N (%) with reoperations due to progressive disc degeneration with segmental instability** Overall: 2 (2%) SG1: 1(1%) SG2: 1 (1%) Calculated P=1.0	NR
Gerszten (2010)[41] Some concerns (6w outcomes) High (12w and later outcomes)	**SG1:** Plasma disc decompression with coblation technology **N randomized:** 46 **N analyzed:** 29 (64% of ITT sample) at 26w **NS1:** Epidural steroid injection **N randomized:** 44 **N analyzed:** 28 (70% of ITT sample) at 26w	**All-cause mortality** At 26w SG1: 1 (2.2%) NS1: 1 (2.3%) (Causes of death: myocardial infarction, acute pyelonephritis) **Surgical mortality** SG1: 0 (0%) NS1: 0 (0%)	**N (%) with procedure-related adverse events** SG1: 5 (11%) NS1: 7 (18%) Calculated P=0.55	NR	**Reduction in use of narcotics, at 26w** Reported as not significantly different between groups

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Haines (2002)[42] High	**SG1:** Automated percutaneous discectomy, endoscopic percutaneous discectomy **N randomized:** 21 **N analyzed:** 17 (81.0%) at 26w **SG2:** Discectomy **N randomized:** 13 **N analyzed:** 10 (76.9%) at 26w	**All-cause mortality** NR **Surgical mortality** NR	NR	NR	NR
Henriksen (1996)[35] Some concerns	**SG1:** Microdiscectomy **N randomized:** 40 **N analyzed:** 39 (97.5%) **SG2:** Standard discectomy **N randomized:** 40 **N analyzed:** 40 (100%)	**All-cause mortality** NR **Surgical mortality** NR	**N (%) with surgical infection** Overall: 5 (6.3%) NR by group	NR	NR
Hermantin (1999)[43] Some concerns	**SG1:** Video-assisted arthroscopic microdiscectomy **N randomized:** 30 **N analyzed:** 30 (100%) **SG2:** Discectomy, with laminotomy **N randomized:** 30 **N analyzed:** 30 (100%)	**All-cause mortality** NR **Surgical mortality** SG1: 0 (0%) SG2: 0 (0%)	**N (%) with spinal fluid leakage** SG1: 0 (0%) SG2: 1 (3.3%) **N (%) with infection** SG1: 0 (0%) SG2: 0 (0%) **N (%) with neurovascular injuries** SG1: 0 (0%) SG2: 0 (0%)	**N (%) with reoperations** SG1: 1 (3.3%) (for the treatment of mild lateral stenosis that had not been recognized at the time of the index operation) SG2: 2 (6.7%) (1 for repair of dural sac/spinal fluid leak and 1 for persistent radicular symptoms) Calculated P=1.0	**Duration of use of narcotics postoperatively, mean in weeks (range)** SG1: 1 (0.43 to 2) SG2: 3.65 (1 to 8)

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Huang (2005)[24] Some concerns	**SG1:** Microendoscopic discectomy **N randomized:** 10 **N analyzed:** 10 (100%) **SG2:** Discectomy **N randomized:** 12 **N analyzed:** 12 (100%)	**All-cause mortality** NR **Surgical mortality** SG1: 0 (0%) SG2: 0 (0%)	**N (%) with nerve root sleeve tear** SG1: 1 (10%) SG2: 0 (0%) **N (%) with superficial wound infection** SG1: 0 (0%) SG2: 1 (8.3%)	NR	NR
Mayer (1993)[34] High	**SG1:** Percutaneous endoscopic discectomy **N randomized:** 20 **N analyzed:** 20 (100%) **SG2:** Microdiscectomy **N randomized:** 20 **N analyzed:** 20 (100%)	**All-cause mortality** NR **Surgical mortality** SG1: 0 (0%) SG2: 0 (0%)	**N (%) with complications** SG1: 0 (0%) SG2: 0 (0%)	**N (%) with reoperations** SG1: 3 (15%) (had percutaneous endoscopic discectomy and later microsurgical discectomy, 2 because of lack of improvement and 1 due to recurrence of symptoms) SG2: 1 (5%) (due to epidural scar tissue and progressive neurological symptoms) Calculated P=0.61	NR

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
McMorland (2010)[23] Some concerns	**SG1:** Microdiscectomy **N randomized:** 20 **N analyzed:** 12w: 20 (100%) (outcomes reported only to 12w for ITT analysis) 24w: 20 (100%) 52w: 15 (75%) **NS1:** Spinal manipulation **N randomized:** 20 **N analyzed:** 12w: 20 (100%) (outcomes reported only to 12w for ITT analysis) 24w: 20 (100%) 52w: 17 (85%)	**All-cause mortality** NR **Surgical mortality** SG1: 0 (0%) NS1: NA	**N (%) with new neurologic deficits** SG1: 0 (0%) NS1: 0 (0%) **N (%) with significant adverse events** SG1: 0 (0%) NS1: 0 (0%)	**N (%) with reoperations** SG1: 0 (0%) NS1: NA	NR
North (2005)[46] High	**SG1:** Repeat lumbosacral decompression **N randomized:** 26 **N analyzed:** 26 (100%) **NS1:** Spinal cord stimulation **N randomized:** 24 **N analyzed:** 19 (79.2%)	**All-cause mortality** At 26w: SG1: 0 (0%) NS1: 1 (2%) **Surgical mortality** SG1: 0 (0%) NS1: 0 (0%)	**N (%) with site infection** SG1: 0 (0%) NS1: 1 (4.2%)	**N (%) with reoperations (SG1) or hardware revisions (NS1)** SG1: None explicitly reported NS1: 3 (12.5%)	Long-term follow-up occurred at a mean of 2.9y (range 1.8 to 5.7) **N (%) with opioid use stable or decreased at long-term follow-up** SG1: 15 (58%) NS1: 20 (87%) P = 0.025

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Osterman (2003)[33] High	**SG1:** Microdiscectomy **N randomized:** 28 **N analyzed:** 6w 26 (93%) 12w 26 (93%) 26w 26 (93%) 52w 21 (75%) 2y 26 (93%) **NS1:** Physiotherapy **N randomized:** 28 **N analyzed:** 6w 26 (93%) 12w 26 (93%) 26w 22 (78.6%) 52w 20 (71.4%) 2y 24 (86%)	**All-cause mortality** NR **Surgical mortality** SG1: 0 (0%) NS1: NA	**N (%) with urosepsis** SG1: 1 (3.6%) No other operative complications were noted	**N (%) with reoperations** SG1: 2 (7.1%) (due to recurring symptoms on the same side and level) NS1: NA	NR
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98] High	**SG1:** Microdiscectomy **N randomized:** 141 **N analyzed:** 52w: 140 (99.3%) 2y: 130 (92.2%) 5y: 115 (81.6%) **NS1:** Conservative management **N randomized:** 142 **N analyzed:** 52w: 141 (99.3%) 2y: 130 (91.5%) 5y: 116 (81.7%)	**All-cause mortality** NR **Surgical mortality** SG1: 0 (0%) NS1: NA	**N (%) with surgical complications** SG1: 3 (1.6%) (2 dural tears and 1 wound hematoma) NS1: NA	**N (%) with reoperation for recurrent sciatica** 52w: S1 NR (3.2%); NS 1 NR (1.8% in crossovers who underwent surgery) 2y: SG1 7 (6%); NS1 4 (6% in crossovers who underwent surgery); Calculated P=0.54 5y: SG1 9 (7%); NS1 8 (12% in crossovers who underwent surgery); Calculated P=0.81	NR

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Ruetten (2008)[31] High	**SG1:** Endoscopic (interlaminar or transforaminal) discectomy **N randomized:** 100 **N analyzed:** 91 (91%) **SG2:** Microdiscectomy **N randomized:** 100 **N analyzed:** 87 (87%)	**All-cause mortality** Overall: 1 (0.5%) (unrelated to operation) **Surgical mortality** SG1: 0 (0%) SG2: 0 (0%)	Overall, complication rate reported to be significantly elevated in SG2 as compared to SG1 (P < 0.05). **N (%) with dural/nerve injury** SG1: 0 (0%) SG2: 0 (0%) **N (%) with cauda equina syndrome** SG1: 0 (0%) SG2: 0 (0%) **N (%) with transient postoperative dysesthesia** SG1: 3 (3.3%) SG2: 5 (5.7%) **N (%) with postoperative bleeding** SG1: 0 (0%) SG2: 2 (2.3%) **N (%) with delayed wound healing** SG1: 0 (0%) SG2: 1 (1.1%) **N (%) with soft tissue infection** SG1: 0 (0%) SG2: 1 (1.1%) **N (%) with thrombosis** SG1: 0 (0%) SG2: 0 (0%) **N (%) with spondylodiscitis** SG1: 0 (0%) SG2: 0 (0%)	**N (%) with revision surgery** SG1: 7 (7.7%) SG2: 10 (11.5%) Calculated P=0.45	NR

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Ruetten (2009)[47] High	**SG1:** Revision endoscopic discectomy **N randomized:** 50 **N analyzed:** 45 (90%) **SG2:** Revision microdiscectomy **N randomized:** 50 **N analyzed:** 42 (84%)	**All-cause mortality** SG1: 0 (0%) SG2: 0 (0%) **Surgical mortality** SG1: 0 (0%) SG2: 0 (0%)	**N (%) with dural injury** SG1: 1 (2%) SG2: 3 (6%) **N (%) with nerve injury, cauda-equina syndrome, spondylodiscitis, or thrombosis** SG1: 0 (0%) SG2: 0 (0%) **N (%) with transient postoperative dysesthesia** SG1: 2 (4%) SG2: 5 (10%) **N (%) with delayed wound healing** SG1: 0 (0%) SG2: 2 (4%) **N (%) with soft tissue infection** SG1: 0 (0%) SG2: 1 (2%) All morbidity above, calculated P=0.02 **Serious complication (not further described)** SG1: NR (6%) SG2: NR (21%) P < 0.05	**N (%) with revision with conventional spinal canal decompression due to persistent leg pain** SG1: 2 (4%) SG2: 1 (2%) Calculated P=1.0 **N (%) with revision with fusion due to progradient back pain** SG1: 0 (0%) SG2: 2 (4%) Calculated P=0.23 All reoperations, calculated P=0.67	NR

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Ryang (2008)[30] Gempt (2013)[103] High	**SG1:** Trocar microdiscectomy **N randomized:** 30 **N analyzed:** unclear **SG2:** Microdiscectomy **N randomized:** 30 **N analyzed:** unclear	**All-cause mortality** NR **Surgical mortality** NR	**N (%) with dural tear** SG1: 0 (0%) SG2: 2 (6.6%) Calculated P=0.49	**N (%) with reoperations for recurrent herniation** SG1: 2 (6.6%) (one during initial hospital stay and one at 6w) SG2: 4 (13.3%) (one during initial hospital stay and 1 at 8w, 28w, and 1.2y) Calculated P=0.67	NR
Sasaoka (2006)[25] High	**SG1:** Microendoscopic discectomy **N randomized:** 15 **N analyzed:** unclear **SG2:** Microdiscectomy **N randomized:** 11 **N analyzed:** unclear	**All-cause mortality** NR **Surgical mortality** NR	NR	NR	NR

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Teli (2010)[29] Some concerns	**SG1:** Microendoscopic discectomy **N randomized:** NR **N analyzed:** 70 **SG2:** Microdiscectomy **N randomized:** NR **N analyzed:** 72 **SG3:** Discectomy **N randomized:** NR **N analyzed:** 70	**All-cause mortality** NR **Surgical mortality** SG1: 0 (0%) SG2: 0 (0%) SG3: 0 (0%)	**N (%) with dural tear** SG1: 6 (8.7%) SG2: 2 (2.7%) SG3: 2 (3%) P=0.37 **N (%) with root injury** SG1: 2 (3%) SG2: 0 (0%) SG3: 0 (0%) P=0.45 **N (%) with wound infection** SG1: 0 (0%) SG2: 4 (5.5%) SG3: 3 (4.2%) P= 0.29 **N (%) with spondylodiscitis** SG1: 1 (1.4%) SG2: 0 (0%) SG3: 0 (0%) P=0.56 **N (%) with worsening motor deficit** SG1: 2 (1.4%) SG2: 1 (1%) SG3: 0 (0%) P=0.47	**N (%) with reoperations at 2y** Overall: 15 (7%) (13 for recurrences and 2 for repair of pseudomeningocele) **N (%) reoperations for recurrence at 2y** SG1: 8 (11.4%) SG2: 3 (4.2%) SG2: 2 (3%) Calculated SG1 vs SG2 P=0.13 Calculated SG2 vs SG3 P=1.00	NR

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Thome (2005)[28] Barth (2008)[104] Some concerns	**SG1:** Sequestrectomy **N randomized:** 42 **N analyzed:** 42 (100%) **SG2:** Microdiscectomy **N randomized:** 42 **N analyzed:** 42 (100%)	**All-cause mortality** NR **Surgical mortality** NR	**N (%) with intraoperative complications** SG1: 0 (0%) SG2: 0 (0%) **N (%) with superficial wound infection** SG1: 0 (0%) SG2: 1 (2.4%)	**N (%) with reoperation for reherniation at 1.5y** SG1: 2 (5%) SG2: 4 (10%) Calculated P=0.68 [Note: Reherniations (with or without reoperations) were reported in 12.5% of SG1 and 10.5% in SG2 by 2y]	NR
Tullberg (1993)[27] Some concerns	**SG1:** Microdiscectomy **N randomized:** 30 **N analyzed:** 29 (97%) **SG2:** Discectomy **N randomized:** 30 **N analyzed:** 29 (97%)	**All-cause mortality** NR **Surgical mortality** NR	**N (%) with nerve root sheath tear** SG1: 1 (3.3%) SG2: 1 (3.3%) Calculated P=1.0 **N (%) with dural leak** SG1: 1 (3.3%) SG2: 0 (0%) Calculated P=1.0 **N (%) with discitis** SG1: 1 (3.3%) SG2: 0 (0%) Calculated P=1.0 All morbidity, calculated P=0.61	**N (%) with reoperation at 52w** SG1: 1 (3.3%) at 16w SG2: 1 (3.3%) at <52w Calculated P=1.0	NR
Weber (1983)[26] High	**SG1:** Discectomy **N randomized**: 60 **N analyzed:** 60 (100%) **NS1:** Conservative management **N randomized:** 66 **N analyzed**: 66 (100%)	**All-cause cumulative mortality** 52w: NR 4y: SG1 1 (1.7%); NS1 0 (0%) 10y: SG1 3 (5.0%); NS1 0 (0%) **Surgical mortality** NR	NR	NR	NR

(continued)

Table D-5. Individual study findings related to safety outcomes (continued)

Main Study Author (Year); Follow-up Studies Author (Year); Risk of Bias	Intervention Names N Randomized N Analyzed	Mortality	Surgical morbidity	Reoperations	Persistent Opioid Use
Weinstein (2006)[22] Weinstein (2008)[99] Lurie (2014)[100] SPORT High	**SG1:** Discectomy/ microdiscectomy **N randomized**: 245 **N analyzed:** 232 in main study's primary analyses. 52w: 202 (82.47%) 2y: 186 (75.9%) 3y: 180 (73.5%) 4y: 149 (60.8%) 8y: 157 (64.1%) **NS1:** Conservative management **N randomized**: 256 **N analyzed**: 240 included in main study's primary analyses. 52w: 213 (83.2%) 2y: 187 (73.0%) 3y: 170 (66.4%) 4y: 150 (58.6%) 8y: 152 (59.4%)	**N (%) cumulative all-cause mortality** 6w, 12w, 26w: SG1 0 (0%); NS1 0 (0%) 52w: SG1 0 (0%); NS1 1 (0.47%) 2y: SG1 0 (0%); NS1 2 (1.07%) 4y: SG1 1 (0.67%); NS1 2 (1.33%) 8y: SG1 3 (1.91%); NS1 4 (2.63%) **N (%) with perioperative death** SG1: 0 (0%) NS1: NA	Based on primary analyses in main study for N=247 (SG1 140; NS1 107) participants that had surgery by 2y **N (%) with dural tear/spinal fluid leak** SG1: 10 (4.0%) **N (%) with vascular injury** SG1: 1 (0.40%) **N (%) with other intraoperative complication** SG1: 2 (0.81%) **N (%) with postoperative wound infection, superficial** SG1: 4 (1.6%) **N (%) with other complication (unspecified)** SG1: 9 (3.6%)	**N (%) with reoperations by 2y** 1y: 18 (7.3%) 2y: 25 (10.1%) Reoperations for additional surgery, recurrent herniation, complication, or other reason in the N=247 participants (SG1 140; NS1 107) that had initial surgery by 2y.	NR

Abbreviations: N = number; NR = not reported; NS = not significant; SD = standard deviation; CSF = cerebrospinal fluid; SG = surgical group; NS = nonsurgical group; SPORT = Spine Patient Outcomes Research Trial; w = week(s); y = year(s).

Table D-6. Individual study findings related to cost outcomes

Cost Study Author (Year); Main Study Author (Year); Country	Intervention [SG1] (N randomized); Comparator(s) [SG2, SG3, NS1] (N randomized)	Study Methods	Results (As Reported by Study)	Results (Converted to 2010 U.S. Dollars)[a]
Malter (1996)[44] United States, but used efficacy estimates from non-U.S. trial.	NA	Study design: CEA Year/unit of currency reported: 1993 USD Discount rate: 5% Time horizon: 10y Costs included: Direct medical costs (inpatient, outpatient, medication, diagnostic services, other health care services) obtained between 1987 and 1989 from a commercial U.S. database. QOL measure(s) used: Author developed time-tradeoff utility measure Other: Efficacy estimates based on an RCT comparing surgery with nonsurgical treatment (Weber et al (1983)[26]), and an RCT comparing surgery with chemonucleolysis (Javid et al. (1998)[111]) and a cohort study comparing surgery with nonsurgical treatment [Atlas et al. (1993)[112]].	**QALY at 10y (undiscounted)** SG1: 8.70 (95% CI, NR) NS1: 8.27 (95% CI, NR) AMD: 0.43 (95% CI, NR) **Costs at 10y** SG1: $17,020 (95% CI, NR) NS1: $4,470 (95% CI, NR) AMD: $12,550 (95% CI, NR) **Cost/QALY gained at 10y (undiscounted, payor perspective)** $29,200 (95% CI, NR) **Cost/QALY gained at 10y (discounted, payor perspective)** $33,900 (95% CI, NR)	**QALY at 10y (undiscounted)** SG1: 8.70 (95% CI, NR) NS1: 8.27 (95% CI, NR) AMD: 0.43 (95% CI, NR) **Costs at 10y** SG1: $25,684 (95% CI, NR) NS1: $6,745 (95% CI, NR) AMD: $18,938 (95% CI, NR) **Cost/QALY gained at 10y (undiscounted, payor perspective)** $44,064 (95% CI, NR) **Cost/QALY gained at 10y (discounted, payor perspective)** $51,156 (95% CI, NR)

(continued)

Table D-6. Individual study findings related to cost outcomes (continued)

Cost Study Author (Year); Main Study Author (Year); Country	Intervention [SG1] (N randomized); Comparator(s) [SG2, SG3, NS1] (N randomized)	Study Methods	Results (As Reported by Study)	Results (Converted to 2010 U.S. Dollars)[a]
Stevenson (1995)[53] Chatterjee (1995)[38] United Kingdom	Automated percutaneous lumbar discectomy (31); Microdiscectomy (40)	Study design: CEA concurrent to RCT Year/unit of currency reported: 1992 GBP Discount rate: NR Time horizon: 26w Costs included: Direct medical costs of procedures (inpatient, outpatient, medication, capital equipment), patient-reported direct and indirect costs and social service usage (e.g., missed work, travel, paid and unpaid caregivers or domestic help) QOL measure(s) used: NR Other: effectiveness was assessed on a 4-pt Likert scale by two clinicians (4=excellent, 1=poor). "Successful outcome" was defined as a 3 or 4.	**Mean total cost at 26w** SG1: £2,317 (95% CI NR) SG2: £1,567 (95% CI NR) AMD: NR Calculated AMD: £750 Note: SG1 includes cost of additional microdiscectomy in failed cases and SG2 includes cost of repeat microdiscectomy in failed case. **Cost per successful outcome at 26w** SG1: £3,264 SG2: £1,958 **Cost per point gained on 4-pt Likert scale of Effectiveness** SG1: £1,381 SG2: £764	**Mean total cost at 26w** SG1: $6,340 (95% CI NR) SG2: $4,288 (95% CI NR) AMD: NR Calculated AMD: $2,052 Note: SG1 includes cost of additional microdiscectomy in failed cases and SG2 includes cost of repeat microdiscectomy in failed case. **Cost per successful outcome at 26w** SG1: $8,931 SG2: $5,358 Calculated AMD: $3,573 **Cost per point gained on 4-pt Likert scale of effectiveness** SG1: $3,779 SG2: $2,091 Calculated AMD: $1,688 (95% CI cannot be calculated)
Teli (2010)[29] (main study includes cost) Italy	Microendoscopic discectomy (70) Microdiscectomy (72) Open discectomy (70)	Study design: Cost analysis concurrent to RCT Year/unit of currency reported: Euros, Year NR[b] Discount rate: NR Time horizon: NA Costs included: direct surgical equipment costs (equipment, tools), operating times, rehospitalizations QOL measure(s) used: NA	**Mean surgical costs (SD)** SG1: €3,010 (450) SG2: € 2,450 (340) SG3: €2,310 (260) Two P values were provided (P=0.002 and P=0.012), but unclear what comparison they are referring to.	**Mean surgical costs (SD)** SG1: $3,878 ($580) SG2: $3,156 ($438) SG3: $2,976 ($322) Calculated AMD SG1 vs. SG2: $722 (95% CI, $551 to $892) Calculated AMD SG2 vs. SG3: $65 (95% CI $52 to $307) Calculated AMD SG1 vs. SG3: $902 (95% CI $745 to $1059

(continued)

Table D-6. Individual study findings related to cost outcomes (continued)

Cost Study Author (Year); Main Study Author (Year); Country	Intervention [SG1] (N randomized); Comparator(s) [SG2, SG3, NS1] (N randomized)	Study Methods	Results (As Reported by Study)	Results (Converted to 2010 U.S. Dollars)[a]
Tosteson (2008)[50] Weinstein (2006)[22] SPORT United States	Discectomy (245); Conservative management (256)	Study design: CEA concurrent to RCT Year/unit of currency: 2004 USD Discount rate: 3% Time horizon: 2y Costs included: Direct medical costs (inpatient, outpatient, medication, diagnostic services, other health care services), patient-reported indirect costs (e.g., missed work, unpaid caregivers) QOL measure(s) used: EQ-5D with U.S. scoring Other: Based on pooled data from SPORT RCT and observational cohort. Crossovers considered in the 'as treated' group. Total N 1,191 (775 surgery, 416 no surgery)	**Mean discounted QALYs at 2y** SG1: 1.64 (95% CI, 1.62 to 1.67) NS1: 1.44 (95% CI, 1.41. to 1.47) AMD: 0.21 (95% CI, 0.16 to 25) **Mean total costs at 2y** SG1: $27,341 (95% CI, $25,882 to $28,799) NS1: $13,135 (95% CI, 11,244 to $14,902) Calculated AMD: $14,206 **Cost/QALY gained (societal perspective) at 2y** $69,403 (95% CI, $49,523 to $94,999) **Direct medical costs at 2y** SG1: $20,237 ($19,314 to $21,160) NS1: $5,804 (95% CI, $4,639 to $6,969) Calculated AMD: $14,433 **Direct medical costs/QALY gained (payor perspective) at 2y** $72,181 (95% CI, $56,473 to $92,394)	**Mean discounted QALYs at 2y** SG1: 1.64 (95% CI, 1.62 to 1.67) NS1: 1.44 (95% CI, 1.41. to 1.47) AMD: 0.21 (95% CI, 0.16 to 25) **Mean total costs at 2y** SG1: $31,561 (95% CI, $29,877 to $33,244) NS1: $15,162 (95% CI, $12,979 to $17,202) Calculated AMD: $16,399 (95% CI, $16,289 to $16,509) **Cost/QALY gained (societal perspective) at 2y** $80,115 (95% CI, $57,167 to $109,662) **Direct medical costs at 2y** SG1: $23,361 ($22,295 to $24,426) NS1: $6,700 (95% CI, $5,355 to $8,045) Calculated AMD: $16,661 (95% CI, $16,590 to $16,732) **Direct medical costs/QALY gained (payor perspective) at 2y** $83,322 (95% CI, $65,189 to $106,655)

(continued)

Table D-6. Individual study findings related to cost outcomes (continued)

Cost Study Author (Year); Main Study Author (Year); Trial Name; Country	Intervention [SG1] (N randomized); Comparator(s) [SG2, SG3, NS1] (N randomized)	Study Methods	Results (As Reported by Study)	Results (Converted to 2010 U.S. Dollars)[a]
Van den Akker (2011)[51] Arts (2009)[40] The Sciatica Micro-Endoscopic Discectomy Randomized Controlled Trial The Netherlands	Tubular discectomy (167); Microdiscectomy (161)	Study design: CEA concurrent with RCT Year/unit of currency reported: 2008 USD Discount rate: 0% Time horizon: 52w Costs included: Direct medical costs (inpatient, outpatient, medication, diagnostic services, other health care services), patient-reported indirect costs (e.g., missed work, travel, paid and unpaid caregivers or domestic help) QOL measure(s) used: EQ-5D	**QALYs at 52w** SG1: NR SG2: NR AMD: -0.012 (95% CI, -0.046 to 0.021) **Mean total costs at 52w** SG1: $16,858 (SD $12,759) SG2: $15,637 (SD $12,165) AMD: $1,491 (95% CI, $-1,335 to $4,318) **Cost/QALY gained (societal perspective) at 52w** NR. Differences in costs and QALYs underpinning this calculated value were not statistically significant, but point estimates suggest microdiscectomy dominates minimally-invasive surgery (i.e., is more effective and costs less). **Mean health care costs at 52w** SG1: $5,529 (SD $3,020) SG2: $5,070 ($3,375) AMD: $460 (95% CI, $-243 to $1,163) **Health care costs/QALY gained (payor perspective) at 52w** NR. Differences in costs and QALYs underpinning this calculated value were not statistically significant, but point estimates suggest microdiscectomy dominates minimally-invasive surgery (i.e., is more effective and costs less).	**QALYs at 52w** SG1: NR SG2: NR AMD: -0.012 (95% CI, -0.046 to 0.021) **Mean total costs at 52w** SG1: $17,074 (SD $12,922) SG2: $15,837 (SD $12,321) AMD: $1,510 (95% CI, -$1,352 to $4,373) **Cost/QALY gained (societal perspective) at 52w** NR. Differences in costs and QALYs underpinning this calculated value were not statistically significant, but point estimates suggest microdiscectomy dominates minimally-invasive surgery (i.e., is more effective and costs less). Calculated cost per QALY gained $-125,833 (95% CI cannot be calculated) **Mean health care costs at 52w** SG1: $5,600 (SD $3,059) SG2: $5,135 (SD $3,418) AMD: $466 (95% CI, $-246 to $1,178) **Health Care Costs/QALY Gained (payor perspective) at 52w** NR. Differences in costs and QALYs underpinning this calculated value were not statistically significant, but point estimates suggest microdiscectomy dominates minimally-invasive surgery (i.e., is more effective and costs less). Calculated cost per QALY gained $-38,833 (95% CI cannot be calculated).

(continued)

Table D-6. Individual study findings related to cost outcomes (continued)

Cost Study Author (Year); Main Study Author (Year); Trial Name; Country	Intervention [SG1] (N randomized); Comparator(s) [SG2, SG3, NS1] (N randomized)	Study Methods	Results (As Reported by Study)	Results (Converted to 2010 U.S. Dollars)[a]
Van den Akker (2017)[52] Brouwer (2015)[39] The Netherlands	Percutaneous laser disc decompression (57); Microdiscectomy (58)	Study design: CEA concurrent to RCT Year/unit of currency reported: 2010 Euros Discount rate: 0% Time horizon: 52w Costs included: Direct medical costs (inpatient, outpatient, medication, diagnostic services, other health care services), patient-reported indirect costs (e.g., missed work, travel, paid and unpaid caregivers or domestic help) QOL measure(s) used: EQ-5D with U.S. Scoring	**QALY at 52w** SG1: 0.733 (SD 0.172) SG2: 0.766 (0.133) AMD: -0.033 (95% CI, NR, P=0.27) **Mean total costs at 52w** SG1: €18,071 (SD €14,351) SG2: €20,451 (SD €13,080) AMD: €-2,379 (95% CI, €-7,618 to €2,860) **Costs/QALY gained (societal perspective) at 52w** NR. Differences in cost and QALYs underpinning this calculated value were not statistically significant, but point estimates suggest minimally-invasive surgery may be less effective but also costs less. **Mean health care costs at 52w** SG1: €5,325 (SD €4,395) SG2: €7,095 (SD €3,109) AMD: €-1,771 (95% CI, €-3,238 to €-303) **Health care costs/QALY gained (payor perspective) at 52w** NR. Difference in QALYs underpinning this calculated value was not statistically significant, but point estimates suggest minimally-invasive surgery may be less effective but also costs less.	**QALY at 52w** SG1: 0.733 (SD 0.172) SG2: 0.766 (SD 0.133) AMD: -0.033 (95% CI, NR, P=0.27) **Mean total costs at 52w** SG1: $24,420 (SD $19,393) SG2: $27,636 (SD $17,676) AMD: $-3,215 (95% CI, $-10,294 to $3,865) **Costs/QALY gained (societal perspective) at 52w** NR. Differences in cost and QALYs underpinning this calculated value were not statistically significant, but point estimates suggest minimally-invasive surgery may be less effective but also costs less. Calculated cost per QALY gained for microdiscectomy compared to minimally-invasive surgery $97,424 (95% CI cannot be calculated) **Mean health care costs at 52w** SG1: $7,196 (SD $5,939) SG2: $9,588 (SD $4,201) AMD: $-2,393 (95% CI, $-4,376 to $-409) **Health care costs/QALY gained (payor perspective) at 52w** NR. Difference in QALYs underpinning this calculated value was not statistically significant, but point estimates suggest minimally-invasive surgery may be less effective but also costs less. Calculated health care costs cost per QALY gained for microdiscectomy compared to minimally-invasive surgery $72,515 95% CI cannot be calculated).

(continued)

Table D-6. Individual study findings related to cost outcomes (continued)

Cost Study Author (Year); Main Study Author (Year); Trial Name; Country	Intervention [SG1] (N randomized); Comparator(s) [SG2, SG3, NS1] (N randomized)	Study Methods	Results (As Reported by Study)	Results (Converted to 2010 U.S. Dollars)[a]
Van den Hout (2008)[49] Peul (2007)[32] Sciatica Trial The Netherlands	Discectomy (141); Conservative management (142)	Study design: CEA concurrent to RCT Year/unit of currency reported: 2008 Euros Discount rate: 0% Time horizon: 52w Costs included: Direct medical costs (inpatient, outpatient, medication, diagnostic services, other health care services), patient-reported indirect costs (e.g., missed work, travel, paid and unpaid caregivers or domestic help) QOL measure(s) used: EQ-5D with UK scoring	**Mean QALY at 52w** SG1: 0.78 (SD 0.17) NS1: 0.73 (SD 0.16) AMD: 0.044 (95% CI, 0.005 to 0.083) **Mean total costs at 52w** SG1: €18,493 (SD €14,548) NS1: €18,506 (SD €18,102) AMD: €-12 (95% CI, €-4,029 to €4,006) **Cost/QALY gained (societal perspective) at 52w** NR, but SG1 dominates NS1 (more effective and less costs) but differences in cost underpinning this calculated value were not statistically significant. Calculated to be €-272. Using lower and upper 95% CI on AMDs for QALY and cost, best case ICER is calculated to be €-48,542 and worst case ICER is calculated to be €801,200 **Mean health care costs at 52w** SG1: €5,626 (SD €3,875) NS1: €3,807 (SD€4,237) AMD: €1,819 (95% CI, €842 to €2,790) **Health care costs/QALY gained (payor perspective) at 52w** €41,000 (95% CI, €14,000 to €430,000)	**Mean QALY at 52w** SG1: 0.78 (SD 0.17) NS1: 0.73 (SD 0.16) AMD: 0.044 (95% CI, 0.005 to 0.083) **Mean total costs at 52w** SG1: $28,421 (SD $22,358) NS1: $28,441 (SD $27,820) AMD: $-18.44 (95% CI, $-6,192 to $6,157) **Cost/QALY gained (societal perspective) at 52w** NR, but SG1 dominates NS1 (more effective and less costs) but differences in cost underpinning this calculated value were not statistically significant. Calculated to be $-419. Using lower and upper 95% CI on AMDs for QALY and cost, best case ICER is calculated to be $-74,602 and worst case ICER is calculated to be $1,231,400. **Mean health care costs at 52w** SG1: $8,646 (SD $5,955) NS1: $5,851 (SD $6,512) AMD: $2,796 (95% CI, $1,294 to $4,288) **Health care costs/QALY gained (payor perspective) at 52w** $63,011 (95% CI, $21,516 to $660,847)

[a] See Appendix C for description of methods used to convert costs to 2010 U.S. dollars.

[b] The study did not provide the year; we assumed 2009 for purposes of converting to 2010 U.S. dollars.

Abbreviations: AMD = absolute mean difference; CEA = cost-effectiveness analysis; CI = confidence interval; NA = not applicable; ICER = incremental cost-effectiveness ratio; NR = not reported; NS = nonsurgical group; QALY = quality-adjusted life year; w = weeks(s); QOL = quality of life; RCT = randomized controlled trial; SD = standard deviation; SG = surgical group; USD = United States Dollar; y = year(s).

Appendix E. Excluded Articles

List of Exclusion Codes

X1: Systematic review for hand search

X2: Ineligible publication type

X3: Ineligible country

X4: Ineligible population

X5: Ineligible intervention

X6: Ineligible comparator

X7: Ineligible outcome

X8: Ineligible study design

X9: Duplicate or superseded

X10: Study protocol or in progress

X11: Abstract only

X12: Non-English full text

X13: Data uninterpretable

1. Anulex Technologies I. Randomized Study of Anular Repair With the Xclose Tissue Repair System. 2008. Exclusion Code: X10.
2. University of Sao Paulo. Percutaneous Diskectomy SpineJet x Open Microdiskectomy in Treatment of Lumbar Radiculopathy (PDOP_TLR). 2011. Exclusion Code: X10.
3. Abdu RW, Abdu WA, Pearson AM, et al. Reoperation for Recurrent Intervertebral Disc Herniation in the Spine Patient Outcomes Research Trial: Analysis of Rate, Risk Factors, and Outcome. Spine (Phila Pa 1976). 2017 Jul 15;42(14):1106-14. doi: 10.1097/brs.0000000000002088. PMID: 28146015. Exclusion Code: X8.
4. Abramovitz JN, Neff SR. Lumbar disc surgery: results of the Prospective Lumbar Discectomy Study of the Joint Section on Disorders of the Spine and Peripheral Nerves of the American Association of Neurological Surgeons and the Congress of Neurological Surgeons. Neurosurgery. 1991 Aug;29(2):301-7; discussion 7-8. PMID: 1886676. Exclusion Code: X8.
5. Adakli B, Cakar Turhan KS, Asik I. The comparison of the efficacy of radiofrequency nucleoplasty and targeted disc decompression in lumbar radiculopathy. Bosn J Basic Med Sci. 2015 Apr 25;15(2):57-61. doi: 10.17305/bjbms.2015.427. PMID: 26042514. Exclusion Code: X8.
6. Adam D, Pevzner E, Gepstein R. Comparison of percutaneous nucleoplasty and open discectomy in patients with lumbar disc protrusions. Chirurgia (Bucur). 2013 Jan-Feb;108(1):94-8. PMID: 23464777. Exclusion Code: X8.
7. Ahn SS, Kim SH, Kim DW, et al. Comparison of Outcomes of Percutaneous Endoscopic Lumbar Discectomy and Open Lumbar Microdiscectomy for Young Adults: A Retrospective Matched Cohort Study. World Neurosurg. 2016 Feb;86:250-8. doi: 10.1016/j.wneu.2015.09.047. PMID: 26409086. Exclusion Code: X8.
8. Amoretti N, Huwart L, Marcy PY, et al. CT- and fluoroscopy-guided percutaneous discectomy for lumbar radiculopathy related to disc herniation: a comparative prospective study comparing lateral to medial herniated discs. Skeletal Radiol. 2013 Jan;42(1):49-53. doi: 10.1007/s00256-012-1422-5. PMID: 22644540. Exclusion Code: X8.
9. Anderson DG, Patel A, Maltenfort M, et al. Lumbar decompression using a traditional midline approach versus a tubular retractor system: comparison of patient-based clinical outcomes. Spine (Phila Pa 1976). 2011 Mar 01;36(5):E320-5. doi: 10.1097/BRS.0b013e3181db1dfb. PMID: 21178844. Exclusion Code: X8.
10. Anderson PA, McCormick PC, Angevine PD. Randomized controlled trials of the treatment of lumbar disk herniation: 1983-2007. J Am Acad Orthop Surg. 2008 Oct;16(10):566-73. PMID: 18832600. Exclusion Code: X2.

11. Anichini G, Landi A, Caporlingua F, et al. Lumbar endoscopic microdiscectomy: where are we now? An updated literature review focused on clinical outcome, complications, and rate of recurrence. Biomed Res Int. 2015;2015:417801. doi: 10.1155/2015/417801. PMID: 26688809. Exclusion Code: X2.
12. Arai Y, Hirai T, Yoshii T, et al. A prospective comparative study of 2 minimally-invasive decompression procedures for lumbar spinal canal stenosis: unilateral laminotomy for bilateral decompression (ULBD) versus muscle-preserving interlaminar decompression (MILD). Spine (Phila Pa 1976). 2014 Feb 15;39(4):332-40. doi: 10.1097/brs.0000000000000136. PMID: 24299721. Exclusion Code: X4.
13. Aronsohn J, Chapman K, Soliman M, et al. Percutaneous microdiscectomy versus epidural injection for management of chronic spinal pain. Proc West Pharmacol Soc. 2010;53:16-9. PMID: 22128444. Exclusion Code: X13.
14. Arts M, Brand R, van der Kallen B, et al. Does minimally invasive lumbar disc surgery result in less muscle injury than conventional surgery? A randomized controlled trial. Eur Spine J. 2011 Jan;20(1):51-7. doi: 10.1007/s00586-010-1482-y. PMID: 20556439. Exclusion Code: X9.
15. Arts MP, Brand R, Koes BW, et al. Effect modifiers of outcome of surgery in patients with herniated disc related sciatica? A subgroup analysis of a randomised clinical trial. J Neurol Neurosurg Psychiatry. 2010 Nov;81(11):1265-74. doi: 10.1136/jnnp.2009.192906. Epub 2010 Jun 14. PMID: 20547620. Exclusion Code: X8.
16. Arts MP, Peul WC, Brand R, et al. Cost-effectiveness of microendoscopic discectomy versus conventional open discectomy in the treatment of lumbar disc herniation: a prospective randomised controlled trial [ISRCTN51857546]. BMC Musculoskelet Disord. 2006 May 13;7:42. doi: 10.1186/1471-2474-7-42. PMID: 16696861. Exclusion Code: X10.
17. Atlas SJ, Tosteson TD, Blood EA, et al. The impact of workers' compensation on outcomes of surgical and nonoperative therapy for patients with a lumbar disc herniation: SPORT. Spine (Phila Pa 1976). 2010 Jan 01;35(1):89-97. doi: 10.1097/BRS.0b013e3181c68047. PMID: 20023603. Exclusion Code: X8.
18. Ay S, Dogan SK, Evcik D. Is low-level laser therapy effective in acute or chronic low back pain? Clin Rheumatol. 2010 Aug;29(8):905-10. doi: 10.1007/s10067-010-1460-0. PMID: 20414695. Exclusion Code: X5.
19. Aykac B, Copuroglu C, Ozcan M, et al. Postoperative evaluation of quality of life in lumbar spinal stenosis patients following instrumented posterior decompression. Acta Orthop Traumatol Turc. 2011;45(1):47-52. doi: 10.3944/aott.2011.2387. PMID: 21478662. Exclusion Code: X5.
20. Azarhomayoun A, Chou R, Shirdel S, et al. Sequestrectomy Versus Conventional Microdiscectomy for the Treatment of a Lumbar Disc Herniation: A Systematic Review. Spine (Phila Pa 1976). 2015 Dec;40(24):E1330-9. doi: 10.1097/brs.0000000000001174. PMID: 26655808. Exclusion Code: X1.
21. Azzazi A, AlMekawi S, Zein M. Lumbar disc nucleoplasty using coblation technology: clinical outcome. J Neurointerv Surg. 2011 Sep;3(3):288-92. doi: 10.1136/jnis.2010.002402. PMID: 21990844. Exclusion Code: X8.
22. Bailey A, Araghi A, Blumenthal S, et al. Prospective, multicenter, randomized, controlled study of anular repair in lumbar discectomy: two-year follow-up. Spine (Phila Pa 1976). 2013 Jun 15;38(14):1161-9. doi: 10.1097/BRS.0b013e31828b2e2f. PMID: 23392414. Exclusion Code: X6.
23. Banczerowski P, Czigleczki G, Papp Z, et al. Minimally invasive spine surgery: systematic review. Neurosurg Rev. 2015 Jan;38(1):11-26; discussion doi: 10.1007/s10143-014-0565-3. PMID: 25199809. Exclusion Code: X2.
24. Barth M, Diepers M, Weiss C, et al. Two-year outcome after lumbar microdiscectomy versus microscopic sequestrectomy: part 2: radiographic evaluation and correlation with clinical outcome. Spine (Phila Pa 1976). 2008 Feb 01;33(3):273-9. doi: 10.1097/BRS.0b013e31816201a6. PMID: 18303459. Exclusion Code: X7.

25. Bokov A, Skorodumov A, Isrelov A, et al. Differential treatment of nerve root compression pain caused by lumbar disc herniation applying nucleoplasty. Pain Physician. 2010 Sep-Oct;13(5):469-80. PMID: 20859316. Exclusion Code: X8.
26. Boult M, Fraser RD, Jones N, et al. Percutaneous endoscopic laser discectomy. Aust N Z J Surg. 2000 Jul;70(7):475-9. PMID: 10901572. Exclusion Code: X8.
27. Bron JL, Helder MN, Meisel HJ, et al. Repair, regenerative and supportive therapies of the annulus fibrosus: achievements and challenges. Eur Spine J. 2009 Mar;18(3):301-13. doi: 10.1007/s00586-008-0856-x. PMID: 19104850. Exclusion Code: X2.
28. Brouwer PA, Peul WC, Brand R, et al. Effectiveness of percutaneous laser disc decompression versus conventional open discectomy in the treatment of lumbar disc herniation; design of a prospective randomized controlled trial. BMC Musculoskelet Disord. 2009 May 13;10:49. doi: 10.1186/1471-2474-10-49. PMID: 19439098. Exclusion Code: X10.
29. Buttermann GR. Treatment of lumbar disc herniation: epidural steroid injection compared with discectomy. A prospective, randomized study. J Bone Joint Surg Am. 2004 Apr;86-a(4):670-9. PMID: 15069129. Exclusion Code: X8.
30. Chang X, Chen B, Li HY, et al. The safety and efficacy of minimally invasive discectomy: a meta-analysis of prospective randomised controlled trials. Int Orthop. 2014 Jun;38(6):1225-34. doi: 10.1007/s00264-014-2331-0. PMID: 24722785. Exclusion Code: X1.
31. Chitragran R, Poopitaya S, Tassanawipas W. Result of percutaneous disc decompression using nucleoplasty in Thailand: a randomized controlled trial. J Med Assoc Thai. 2012 Oct;95 Suppl 10:S198-205. PMID: 23451463. Exclusion Code: X3.
32. Choi G, Lee SH, Deshpande K, et al. Working channel endoscope in lumbar spine surgery. J Neurosurg Sci. 2014 Jun;58(2):77-85. PMID: 24819484. Exclusion Code: X2.
33. 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 May 01;34(10):1094-109. doi: 10.1097/BRS.0b013e3181a105fc. PMID: 19363455. Exclusion Code: X1.
34. Chou R, Loeser JD, Owens DK, et al. Interventional therapies, surgery, and interdisciplinary rehabilitation for low back pain: an evidence-based clinical practice guideline from the American Pain Society. Spine (Phila Pa 1976). 2009 May 01;34(10):1066-77. doi: 10.1097/BRS.0b013e3181a1390d. PMID: 19363457. Exclusion Code: X2.
35. Clark AJ, Safaee MM, Khan NR, et al. Tubular microdiscectomy: techniques, complication avoidance, and review of the literature. Neurosurg Focus. 2017 Aug;43(2):E7. doi: 10.3171/2017.5.focus17202. PMID: 28760036. Exclusion Code: X1.
36. Cole JSt, Jackson TR. Minimally invasive lumbar discectomy in obese patients. Neurosurgery. 2007 Sep;61(3):539-44; discussion 44. doi: 10.1227/01.neu.0000290900.23190.c9. PMID: 17881966. Exclusion Code: X8.
37. Cong L, Zhu Y, Tu G. A meta-analysis of endoscopic discectomy versus open discectomy for symptomatic lumbar disk herniation. Eur Spine J. 2016 Jan;25(1):134-43. doi: 10.1007/s00586-015-3776-6. PMID: 25632840. Exclusion Code: X1.
38. Couto JM, Castilho EA, Menezes PR. Chemonucleolysis in lumbar disc herniation: a meta-analysis. Clinics (Sao Paulo). 2007 Apr;62(2):175-80. PMID: 17505703. Exclusion Code: X6.
39. Crawshaw C, Frazer AM, Merriam WF, et al. A comparison of surgery and chemonucleolysis in the treatment of sciatica A prospective randomized trial. Spine (Phila Pa 1976). 1984 Mar;9(2):195-8. PMID: 6374921. Exclusion Code: X5.
40. Cristante AF, Rocha ID, MartusMarcon R, et al. Randomized clinical trial comparing lumbar percutaneous hydrodiscectomy with lumbar open microdiscectomy for the treatment of lumbar disc protrusions and herniations. Clinics (Sao Paulo). 2016 May;71(5):276-80. doi:

10.6061/clinics/2016(05)06. PMID: 27276397. Exclusion Code: X3.
41. Dasenbrock HH, Juraschek SP, Schultz LR, et al. The efficacy of minimally invasive discectomy compared with open discectomy: a meta-analysis of prospective randomized controlled trials. J Neurosurg Spine. 2012 May;16(5):452-62. doi: 10.3171/2012.1.spine11404. PMID: 22404142. Exclusion Code: X1.
42. Duarte R, Costa J. Percutaneous laser disc decompression for lumbar discogenic radicular pain. Radiología (English Edition). 2012;54(4):336-41. Exclusion Code: X8.
43. Eichen PM, Achilles N, Konig V, et al. Nucleoplasty, a minimally invasive procedure for disc decompression: a systematic review and meta-analysis of published clinical studies. Pain Physician. 2014 Mar-Apr;17(2):E149-73. PMID: 24658486. Exclusion Code: X1.
44. Ejeskar A, Nachemson A, Herberts P, et al. Surgery versus chemonucleolysis for herniated lumbar discs. A prospective study with random assignment. Clin Orthop Relat Res. 1983 Apr(174):236-42. PMID: 6339137. Exclusion Code: X6.
45. Epstein NE. Should anyone perform percutaneous endoscopic laser diskectomy and percutaneous lumbar disc decompressions? Surg Neurol Int. 2016;7(Suppl 42):S1080-s4. doi: 10.4103/2152-7806.196764. PMID: 28144489. Exclusion Code: X1.
46. National Institute for Clinical Excellence (NICE). Interventional Procedure Overview of Endoscopic Laser Foraminoplasty. London, United Kingdom: National Institute for Clinical Excellence; 2003. p. IP Guidance Number IPG031. Exclusion Code: X4.
47. National Institute for Clinical Excellence (NICE). Interventional Procedure Overview of Laser Lumbar Discectomy. London, United Kingdom: National Institute for Clinical Excellence; 2003. p. IP Guidance Number IPG027. Exclusion Code: X1.
48. Fakouri B, Patel V, Bayley E, et al. Lumbar microdiscectomy versus sequesterectomy/free fragmentectomy: a long-term (> 2 y) retrospective study of the clinical outcome. Clinical Spine Surgery. 2011;24(1):6-10. Exclusion Code: X8.
49. Fakouri B, Shetty NR, White TC. Is sequestrectomy a viable alternative to microdiscectomy? A systematic review of the literature. Clin Orthop Relat Res. 2015 Jun;473(6):1957-62. doi: 10.1007/s11999-014-3904-3. PMID: 25183219. Exclusion Code: X1.
50. Findlay GF, Hall BI, Musa BS, et al. A 10-year follow-up of the outcome of lumbar microdiscectomy. Spine (Phila Pa 1976). 1998 May 15;23(10):1168-71. PMID: 9615370. Exclusion Code: X8.
51. Fisher C, Noonan V, Bishop P, et al. Outcome evaluation of the operative management of lumbar disc herniation causing sciatica J Neurosurg. 2004 Apr;100(4 Suppl Spine):317-24. PMID: 15070138. Exclusion Code: X8.
52. Fitzsimmons D, Phillips CJ, Bennett H, et al. Cost-effectiveness of different strategies to manage patients with sciatica Pain. 2014 Jul;155(7):1318-27. doi: 10.1016/j.pain.2014.04.008. PMID: 24726924. Exclusion Code: X9.
53. Fourney DR, Dettori JR, Norvell DC, et al. Does minimal access tubular assisted spine surgery increase or decrease complications in spinal decompression or fusion? Spine (Phila Pa 1976). 2010 Apr 20;35(9 Suppl):S57-65. doi: 10.1097/BRS.0b013e3181d82bb8. PMID: 20407352. Exclusion Code: X1.
54. Fukui S, Iwashita N, Nitta K, et al. The results of percutaneous intradiscal high-pressure injection of saline in patients with extruded lumbar herniated disc: comparison with microendoscopic discectomy. Pain Med. 2012 Jun;13(6):762-8. doi: 10.1111/j.1526-4637.2012.01400.x. Epub 2012 May 23. PMID: 22621436. Exclusion Code: X8.
55. Garg B, Nagraja UB, Jayaswal A Microendoscopic versus open discectomy for lumbar disc herniation: a prospective randomised study. J Orthop Surg (Hong Kong). 2011 Apr;19(1):30-4. doi: 10.1177/230949901101900107. PMID: 21519072. Exclusion Code: X3.
56. Gerges FJ, Lipsitz SR, Nedeljkovic SS. A systematic review on the effectiveness of the Nucleoplasty procedure for discogenic pain. Pain Physician. 2010 Mar-Apr;13(2):117-32. PMID: 20309378. Exclusion Code: X1.

57. Gibson JA, Waddell G. Surgical interventions for lumbar disc prolapse. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2007. Exclusion Code: X9.
58. Gibson JN, Waddell G. Surgical interventions for lumbar disc prolapse: updated Cochrane Review. Spine (Phila Pa 1976). 2007 Jul 15;32(16):1735-47. doi: 10.1097/BRS.0b013e3180bc2431. PMID: 17632394. Exclusion Code: X1.
59. Gibson JN, Waddell G. Surgical interventions for lumbar disc prolapse. Cochrane Database Syst Rev. 2007 Apr 18(2):Cd001350. doi: 10.1002/14651858.CD001350.pub4. PMID: 17443505. Exclusion Code: X1.
60. Gotfryd A, Avanzi O. A systematic review of randomised clinical trials using posterior discectomy to treat lumbar disc herniations. Int Orthop. 2009 Feb;33(1):11-7. doi: 10.1007/s00264-008-0559-2. PMID: 18500517. Exclusion Code: X1.
61. . Microdiscectomy and conservative treatment for lumbar disc herniation with back pain and sciatica: a randomised clinical trial. Proceedings of the 30th meeting of the International Society for the Study of the Lumbar Spine; 2003; Vancouver. Exclusion Code: X11.
62. Guyatt GH, Cook DJ, Jaeschke R, et al. Grades of recommendation for antithrombotic agents: American College of Chest Physicians evidence-based clinical practice guidelines. CHEST Journal. 2008;133(6_suppl):123S-31S. Exclusion Code: X2.
63. Helm Ii S, Deer TR, Manchikanti L, et al. Effectiveness of thermal annular procedures in treating discogenic low back pain. Pain Physician. 2012 May-Jun;15(3):E279-304. PMID: 22622914. Exclusion Code: X4.
64. Hirsch JA, Singh V, Falco FJ, et al. Automated percutaneous lumbar discectomy for the contained herniated lumbar disc: a systematic assessment of evidence. Pain Physician. 2009 May-Jun;12(3):601-20. PMID: 19461826. Exclusion Code: X9.
65. Hussein M. Minimal Incision, Multifidus-sparing Microendoscopic Diskectomy Versus Conventional Microdiskectomy for Highly Migrated Intracanal Lumbar Disk Herniations. J Am Acad Orthop Surg. 2016 Nov;24(11):805-13. doi: 10.5435/jaaos-d-15-00588. PMID: 27673378. Exclusion Code: X3.
66. Hussein M, Abdeldayem A, Mattar MM. Surgical technique and effectiveness of microendoscopic discectomy for large uncontained lumbar disc herniations: a prospective, randomized, controlled study with 8 years of follow-up. Eur Spine J. 2014 Sep;23(9):1992-9. doi: 10.1007/s00586-014-3296-9. PMID: 24736930. Exclusion Code: X3.
67. Iwamoto J, Sato Y, Takeda T, et al. The return to sports activity after conservative or surgical treatment in athletes with lumbar disc herniation. Am J Phys Med Rehabil. 2010 Dec;89(12):1030-5. doi: 10.1097/PHM.0b013e3181f71044. PMID: 20881585. Exclusion Code: X4.
68. Jacobs WC, Arts MP, van Tulder MW, et al. Surgical techniques for sciatica due to herniated disc, a systematic review. Eur Spine J. 2012 Nov;21(11):2232-51. doi: 10.1007/s00586-012-2422-9. PMID: 22814567. Exclusion Code: X1.
69. Jacobs WC, van Tulder M, Arts M, et al. Surgery versus conservative management of sciatica due to a lumbar herniated disc: a systematic review. Eur Spine J. 2011 Apr;20(4):513-22. doi: 10.1007/s00586-010-1603-7. PMID: 20949289. Exclusion Code: X1.
70. Johans SJ, Amin BY, Mummaneni PV. Minimally invasive lumbar decompression for lumbar stenosis: review of clinical outcomes and cost effectiveness. J Neurosurg Sci. 2015 Mar;59(1):37-45. PMID: 25370820. Exclusion Code: X2.
71. Jordan J, Konstantinou K, O'Dowd J. Herniated lumbar disc. BMJ Clin Evid. 2009 Mar 26;2009 PMID: 19445754. Exclusion Code: X1.
72. Jun-liang LIU W-xZ, Guo-yong GAO et al. A prospective and controlled study of percutaneous transforaminal endoscopic discectomy versus fenestration discectomy for lumbar disc herniation. Chin J Bone Joint. 2014;3:245–50. Exclusion Code: X12.
73. Kamper SJ, Ostelo RW, Rubinstein SM, et al. Minimally invasive surgery for lumbar disc herniation: a systematic review and meta-analysis. Eur Spine J. 2014 May;23(5):1021-

43. doi: 10.1007/s00586-013-3161-2. PMID: 24442183. Exclusion Code: X1.
74. Kelly RE, Dinner MH, Lavyne MH, et al. The effect of lumbar disc surgery on postoperative pulmonary function and temperature. A comparison study of microsurgical lumbar discectomy with standard lumbar discectomy. Spine (Phila Pa 1976). 1993 Feb;18(2):287-90. PMID: 8441946. Exclusion Code: X7.
75. Kerr D, Zhao W, Lurie JD. What Are Long-term Predictors of Outcomes for Lumbar Disc Herniation? A Randomized and Observational Study. Clin Orthop Relat Res. 2015 Jun;473(6):1920-30. doi: 10.1007/s11999-014-3803-7. PMID: 25057116. Exclusion Code: X9.
76. Koerner JD, Glaser J, Radcliff K. Which Variables Are Associated With Patient-reported Outcomes After Discectomy? Review of SPORT Disc Herniation Studies. Clin Orthop Relat Res. 2015 Jun;473(6):2000-6. doi: 10.1007/s11999-014-3671-110.1007/s11999-014-3671-1. PMID: 24818737. Exclusion Code: X2.
77. Krappel F, Schmitz R, Bauer E, et al. Open or endoscopic nucleotomy? Results of a prospective, controlled clinical trial with independent follow up, MRI and special reference to cost effectiveness. Offene oder endoskopische nukleotomie—Ergebnisse einer kontrollierten klinischen studie mit unabhangiger nachuntersuchung, MRT und unter besonderer berucksichtigung der kosten-nutzen-relation. Orthop Prax. 2001;37:164-9. Exclusion Code: X12.
78. Kreiner DS, Hwang SW, Easa JE, et al. An evidence-based clinical guideline for the diagnosis and treatment of lumbar disc herniation with radiculopathy. Spine J. 2014 Jan;14(1):180-91. doi: 10.1016/j.spinee.2013.08.003. PMID: 24239490. Exclusion Code: X2.
79. Krugluger J, Knahr K. Chemonucleolysis and automated percutaneous discectomy--a prospective randomized comparison. Int Orthop. 2000;24(3):167-9. PMID: 10990391. Exclusion Code: X6.
80. Lagarrigue J, Chaynes P. [Comparative study of disk surgery with or without microscopy. A prospective study of 80 cases]. Neurochirurgie. 1994;40(2):116-20. PMID: 7870244. Exclusion Code: X12.
81. Lavignolle B, Vital JM, Baulny D, et al. [Comparative study of surgery and chemonucleolysis in the treatment of sciatica caused by a herniated disk]. Acta Orthop Belg. 1987;53(2):244-9. PMID: 3687413. Exclusion Code: X12.
82. Lee DY, Lee SH. Carbon dioxide (CO2) laser-assisted microdiscectomy for extraforaminal lumbar disc herniation at the L5-S1 level. Photomed Laser Surg. 2011 Aug;29(8):531-5. doi: 10.1089/pho.2010.2954. PMID: 21309702. Exclusion Code: X6.
83. Lee SH, Chung SE, Ahn Y, et al. Comparative radiologic evaluation of percutaneous endoscopic lumbar discectomy and open microdiscectomy: a matched cohort analysis. Mt Sinai J Med. 2006 Sep;73(5):795-801. PMID: 17008941. Exclusion Code: X8.
84. Leven D, Passias PG, Errico TJ, et al. Risk Factors for Reoperation in Patients Treated Surgically for Intervertebral Disc Herniation: A Subanalysis of Eight-Year SPORT Data J Bone Joint Surg Am. 2015 Aug 19;97(16):1316-25. doi: 10.2106/jbjs.n.01287. PMID: 26290082. Exclusion Code: X7.
85. Lewis R, Williams N, Matar HE, et al. The clinical effectiveness and cost-effectiveness of management strategies for sciatica: systematic review and economic model. Health Technol Assess. 2011 Nov;15(39):1-578. doi: 10.3310/hta15390. PMID: 22078311. Exclusion Code: X1.
86. Lewis RA, Williams NH, Sutton AJ, et al. Comparative clinical effectiveness of management strategies for sciatica: systematic review and network meta-analyses. Spine J. 2015 Jun 01;15(6):1461-77. doi: 10.1016/j.spinee.2013.08.049. PMID: 24412033. Exclusion Code: X1.
87. Li T, Han D, Liu B, et al. Clinical assessment of reformed lumbar microdiscectomy. Eur J Orthop Surg Traumatol. 2014 Jan;24(1):23-7. doi: 10.1007/s00590-012-1123-8. PMID: 23412248. Exclusion Code: X3.
88. Li X, Han Y, Di Z, et al. Percutaneous endoscopic lumbar discectomy for lumbar disc herniation. J Clin Neurosci. 2016 Nov;33:19-27. doi: 10.1016/j.jocn.2016.01.043. PMID: 27475315. Exclusion Code: X1.

89. Li X, Hu Z, Cui J, et al. Percutaneous endoscopic lumbar discectomy for recurrent lumbar disc herniation. Int J Surg. 2016 Mar;27:8-16. doi: 10.1016/j.ijsu.2016.01.034. PMID: 26805569. Exclusion Code: X1.
90. Li XC, Zhong CF, Deng GB, et al. Full-endoscopic procedures versus traditional discectomy surgery for discectomy: a systematic review and meta-analysis of current global clinical trials. Pain Physician. 2016 Mar;19(3):103-18. PMID: 27008284. Exclusion Code: X1.
91. Linzer P, Filip M, Samal F, et al. Comparison of Biochemical Markers of Muscle Damage and Inflammatory Response Between the Open Discectomy, Microsurgical Discectomy, and Microsurgical Discectomy Using Tubular Retractor. J Neurol Surg A Cent Eur Neurosurg. 2015 Sep;76(5):384-91. doi: 10.1055/s-0034-1393929. PMID: 26024471. Exclusion Code: X8.
92. Livesey J, Sundaram S, Foster L, et al. Laser discectomy versus lumbar epidural steroid injection: a randomised comparative study of two treatments for sciatica The Journal of Bone and Joint Surgery-british Volume. 2000;82:74. Exclusion Code: X11.
93. Lurie JD, Faucett SC, Hanscom B, et al. Lumbar discectomy outcomes vary by herniation level in the Spine Patient Outcomes Research Trial. J Bone Joint Surg Am. 2008 Sep;90(9):1811-9. doi: 10.2106/jbjs.g.00913. PMID: 18762639. Exclusion Code: X8.
94. Lurie JD, Tosteson TD, Tosteson AN, et al. Surgical versus nonoperative treatment for lumbar disc herniation: eight-year results for the spine patient outcomes research trial. Spine (Phila Pa 1976). 2014 Jan 01;39(1):3-16. doi: 10.1097/brs.0000000000000088. PMID: 24153171. Exclusion Code: X9.
95. Machado GC, Witzleb AJ, Fritsch C, et al. Patients with sciatica still experience pain and disability 5 years after surgery: A systematic review with meta-analysis of cohort studies. Eur J Pain. 2016 Nov;20(10):1700-9. doi: 10.1002/ejp.893. PMID: 27172245. Exclusion Code: X8.
96. Macnab I. Negative Disc Exploration: an analysis of the causes of nerve-root involvement in sixty-eight patients. JBJS. 1971;53(5):891-903. PMID: 00004623-197153050-00004. Exclusion Code: X8.
97. Manchikanti L, Abdi S, Atluri S, et al. An update of comprehensive evidence-based guidelines for interventional techniques in chronic spinal pain. Part II: guidance and recommendations. Pain Physician. 2013 Apr;16(2 Suppl):S49-283. PMID: 23615883. Exclusion Code: X1.
98. Manchikanti L, Derby R, Benyamin RM, et al. A systematic review of mechanical lumbar disc decompression with nucleoplasty. Pain Physician. 2009 May-Jun;12(3):561-72. PMID: 19461823. Exclusion Code: X9.
99. Manchikanti L, Falco FJ, Benyamin RM, et al. An update of the systematic assessment of mechanical lumbar disc decompression with nucleoplasty. Pain Physician. 2013 Apr;16(2 Suppl):Se25-54. PMID: 23615886. Exclusion Code: X1.
100. Manchikanti L, Singh V, Calodney AK, et al. Percutaneous lumbar mechanical disc decompression utilizing Dekompressor(R): an update of current evidence. Pain Physician. 2013 Apr;16(2 Suppl):Se1-24. PMID: 23615884. Exclusion Code: X8.
101. Manchikanti L, Singh V, Falco FJ, et al. An updated review of automated percutaneous mechanical lumbar discectomy for the contained herniated lumbar disc. Pain Physician. 2013 Apr;16(2 Suppl):Se151-84. PMID: 23615890. Exclusion Code: X8.
102. Marin FZ. CAM versus nucleoplasty. Acta Neurochir Suppl. 2005;92:111-4. PMID: 15830980. Exclusion Code: X6.
103. Martin-Laez R, Martinez-Agueros JA, Suarez-Fernandez D, et al. Complications of endoscopic microdiscectomy using the EASYGO! system: is there any difference with conventional discectomy during the learning-curve period? Acta Neurochir (Wien). 2012 Jun;154(6):1023-32. doi: 10.1007/s00701-012-1321-5. PMID: 22446750. Exclusion Code: X8.
104. Mazanec D, Okereke L. Interpreting the Spine Patient Outcomes Research Trial. Medical vs surgical treatment of lumbar disk herniation: implications for future trials. Cleve Clin J Med. 2007 Aug;74(8):577-83. PMID: 17708129. Exclusion Code: X2.

105. McCormick ZL, Slipman C, Kotcharian A, et al. Percutaneous Lumbar Disc Decompression Using the Dekompressor: A Prospective Long-Term Outcome Study. Pain Medicine. 2016;17(6):1023-30. Exclusion Code: X8.
106. McGirt MJ, Ambrossi GL, Datoo G, et al. Recurrent disc herniation and long-term back pain after primary lumbar discectomy: review of outcomes reported for limited versus aggressive disc removal. Neurosurgery. 2009 Feb;64(2):338-44; discussion 44-5. doi: 10.1227/01.neu.0000337574.58662.e2. PMID: 19190461. Exclusion Code: X1.
107. McGuire KJ, Khaleel MA, Rihn JA, et al. The effect of high obesity on outcomes of treatment for lumbar spinal conditions: subgroup analysis of the spine patient outcomes research trial. Spine (Phila Pa 1976). 2014 Nov 01;39(23):1975-80. doi: 10.1097/brs.0000000000000577. PMID: 25365713. Exclusion Code: X4.
108. Muralikuttan KP, Hamilton A, Kernohan WG, et al. A prospective randomized trial of chemonucleolysis and conventional disc surgery in single level lumbar disc herniation. Spine (Phila Pa 1976). 1992 Apr;17(4):381-7. PMID: 1579871. Exclusion Code: X6.
109. National Guideline Centre. National Institute for Health and Care Excellence: Clinical Guidelines. Low Back Pain and Sciatica in Over 16s: Assessment and Management. London: National Institute for Health and Care Excellence; 2016. Exclusion Code: X1.
110. National Institute for Health and Care Excellence. Automated percutaneous mechanical lumbar discectomy. United Kingdom: National Institute for Health and Care Excellence; 2005. https://www.nice.org.uk/guidance/IPG141. Accessed on November 8 2017. Exclusion Code: X1.
111. National Institute for Health and Care Excellence. Percutaneous transforaminal endoscopic lumbar discectomy for sciatica United Kingdom: National Institute for Health and Care Excellence; 2016. https://www.nice.org.uk/guidance/ipg556/resources/percutaneous-transforaminal-endoscopic-lumbar-discectomy-for-sciatica-pdf-1899871990618309. Accessed on April 2017. Exclusion Code: X1.
112. National Institute for Health and Care Excellence. Percutaneous interlaminar endoscopic lumbar discectomy for sciatica United Kingdom: National Institute for Health and Care Excellence; 2016. https://www.nice.org.uk/guidance/ipg555. Accessed on November 8 2017. Exclusion Code: X1.
113. Nellensteijn J, Ostelo R, Bartels R, et al. Transforaminal endoscopic surgery for symptomatic lumbar disc herniations: a systematic review of the literature. Eur Spine J. 2010 Feb;19(2):181-204. doi: 10.1007/s00586-009-1155-x. PMID: 19756781. Exclusion Code: X1.
114. Nikoobakht M, Yekanineajd MS, Pakpour AH, et al. Plasma disc decompression compared to physiotherapy for symptomatic contained lumbar disc herniation: A prospective randomized controlled trial. Neurol Neurochir Pol. 2016;50(1):24-30. doi: 10.1016/j.pjnns.2015.11.001. Epub 2015 Nov 28. PMID: 26851686. Exclusion Code: X3.
115. North American Spine Society. Clinical Guidelines for Diagnosis and Treatment of Lumbar Disc Herniation with Radiculopathy. Burr Ridge, IL: Society NAS; 2012 2012. https://www.spine.org/Portals/0/Documents/ResearchClinicalCare/Guidelines/LumbarDiscHerniation.pdf. Exclusion Code: X1.
116. Ohtori S, Koshi T, Yamashita M, et al. Surgical versus nonsurgical treatment of selected patients with discogenic low back pain: a small-sized randomized trial. Spine (Phila Pa 1976). 2011 Mar 01;36(5):347-54. doi: 10.1097/BRS.0b013e3181d0c944. PMID: 20838371. Exclusion Code: X4.
117. Ong D, Chua NH, Vissers K. Percutaneous Disc Decompression for Lumbar Radicular Pain: A Review Article. Pain Pract. 2016 Jan;16(1):111-26. doi: 10.1111/papr.12250. PMID: 25354274. Exclusion Code: X2.
118. Osterman H, Seitsalo S, Karppinen J, et al. Effectiveness of microdiscectomy for lumbar disc herniation: a randomized controlled trial with 2 years of follow-up. Spine (Phila Pa 1976). 2006 Oct 01;31(21):2409-14. doi: 10.1097/01.brs.0000239178.08796.52. PMID: 17023847. Exclusion Code: X9.
119. Six-year outcomes of lumbar microdiscectomy: results of a randomised

controlled trial. Proceedings of the International Society for the Study of the Lumbar Spine; 2005. Exclusion Code: X11.
120. Pan L, Zhang P, Yin Q. Comparison of tissue damages caused by endoscopic lumbar discectomy and traditional lumbar discectomy: a randomised controlled trial. Int J Surg. 2014;12(5):534-7. doi: 10.1016/j.ijsu.2014.02.015. PMID: 24583364. Exclusion Code: X3.
121. Pan Z, Ha Y, Yi S, et al. Efficacy of Transforaminal Endoscopic Spine System (TESSYS) Technique in Treating Lumbar Disc Herniation. Med Sci Monit. 2016 Feb 18;22:530-9. PMID: 26887645. Exclusion Code: X3.
122. Pauza KJ, Howell S, Dreyfuss P, et al. A randomized, placebo-controlled trial of intradiscal electrothermal therapy for the treatment of discogenic low back pain. Spine J. 2004 Jan-Feb;4(1):27-35. PMID: 14749191. Exclusion Code: X4.
123. Payer M. "Minimally invasive" lumbar spine surgery: a critical review. Acta Neurochir (Wien). 2011 Jul;153(7):1455-9. doi: 10.1007/s00701-011-1023-4. PMID: 21533888. Exclusion Code: X2.
124. Pearson A, Lurie J, Tosteson T, et al. Who should have surgery for an intervertebral disc herniation? Comparative effectiveness evidence from the spine patient outcomes research trial. Spine (Phila Pa 1976). 2012 Jan 15;37(2):140-9. doi: 10.1097/BRS.0b013e3182276b2b. PMID: 21681140. Exclusion Code: X8.
125. Pearson AM, Blood EA, Frymoyer JW, et al. SPORT lumbar intervertebral disk herniation and back pain: does treatment, location, or morphology matter? Spine (Phila Pa 1976). 2008 Feb 15;33(4):428-35. doi: 10.1097/BRS.0b013e31816469de. PMID: 18277876. Exclusion Code: X8.
126. Peul WC, van Houwelingen HC, van den Hout WB, et al. Surgery versus prolonged conservative treatment for sciatica N Engl J Med. 2007 May 31;356(22):2245-56. doi: 10.1056/NEJMoa064039. PMID: 17538084. Exclusion Code: X9.
127. Phan K, Xu J, Schultz K, et al. Full-endoscopic versus micro-endoscopic and open discectomy: A systematic review and meta-analysis of outcomes and complications. Clin Neurol Neurosurg. 2017 Mar;154:1-12. doi: 10.1016/j.clineuro.2017.01.003. PMID: 28086154. Exclusion Code: X1.
128. Postacchini F, Lami R, Massobrio M. Chemonucleolysis versus surgery in lumbar disc herniations: correlation of the results to preoperative clinical pattern and size of the herniation. Spine (Phila Pa 1976). 1987 Mar;12(2):87-96. PMID: 3296236. Exclusion Code: X6.
129. Ran J, Hu Y, Zheng Z, et al. Comparison of discectomy versus sequestrectomy in lumbar disc herniation: a meta-analysis of comparative studies. PLoS One. 2015;10(3):e0121816. doi: 10.1371/journal.pone.0121816. PMID: 25815514. Exclusion Code: X1.
130. Rasouli MR, Rahimi-Movaghar V, Shokraneh F, et al. Minimally invasive discectomy versus microdiscectomy/open discectomy for symptomatic lumbar disc herniation. Cochrane Database of Systematic Reviews: John Wiley & Sons, Ltd; 2014. Exclusion Code: X1.
131. Rasouli MR, Rahimi-Movaghar V, Shokraneh F, et al. Minimally invasive discectomy versus microdiscectomy/open discectomy for symptomatic lumbar disc herniation. Cochrane Database Syst Rev. 2014 Sep 04(9):Cd010328. doi: 10.1002/14651858.CD010328.pub2. PMID: 25184502. Exclusion Code: X9.
132. Revel M, Payan C, Vallee C, et al. Automated percutaneous lumbar discectomy versus chemonucleolysis in the treatment of sciatica A randomized multicenter trial. Spine (Phila Pa 1976). 1993 Jan;18(1):1-7. PMID: 8434309. Exclusion Code: X6.
133. Reverberi C, Bottoli MG, Pennini M, et al. Disc coablation and epidural injection of steroids: a comparison of strategies in the treatment of mechanical spinal discogenic pain. Acta Neurochir Suppl. 2005;92:127-8. PMID: 15830983. Exclusion Code: X8.
134. Righesso O, Falavigna A, Avanzi O. Comparison of open discectomy with microendoscopic discectomy in lumbar disc herniations: results of a randomized controlled trial. Neurosurgery. 2007 Sep;61(3):545-9; discussion 9. doi:

10.1227/01.neu.0000290901.00320.f5. PMID: 17881967. Exclusion Code: X3.
135. Rihn JA, Hilibrand AS, Radcliff K, et al. Duration of symptoms resulting from lumbar disc herniation: effect on treatment outcomes: analysis of the Spine Patient Outcomes Research Trial (SPORT). J Bone Joint Surg Am. 2011 Oct 19;93(20):1906-14. doi: 10.2106/jbjs.j.00878. PMID: 22012528. Exclusion Code: X8.
136. Rowe SM ME-S, Song EK, Lee KB, Seo HY, Chung JY. A prospective randomized study of percutaneous endosc opicand microscopic lumbar disc ectomy. Proceedings of theAmerican Academy of Orthopaedic Surgeon; 2004. Exclusion Code: X11.
137. Ruan W, Feng F, Liu Z, et al. Comparison of percutaneous endoscopic lumbar discectomy versus open lumbar microdiscectomy for lumbar disc herniation: A meta-analysis. Int J Surg. 2016 Jul;31:86-92. doi: 10.1016/j.ijsu.2016.05.061. PMID: 27260312. Exclusion Code: X1.
138. Ruetten S, Komp M, Merk H, et al. Surgical treatment for lumbar lateral recess stenosis with the full-endoscopic interlaminar approach versus conventional microsurgical technique: a prospective, randomized, controlled study. J Neurosurg Spine. 2009 May;10(5):476-85. doi: 10.3171/2008.7.17634. PMID: 19442011. Exclusion Code: X4.
139. S. A-F. Study of percutaneous intradiscal nucleoplasty efficacy. Exclusion Code: X10.
140. Schenk B. Percutaneous laser disc decompression is more cost-effective than conventional surgical treatment for lumbar disc herniation and allows faster patient rehabilitation, while long-term results are comparable. ISRCTNR25884790. 2006. Exclusion Code: X10.
141. Schenk B, Brouwer PA, Peul WC, et al. Percutaneous laser disk decompression: a review of the literature. AJNR Am J Neuroradiol. 2006 Jan;27(1):232-5. PMID: 16418391. Exclusion Code: X2.
142. Schick U, Dohnert J, Richter A, et al. Microendoscopic lumbar discectomy versus open surgery: an intraoperative EMG study. Eur Spine J. 2002 Feb;11(1):20-6. PMID: 11931060. Exclusion Code: X7.
143. Schizas C, Tsiridis E, Saksena J. Microendoscopic discectomy compared with standard microsurgical discectomy for treatment of uncontained or large contained disc herniations. Neurosurgery. 2005 Oct;57(4 Suppl):357-60; discussion -60. PMID: 16234685. Exclusion Code: X8.
144. Schoenfeld AJ, Bono CM. Does surgical timing influence functional recovery after lumbar discectomy? A systematic review. Clin Orthop Relat Res. 2015 Jun;473(6):1963-70. doi: 10.1007/s11999-014-3505-1. PMID: 24526298. Exclusion Code: X7.
145. Shamji MF, Bains I, Yong E, et al. Treatment of herniated lumbar disk by sequestrectomy or conventional diskectomy. World Neurosurg. 2014 Nov;82(5):879-83. doi: 10.1016/j.wneu.2013.02.066. PMID: 23454687. Exclusion Code: X8.
146. Shih P, Wong AP, Smith TR, et al. Complications of open compared to minimally invasive lumbar spine decompression. J Clin Neurosci. 2011 Oct;18(10):1360-4. doi: 10.1016/j.jocn.2011.02.022. PMID: 21775145. Exclusion Code: X4.
147. Shin DA, Kim KN, Shin HC, et al. The efficacy of microendoscopic discectomy in reducing iatrogenic muscle injury. J Neurosurg Spine. 2008 Jan;8(1):39-43. doi: 10.3171/spi-08/01/039. PMID: 18173345. Exclusion Code: X7.
148. Shriver MF, Xie JJ, Tye EY, et al. Lumbar microdiscectomy complication rates: a systematic review and meta-analysis. Neurosurg Focus. 2015 Oct;39(4):E6. doi: 10.3171/2015.7.focus15281. PMID: 26424346. Exclusion Code: X1.
149. Singh V, Benyamin RM, Datta S, et al. Systematic review of percutaneous lumbar mechanical disc decompression utilizing Dekompressor. Pain Physician. 2009 May-Jun;12(3):589-99. PMID: 19461825. Exclusion Code: X9.
150. Singh V, Manchikanti L, Benyamin RM, et al. Percutaneous lumbar laser disc decompression: a systematic review of current evidence. Pain Physician. 2009 May-Jun;12(3):573-88. PMID: 19461824. Exclusion Code: X9.
151. Singh V, Manchikanti L, Calodney AK, et al. Percutaneous lumbar laser disc

decompression: an update of current evidence. Pain Physician. 2013 Apr;16(2 Suppl):Se229-60. PMID: 23615885. Exclusion Code: X8.
152. Smith N, Masters J, Jensen C, et al. Systematic review of microendoscopic discectomy for lumbar disc herniation. Eur Spine J. 2013 Nov;22(11):2458-65. doi: 10.1007/s00586-013-2848-8. PMID: 23793558. Exclusion Code: X1.
153. Steffen R. A prospective comparative study of chemonucleolysis and laser discectomy. Orthop Trans. 1996;20:388. Exclusion Code: X6.
154. Steffen R. Chemonucleolysis versus laser disc decompression-a prospective randomised trial. J Bone Joint Surg, B. 1997;79:247. Exclusion Code: X11.
155. Steffen R. Laser discotomy/chemonucleolysis ini the therapy of displaced lumbal intervertebral disc—prospective randomised comparison. Hefte zur der Unfallchirurg. 1999;271:215-21. Exclusion Code: X12.
156. Stula D. [Chemonucleolysis with chymopapain in lumbar disk herniationa. Randomized comparative study with operated patients]. Neurochirurgia (Stuttg). 1990 Sep;33(5):169-72. doi: 10.1055/s-2008-1053577. PMID: 2263306. Exclusion Code: X12.
157. Thomas KC, Fisher CG, Boyd M, et al. Outcome evaluation of surgical and nonsurgical management of lumbar disc protrusion causing radiculopathy. Spine (Phila Pa 1976). 2007 Jun 01;32(13):1414-22. doi: 10.1097/BRS.0b013e318060a5d1. PMID: 17545909. Exclusion Code: X8.
158. Thome C, Borm W, Meyer F. Degenerative lumbar spinal stenosis: current strategies in diagnosis and treatment. Dtsch Arztebl Int. 2008 May;105(20):373-9. doi: 10.3238/arztebl.2008.0373. PMID: 19626175. Exclusion Code: X4.
159. Tohmeh AG, Smith WD. Lumbar total disc replacement by less invasive lateral approach: a report of results from two centers in the US IDE clinical trial of the XL TDR(R) device. Eur Spine J. 2015 Apr;24 Suppl 3:331-8. doi: 10.1007/s00586-015-3843-z. PMID: 25794698. Exclusion Code: X5.
160. Tosteson AN, Skinner JS, Tosteson TD, et al. The cost effectiveness of surgical versus nonoperative treatment for lumbar disc herniation over two years: evidence from the Spine Patient Outcomes Research Trial (SPORT). Spine (Phila Pa 1976). 2008 Sep 01;33(19):2108-15. PMID: 18777603. Exclusion Code: X9.
161. Toyoda H, Nakamura H, Konishi S, et al. Clinical outcome of microsurgical bilateral decompression via unilateral approach for lumbar canal stenosis: minimum five-year follow-up. Spine (Phila Pa 1976). 2011 Mar 01;36(5):410-5. doi: 10.1097/BRS.0b013e3181d25829. PMID: 20847714. Exclusion Code: X4.
162. Toyone T, Tanaka T, Kato D, et al. Low-back pain following surgery for lumbar disc herniation. A prospective study. J Bone Joint Surg Am. 2004 May;86-a(5):893-6. PMID: 15118028. Exclusion Code: X8.
163. Tureyen K. One-level one-sided lumbar disc surgery with and without microscopic assistance: 1-year outcome in 114 consecutive patients. J Neurosurg. 2003 Oct;99(3 Suppl):247-50. PMID: 14563140. Exclusion Code: X3.
164. Urrutia G, Kovacs F, Nishishinya MB, et al. Percutaneous thermocoagulation intradiscal techniques for discogenic low back pain. Spine (Phila Pa 1976). 2007 May 01;32(10):1146-54. doi: 10.1097/01.brs.0000261492.55121.93. PMID: 17471101. Exclusion Code: X4.
165. Vaccaro AR, Fehlings MG. The applicability of clinical equipoise and sham surgery in patients with symptomatic lumbar radiculopathy due to a herniated disc: the SPORT trial. LWW; 2007. Exclusion Code: X2.
166. van Alphen HA, Braakman R, Berfelo MW, et al. Chemonucleolysis or discectomy? Results of a randomized multicentre trial in patients with a herniated lumbar intervertebral disc (a preliminary report). Acta Neurochir Suppl (Wien). 1988;43:35-8. PMID: 3063076. Exclusion Code: X6.
167. van Alphen HA, Braakman R, Bezemer PD, et al. Chemonucleolysis versus discectomy: a randomized multicenter trial. J Neurosurg. 1989 Jun;70(6):869-75. doi: 10.3171/jns.1989.70.6.0869. PMID: 2654335. Exclusion Code: X6.

168. Veresciagina K, Spakauskas B, Ambrozaitis KV. Clinical outcomes of patients with lumbar disc herniation, selected for one-level open-discectomy and microdiscectomy. Eur Spine J. 2010 Sep;19(9):1450-8. doi: 10.1007/s00586-010-1431-9. PMID: 20502926. Exclusion Code: X8.
169. A randomised prospective study comparing laser decompression and epidural steroid injection in alleviating radicular pain secondary to prolapsed lumbar discs. Proceedings of the International Society for Study of theLumbar Spine; 1998. Exclusion Code: X11.
170. Wang H, Lü F, Jiang J, et al. Minimally invasive lumbar interbody fusion via MAST Quadrant retractor versus open surgery: a prospective randomized clinical trial. Chinese medical journal; 2011. p. 3868-74. Exclusion Code: X5.
171. Wang XS, Sun RF, Ji Q, et al. A meta-analysis of interlaminar minimally invasive discectomy compared to conventional microdiscectomy for lumbar disk herniation. Clin Neurol Neurosurg. 2014 Dec;127:149-57. doi: 10.1016/j.clineuro.2014.10.001. PMID: 25459263. Exclusion Code: X1.
172. Watters WC, 3rd, McGirt MJ. An evidence-based review of the literature on the consequences of conservative versus aggressive discectomy for the treatment of primary disc herniation with radiculopathy. Spine J. 2009 Mar;9(3):240-57. doi: 10.1016/j.spinee.2008.08.005. PMID: 18809356. Exclusion Code: X1.
173. Watters WC, 3rd, Mirkovic S, Boss J. Treatment of the isolated lumbar intervertebral disc herniation: microdiscectomy versus chemonucleolysis. Spine (Phila Pa 1976). 1988 Mar;13(3):360-2. PMID: 3388122. Exclusion Code: X6.
174. Weinstein JN, Lurie JD, Tosteson TD, et al. Surgical vs nonoperative treatment for lumbar disk herniation: the Spine Patient Outcomes Research Trial (SPORT) observational cohort. Jama 2006 Nov 22;296(20):2451-9. doi: 10.1001/jama.296.20.2451. PMID: 17119141. Exclusion Code: X8.
175. Weinstein JN, Tosteson TD, Lurie JD, et al. Surgical versus nonoperative treatment for lumbar spinal stenosis four-year results of the Spine Patient Outcomes Research Trial. Spine (Phila Pa 1976). 2010 Jun 15;35(14):1329-38. doi: 10.1097/BRS.0b013e3181e0f04d. PMID: 20453723. Exclusion Code: X8.
176. Wu S, Li X, Lin C, et al. CT-guided nucleoplasty with radiofrequency energy for the treatment of lumbar disk herniation. J Spinal Disord Tech. 2015 Feb;28(1):E9-16. doi: 10.1097/bsd.0000000000000132. PMID: 25023711. Exclusion Code: X3.
177. Wu Z, Wei LX, Li J, et al. Percutaneous treatment of non-contained lumbar disc herniation by injection of oxygen-ozone combined with collagenase. Eur J Radiol. 2009 Dec;72(3):499-504. doi: 10.1016/j.ejrad.2008.07.029. PMID: 18778905. Exclusion Code: X3.
178. Xu HH, Wang XQ, Zhang YL, et al. Efficacy evaluation for the treatment of unstable lumbar disc herniation by traditional and modified lamina osteotomy replantation. Chin Med J (Engl). 2013;126(15):2840-5. PMID: 23924453. Exclusion Code: X5.
179. Yao Y, Zhang H, Wu J, et al. Comparison of Three Minimally Invasive Spine Surgery Methods for Revision Surgery for Recurrent Herniation After Percutaneous Endoscopic Lumbar Discectomy. World Neurosurg. 2017 Apr;100:641-7.e1. doi: 10.1016/j.wneu.2017.01.089. PMID: 28153616. Exclusion Code: X3.
180. Ying J, Huang K, Zhu M, et al. The Effect and Feasibility Study of Transforaminal Percutaneous Endoscopic Lumbar Discectomy Via Superior Border of Inferior Pedicle Approach for Down-Migrated Intracanal Disc Herniations. Medicine (Baltimore). 2016 Feb;95(8):e2899. doi: 10.1097/md.0000000000002899. PMID: 26937930. Exclusion Code: X3.
181. Yoshihara H, Chatterjee D, Paulino CB, et al. Revision Surgery for "Real" Recurrent Lumbar Disk Herniation: A Systematic Review. Clin Spine Surg. 2016 Apr;29(3):111-8. doi: 10.1097/bsd.0000000000000365. PMID: 27002374. Exclusion Code: X7.
182. Zhang L, Ding XL, Zhao XL, et al. Fluoroscopy-guided Bipolar Radiofrequency Thermocoagulation Treatment for Discogenic Low Back Pain. Chin Med J (Engl). 2016 Oct 05;129(19):2313-8. doi: 10.4103/0366-

6999.190682. PMID: 27647190. Exclusion Code: X4.

183. Zhang Q, Qian J, Zhu Y. Meta-Analysis on Microdiscectomy and Sequestrectomy for Lumbar Disc Herniation. J Invest Surg. 2015;28(4):225-9. PMID: 26268422. Exclusion Code: X1.

184. Zhu H, Zhou XZ, Cheng MH, et al. The efficacy of coblation nucleoplasty for protrusion of lumbar intervertebral disc at a two-year follow-up. Int Orthop. 2011 Nov;35(11):1677-82. doi: 10.1007/s00264-010-1196-0. PMID: 21240606. Exclusion Code: X8.

Appendix F. Individual Study Risk of Bias Assessments

Table F-1. Risk of bias ratings-Overall rating and randomization process

Main Study Author (Year); Follow-up Studies Author (Year)	Overall Bias		Randomization Process Bias				
	Overall Quality Rating	Overall Rationale for Quality Rating	Was the allocation sequence random?	Was allocation sequence concealed until participants were recruited and assigned to interventions?	Were there baseline imbalances that suggest a problem with the randomization process?	Bias arising from randomization or selection?	Comments
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48]	Low for outcomes up to 2y; High for outcomes longer than 2y	This risk of bias assessment includes both 1y, 2y, and 5y results.	Yes	Yes	No	Low	None
Brouwer (2015)[39] Brouwer (2017)[102]	Some concerns	Some concerns for bias because of lack of blinding of study participants, who also served as outcome assesors.	Yes	Yes	No	Low	None
Chatterjee (1995)[38]	Some concerns	Some concerns for bias because of no information on randomization methods, patients were not blinded, and unclear whether outcome assessors were fully blinded or if patient self-reported outcomes were used.	No information	No information	No information	Some concerns	No information provided at all about how participants were randomized and how allocation was concealed, and the study did not provide any comparison of baseline characteristics between groups to assess adequancy of randomization.

(continued)

Table F-1. Risk of bias ratings-Overall rating and randomization process (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Overall Bias		Randomization Process Bias				
	Overall Quality Rating	Overall Rationale for Quality Rating	Was the allocation sequence random?	Was allocation sequence concealed until participants were recruited and assigned to interventions?	Were there baseline imbalances that suggest a problem with the randomization process?	Bias arising from randomization or selection?	Comments
Erginousakis (2011)[37]	High	High risk of bias concerns due to inadequate method of randomization. Some concerns for risk of bias due to lack of participant blinding and use of patient self-reported outcomes. Low risk of bias for some safety outcomes.	No information	Probably no	No information	High	Evenly numbered referral patients were assigned to surgery and odd-numbered referral patients were assigned to conservative care. This is an inadequate method of randomization. Only gender and age are reported at basleline; there were differences in baseline pain between the groups, but these differences were small.

(continued)

Table F-1. Risk of bias ratings-Overall rating and randomization process (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Overall Bias		Randomization Process Bias				
	Overall Quality Rating	Overall Rationale for Quality Rating	Was the allocation sequence random?	Was allocation sequence concealed until participants were recruited and assigned to interventions?	Were there baseline imbalances that suggest a problem with the randomization process?	Bias arising from randomization or selection?	Comments
Franke (2009)[36]	Some concerns	Some concerns for bias due to no information about baseline balance between groups, lack of participant blinding, including outcome assessment with patient self-reported outcomes.	Probably yes	No information	No information	Some concerns	Authors report no significant group difference for preop parameters but don't provide data.
Gerszten (2010)[41]	Some concerns for outcomes at 6w; high for outcomes at 12w or later.	Some concerns for bias because patients were not blinded and self-reported outcomes used at all time points. High risk of bias for outcomes reported at 12w and later because of second procedures provided to participants in both groups, and high attrition at 12w and later. Low risk of bias for some safety outcomes.	Probably yes	Yes	Probably no	Low	Mean duration of radicular pain was double in the epidural injection group, but the sample size was small and range was large.

(continued)

Table F-1. Risk of bias ratings-Overall rating and randomization process (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Overall Bias		Randomization Process Bias				
	Overall Quality Rating	Overall Rationale for Quality Rating	Was the allocation sequence random?	Was allocation sequence concealed until participants were recruited and assigned to interventions?	Were there baseline imbalances that suggest a problem with the randomization process?	Bias arising from randomization or selection?	Comments
Haines (2002)[42]	High	High risk of bias because of randomization process and high attrition in both groups, also some concerns for bias because participants and clinician outcome assessors not blinded.	No information	No information	Probably yes	High	No information about randomization process and allocation concealment combined with imbalances in randomization (21 randomized to APLD 13 randomized to discectomy); average age was 7 years less in the discectomy group, but this difference was not statistically significant.
Henriksen (1996)[35]	Some concerns	Some concerns because of randomization process/allocation concealment unconventional and no baseline characteristic to assess balance and use of patient self-reported outcomes for unblinded participants.	Probably yes	Probably yes	No information	Some concerns	Not a conventional approach to generating random sequence, and no table of baseline characteristics to assess whether groups were balanced at baseline.

(continued)

Table F-1. Risk of bias ratings-Overall rating and randomization process (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Overall Bias		Randomization Process Bias				
	Overall Quality Rating	Overall Rationale for Quality Rating	Was the allocation sequence random?	Was allocation sequence concealed until participants were recruited and assigned to interventions?	Were there baseline imbalances that suggest a problem with the randomization process?	Bias arising from randomization or selection?	Comments
Hermantin (1999)[43]	Some concerns	Some concerns for bias because of use of patient-self reported outcomes with participants that were not blinded.	No information	Probably yes	Probably no	Low	None
Huang (2005)[24]	Some concerns	Some concerns for bias because of randomication/allocation concealment process and use of patient-reported outcomes when participants were not blinded.	No information	No information	No information	Some concerns	None
Mayer (1993)[34]	High	Some concerns for bias over randomization process, deviations from interventions, and use of patient-reported outcomes among participants that were not blinded. Because of some concerns in 3 of the 5 domains, this increases the level of concern overall to high.	No information	No information	Probably no	Some concerns	None

(continued)

Table F-1. Risk of bias ratings-Overall rating and randomization process (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Overall Bias		Randomization Process Bias				
	Overall Quality Rating	Overall Rationale for Quality Rating	Was the allocation sequence random?	Was allocation sequence concealed until participants were recruited and assigned to interventions?	Were there baseline imbalances that suggest a problem with the randomization process?	Bias arising from randomization or selection?	Comments
McMorland (2010)[23]	Some concerns	Some concerns for risk of bias because study used patient self-reported outcomes and participants were not blinded to treatment allocation. Note: this risk of bias rating only applies to outcomes up to 12 weeks for the intent-to-treat analysis. Crossovers were allowed after 12 weeks and raise the risk of bias to high for outcomes after 12 weeks. Low risk of bias for some safety outcomes.	Yes	Probably yes	No	Low	None

(continued)

Table F-1. Risk of bias ratings-Overall rating and randomization process (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Overall Bias		Randomization Process Bias				
	Overall Quality Rating	Overall Rationale for Quality Rating	Was the allocation sequence random?	Was allocation sequence concealed until participants were recruited and assigned to interventions?	Were there baseline imbalances that suggest a problem with the randomization process?	Bias arising from randomization or selection?	Comments
North (2005)[46]	High	High risk of bias because of extensive deviations from intended interventions as a result of crossovers, also differential attrition. Some concerns for patient self-reported outcomes when participants were not blinded.	Yes	Yes	No information	Low	None
Osterman (2003)[33]	High	High risk of bias because of extensive differential crossover from physiotherapy to surgery, and differential co-interventions related to the outcome. Also, some concerns for patient-reported outcomes when participants were not blinded. Low risk of bias for some safety outcomes.	Yes	Yes	No	Low	None

(continued)

Table F-1. Risk of bias ratings-Overall rating and randomization process (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Overall Bias		Randomization Process Bias				
	Overall Quality Rating	Overall Rationale for Quality Rating	Was the allocation sequence random?	Was allocation sequence concealed until participants were recruited and assigned to interventions?	Were there baseline imbalances that suggest a problem with the randomization process?	Bias arising from randomization or selection?	Comments
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98]	High	Moderate to extensive deviations from intended interventions as a result of crossovers, which are likely non-random. This most likely biases the effect toward the null. Some concerns from use of patient-reported outcomes when participants were not blinded. Low risk of bias for some safety outcomes.	Yes	Yes	No	Low	None
Ruetten (2008)[31]	High	Randomization and allocation concealment process was inadequate and thus, subjects study to high risk of bias. Some concerns for bias because of use of self-reported outcomes among participants that were not blinded.	Probably no	Probably no	No information	High	Randomization was peformed by alternate assignment of participants. No baseline characteristics are provided to assess balance between groups.

(continued)

Table F-1. Risk of bias ratings-Overall rating and randomization process (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Overall Bias		Randomization Process Bias				
	Overall Quality Rating	Overall Rationale for Quality Rating	Was the allocation sequence random?	Was allocation sequence concealed until participants were recruited and assigned to interventions?	Were there baseline imbalances that suggest a problem with the randomization process?	Bias arising from randomization or selection?	Comments
Ruetten (2009)[47]	High	Randomization and allocation concealment process was inadequate and thus, subjects study to high risk of bias. Some concerns for bias because of use of self-reported outcomes among participants that were not blinded.	Probably no	Probably no	No information	High	Randomization was peformed by alternate assignment of participants. No baseline characteristics are provided to assess balance between groups.
Ryang (2008)[30] Gempt (2013)[103]	Some concerns	Some concerns for bias in three of the five domains elevates this to level of concern for bias to high. Some concerns over randomization process, information about missing data, and use of patient-reported outcomes among participants that were not blinded.	No information	No information	Probably no	Some concerns	Method of randomization and allocation concealment not reported; only age and gender reported at baseline to assess baseline comparability.

(continued)

Table F-1. Risk of bias ratings-Overall rating and randomization process (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Overall Bias		Randomization Process Bias				
	Overall Quality Rating	Overall Rationale for Quality Rating	Was the allocation sequence random?	Was allocation sequence concealed until participants were recruited and assigned to interventions?	Were there baseline imbalances that suggest a problem with the randomization process?	Bias arising from randomization or selection?	Comments
Sasaoka (2006)[25]	High	Some concerns for bias across 4 of the 5 domains increase the overall risk of bias to high. Some concerns for bias in randomization, deviations from interventions, missing data and measurement of outcome.	No information	No information	No information	Some concerns	No details on methods of randomization and allocation concealment, only sex and age, and location of herniation was reported by group so unable to tell if groups were balanced at baseline.
Teli (2010)[29]	Some concerns	Some concerns for bias because patient-reported outcomes were used and participants were not blinded to treatment allocation.	Probably yes	No information	No	Low	No details on allocation concealment.
Thome (2005)[28] Barth (2008)[104]	Some concerns	Some concerns for bias because patient-reported outcomes were used and participants were likely not blinded to treatment allocation.	Yes	Probably yes	No	Low	Randomization was reported as "concealed" but no specific details as to how it was concealed were provided.

(continued)

Table F-1. Risk of bias ratings-Overall rating and randomization process (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Overall Bias		Randomization Process Bias				
	Overall Quality Rating	Overall Rationale for Quality Rating	Was the allocation sequence random?	Was allocation sequence concealed until participants were recruited and assigned to interventions?	Were there baseline imbalances that suggest a problem with the randomization process?	Bias arising from randomization or selection?	Comments
Tullberg (1993)[27]	Some concerns	Some concerns for bias because methods of randomization and allocation concealment were not reported and not enough information was available to judge baseline balance among groups. Also some concerns for bias because patient-reported outcomes were used and participants were not blinded to treatment allocation.	No information	No information	No information	Some concerns	Method of randomization and allocation concealment not reported. Very little information was provided except sex, age, length of time off work prior to surgery, and workloads to assess balance between groups. The discectomy group was biased toward heavier workloads, but sample size is small so difference may not be significant.

(continued)

Table F-1. Risk of bias ratings-Overall rating and randomization process (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Overall Bias		Randomization Process Bias				
	Overall Quality Rating	Overall Rationale for Quality Rating	Was the allocation sequence random?	Was allocation sequence concealed until participants were recruited and assigned to interventions?	Were there baseline imbalances that suggest a problem with the randomization process?	Bias arising from randomization or selection?	Comments
Weber (1983)[26]	High	High risk of bias because extensive crossovers, also some concerns because patient-reported outcomes were used and participants were not blinded.	Probably yes	Probably yes	No information	Low	Age and gender are the only characteristics reported
Weinstein (2006) [22] Weinstein (2008)[99] Lurie Jon (2014)[100]	High	High risk of bias because of extensive deviations from intended interventions as a result of crossovers, which are likely non-random. This most likely biases the effect toward the null. High risk of bias for outcomes reported at 2y or longer because of high attrition. Some concerns for bias because of patient-reported outcomes when participants were not blinded, and selective outcome reporting bias. Low risk of bias for some safety outcomes.	Yes	Probably yes	No	Low	None

Table F-2. Risk of bias —deviations from intended interventions

Main Study Author (Year); Follow-up Studies Author (Year)	Were the participants aware of their assigned intervention during the trial?	Were carers and trial personnel aware of participants' assigned intervention during the trials?	Were there deviations from the intended intervention beyond what would be expected in usual practice?	Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome?	Were any participants analyzed in a group different from the one to which they were assigned?	Was there potential for a substantial impact of analyzing participants in the wrong group?	Bias arising from deviations from intended interventions?	Comments
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48]	No	Yes	No	NA	No	NA	Low	Not feasible to blind surgeons and other caregivers to the intervention assignment.
Brouwer (2015)[39] Brouwer (2017)[102]	Yes	Yes	No	NA	No	NA	Low	None
Chatterjee (1995)[38]	Yes	Yes	No	NA	No	NA	Low	It was not reported whether participants were blinded, but the study treatments are different enough that it would be obvious.
Erginousakis (2011)[37]	Yes	Probably no	Probably no	NA	Probably no	NA	Low	None
Franke (2009)[36]	No information	No information	Probably no	NA	No	NA	Low	None

(continued)

Table F-2. Risk of bias —deviations from intended interventions (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Were the participants aware of their assigned intervention during the trial?	Were carers and trial personnel aware of participants' assigned intervention during the trials?	Were there deviations from the intended intervention beyond what would be expected in usual practice?	Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome?	Were any participants analyzed in a group different from the one to which they were assigned?	Was there potential for a substantial impact of analyzing participants in the wrong group?	Bias arising from deviations from intended interventions ?	Comments
Gerszten (2010)[41]	Yes	Yes	Probably yes	Yes	Probably no	NA	Some concerns for 6w outcomes; High for outcomes at 12w or later.	5 patients were randomized but did not receive treatment (1 in surgical group, 4 in nonsurgical group). 30 of 40 patients randomized to epidural injections received both injections; 10 opted to not receive both injections. By 26w, 12 participants in surgery group received additional unspecfied second procedure. By 26w, 8 participants in the epidural injection received additional unspecfiied second procedures. These numbers are 9 and 5, respectively, at 12w.

(continued)

Table F-2. Risk of bias —deviations from intended interventions (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Were the participants aware of their assigned intervention during the trial?	Were carers and trial personnel aware of participants' assigned intervention during the trials?	Were there deviations from the intended intervention beyond what would be expected in usual practice?	Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome?	Were any participants analyzed in a group different from the one to which they were assigned?	Was there potential for a substantial impact of analyzing participants in the wrong group?	Bias arising from deviations from intended interventions?	Comments
Haines (2002)[42]	No information	Yes	No	NA	No	NA	Low	None
Henriksen (1996)[35]	No	Probably yes	Probably no	NA	Probably no	NA	Low	None
Hermantin (1999)[43]	Yes	Yes	No	NA	No	NA	Low	None
Huang (2005)[24]	No information	Yes	No	NA	No	NA	Low	None
Mayer (1993)[34]	No information	Yes	Probably yes	Probably yes	No	NA	Some concerns	3 patients in percutaneous surgery group also received microdiscectomy as a second procedure.
McMorland (2010)[23]	Yes	Yes	Probably no	NA	No	NA	Low	Pertains to outcomes at 12w

(continued)

Table F-2. Risk of bias —deviations from intended interventions (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Were the participants aware of their assigned intervention during the trial?	Were carers and trial personnel aware of participants' assigned intervention during the trials?	Were there deviations from the intended intervention beyond what would be expected in usual practice?	Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome?	Were any participants analyzed in a group different from the one to which they were assigned?	Was there potential for a substantial impact of analyzing participants in the wrong group?	Bias arising from deviations from intended interven-tions?	Comments
North (2005)[46]	Yes	Yes	Yes	Yes	No information	NA	High	10 patients (4 in reoperation group, 6 in spinal cord stimulation group) were excluded postrandomization because of failure to receive authorization. 5 (21%) of those allocated to spinal cord stimulation also received reoperation. 14 (54%) of those allocated to reoperation also received spinal cord stimulation.

(continued)

Table F-2. Risk of bias —deviations from intended interventions (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Were the participants aware of their assigned intervention during the trial?	Were carers and trial personnel aware of participants' assigned intervention during the trials?	Were there deviations from the intended intervention beyond what would be expected in usual practice?	Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome?	Were any participants analyzed in a group different from the one to which they were assigned?	Was there potential for a substantial impact of analyzing participants in the wrong group?	Bias arising from deviations from intended interven-tions?	Comments
Osterman (2003)[33]	Yes	Yes	Yes	Yes	No	NA	High	0% allocated to surgery received physiotherapy; 39% allocated to physiotherapy received surgery. In addition, 15 participants in the physiotherapy group and 8 participants in the surgery group reported receiving additional physical therapy outside of the study during follow-up.

(continued)

Table F-2. Risk of bias —deviations from intended interventions (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Were the participants aware of their assigned intervention during the trial?	Were carers and trial personnel aware of participants' assigned intervention during the trials?	Were there deviations from the intended intervention beyond what would be expected in usual practice?	Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome?	Were any participants analyzed in a group different from the one to which they were assigned?	Was there potential for a substantial impact of analyzing participants in the wrong group?	Bias arising from deviations from intended interven-tions?	Comments
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98]	Yes	Probably yes	Yes	Probably yes	No	NA	High	Moderate deviation in surgical group, 11.3% did not receive surgery; extensive deviation in nonsurgical group, 39% received surgery by 1 year and 44% received surgery by 2 years. These deviations make the groups more similar to each other and are very likely to bias results toward the null.
Ruetten (2008)[31]	Yes	Probably yes	Probably no	NA	Probably no	NA	Low	The physicians performing the 2-year exams were not involved in the surgeries, but it is probably visually clear what surgery may have been performed.
Ruetten (2009)[47]	Yes	Probably yes	Probably no	NA	Probably no	NA	Low	None

(continued)

Table F-2. Risk of bias for surgery—deviations from intended interventions (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Were the participants aware of their assigned intervention during the trial?	Were carers and trial personnel aware of participants' assigned intervention during the trials?	Were there deviations from the intended intervention beyond what would be expected in usual practice?	Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome?	Were any participants analyzed in a group different from the one to which they were assigned?	Was there potential for a substantial impact of analyzing participants in the wrong group?	Bias arising from deviations from intended interven-tions?	Comments
Ryang (2008)[30] Gempt (2013)[103]	No information	Yes	Probably no	NA	No	NA	Low	None
Sasaoka (2006)[25]	No information	No information	Probably no	NA	Probably no	NA	Some concerns	No information about blinding of participants or clinicians, no information about deviations from interventions though likely crossovers would have been reported if they occurred.
Teli (2010)[29]	Yes	Probably yes	Probably no	NA	No	NA	Low	None
Thome (2005)[28] Barth (2008)[104]	No information	No information	No	NA	Probably no	NA	Low	None
Tullberg (1993)[27]	No information	No information	Probably no	NA	Yes	Probably no	Low	No information about blinding of participants or clinicians, no information about deviations from interventions.

(continued)

Table F-2. Risk of bias—deviations from intended interventions (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Were the participants aware of their assigned intervention during the trial?	Were carers and trial personnel aware of participants' assigned intervention during the trials?	Were there deviations from the intended intervention beyond what would be expected in usual practice?	Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome?	Were any participants analyzed in a group different from the one to which they were assigned?	Was there potential for a substantial impact of analyzing participants in the wrong group?	Bias arising from deviations from intended interven-tions?	Comments
Weber (1983)[26]	Yes	Probably yes	Probably yes	Probably yes	No	NA	High	Over a quarter of participants allocated to conservative management were referred for surgery within a year of randomization. Only 2 surgical patients did not have the operation as randomized. Authors report the data so an ITT analysis can be performed.

(continued)

Table F-2. Risk of bias —deviations from intended interventions (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Were the participants aware of their assigned intervention during the trial?	Were carers and trial personnel aware of participants' assigned intervention during the trials?	Were there deviations from the intended intervention beyond what would be expected in usual practice?	Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome?	Were any participants analyzed in a group different from the one to which they were assigned?	Was there potential for a substantial impact of analyzing participants in the wrong group?	Bias arising from deviations from intended interven-tions?	Comments
Weinstein (2006) [22] Weinstein (2008)[99] Lurie Jon (2014)[100]	Yes	Yes	Yes	Yes	No	NA	High	Extensive deviation. Of those assigned to surgery, 50% received surgery by 3 months (60% by 2 years); of those assigned to no surgery, 30% received surgery by 3 months (45% by 2 years). These deviations make the groups more similar to each other and are very likely to bias results toward the null.

Table F-3. Risk of bias —missing outcome data

Main Study Author (Year); Follow-up Studies Author (Year)	Were outcome data available for all, or nearly all, participants randomized?	Are the proportions of missing outcome data and reasons for missing outcome data similar across intervention groups?	Is there evidence that results were robust to the presence of missing outcome data?	Bias arising from missing outcome data?	Comments
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48]	Yes for outcomes up to 2y; No for outcomes longer than 2y.	Yes	NR	Low for outcomes up to 2y High for outcomes longer than 2y.	SG1: 93% had data available at 52w, 92% had data available at 2y. 64% had data available at 5y. SG2: 94% had data available at 52w, 89% had data available at 2y. 61% had data available at 5y.
Brouwer (2015)[39] Brouwer (2017)[102]	Yes	NA	NA	Low	None
Chatterjee (1995)[38]	Yes	NA	NA	Low	None
Erginousakis (2011)[37]	Probably yes	NA	NA	Low	None
Franke (2009)[36]	Yes	NA	NA	Low	None
Gerszten (2010)[41]	No	Probably yes	Probably yes	Some concerns at 6w, high for outcomes at 12w or later.	High attrition in both groups at 12w and 26w (surgery attrition 34%; epidural steroid group attrition 30%)
Haines (2002)[42]	No	Probably no	No information	High	Conflicting information about number of participants lost to follow-up; 24% attrition in APLD group; 31% attrition in discectomy group; overall attrition at 52w was 44%.
Henriksen (1996)[35]	Yes	NA	NA	Low	None
Hermantin (1999)[43]	Yes	NA	NA	Low	None
Huang (2005)[24]	Yes	NA	NA	Low	None
Mayer (1993)[34]	Yes	NA	NA	Low	None
McMorland (2010)[23]	Yes	NA	NA	Low	None

(continued)

Table F-3. Risk of bias —missing outcome data (continued)

	Were outcome data available for all, or nearly all, participants randomized?	Are the proportions of missing outcome data and reasons for missing outcome data similar across intervention groups?	Is there evidence that results were robust to the presence of missing outcome data?	Bias arising from missing outcome data?	Comments
North (2005)[46]	Probably no	No	No information	High	Loss of follow-up was exclusively in the spinal stimulation group. 82% of randomized patients provided short-term results; 75% of randomized provided long-term results. 98% of the randomized and treated patients provided short-term results; 90% of the randomized and treated patients provided long-term results.
Osterman (2003)[33]	Yes	NA	NA	Low	None
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98]	Yes	NA	NA	Low	None
Ruetten (2008)[31]	Yes	NA	NA	Low	None
Ruetten (2009)[47]	Yes	NA	NA	Low	None
Ryang (2008)[30] Gempt (2013)[103]	No information	No information	No information	Some concerns	No information provided as to how many participants contributed data at follow-up; no way to ascertain whether missing data were present.
Sasaoka (2006)[25]	No information	No information	No information	Some concerns	No information about how many participants contributed follow-up data.
Teli (2010)[29]	Yes	NA	NA	Low	None
Thome (2005)[28] Barth (2008)[104]	Yes	NA	NA	Low	None

(continued)

Table F-3. Risk of bias —missing outcome data (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Were outcome data available for all, or nearly all, participants randomized?	Are the proportions of missing outcome data and reasons for missing outcome data similar across intervention groups?	Is there evidence that results were robust to the presence of missing outcome data?	Bias arising from missing outcome data?	Comments
Tullberg (1993)[27]	Yes	NA	NA	Low	2 patients, 1 in each group, underwent reoperations for recurrence within a year of the first surgery and were not included among the overall study's results.
Weber (1983)[26]	Probably yes	NA	NA	Low	None
Weinstein (2006)[22] Weinstein (2008)[99] Lurie Jon (2014)[100]	Probably yes for follow-up at 52w or less; no for follow-up at outcomes at 2y or longer	Yes	NR	Low for outcomes at 52w or less; High for outcomes at 2y or longer	SG1: 82% had data available at 52w, 76% had data available at 2y; 64% had data available at 4y and 8y SG2: 83% had data available at 52w; 75% had data available at 2y; 62% had data available at 4y; 59% had data available at 8y

Table F-4. Risk of bias —measurement of the outcome

Main Study Author (Year); Follow-up Studies Author (Year)	Were outcome assessors aware of the intervention received by study participants?	Was the assessment of the outcome likely to be influenced by knowledge of intervention received?	Bias arising from measurement of the outcome?	Comments
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48]	Probably no	NA	Low	Participants and observers were blinded to allocated treatment during the follow-up period.
Brouwer (2015)[39] Brouwer (2017)[102]	Yes	Probably yes	Some concerns	Many patient-reported outcomes used as appropriate and not feasible to blind participants to study intervention, though most participants had no stated preference for treatment suggesting that risk of bias is probably not high.
Chatterjee (1995)[38]	Probably no	NA	Some concerns	Not enough information to judge were outcome assessors were truly blinded or if patient contributed towards outcomes.
Erginousakis (2011)[37]	Probably yes	Probably yes	Some concerns	Uses some patient self-reported outcomes, and intervention was not blinded to participants.
Franke (2009)[36]	No information	Probably yes	Some concerns	Since most outcomes are patient-reported and not clear that allocation was blinded, there is some concern for bias, though probably only small given the comparison in this study is between two surgical treatments.
Gerszten (2010)[41]	Yes	Probably yes	Some concerns	Study uses patient-reported outcomes, thus some risk of bias is possible.
Haines (2002)[42]	No information	Probably yes	Some concerns	Assessment matrix included both clinician and patient-reported outcomes; neither were blinded.
Henriksen (1996)[35]	Probably yes	Probably yes	Some concerns	Use of patient self-reported outcomes, participants weren't blinded.

(continued)

Table F-4. Risk of bias —measurement of the outcome (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Were outcome assessors aware of the intervention received by study participants?	Was the assessment of the outcome likely to be influenced by knowledge of intervention received?	Bias arising from measurement of the outcome?	Comments
Hermantin (1999)[43]	Yes	Probably yes	Some concerns	Study uses patient-reported outcomes, so some concern for bias as may be influenced by knowledge of treatment allocation.
Huang (2005)[24]	No information	Probably yes	Some concerns	Study used patient-reported outcomes, which may be influenced by knowledge of treatment allocation.
Mayer (1993)[34]	Probably yes	Probably yes	Some concerns	Use of patient self-reported outcomes, participants weren't blinded.
McMorland (2010)[23]	No information	Probably yes	Some concerns	Study used patient-reported outcomes, which have some concern for bias given that participants were not blinded to the treatment allocation.
North (2005)[46]	Yes	Probably yes	Some concerns	Uses some patient self-reported outcomes, and treatment allocation was not blinded.
Osterman (2003)[33]	Yes	Probably yes	Some concerns	Study used patient-reported outcomes, and there is some concern that these can be influenced by knowledge of treatment allocation. Study evaluated patient and surgeon's expectations of improvement immediately after randomization and showed higher expectations.
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98]	Yes	Probably yes	Some concerns	Study used patient-reported outcomes, which are likely to be influenced by knowledge of treatment allocation since participants were not blinded.
Ruetten (2008)[31]	Probably yes	Probably yes	Some concerns	Most outcome measures were patient self-report; participants were not blind to treatment allocation.

(continued)

Table F-4. Risk of bias —measurement of the outcome (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Were outcome assessors aware of the intervention received by study participants?	Was the assessment of the outcome likely to be influenced by knowledge of intervention received?	Bias arising from measurement of the outcome?	Comments
Ruetten (2009)[47]	Probably yes	Probably yes	Some concerns	Most outcome measures were patient self-report; participants were not blind to treatment allocation.
Ryang (2008)[30] Gempt (2013)[103]	No information	Probably yes	Some concerns	Use of patient-reported outcomes among participants,who were likely not blinded to treatment allocation.
Sasaoka (2006)[25]	No information	No information	Some concerns	Use of patient-reported outcomes by participants who were likely not blinded to treatment allocation.
Teli (2010)[29]	Yes	Probably yes	Some concerns	Patient self-reported outcomes; patients were not blind to treatment allocation.
Thome (2005)[28] Barth (2008)[104]	No information	Probably yes	Some concerns	Use of patient-reported outcomes among participants who were likely not blinded to treatment allocation.
Tullberg (1993)[27]	No information	No information	Some concerns	Use of patient-reported outcomes by participants who were likely not blinded to treatment allocation.
Weber (1983)[26]	Yes	Probably yes	Some concerns	Patient-reported outcomes were used among participants who were not blinded.
Weinstein (2006) [22] Weinstein (2008)[99] Lurie Jon (2014)[100]	No information	No information	Some concerns	Study used patient-reported outcomes, which are likely to be influenced by knowledge of treatment allocation as participants were not blinded.

Table F-5. Risk of bias —selection of the reported result

Main Study Author (Year); Follow-up Studies Author (Year)	Are the reported outcome data likely to have been selected on the basis of results from multiple outcome measurements within the outcome domain?	Are the reported outcome data likely to have been selected on the basis of results from multiple analyses of the data?	Bias arising from selection of reported results?	Comments
Arts (2009)[40] Arts (2011)[101] Overdevest (2017)[48]	No at follow-up points up through 2 years, probably yes at later follow-up points	No	Some concerns	The publication reporting 5 year outcomes does not report SF-36, Sciatica Index, or Prolo Scale, all outcomes that had been reported at earlier timepoints.
Brouwer (2015)[39] Brouwer (2017)[102]	No	No	Low	None
Chatterjee (1995)[38]	No	No	Low	None
Erginousakis (2011)[37]	Probably no	Probably no	Low	None
Franke (2009)[36]	No	No	Low	None
Gerszten (2010)[41]	No	No	Low	None
Haines (2002)[42]	No	No	Low	None
Henriksen (1996)[35]	No	No	Low	None
Hermantin (1999)[43]	No	No	Low	None
Huang (2005)[24]	No	No	Low	None
Mayer (1993)[34]	No	No	Low	None
McMorland (2010)[23]	No	No	Low	None
North (2005)[46]	Probably no	Probably no	Low	None
Osterman (2003)[33]	No	No	Low	None
Peul (2007)[32] Peul (2008)[97] Lequin (2013)[98]	No	No	Low	None
Ruetten (2008)[31]	Probably no	Probably no	Low	None
Ruetten (2009)[47]	Probably no	Probably no	Low	None
Ryang (2008)[30] Gempt (2013)[103]	No	No	Low	None
Sasaoka (2006)[25]	No	No	Low	None
Teli (2010)[29]	Probably no	Probably no	Low	None
Thome (2005)[28] Barth (2008)[104]	No	No	Low	None
Tullberg (1993)[27]	Probably no	Probably no	Low	None
Weber (1983)[26]	Probably no	Probably no	Low	None

(continued)

Table F-5. Risk of bias —selection of the reported result (continued)

Main Study Author (Year); Follow-up Studies Author (Year)	Are the reported outcome data likely to have been selected on the basis of results from multiple outcome measurements within the outcome domain?	Are the reported outcome data likely to have been selected on the basis of results from multiple analyses of the data?	Bias arising from selection of reported results?	Comments
Weinstein (2006) [22] Weinstein (2008)[99] Lurie Jon (2014)[100]	Probably yes	No	Some concerns	Trials registry indicates the SF-36 was the primary outcome, but study results only report Bodily Pain and physical function subscales. The trial registry does not list the Oswestry Disability Index as a primary or secondary outcome, yet the study reports it as a primary outcome. The trial registry does not list the Sciatica Index as an outcome, yet the study lists it as a secondary outcome. Authors also conducted an 'as treated' analysis in addition to the intent-to-treat analysis, but this risk of bias assessment is only focused on the intent-to-treat analysis.

Table F-6. Quality of health economic studies —part I

Cost Study Author (Year); Main Study Author (Year)	Overall Quality Rating (Score[a])	Was the study objective presented in a clear, specific, and measurable manner?	Were the perspective of the analysis (societal, third-party payer, and so on) and reasons for its selection stated?	Were variable estimates used in the analysis from the best available source (i.e., Randomized Control Trial-Best, Expert Opinion-Worst)?	If estimates came from a subgroup analysis, were the groups pre-specified at the beginning of the study?	Was uncertainty handled by: (i) statistical analysis to address random events; (ii) sensitivity analysis to cover a range of assumptions?
Malter (1996)[44]	Fair (79)	Yes	Yes	No	NA	Yes
Stevenson (1995)[53] Chatterjee (1995)[38]	Poor (53)	Yes	No	Yes	NA	No
Teli (2010)[29] (main study includes cost)	NA					
Tosteson (2008)[50] Weinstein (2006)[22]	Good (94)	Yes	Yes	Yes	NA	Yes
Van den Akker (2011)[51] Arts (2009)[40]	Good (94)	Yes	Yes	Yes	NA	Yes
Van den Akker (2017)[52] Brouwer (2015)[39]	Good (94)	Yes	Yes	Yes	NA	Yes
Van den Hout (2008)[49] Peul (2007)[32]	Good (94)	Yes	Yes	Yes	NA	Yes

[a] Based on scale of 0 (worst quality) to 100 (best quality).

Table F-7 Quality of health economic studies —part 2

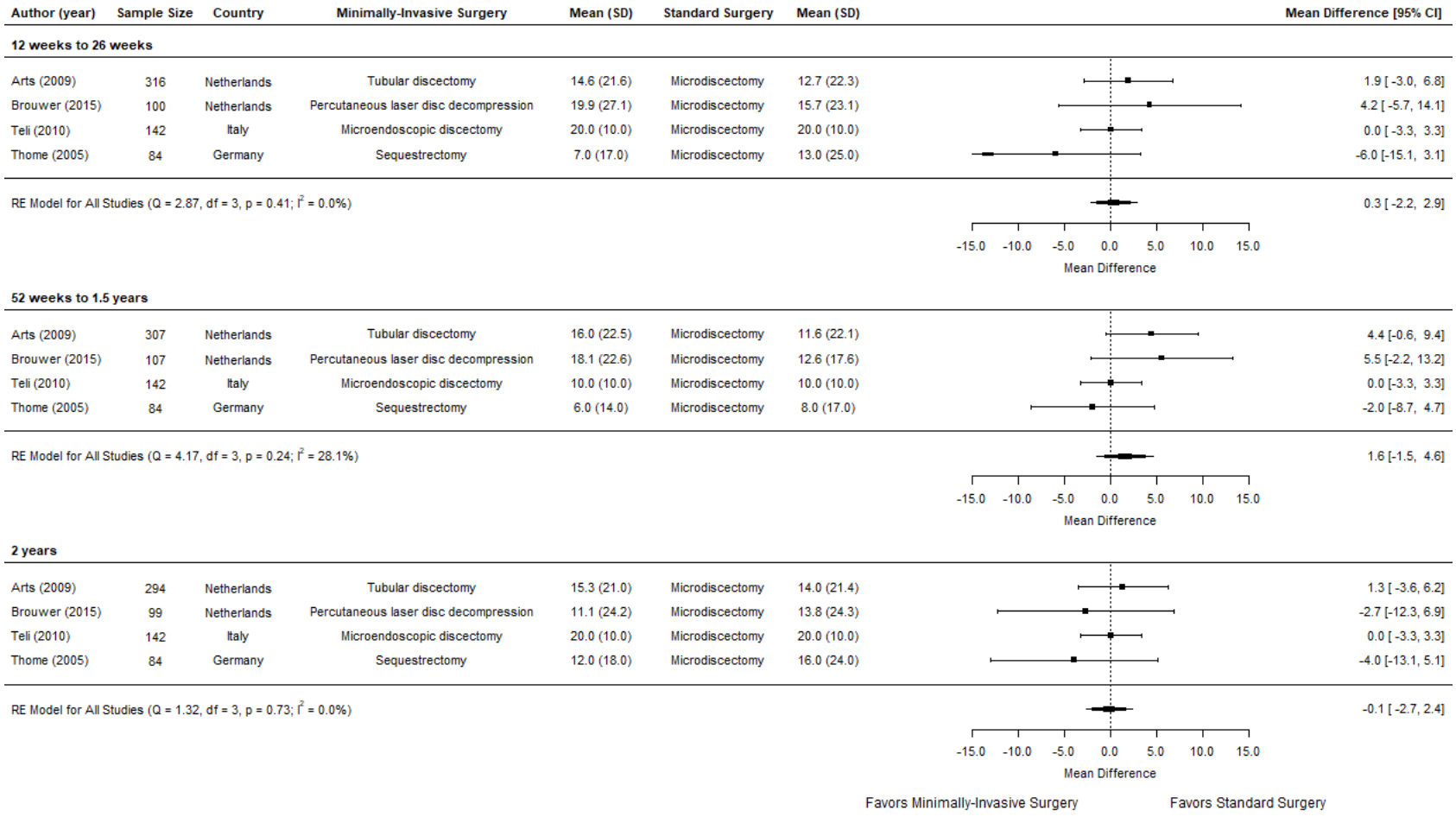
Cost Study Author (Year); Main Study Author (Year)	Was incremental analysis performed between alternatives for resources and costs?	Was the methodology for data abstraction (including value health states and other benefits) stated?	Did the analytic horizon allow time for all relevant and important outcomes? Were benefits and costs that went beyond 1 year discounted (3–5%) and justification given for the discount rate?	Was the measurement of costs appropriate and the methodology for the estimation of quantities and unit costs clearly described?	Was the primary outcome measure(s) for the economic evaluation clearly stated and were the major short-term, long-term and negative outcomes included?
Malter (1996)[44]	Yes	Yes	Yes	Yes	Yes
Stevenson (1995)[53] Chatterjee (1995)[38]	Yes	Yes	Yes	Yes	No
Teli (2010)[29] (main study includes cost)					
Tosteson (2008)[50] Weinstein (2006)[22]	Yes	Yes	Yes	Yes	Yes
Van den Akker (2011)[51] Arts (2009)[40]	Yes	Yes	Yes	Yes	Yes
Van den Akker (2017)[52] Brouwer (2015)[39]	Yes	Yes	Yes	Yes	Yes
Van den Hout (2008)[49] Peul (2007)[32]	Yes	Yes	Yes	Yes	Yes

Table F-8 Quality of health economic studies —part 3

Cost Study Author (Year); Main Study Author (Year)	Were the health outcomes measures/scales valid and reliable? If previously tested valid and reliable measures were not available, was justification given for the measures/scales used?	Were the economic model (including structure), study methods and analysis, and the components of the numerator and denominator displayed in a clear transparent manner?	Were the choice of economic model, main assumptions and limitations of the study stated and justified?	Did the author(s) explicitly discuss direction and magnitude of potential biases?	Were the conclusions/recommendations of the study justified and based on the study results?	Was there a statement disclosing the source of funding for the study?
Malter (1996)[44]	No	Yes	Yes	No	Yes	Yes
Stevenson (1995)[53] Chatterjee (1995)[38]	Can't determine	Can't determine	No	No	Yes	Yes
Teli (2010)[29] (main study includes cost)						
Tosteson (2008)[50] Weinstein (2006)[22]	Yes	Yes	Yes	No	Yes	Yes
Van den Akker (2011)[51] Arts (2009)[40]	Yes	Yes	Yes	No	Yes	Yes
Van den Akker (2017)[52] Brouwer (2015)[39]	Yes	Yes	Yes	No	Yes	Yes
Van den Hout (2008)[49] Peul (2007)[32]	Yes	Yes	Yes	No	Yes	Yes

Appendix G. Meta-analyses

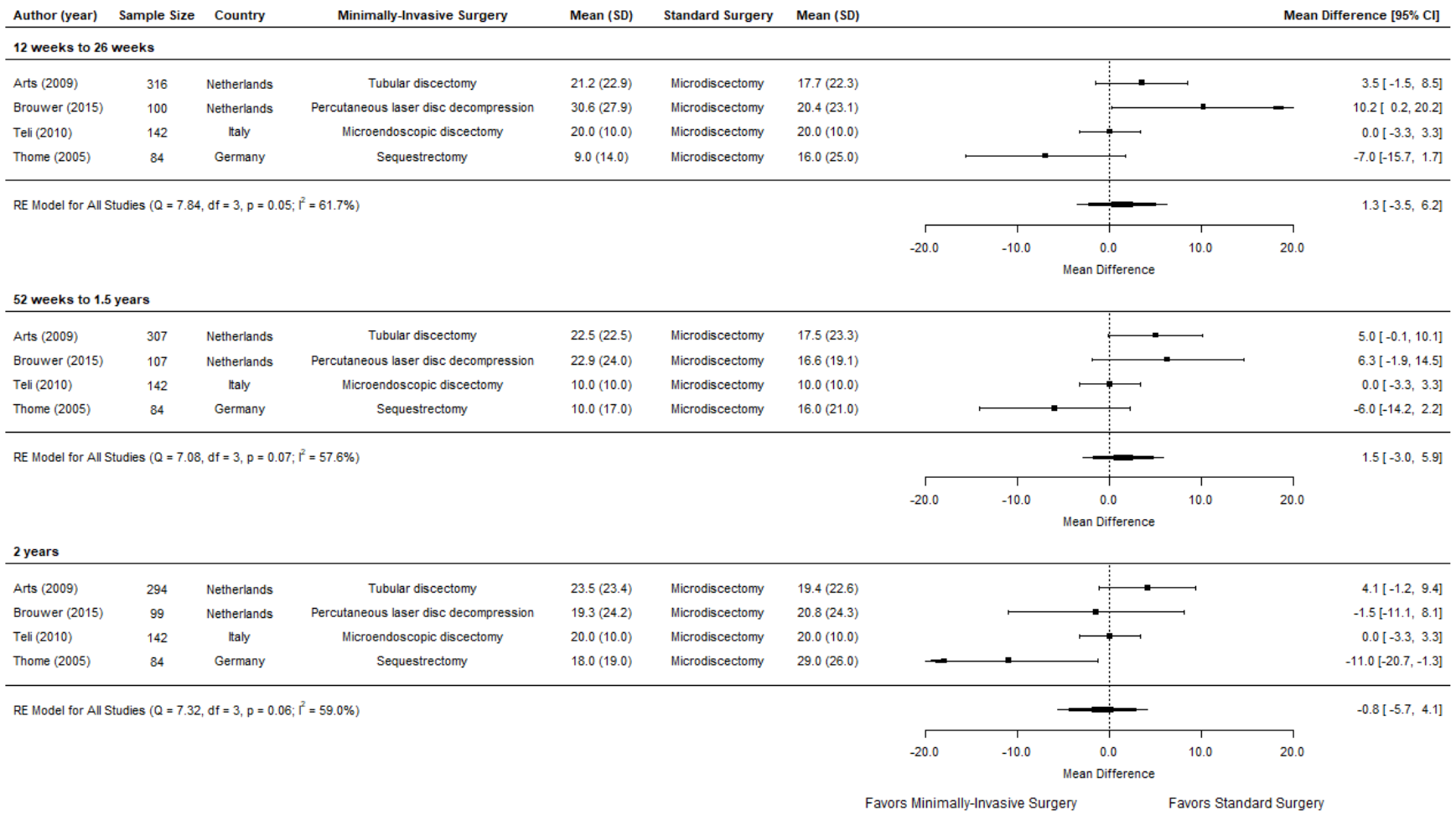
Figure G-1. Between-group differences in visual analog scale for leg pain in randomized controlled trials comparing minimally- invasive surgery with standard surgery at 12 weeks to 26 weeks, 52 weeks to 1.5 years, and 2 years.



Note: This figure only depicts the microendoscopic discectomy compared with microdiscectomy comparison reported by Teli et al.[29] The VAS 10 cm scores reported by Thome et al.[28] were converted to 100 mm for this analysis. Mean follow-up scores for Arts et al.[40] and Brouwer et al.[39] are unadjusted for baseline, so the AMDs reported here may differ from the AMDs reported by the study publications.

Abbreviations: VAS = visual analog scale; SD = standard deviation; df = degrees of freedom; AMD = absolute mean difference; RE = random effects; CI = confidence interval.

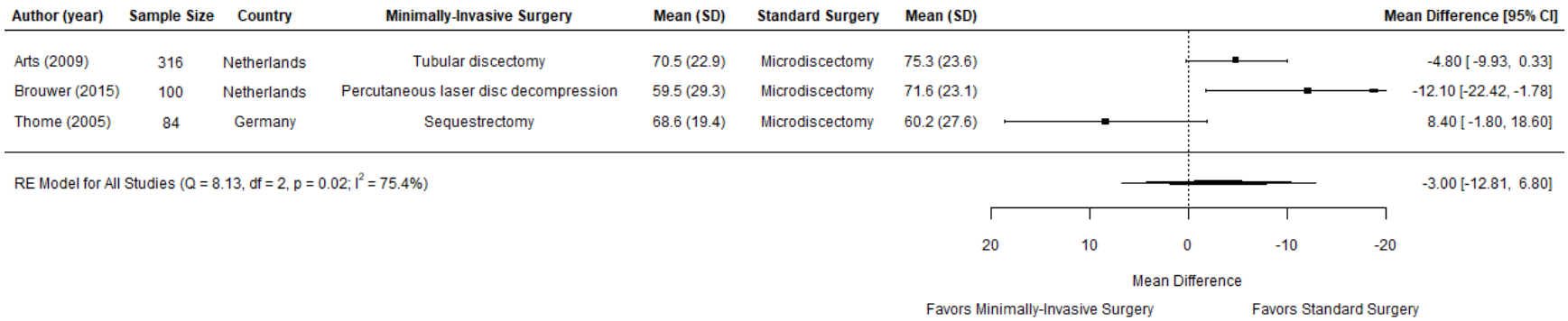
Figure G-2. Between-group differences in visual analog scale for back pain in randomized controlled trials comparing minimally-invasive surgery with standard surgery at 12 weeks to 26 weeks, 52 weeks to 1.5 years, and 2 years.



Note: This figure only depicts the microendoscopic discectomy compared with microdiscectomy comparison reported by Teli et al.[29] The VAS 10 cm scores reported by Thome et al.[28] were converted to 100 mm for this analysis. Mean follow-up scores for Arts et al.[40] and Brouwer et al.[39] are unadjusted for baseline, so the AMDs reported here may differ from the AMDs reported by the study publications.

Abbreviations: VAS = visual analog scale; SD = standard deviation; df = degrees of freedom; AMD = absolute mean difference; RE = random effects; CI = confidence interval.

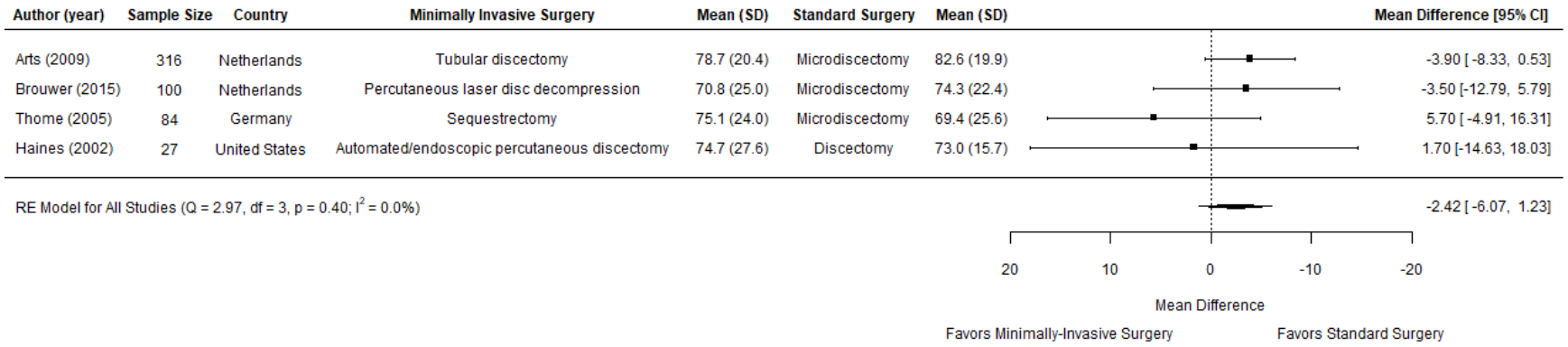
Figure G-3. Between-group differences in SF-36 Bodily Pain subscale in randomized controlled trials comparing minimally-invasive surgery with standard surgery at 12 weeks to 26 weeks.



Note: Mean follow-up scores for all studies are unadjusted for baseline, so the AMDs reported here may differ from the AMDs reported by the study publications.

Abbreviations: N = number; SF-36 = Short Form 36; SD = standard deviation; df = degrees of freedom; AMD = absolute mean difference; RE = random effects; CI = confidence interval.

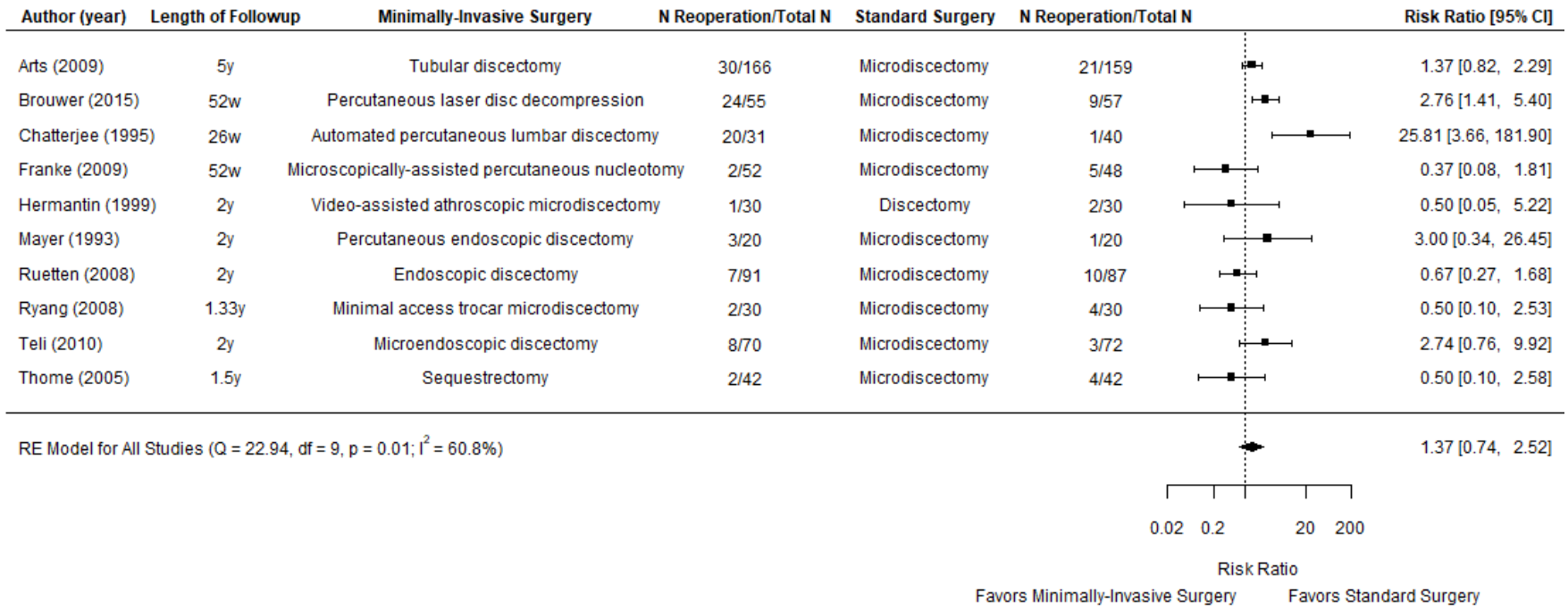
Figure G-4. Between-group differences in SF-36 Physical Functioning subscale in randomized controlled trials comparing minimally-invasive surgery with standard surgery at 12 weeks to 26 weeks.



Note: Mean follow-up scores for all studies are unadjusted for baseline, so the AMDs reported here may differ from the AMDs reported by the study publications.

Abbreviations: N = number; SF-36 = Short Form 36; SD = standard deviation; df = degrees of freedom; AMD = absolute mean difference; RE = random effects; CI = confidence interval.

Figure G-5. Relative risk for reoperation in randomized controlled trials comparing minimally-invasive surgery with standard surgery.

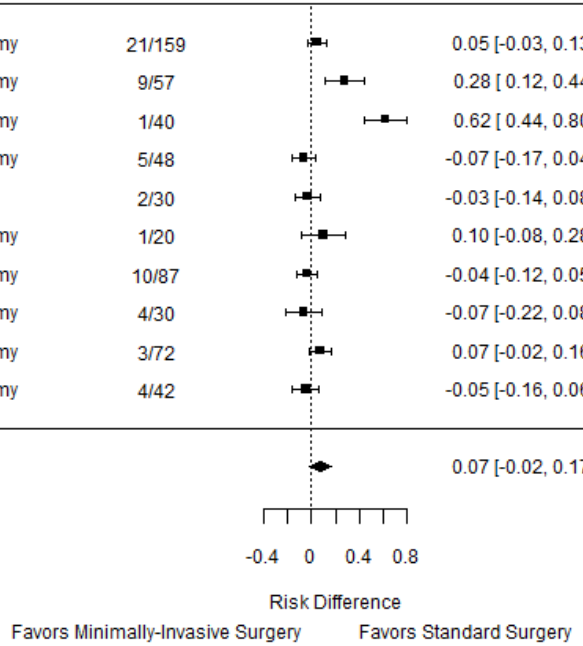


Note: Because of the unique circumstances in the Chatterjee et al [38] study as discussed in the Full Report, we excluded it from the pooled estimate in a sensitivity analysis. The pooled RR without it was 1.17 (95% CI, 0.70 to 1.97; 9 RCTs, 1,101 participants; I^2= 44.4%).

Abbreviations: N = number; RCT = randomized controlled trial; df = degrees of freedom; CI = confidence interval; y = year(s); w = week(s); RE = random effects; RR = relative risk ratio.

Figure G-6. Absolute risk difference for reoperation in randomized controlled trials comparing minimally-invasive surgery with standard surgery.

Author (year)	Length of Followup	Minimally-Invasive Surgery	N Reoperation/Total N	Standard Surgery	N Reoperation/Total N	Risk Difference [95% CI]
Arts (2009)	5y	Tubular discectomy	30/166	Microdiscectomy	21/159	0.05 [-0.03, 0.13]
Brouwer (2015)	52w	Percutaneous laser disc decompression	24/55	Microdiscectomy	9/57	0.28 [0.12, 0.44]
Chatterjee (1995)	26w	Automated percutaneous lumbar discectomy	20/31	Microdiscectomy	1/40	0.62 [0.44, 0.80]
Franke (2009)	52w	Microscopically-assisted percutaneous nucleotomy	2/52	Microdiscectomy	5/48	-0.07 [-0.17, 0.04]
Hermantin (1999)	2y	Video-assisted athroscopic microdiscectomy	1/30	Discectomy	2/30	-0.03 [-0.14, 0.08]
Mayer (1993)	2y	Percutaneous endoscopic discectomy	3/20	Microdiscectomy	1/20	0.10 [-0.08, 0.28]
Ruetten (2008)	2y	Endoscopic discectomy	7/91	Microdiscectomy	10/87	-0.04 [-0.12, 0.05]
Ryang (2008)	1.33y	Minimal access trocar microdiscectomy	2/30	Microdiscectomy	4/30	-0.07 [-0.22, 0.08]
Teli (2010)	2y	Microendoscopic discectomy	8/70	Microdiscectomy	3/72	0.07 [-0.02, 0.16]
Thome (2005)	1.5y	Sequestrectomy	2/42	Microdiscectomy	4/42	-0.05 [-0.16, 0.06]
RE Model for All Studies (Q = 64.81, df = 9, p = 0.00; I^2 = 86.1%)						0.07 [-0.02, 0.17]



Note: The risk difference is expressed in this plot as a proportion; multiply by 100 to express this risk difference as a percentage (e.g., 0.07 is a 7% difference in absolute risk). Because of the unique circumstances in the Chatterjee et al [38] study as discussed in the Full Report, we excluded it from the pooled estimate in a sensitivity analysis. The pooled ARD without it was 2% (95% CI, -4% to 8%; 9 RCTs, 1,101 participants; I^2= 60.7%).

Abbreviations: N = number; RCT = randomized controlled trial; df = degrees of freedom; CI = confidence interval; y = year(s); w = week(s); RE = random effects; ARD = absolute risk difference.