Health Quality Ontario

Let's make our health system healthier

ONTARIO HEALTH TECHNOLOGY ASSESSMENT SERIES

Cervical Artificial Disc Replacement Versus Fusion for Cervical Degenerative Disc Disease: A Health Technology Assessment

KEY MESSAGES

What Is This Health Technology Assessment About?

Cervical degenerative disc disease occurs in the cervical spine (the part of the spine in the neck) when the discs between the vertebrae (the bones of the spine) start to deteriorate. It causes painful and disabling symptoms that impact people's quality of life and ability to function.

When treatments such as medication and physical therapy are insufficient, surgery is an option. The most common surgery is anterior cervical discectomy and fusion (often simply called "fusion"). However, this surgery sometimes has a negative effect on the discs next to the one being treated. Another surgical option is cervical artificial disc replacement (C-ADR).

This health technology assessment looked at the effectiveness, safety, durability, and cost-effectiveness of C-ADR compared with fusion for treating cervical degenerative disc disease. We also looked at the budget impact of publicly funding C-ADR and the preferences, values, and experiences of people with cervical degenerative disc disease.

What Did This Health Technology Assessment Find?

C-ADR and fusion are relatively safe, and both decrease pain and improve symptom-related disability and health-related quality of life. Clinical trials show that C-ADR is an effective and safe alternative to fusion. Unlike fusion, C-ADR also allows the neck to move more normally and likely results in better outcomes in terms of recovery, return to work, technical failures, and need for re-operation at the original surgery site. Although further surgeries for degeneration at other spinal levels might be needed later for people having either type of surgery, we don't yet know if the need for additional surgeries differs between C-ADR and fusion.

C-ADR appears to be cost-effective for both one-level and two-level cervical disc degeneration. In Ontario, publicly funding C-ADR could result in extra costs of about $900,000 for one-level procedures and about $700,000 for two-level procedures over the next 5 years.

People who had undergone C-ADR reported positively on its effect on their symptoms, their quality of life, and their ability to move their neck following surgery. Limited access to C-ADR in Ontario was viewed as a barrier to receiving this treatment.

Published February 2019
Volume 19, Number 3

Ontario
Health Quality Ontario

HEALTH TECHNOLOGY ASSESSMENT AT HEALTH QUALITY ONTARIO

This report was developed by a multidisciplinary team from Health Quality Ontario. The clinical epidemiologist was Gaylene Pron; the health economist was Hossein Zivaripiran; the Patient, Caregiver, and Public Engagement analyst was David Wells; and the medical librarian was Melissa Walter.

The medical editor was Elizabeth Jean Betsch. Others involved in the development and production of this report were Paul Kolodziej, Kellee Kaulback, Ana Laing, Kathryn Schwarz, Saleemeh Abdolzahraei, Jeanne McKane, Timothy Maguire, Kara Cowan, Claude Soulodre, Sarah McDowell, Vivian Ng, Andrée Mitchell, Nancy Sikich, and Irfan Dhalla.

We are grateful to the following experts for generously giving their time to provide input into the development of this report: Drew A. Bednar (orthopedic surgeon, Department of Orthopedic Surgery, McMaster University), Simon Harris (orthopedic surgeon, Department of Orthopedic Surgery, The Scarborough and Rouge General Hospital), Sean Christie (neurosurgeon, QEII Health Sciences Centre, Dalhousie University), and Neil Duggal (neurosurgeon, Department of Clinical Neurosciences, Division of Neurosurgery, Western University).

We thank the people with cervical degenerative disc disease and the family members who generously gave their time to share their experiences of cervical degenerative disc disease and treatment options with us.

The statements, conclusions, and views expressed in this report do not necessarily represent the views of those we consulted.

Citation

Health Quality Ontario. Cervical artificial disc replacement versus fusion for cervical degenerative disc disease: a health technology assessment. Ont Health Technol Assess Ser [Internet]. 2019 Feb;19(3): 1–223. Available from: http://www.hqontario.ca/evidence-to-improve-care/journal-ontario-health-technology-assessment-series

ABSTRACT

Background

Cervical degenerative disc disease is a multifactorial condition that begins with deterioration of the intervertebral disc and results in further degeneration within the spine involving the facet joints and ligaments. This health technology assessment examined the effectiveness, safety, durability, and cost-effectiveness of cervical artificial disc replacement (C-ADR) versus fusion for treating cervical degenerative disc disease.

Methods

We performed a systematic literature search of the clinical evidence comparing C-ADR with fusion. We assessed the risk of bias in each study and the quality of the body of evidence according to the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) Working Group criteria. We performed a systematic review of the economic literature and assessed the cost-effectiveness of C-ADR compared with fusion. We also estimated the budget impact of publicly funding C-ADR in Ontario over the next 5 years. To contextualize the potential value of C-ADR, we spoke with people with cervical degenerative disc disease.

Results

Eight studies of C-ADR for one-level cervical degenerative disc disease and two studies of C-ADR for two-level disease satisfied the criterion of statistical noninferiority compared with fusion on the primary outcome of 2-year overall treatment success (GRADE: Moderate). In two studies of C-ADR for two-level disease, C-ADR was statistically superior to fusion surgery for the same primary outcome (GRADE: Moderate). C-ADR was also noninferior to fusion for perioperative outcomes (e.g., operative time, blood loss), patient satisfaction, and health-related quality of life (GRADE: Moderate). C-ADR was superior to fusion for recovery and return to work, had higher technical success, and had lower rates of re-operation at the index site (GRADE: Moderate). C-ADR also maintained motion at the index-treated cervical level (GRADE: Moderate), but evidence was insufficient to determine if adjacent-level surgery rates differed between C-ADR and fusion. Current evidence is also insufficient to determine the long-term durability of C-ADR.

The primary economic analysis shows that C-ADR is likely to be cost-effective compared with fusion for both one-level ($11,607/quality-adjusted life-year [QALY]) and two-level ($16,782/QALY) degeneration. Various sensitivity and scenario analyses confirm the robustness of the results. The current uptake for one-level and two-level C-ADR in Ontario is about 8% of the total eligible. For one-level involvement, the estimated net budget impact increases from $7,243 (18 procedures) in the first year to $395,623 (196 procedures) in the fifth year following public funding, for a total budget impact over 5 years of $916,326. For two-level involvement, the corresponding values are $5,460 (7 procedures) in the first year and $283,689 (76 procedures) in the fifth year, for an estimated total budget impact of $705,628 over 5 years.

People with cervical degenerative disc disease reported that symptoms of pain and numbness can have a negative impact on their quality of life. People with whom we spoke had tried a variety of treatments with minor success; surgery was perceived as the most effective and permanent solution. Those who had undergone C-ADR spoke positively of its impact on their quality of life and ability to move their neck after surgery. The limited availability of C-ADR in Ontario was viewed as a barrier to receiving this treatment.

Conclusions

For carefully selected patients with cervical degenerative disc disease, C-ADR provides patient-important and statistically significant reductions in pain and disability. Further, unlike fusion, C-ADR allows people to maintain relatively normal cervical spine motion.

Compared with fusion, C-ADR appears to represent good value for money for adults with one-level cervical degenerative disc disease ($11,607/QALY) and for adults with two-level disease ($16,782/QALY). In Ontario, publicly funding C-ADR could result in total additional costs of $916,326 for one-level procedures and $705,628 for two-level procedures over the next 5 years.

People with whom we spoke who had undergone C-ADR surgery spoke positively of its impact on their quality of life and ability to move their neck after surgery. The limited availability of C-ADR in Ontario was viewed as a barrier to receiving this treatment.

TABLE OF CONTENTS

LIST OF TABLES **8**
LIST OF FIGURES **10**
OBJECTIVE **11**
BACKGROUND **11**
Clinical Need and Target Population 11
Current Surgical Treatment *12*
Health Technology Under Review *13*
Regulatory Information 18
Ontario Context 18
CLINICAL EVIDENCE **19**
Research Question 19
Methods 19
Clinical Literature Search *19*
Literature Screening *19*
Inclusion Criteria *19*
Exclusion Criteria *19*
Outcomes of Interest *20*
Data Extraction *20*
Statistical Analysis *20*
Critical Appraisal of Evidence *21*
Quality of Evidence *21*
Expert Consultation *21*
Results 21
Literature Search *21*
Included Studies *22*
Short-Term Treatment Outcomes *33*
Longer-Term Treatment Outcomes *73*
Discussion 96
C-ADR Advantages *97*
Limitations *98*
Conclusions 98
ECONOMIC EVIDENCE **99**
Research Questions 99
Methods 99
Economic Literature Search *99*
Literature Screening *99*
Inclusion Criteria *99*
Exclusion Criteria *99*
Outcomes of Interest *99*
Data Extraction *100*
Study Applicability and Limitations *100*

Results……100
Literature Search……*100*
Review of Included Economic Studies……*101*
Applicability of Included Studies……*111*
Discussion……111
Conclusions……112
PRIMARY ECONOMIC EVALUATION……113
Research Questions……113
Methods……113
Type of Analysis……*113*
Target Population……*113*
Perspective……*114*
Interventions……*114*
Discounting and Time Horizon……*114*
Main Assumptions……*114*
Model Structure and Structure of Analysis……*115*
Clinical Outcome and Utility Parameters……*116*
Cost Parameters……*123*
Analysis……*125*
Generalizability……*127*
Expert Consultation……*127*
Results……127
Reference Case Analysis……*127*
Sensitivity Analysis……*127*
Discussion……134
Conclusions……135
BUDGET IMPACT ANALYSIS……136
Research Questions……136
Methods……136
Analytic Framework……*136*
Key Assumptions……*137*
Target Population……*137*
Current Intervention Mix……*138*
Uptake of the New Intervention and Future Intervention Mix……*140*
Resources and Costs……*142*
Analysis……*142*
Results……142
One-Level C-ADR Budget Impact……*142*
Two-Level C-ADR Budget Impact……*143*
Discussion……143
Conclusions……144
PATIENT PREFERENCES AND VALUES……145

Objective 145
Background 145
Methods 145
Engagement Plan *145*
Participant Outreach *146*
Approach *146*
Data Extraction and Analysis *147*
Results 147
Lived Experience of Cervical Degenerative Disc Disease *147*
Treatments for Degenerative Disc Disease *149*
Fusion Versus Cervical Artificial Disc Replacement *150*
Discussion 153
Conclusions 154
CONCLUSIONS OF THE HEALTH TECHNOLOGY ASSESSMENT 155
ABBREVIATIONS 156
GLOSSARY 157
APPENDICES 159
Appendix 1: Literature Search Strategies 159
Clinical Evidence Search *159*
Economic Evidence Search *161*
Grey Literature Search *163*
Appendix 2: Randomized Controlled Trials Comparing C-ADR to Fusion Surgery 164
Appendix 3: Secondary Surgeries at Index and Adjacent Cervical Levels in Longer-Term Follow-Up 173
Appendix 4: Risk of Bias and GRADE Tables 181
Appendix 5: Selected Excluded Studies 186
Appendix 6: Results of Applicability Checklists for Studies Included in Economic Literature Review 192
Appendix 7: Artificial Disc Devices Considered in Economic Evaluation 194
Appendix 8: Extracted Natural History Parameters 195
Appendix 9: Cost Details 197
Appendix 10: Primary Economic Evaluation, Scenario Analysis 204
Appendix 11: Budget Impact, Scenario Analysis 205
Appendix 12: Letter of Information 208
Appendix 13: Interview Guide 209
REFERENCES 210

LIST OF TABLES

Table 1: Characteristics of Cervical Artificial Disc Replacement Systems 15
Table 2: Randomized Controlled Trials of Cervical Artificial Disc Replacement Versus Fusion for Cervical Degenerative Disc Disease.....23
Table 3: FDA Regulatory Trials of C-ADR Versus Fusion for Cervical Degenerative Disc Disease 24
Table 4: Statistical Noninferiority and Superiority Status of Two-Year Overall Treatment Success of C-ADR Versus Fusion for Cervical Degenerative Disc Disease 34
Table 5: Second Index-Level Surgeries for Technical Failures at Two Years 38
Table 6: Cervical Range of Motion at Two Years 40
Table 7: Fusion Radiologic Success at Two Years.....41
Table 8: Adverse Events—C-ADR Versus Fusion for Cervical Degenerative Disc Disease.....43
Table 9: Dysphagia Following C-ADR or Fusion for Cervical Degenerative Disc Disease 45
Table 10: Neck Disability Index Scores—C-ADR Versus Fusion in Nonregulatory Randomized Controlled Trials 48
Table 11: Perioperative Outcomes of C-ADR Versus Fusion for Cervical Degenerative Disc Disease 50
Table 12: Return-to-Work Status After C-ADR or Fusion for Cervical Degenerative Disc Disease 53
Table 13: Health-Related Quality of Life With C-ADR Versus Fusion for Cervical Degenerative Disc Disease 58
Table 14: Satisfaction With C-ADR Versus Fusion for Cervical Degenerative Disc Disease.....67
Table 15: Cervical Motion at Treated and Adjacent Sites After C-ADR or Fusion.....75
Table 16: Heterotopic Ossification at C-ADR-Treated Index Level on Follow-Up 81
Table 17: Fusion Failure at Index Level on Follow-Up.....82
Table 18: Adjacent-Segment Disease at Longer-Term Follow-Up of C-ADR.....85
Table 19: Secondary Surgeries at Longer-Term Follow-Up of C-ADR Versus Fusion 88
Table 20: Longer-Term Safety of C-ADR Versus Fusion for Cervical Degenerative Disc Disease 91
Table 21: Economic Literature Review for One-Level Symptomatic Cervical Degenerative Disc Disease 103
Table 22: Economic Literature Review for Two-Level Symptomatic Cervical Degenerative Disc Disease 110
Table 23: Disease Interventions and Comparators Evaluated in the Primary Economic Model 114
Table 24: Probability (Cumulative Incidence) of Adverse Events in 2-Year Follow-Up.....118
Table 25: Probability (Cumulative Incidence) of Adverse Events in 7-Year Follow-Up.....119
Table 26: Probabilities (Cumulative Incidence) of Index-Level Corrective Surgeries and Adjacent-Segment Surgeries in 2-Year Follow-Up.....120
Table 27: Probabilities (Cumulative Incidence) of Index-Level Corrective Surgeries and Adjacent-Segment Surgeries in 7-Year Follow-Up.....120
Table 28: Utilities (Health-Related Quality of Life) for Artificial Disc and Fusion Surgeries at 5- or 7-Year Follow-Up 122
Table 29: Multiplicative Factor to Calculate Disutilities for Patients Experiencing Adverse Events 123
Table 30: Unit Prices and Sales for Artificial Discs 124
Table 31: Price Details for Instruments Used in Fusion Surgery in Ontario 124
Table 32: 2018 Surgeon, Anesthetist, and Surgical Assistant Fees for C-ADR Versus Fusion124
Table 33: Approximation of Hospital-Only Costs for C-ADR and Fusion in Ontario 125
Table 34: Parameters Varied in One-Way and Probabilistic Sensitivity Analysis 126

Table 35: Reference Case Analysis for One-Level and Two-Level Surgeries[a] 127
Table 36: Scenarios With Longer Follow-Up Times and for Various C-ADR Devices 134
Table 37: Yearly Number of C-ADR and Fusion Surgeries in Ontario 138
Table 38: Predicted Yearly Number of Total Surgeries Eligible for Fusion in Ontario 138
Table 39: Number and Relative Distribution of Fusion Surgeries Performed in Alberta, 2004–2007 139
Table 40: Predicted Yearly Number of Patients Eligible for One-Level and Two-Level Fusion in Ontario 139
Table 41: Anticipated Yearly Percentage of Patients Receiving C-ADR and Fusion in Ontario Among Patients Eligible for Both Surgeries, 2018–2022.....140
Table 42: Predicted Yearly Number of Patients Receiving One-Level and Two-Level C-ADR and Fusion Surgery in Ontario Among Patients Eligible for Both Surgeries 141
Table 43: Undiscounted Yearly Costs of One-Level and Two-Level C-ADR and Fusion Surgeries in Ontario.....142
Table 44: Total and Net Budget Impact for Adoption of C-ADR Versus Fusion for One- and Two-Level Cervical Degenerative Disc Disease 143
Table A1: Cervical Implant Devices Used in Randomized Controlled Trials of C-ADR Versus Fusion for Degenerative Disc Disease.....164
Table A2: Secondary Surgeries in Longer-Term Follow-Up of C-ADR– Versus Fusion–Treated One-Level Cervical Degenerative Disc Disease 173
Table A3: Secondary Surgery in Longer-Term Follow-Up of C-ADR– Versus Fusion–Treated Two-Level Degenerative Disc Disease 179
Table A4: Risk of Bias of C-ADR Versus Fusion RCTs 181
Table A5: GRADE Evidence Profile for Anterior Cervical Discectomy and Artificial Cervical Disc Replacement Versus Fusion for Cervical Degenerative Disc Disease 183
Table A6: AMSTAR Scores of Included Systematic Reviews 185
Table A7: Excluded Systematic Reviews 186
Table A8: Applicability of Studies Assessing Cost-Effectiveness of C-ADR.....192
Table A9: Devices Considered.....194
Table A10: Devices Excluded from Consideration.....194
Table A11: Incidence of Index-Level Corrective Surgeries by Type in 7-Year Follow-Up 196
Table A12: Other Sources Explored for Unit Prices of Artificial Discs.....197
Table A13: Other Sources Explored for Unit Prices of Instruments Used in Fusion.....197
Table A14: Surgeon, Anesthetist, and Surgical Assistant Fees for Index and Revision Surgeries, 2018 198
Table A15: Surgeon, Anesthetist, and Surgical Assistant Fees for Removal Surgery, 2018....199
Table A16: Surgeon, Anesthetist, and Surgical Assistant Fees for Cervical Supplemental Fixation Surgery, 2018 199
Table A17: Surgeon, Anesthetist, and Surgical Assistant Fees for Reoperation Surgery, 2018 199
Table A18: Hospital Cost for C-ADR and Fusion Surgeries in Ontario 200
Table A19: Hospital Cost for Cervical Posterior (Supplemental) Fixation and Reoperation Surgeries in Ontario.....200
Table A20: Multi-level (Lumbar + Cervical) Artificial Disc Implantation 200
Table A21: Number and Relative Distribution of Cervical Fusion Procedures in Alberta, 2004–2007 201
Table A22: Cost for C-ADR and Fusion Surgeries in Ontario 201
Table A23: Other Sources Explored for Hospital-Only Cost 201
Table A24: Consultation Fees for Surgeons.....202
Table A25: Consultations Needed for Each Intervention 202
Table A26: Imaging Needed for Each Intervention 202

Table A27: Sources Explored for Unit Prices of MRI in Ontario 202
Table A28: Sources Explored for Unit Prices of X-Ray in Ontario 203
Table A29a: Sources Explored for Cost of Adverse Events 203
Table A29b: Adjusted Canadian Adverse Event Cost 203
Table A30: Reference Case Analysis Results for One-Level and Two-Level, Hospital-Payer Perspective[a] 204
Table A31: Predicted Yearly Number of Surgeries in Ontario Eligible for Fusion 205
Table A32: Predicted Yearly Number of Surgeries in Ontario Eligible for C-ADR 205
Table A33: Predicted Number of Patients Receiving C-ADR or Fusion in Ontario Annually Among Patients Eligible for Either Surgery 206
Table A34: Net Budget Impact for C-ADR Versus Fusion for Cervical Degenerative Disc Disease 207
Table A35: Net Budget Impact for C-ADR Versus Fusion for Cervical Degenerative Disc Disease 207

LIST OF FIGURES

Figure 1: PRISMA Flow Diagram—Clinical Search Strategy 22
Figure 2: Graphical Display of Two-Year Mean OTS for One-level Disease—Proportion of C-ADR Versus Fusion OTS and 10% Noninferiority Margin 35
Figure 3: PRISMA Flow Diagram—Economic Search Strategy 101
Figure 4: Decision Tree and Markov Models for One-Level and Two-Level C-ADR Versus Fusion 116
Figure 5: Tornado Diagram for One-Level Surgery 128
Figure 6: Tornado Diagram for Two-Level Surgery 129
Figure 7: Incremental Cost-Effectiveness Plane for C-ADR Versus Fusion With a Willingness-to-Pay of $50,000 per QALY for One-Level Cervical Degenerative Disc Disease 130
Figure 8: Incremental Cost-Effectiveness Plane for C-ADR Versus Fusion With a Willingness-to-Pay of $50,000 per QALY for Two-Level Cervical Degenerative Disc Disease 131
Figure 9: Cost-Effectiveness Acceptability Curve for C-ADR Versus Fusion for One-Level Cervical Degenerative Disc Disease 132
Figure 10: Cost-Effectiveness Acceptability Curve for C-ADR Versus Fusion for Two-Level Cervical Degenerative Disc Disease 133
Figure 11: Model Schematic of Budget Impact 137

OBJECTIVE

This health technology assessment examined the effectiveness, safety, durability, and cost-effectiveness of cervical artificial disc replacement (C-ADR) versus fusion for treating cervical degenerative disc disease. It also examined the preferences, values, and experiences of people with cervical degenerative disc disease.

BACKGROUND

Cervical degenerative disc disease refers to deterioration of the intervertebral discs in the spine. Although all discs can be at risk, the condition mainly affects the lumbar (L1–L5) and, less frequently, the cervical level (C3–C7) of the spine. Discs are located between the vertebrae and have several functions, including allowing for flexible movement and acting as shock absorbers for stresses on the spine. Degeneration involves changes in disc characteristics, such as loss of moisture, which reduces disc elasticity and increases disc brittleness.[1] Alterations in shape can also lead to disc collapse, with a resultant decrease in intervertebral space or height, and produce abnormal spinal movement.

Deteriorating discs may bulge toward the spinal canal, which can cause radiculopathy (pressure on the nerve roots) or myelopathy (pressure on the spinal cord), causing a range of painful and disabling symptoms. Nerve root compression can cause disabling pain in the neck and/or arm. Spinal cord compression can cause a range of neurologic symptoms, including weakness, paresthesia (numbness in the arms or legs), tingling in the arms or hands, and loss of balance and coordination.

Cervical degenerative disc disease is a multifactorial condition that begins with disc deterioration and proceeds to a degenerative cascade of interrelated adverse spinal events involving the facet joints and ligaments.[2] A decrease in the height of the intervertebral disc can lead to an increase in the sagittal diameter, with various degrees of disc bulging or protruding into the spinal canal. Osteophytes (bone spurs) can also project into the canal, further reducing space for the spinal cord and blood supply. The narrowing of the disc space can also increase stress on other spinal joints, such as the uncovertebral and facet joints. Osteophytes can also project from these joints, further reducing space.

Clinical Need and Target Population

Cervical degenerative disc disease involving radiculopathy (nerve compression) consists of a group of symptoms and signs related to dysfunction of cervical spinal nerves. One American population-based survey[3] reported an average annual incidence rate of 83.2 per 100,000 people (107.3 per 100,000 for males; 63.5 per 100,00 for females) for cervical degenerative disc disease with radiculopathy (nerve root compression). In this survey, age-specific incidence rates peaked at 202.9 per 100,000 people for adults 50 to 54 years of age.

Cervical degenerative disc disease involving myelopathy (cord compression) occurs much less often compared with cervical nerve root compression. However, one study reported cervical myelopathy was the most common spinal cord disorder in older people; 24% of 2,104 patients with nontraumatic paraparesis (paralysis or loss of motor function) had cervical myelopathy.[4] Nouri et al[5] cite two studies estimating prevalence of cervical myelopathy on the basis of hospital admissions. A cervical myelopathy prevalence of 1.6 per 100,000 inhabitants was reported on the basis of surgical cases at a hospital in the Netherlands[6] and a cervical

myelopathy related hospitalization rate of 4.04 per 100,000 based on a Taiwan nationwide database.[7]

The incidence of both forms of cervical degenerative disc disease, radiculopathy and myelopathy, increases with age, and, in most cases, the condition is asymptomatic.[5,8] Spinal degenerative changes have been commonly reported in various imaging investigations of asymptomatic volunteers. Degenerative cervical changes have been examined with x-ray investigations[9,10,11] with a 10-year follow-up. Degenerative cervical changes involving disc space narrowing, vertebral body endplate sclerosis, and osteophyte formation were reported in 63% of 159 asymptomatic participants.[9,10]

Current Surgical Treatment

When symptoms of cervical degenerative disc disease are refractory to conservative management, which can include medical management (e.g., medication) or physiotherapy to manage pain, surgical treatment is an option. The most common form of surgery for cervical degenerative disc disease is anterior cervical discectomy and fusion, often simply called "fusion." Fusion has broad indications not only for symptom relief from nerve or spinal cord compression but also for spinal instability. However, surgery would usually be considered the first choice for treatment if symptoms of cord compression and neurologic deficits are serious or worsen. Fusion can involve multiple procedures, including spinal cord decompression, removal of posterior osteophytes (bone spurs), partial corpectomy (removal of vertebral endplates), or removal of posterior longitudinal ligaments.[12]

Fusion failure, or pseudarthrosis, is reported to vary widely from 0% to 20% for one-level fusion and to be as high as 60% for multi-level fusion. A review of pseudarthrosis rates for fusion with allograft procedures reported an average rate of 4.8% (95% confidence interval [CI] 1.7%–7.9%) among clinical trials, with rates ranging from 0% to 15.2%.[13] Fusion failure rates have been reported to decline substantially with plating techniques for both one-level disease (from 9% to 4%) and for two-level disease (from 28% to 10%).[14] Variation in surgical techniques, levels treated, radiologic definitions of pseudarthrosis, length of follow-up, and patient differences all influence fusion failure rates.

The most important disadvantage of fusion is its potential effect on adjacent levels of the treated spine. It remains controversial whether adjacent-level deterioration reported after fusion can be attributed to the increased motion and biomechanical stresses offloaded by fusion, to the natural progression of cervical degenerative disc disease, or to both factors. Several large cohort studies with long-term follow-up have reported symptomatic and radiologically detected adjacent-segment disease.[15-19]

The study by Hillibrand et al[16] is most frequently cited; this study included a 10-year follow-up of an American cohort of 374 people undergoing fusion. The annual incidence of symptomatic adjacent-segment disease was relatively constant at 2.9%. The follow-up showed 25.6% of people would develop symptoms of adjacent-segment disease after fusion surgery. No differences were found between one-level and two-level index-level (initially) treated cases.

Surgery for adjacent-segment disease after cervical fusion was reported in two American cohorts. In a 1,038 patient-cohort the 5-year adjacent-level operation rate was 8.3% and the 10-year rate was 22.2%.[17] In a second 888 patient-cohort with 7.8-year mean follow-up, a 12% rate of adjacent-level surgery for symptomatic degenerative disease was reported.[19] In European studies, a 6.1% surgery rate for adjacent-level degeneration was reported for a cohort

with 8-year mean follow-up[15] and a 5.9% surgery rate for a cohort with 14.5-year mean follow-up.[18]

Health Technology Under Review

Cervical artificial disc replacement (C-ADR), also called arthroplasty, is an alternative to fusion for cervical degenerative disc disease. The overall treatment objectives of C-ADR are similar to those of fusion: to remove the diseased or damaged intervertebral discs; to restore normal disc height; to reduce disc-related neck pain and associated arm pain or weakness; and improve function. However, unlike fusion, C-ADR is also intended to preserve spine mobility. By resulting in more normal biomechanics with less stress on adjacent disc levels of the spine, a potential for less adjacent-level degeneration or disease and for fewer subsequent surgeries is anticipated. In addition, potential complications, such as the need for bone grafts and the need for hardware associated with fusion, could be avoided. The smaller profile of the C-ADR implants might also lead to significantly less morbidity after surgery; for example, in terms of dysphagia (difficulty swallowing) and dysphonia (vocal abnormalities).

Both fusion and C-ADR are performed under general anesthesia and use an anterior approach (from the front of the neck) for removal of the affected disc. For C-ADR procedures, surgeons can choose from among various C-ADR devices. For fusion, surgeons can choose from a variety of bone grafting materials, plating materials (for the metal plates used to connect the vertebral bodies), and interbody devices (solid structures to fill the space between vertebrae; e.g., metal blocks, cages).

The potential surgical complications of C-ADR are similar to those of other spine surgeries and include side effects from anesthesia, an allergic reaction to the implant materials, hemorrhage requiring transfusion, incision problems, infection, numbness or tingling in the extremities, spinal fluid leakage, tears of the dura (fibrous membrane covering the spinal cord), paralysis, or even death. The rate of longer-term complications following C-ADR, such as disease progression to other cervical levels and the need for additional surgery, might be reduced compared with surgical fusion. It is unclear whether revision surgery after C-ADR is more complicated than after fusion.

C-ADR devices have evolved considerably since their introduction in Europe in the late 1990s. They differ in biomaterials, bearing design, articulating system, and methods of fixation to vertebral endplates (the top and bottom portions of vertebral bodies that connect to the vertebral discs). Table 1 outlines the properties of C-ADR devices with regulatory approvals in Canada or the United States. The Discover system has regulatory approval only in Europe. C-ADR devices are typically designed either as metal-on-metal (MoM) devices, consisting of two metal pieces, or metal-on-polyurethane (MoP) devices, consisting of two metal endpieces and a centre core of various commercial polyurethane preparations. MoM and MoP devices differ in the type of metal or metal alloy used for the endplates, but it is often titanium (or, more recently, a ceramic composite material, as in the Prestige-LP device). The various materials have different biological reactivities, strengths, MRI compatibilities.

Most of the currently available C-ADR devices do not reproduce the normal visco-elastic (compressibility) properties of the native disc. Various implant designs have different kinematic properties and vary in the degree of movement the device allows; these factors represent the potential for important differences in device performance. The full range of motion of the normal cervical spine involves translation and rotation. Sliding rotational discs without inherent elasticity do not replicate the kinematics of the native intervertebral disc. They have been characterized

as constrained (furthest from normal anatomy and physiologic movement), semiconstrained, and unconstrained (closest to normal anatomy and physiologic movement). Several criteria drive device selection: device size, height of device joint line, rotating axis, and position of rotating centre. The device's centre of rotation is the most important characteristic. Devices are also made in a range of heights to avoid "over-stuffing" or "under-stuffing" the intervertebral space.

Table 1: Characteristics of Cervical Artificial Disc Replacement Systems

C-ADR Device, Manufacturer	Regulatory Approvals[a]	Biomaterials	Endplate Fixation Method	Bearing Design and Articulation and Motion	Comments
Bryan, Medtronic	• EMA CE mark • FDA • Health Canada	Metal on polyurethane/ polyethylene • Titanium alloy (titanium, aluminum, vandolinium) endplates • Polycarbonate polyurethane nucleus with surrounding flexible polyether urethane sheath	• Press-fit endplates screwed to adjacent vertebral body • Porous, coated titanium endplate alloy	• Semiconstrained • Mobile centre of rotation • Mobile, bi-articulating surface • Inward-facing posts on inferior and superior endplates fit into flared holes in nucleus	• One of first approved C-ADR devices • Requires precise concave milling of bony endplates • Lengthy procedure steps • MRI compatible
Discover, DePuy Synthes Spine	• EMA CE mark (no longer available)	Metal on polyurethane/ polyethylene • Titanium endplates • Ultrahigh-molecular-weight polyethylene core	• Teeth on endplates • Titanium plasma-sprayed endplate coating	• Semiconstrained • Fixed ball-in-socket design	• Endplate design improves lordosis or sagittal alignment • Ridges require vertebral bony body preparation • MRI compatible
Kineflex-C, Spinal Motion	• EMA CE mark (no longer available)	Metal on metal • Cobalt–chromium endplates	• Midline keel on inferior and superior endplates • Titanium plasma-sprayed endplate coating	• Semiconstrained • Ball-in-trough design • Mobile centre core translated with retention ring	• Potential metal-on-metal wear • Keels require bony vertebral body preparation • Immediate, secure press-fit in place • MRI artifacts and scatter
M6-C, Spinal Kinetics	• EMA CE mark • FDA (IDE trial in progress) • Health Canada	Metal on polyurethane/ polyethylene • Titanium endplates • Polyurethane nucleus surrounded by ultrahigh-molecular-weight polyethylene annulus	• Three fins on superior and inferior endplates • Titanium plasma-sprayed endplate coating	• Unconstrained • Bi-articulating surface • Mobile centre of rotation	• Designed to provide axial compression • Design allows for 6° of freedom • Simplified assembly owing to a one-piece design with numerous complex components • MRI compatible

C-ADR Device, Manufacturer	Regulatory Approvals[a]	Biomaterials	Endplate Fixation Method	Bearing Design and Articulation and Motion	Comments
Mobi-C, Zimmer (former manufacturer: LDR Spine)	• EMA CE mark • FDA • Health Canada (submission in progress for up to 2 levels)	Metal on polyurethane/ polyethylene • Cobalt–chromium–molybdenum endplates Ultrahigh-molecular- weight polyethylene core	• Two lateral rows of serrated teeth on inferior and superior plates • Lateral stops (tabs) on inferior plates to hold the core • Titanium and hydroxyapatite plasma–sprayed endplate coating	• Semiconstrained • Ball-in-socket design • Mobile centre of rotation	• First C-ADR device to receive 2-level regulatory approval from FDA • Centre core self-centring maintains vertebral alignment • No preparation, removal, or chiselling of bony endplates required • Teeth require time to settle into vertebral endplates • MRI artifacts and scatter
PCM, NuVasive	• FDA • Health Canada	Metal on polyurethane/ polyethylene • Cobalt–chromium–molybdenum endplates Ultrahigh-molecular-weight polyethylene core	• Transverse ridges on inferior and superior endplates • Two layers of titanium plasma spray; calcium phosphate endplate coating	• Semiconstrained • Ball-in-socket design (concave endplate surface geometry and convex polyethylene core surface geometry)	• Design closest to normal anatomy and kinetics • Teeth require time to settle into endplates • MRI artifacts and scatter •
Prestige-LP, Medtronic	• EMA CE mark • FDA • Health Canada	Metal on metal Titanium–ceramic composite (titanium–6aluminum–4vanadium with 10% titanium carbide) endplates	• Endplates fixed to adjacent vertebral bodies by 2 rail geometries incorporating anti-migration teeth press-fit into 2 pre-drilled holes in vertebral bone • Titanium plasma–sprayed endplate coating	• Unconstrained • Ball-in-trough design	• 2-level indication approved by FDA • Requires precise concave milling of bony vertebral body • Lengthy procedure steps • MRI compatible
Prestige-ST, Medtronic	• EMA CE mark • FDA • Health Canada	Metal on metal Stainless steel endplates	• Endplates fixed to adjacent vertebral body by 2 bone screws per endplate through anterior flange held in place by lock-screw mechanism • Endplates in contact with vertebral body are aluminum oxide grit blasted	• Unconstrained (axial rotation) • Ball-in-socket (superior endplate), in-trough (inferior endplate) design	• Secure, immediate fixation • Requires preparation of vertebral body • Lengthy procedure steps • MRI artifacts and scatter

C-ADR Device, Manufacturer	Regulatory Approvals[a]	Biomaterials	Endplate Fixation Method	Bearing Design and Articulation and Motion	Comments
ProDisc-C, Centinel Spine (former manufacturer: DePuy Synthes Spine)	• FDA • Health Canada	Metal on polyurethane/ polyethylene • Cobalt–chromium endplates Ultrahigh-molecular-weight polyethylene core	• Endplates anchor to vertebral body with midline keel • Titanium plasma–sprayed endplate coating	• Semiconstrained • Unconstrained in axial rotation • Ball-in-socket (concave, polished endplate surface geometry; convex spherical core dome) design	• Keels require cuts into cortical and cancellous bone of vertebral body • Immediate, secure press-fit in place • MRI artifacts and scatter
ProDisc-C Nova, Centinel Spine (former manufacturer: DePuy Synthes Spine)	• Health Canada	Metal on polyurethane/ polyethylene • Cobalt–chromium–titanium endplates Ultrahigh-molecular-weight polyethylene core	• Endplates have staggered keel design: 2 keels on inferior endplate, 1 keel on superior endplate • Titanium plasma–sprayed endplate coating	• Semiconstrained • Unconstrained in axial rotation • Ball-in-socket design	• Keels require cuts into cortical and cancellous bone of vertebral body • Tripod keel shape allows easy multi-level application • Reduced MRI artifacts and scatter
ProDisc-C Vivo, Centinel Spine (former manufacturer: DePuy Synthes Spine)	• Health Canada	Metal on polyurethane/ polyethylene • Cobalt–chromium–molybdenum endplates Ultrahigh-molecular-weight polyethylene	• Endplates have teeth • Titanium plasma–sprayed endplate coating	• Semiconstrained • Unconstrained in axial rotation • Ball-in-socket design	• No vertebral bone cuts to accommodate keel • MRI artifacts and scatter
Secure-C, Globus Medical	• FDA	Metal on polyurethane/ polyethylene • Cobalt–chromium–molybdenum endplates Ultrahigh-molecular-weight polyethylene core	• Serrated keels on inferior and superior endplates • Titanium plasma–sprayed endplate coating	• Semiconstrained • Mobile core • Ball-in-socket design	• Keels require cuts into cortical and cancellous bone of vertebral body • Immediate, secure press-fit in place • MRI artifacts and scatter

Abbreviations: C-ADR, cervical artificial disc replacement; CE, Conformité Européenne; EMA, European Medicine Agency; FDA, U.S. Food and Drug Administration; IDE, investigational device exemption; MRI, magnetic resonance imaging.

[a]As of February 2018.

Regulatory Information

C-ADR was first introduced in Europe in the late 1990s. In the United States, randomized U.S. Food and Drug Administration (FDA)–regulated investigational device exemption (IDE) trials led to the first three FDA approvals for C-ADR devices for cervical degenerative disc disease: Bryan,[20] Prestige-ST,[21] and ProDisc-C.[22] In 2012 and 2013, three more devices received FDA approval. Two further devices, Mobi-C and Prestige-LP, have received FDA regulatory approval for two-level cervical disc replacement.

In Canada, as of February 2018, four manufacturers have received regulatory approval for C-ADR devices licensed as Class III devices by Health Canada: Medtronic (Bryan, 2009; Prestige-ST, 2007; Prestige-LP, 2010); Centinal Spine (formerly J&J Depuy Synthes) (ProDisc-C, 2007; ProDisc-C Nova, 2015; ProDisc-C Vivo, 2017); Nuvasive (PCM, 2007); and Spinal Kinetics (M6-C, 2014). The licensing status of C-ADR devices in Canada was verified by a Health Canada regulatory official (personal communication, Health Canada Device Licensing Services Division, Medical Devices Bureau, Health Canada).

Ontario Context

In Ontario, spine surgeries are specialized services performed in university-affiliated hospitals and community hospitals serving as regional centres for spine services. Surgery is performed by spine surgeons, usually neurosurgeons or orthopedic surgeons. In only a few centres are Ontario surgeons trained to perform C-ADR in either the lumbar or cervical spine. In Ontario, 900 cervical fusions for non-emergency cervical spine conditions were performed in 2013/14.[23]

Spine surgeons reported that 50% of patients who were referred to them had lived with cervical symptoms related to nerve or cord compression for 2 years or longer.[24] Canadian spine surgeons have also reported that the referral criteria for spinal surgery in the neck and back have not been consistently established; criteria conflict within regions and within institutions. Forty-two percent of spine surgeons reported being referred more than 10 people for surgery, of whom only one was eligible for surgery.[25]

In Ontario, device costs for both fusion and C-ADR devices are covered by hospitals' overall global budgets or through specialized Ministry of Health and Long-Term Care funding programs. Physician remuneration is provided by the Ontario Health Insurance Plan (OHIP) and determined by specific OHIP fee codes.

The Ministry of Health and Long-Term Care and an expert panel developed the Quality-Based Pathway Clinical Handbook for Non-emergent Integrated Spine Care in 2017.[23] Cervical fusion is included in this handbook, but C-ADR is not.

CLINICAL EVIDENCE

Research Question

What are the clinical effectiveness, safety, and durability of C-ADR compared with fusion for one-level and two-level cervical degenerative disc disease?

Methods

We developed the research questions in consultation with patients, health care providers, and clinical experts.

Clinical Literature Search

We performed a literature search on July 11, 2017, to retrieve studies published from inception to the search date. We used the Ovid interface to search the following databases: MEDLINE, Embase, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, Health Technology Assessment, and National Health Service Economic Evaluation Database (NHS EED).

Search strategies were developed by medical librarians using controlled vocabulary (e.g., Medical Subject Headings) and relevant keywords. Search filters were used to limit results to systematic reviews, meta-analyses, health technology assessments, and randomized controlled trials (RCTs). The final search strategy was peer-reviewed using the PRESS Checklist.[26] Database auto-alerts were created in MEDLINE and Embase and monitored for the duration of the health technology assessment.

We performed targeted grey literature searching of health technology assessment agency sites and clinical trial registries. See Appendix 1 for literature search strategies, including all search terms.

Literature Screening

A single reviewer conducted an initial screening of titles and abstracts and obtained the full text of studies that appeared eligible for the review, according to the inclusion criteria. The author then examined the full-text articles and selected studies that were eligible for inclusion.

Inclusion Criteria

- English-language full-text publications
- Studies published until July 11, 2017
- Randomized controlled trials, systematic reviews, and health technology assessments
- Studies involving interventions with cervical artificial discs for degenerative disc disease
- Studies involving follow-up evaluations including one or more clinical, radiologic, or patient outcomes

Exclusion Criteria

- Animal and in vitro studies
- Observational studies, editorials, letters, or commentaries
- Expert reviews
- Protocol reports

- Abstracts and conference proceedings

Outcomes of Interest

- Clinical outcomes: major adverse events, neurologic status, and overall treatment success (OTS), as defined by study investigators or regulators
- Functional outcomes: Neck Disability Index (NDI)
- Surgical outcomes: perioperative outcomes, secondary surgeries, surgeon satisfaction
- Radiologic outcomes: heterotopic ossification (abnormal growth of bone in nonskeletal tissues), adjacent-segment disease, kinematics, pseudofusion (fusion failure)
- Patient-reported outcomes: health-related quality of life (HRQOL), recovery or return to work, employment status, patient treatment satisfaction
- Durability: C-ADR device wear and biologic reactivity

Data Extraction

We extracted relevant data on study characteristics and risk-of-bias items using a data form to collect information about the following:

- Source (i.e., citation information, study type)
- Methods (i.e., study design, study duration in years, participant allocation, reporting of missing data, reporting of outcomes, and whether the study compared two or more groups)
- Outcomes (i.e., outcomes measured, number of participants for each outcome, outcome definition and source of information, unit of measurement, upper and lower limits [for scales], and time points at which outcomes were assessed)

We contacted study authors to provide clarification as needed.

Statistical Analysis

Given the extensive availability of systematic reviews on this topic, the research question was addressed either by reviewing existing systematic reviews or by de novo processes of a systematic review. Existing systematic reviews were included if they clearly specified a review question, eligibility criteria, and recent search duration, and if they undertook a reproducible search of two or more electronic literature databases. Meta-analysis within existing or de novo systematic reviews was considered when studies were judged to be sufficiently homogeneous in study populations, interventions, outcome measures, and follow-up duration. If meta-analysis was inappropriate because of clinical, methodologic, or statistical heterogeneity, a narrative summary of results was produced.[27]

We reported the results of the included RCTs, most of which were designed as noninferiority trials (trials that attempt to determine if a new treatment is no worse than the current treatment). Within the C-ADR trials, noninferiority of C-ADR to fusion was accepted if the proportion of people achieving the primary outcome measure of overall treatment success in the C-ADR group was within 10% of the lower 95% confidence interval for the fusion group. Overall treatment success, defined by investigators and regulators, was a composite measure based on subcriteria involving technical success, treatment effectiveness, and safety and was variably defined and measured across the C-ADR RCTs. Superiority of C-ADR to fusion for an outcome was determined if the lower limit of the 95% confidence interval did not cross zero.

Statistical analyses from the included studies involved testing noninferiority or superiority using frequentist and/or Bayesian estimations. For studies using the frequentist approach, significant probabilities were reported as $P < .05$ if the data were compatible with a null hypothesis of noninferiority (or superiority). For studies employing Bayesian statistics, a posterior distribution of the proportion achieving success in the study groups was reported, and the probability (P_B) that the difference in proportions was inferior or superior was also reported. Probabilities greater than 95% supported claims of noninferiority or superiority of C-ADR to fusion.

Critical Appraisal of Evidence

We assessed risk of bias for RCTs with the Cochrane Back and Neck guidelines, which consider four types of bias: selection, performance, attrition, and detection bias (Appendix 4, Table A4).[28] A low risk of bias was defined as no serious methodologic flaws in the study. We used a validated tool—A Measurement Tool to Assess Systematic Reviews (AMSTAR)—to perform a critical appraisal of the methodologic quality of the systematic reviews we included (Appendix 4, Table A6).[29,30]

Quality of Evidence

The quality of the overall body of evidence for each outcome was evaluated according to the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) Handbook.[31] The quality reflects our certainty about the evidence. We assessed the body of evidence based on the following considerations: risk of bias, inconsistency, indirectness, imprecision, and publication bias (Appendix 5, Table A5).

Expert Consultation

Consultations with clinical experts began in July 2017. We solicited expert consultation on the appropriate use of C-ADR with neurosurgeons and orthopedic surgeons in the specialty area of spine surgery. The role of the expert advisors was to contextualize the evidence and provide advice on surgical treatments for cervical degenerative disease. Consultations with industry representatives were also held to inform them of this health technology assessment.

Results

Literature Search

The literature search yielded 601 citations published from inception to July 11, 2017, after removing duplicates. We reviewed titles and abstracts to identify potentially relevant articles. Of the 601 records, 380 were excluded based on the title and abstract. We obtained full texts of the remaining 221 articles for further assessment. We hand-searched the reference lists of the included studies, along with health technology assessment websites and other sources, to identify additional relevant studies. For this health technology assessment, 3 systematic reviews and 23 RCTs involving 85 reports were judged to have met the study inclusion criteria. Appendix 5 provides a selected list of systematic reviews excluded after full-text review that includes the primary reason for exclusion.

Figure 1 presents the flow diagram for the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA).

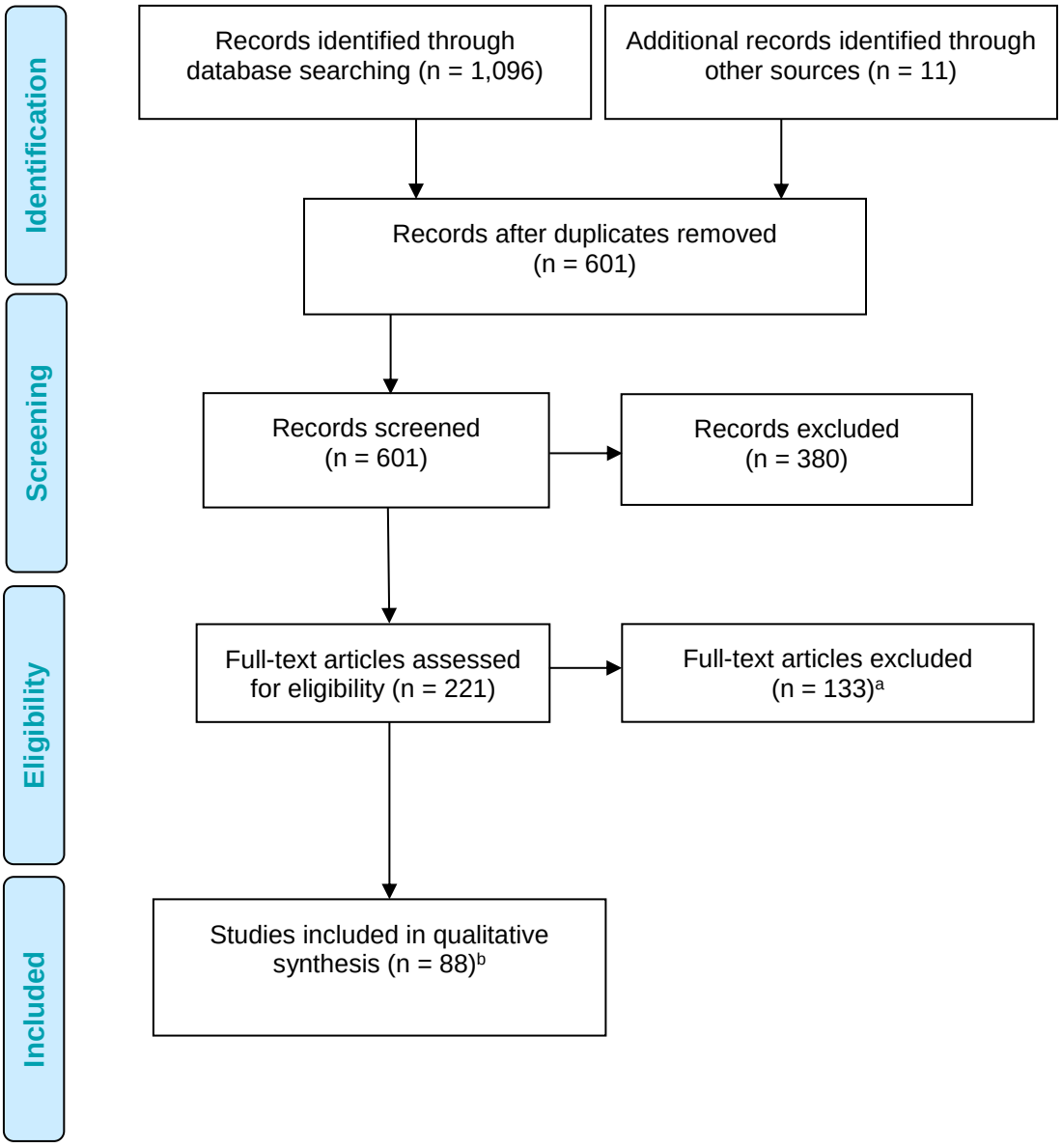

Figure 1: PRISMA Flow Diagram—Clinical Search Strategy

Abbreviations: PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-analyses; RCT, randomized controlled trial.
[a]Table A7 (Appendix 5) describes reasons for excluding 91 systematic reviews; a further 42 reports were excluded for wrong surgical approach, different device, no follow-up, or no target outcome reported.
[b]Three systematic reviews, 23 RCTs cited in 85 reports.
Source: Adapted from Moher et al.[32]

Included Studies

Three systematic reviews were included.[33-35] Two[34,35] involved an assessment of the wear characteristics and biological reactivity of C-ADR devices, and the other[33] assessed adverse events experienced following C-ADR or fusion. The results of these reviews are detailed later in the report.

Table 2 summarizes the 23 RCTs (from which 85 articles were written) of C-ADR versus fusion we included. Characteristics of individual RCTs and their follow-up reports are detailed more fully in Appendix 2, Table A1.

Table 2: Randomized Controlled Trials of Cervical Artificial Disc Replacement Versus Fusion for Cervical Degenerative Disc Disease

	IDE Multicentre RCTs		Non-IDE Multicentre RCTs		Non-IDE Single-Centre RCTs	
Cervical Disc Device	**No. of Trials**	**No. of Reports[a]**	**No. of Trials**	**No. of Reports**	**No. of Trials**	**No. of Reports**
Bryan	1	15	1	1	2	3
Discover	–	–	2	6	2	2
Kineflex-C	1	2	–	–	–	–
Mobi-C	2	14	1	2	–	–
PCM	1	4	–	–	–	–
Prestige-LP[b]	2	4	–	–	2	2
Prestige-ST	1	3	1	1	–	–
ProDisc-C	1	13	–	–	2	4
Secure-C	1	1	–	–	–	–
Total	**10**	**56**	**5**	**10**	**8**	**11**
Mixed-device IDE RCT[c]	–	8	–	–	–	–

Abbreviations: IDE, investigational device exemption; RCT, randomized controlled trial.

[a]Reports on trial involve various follow-up periods and/or outcome assessments.

[b]Two IDE trials were conducted for Prestige-LP, one for one-level disease and one for two-level disease. The one-level study on degenerative disc disease was a prospective controlled clinical trial using data from propensity matched control fusion patients in the FDA RCT of Prestige-ST, the same patient selection criteria, and the same surgeons. The two-level study was an RCT.

[c]The mixed-device IDE RCT combined results from two or more trials from one or more sites to evaluate certain outcomes or outcomes in targeted subpopulations.

Eight regulatory investigational device exemption (IDE) trials from the U.S. Food and Drug Administration (FDA) assessed C-ADR versus fusion for eight C-ADR devices involving one-level cervical degenerative disc disease.[20-22,36-40] FDA IDE trials are trials conducted under an exemption of regulatory approval in order to generate evidence to support regulatory approval to legally commercially market medical devices or products. Two of these C-ADR devices (Mobi-C, Prestige-LP) also had separate FDA regulatory IDE RCTs involving two-level disease.[41,42] All the regulatory trials were large (> 20 sites) multicentre RCTs, except the FDA trial of Prestige-LP for one-level cervical degenerative disc disease.[37] The FDA trial of Prestige-LP, a propensity-matched prospective cohort study, was included in the review because it was also an FDA IDE regulatory trial and had the same oversight, patient selection criteria, and prospective follow-up as the prior FDA RCT of Prestige-ST[21] that was the source of fusion control subjects. These trials all followed a noninferiority design with a priori specified testing procedures (Table 3).

In addition to the FDA IDE trials, 13 nonregulatory RCTs were conducted in Europe, Asia, and Scandinavia.[43-55] These trials generally involved single sites (five were multicenter trials[47,52-55]) studied fewer patients, and were not industry sponsored. Their primary outcome measure, unlike that of the FDA regulatory trials, was the Neck Disability Index; overall treatment success was not reported in these nonregulatory trials.

Table 3: FDA Regulatory Trials of C-ADR Versus Fusion for Cervical Degenerative Disc Disease

C-ADR Device, FDA Approval Date, Post-approval Request	Recruitment Period, No. of Sites, N Randomized (C-ADR/Fusion), % Followed Up at 2 Yr	Inclusion Criteria	Primary Endpoint/ Noninferiority Margin	Secondary Endpoints	Review Committees	Post-operative Management
Bryan[20]						
May 12, 2009 • 5-yr enhanced surveillance study • 10-yr post-approval study	• May 28, 2002–October 8, 2004 • 30 sites • 463 enrolled (242/221) • C-ADR: 97%; fusion: 85%	• Skeletally mature (≥ 21 yr of age) • Radiographically confirmed single-level disease, C3–4 to C6–7 • Radiculopathy (failed 6 wk of conservative therapy) or myelopathy • NDI score ≥ 30	10% Δ overall success rate on 4 criteria at 24 mo: 1. NDI score improvement from baseline of ≥ 15 points 2. Maintenance of or improvement in neurologic status (motor function, sensory function, reflexes) 3. No serious implant- or surgery-associated adverse events 4. No failure-related additional surgeries	• Neck pain • Arm pain • SF-36 score • Radiologic outcomes (adjacent-level stability, functional spinal unit height, fusion, subsidence)	• Data safety monitoring committee • Independent radiographic assessments	• Avoid heavy physical activity and sports for 3 mo • Discretionary advice on orthotic device • NSAIDs for 2 wk
Kineflex-C[36]						
	• Recruitment period not reported • 21 sites • 269 enrolled (136/133) • C-ADR: 88%; fusion: 87%	• 18–60 yr of age • Single-level disease/ spondylosis, C3–C7 • Radiculopathy (failed to benefit from 6 mo of conservative therapy) or myelopathy • NDI score ≥ 40	10% Δ overall success rate on 5 criteria at 24 mo: 1. ≥ 20% NDI score improvement from baseline 2. Maintenance of or improvement in neurologic status (motor function, sensory function, reflexes) 3. No device failure 4. No index-level reoperation 5. No major device-related adverse events	• Neck pain • Arm pain • Radiologic outcomes (ASD, disc height, ROM) • Patient satisfaction	• Clinical events committee • Core radiology lab and independent radiologist assessment	Not reported

C-ADR Device, FDA Approval Date, Post-approval Request	Recruitment Period, No. of Sites, N Randomized (C-ADR/Fusion), % Followed Up at 2 Yr	Inclusion Criteria	Primary Endpoint/ Noninferiority Margin	Secondary Endpoints	Review Committees	Post-operative Management
Mobi-C[38] for one-level disease						
August 7, 2013 • Retrieve all explanted devices • 7-yr post-approval study • 10-yr enhanced surveillance system under general conditions of use	• April 2006–March 2008 • 24 sites • 260 enrolled (179/81) • C-ADR: 94%; fusion: 92%	• 18–69 yr of age • Radiographically confirmed single-level disease, C3–4 to C6–7 • Radiculopathy (failed to benefit from 6 wk of conservative therapy) or myelopathy • NDI score ≥ 15 points or ≥ 30%	10% Δ overall success rate on 3 criteria at 24 mo: 1. NDI score improvement from baseline of ≥ 15 points in patients with baseline score of ≥ 30, or 50% improvement in patients with baseline NDI score < 30 2. No device failure requiring index-level surgery (revision, reoperation, or removal) 3. No major complications, defined as radiographic, neurologic, or failure owing to adverse events (as judged by clinical events committee)	• Neck pain • Arm pain (measured via VAS) • SF-12 score • SF-36 score • Radiologic outcomes (ASD, device migration, fusion, HO, radiolucency, ROM) • Dysphagia (measured via FOSS) • Gait analysis (evaluated by Nurick grade) • Patient satisfaction	• Data monitoring committee • Clinical events committee • Independent radiographic assessments by central core lab	• Physician-managed individual post-operative rehabilitation program; could have included use of cervical collars • Discontinued use of NSAIDs 1 wk before and 3 mo after surgery • Control groups were permitted to use bone growth stimulators

C-ADR Device, FDA Approval Date, Post-approval Request	Recruitment Period, No. of Sites, N Randomized (C-ADR/Fusion), % Followed Up at 2 Yr	Inclusion Criteria	Primary Endpoint/ Noninferiority Margin	Secondary Endpoints	Review Committees	Post-operative Management
Mobi-C[41] for two-level disease						
August 23, 2013 • Retrieve all explanted devices • 7-yr post-approval study • 10-yr enhanced surveillance system under general conditions of use Annual survey of physicians	• April 2006–March 2008 • 24 sites • 330 enrolled (225/105) • C-ADR 98%; fusion 94%	• 18–69 yr of age • Radiographically confirmed 2-level disease, C3–4 to C6–7 • Radiculopathy (failed to benefit from 6 wk of conservative therapy) or myelopathy • NDI score ≥ 30 points	10% Δ overall success rate on 5 criteria at 24 mo: 1. NDI score improvement from baseline of ≥ 15 points in patients with baseline score of ≥ 30, or 50% improvement in patients with baseline NDI score < 30 2. No device failure requiring index-level surgery (revision, reoperation, or removal) 3. No adverse events assessed by CEC as major complications, defined by CEC as radiographic, neurologic, or failure owing to adverse events 4. Maintenance or improvement in neurologic function 5. Radiographic success	• Neck pain • Arm pain (measured via VAS) • SF-12 score • Radiologic outcomes (ASD, device migration, fusion, HO, radiolucency, ROM) • Patient satisfaction	• Data monitoring committee • CEC • Independent radiographic assessments by central core lab	• Physician-managed individual post-operative rehabilitation program; could have included use of cervical collars • Discontinued use of NSAIDs 1 wk before and 3 mo after surgery • Control groups were permitted to use bone growth stimulators
PCM[56]						
October 26, 2012 • 7-yr post-approval study w/ 85% follow-up • 10-yr enhanced surveillance system under general conditions of use and with annual surgeon feedback on device function and safety	• January 19, 2005–December 5, 2007 • 24 sites • 479 enrolled (289 − 75 training/190) • C-ADR: 90%; fusion: 82%	• 18–65 yr of age • Radiographically confirmed single-level disease, C3–4 to C6–7 • Radiculopathy (failed to benefit from 6 wk of conservative therapy) or myelopathy • NDI score ≥ 15	12.5% and 10% (FDA) NI margin or Δ overall success rate on 3 criteria at 24 mo: 1. ≥ 20% NDI score improvement from baseline 2. No device failure requiring revision, reoperation, or removal 3. No major complications	• Neck pain • Arm pain (VAS) • SF-36 score • Radiologic outcomes (ASD, device migration, disc height, fusion, HO, radiolucency, ROM) • Patient satisfaction	• Clinical events committee consisting of 3 board-certified spine surgeons • Radiographic assessments by independent core lab and reviews by independent board-certified radiologist	• Physical therapy program consisting of non-impact and active ROM exercises • Avoid repetitive cervical flexion and extension, bending, lateral bending, and rotation for 6 wk

C-ADR Device, FDA Approval Date, Post-approval Request	Recruitment Period, No. of Sites, N Randomized (C-ADR/Fusion), % Followed Up at 2 Yr	Inclusion Criteria	Primary Endpoint/ Noninferiority Margin	Secondary Endpoints	Review Committees	Post-operative Management
Prestige-ST[21]						
July 16, 2007 • 7-yr extended post-approval study including those in continued-access arm • 5-yr enhanced surveillance study to more fully characterize adverse events when device is used in broader patient population	• January 13, 2005–November 8, 2005 • 32 sites • 541 enrolled (276/265) • A priori interim analyses of 250 implanted cases that reached 24-mo follow-up	• ≥ 18 yr of age • Clinically and radiographically confirmed single-level disease, C3–4 to C6–7 • Radiculopathy (failed to benefit from 6 wk of conservative therapy) or myelopathy • NDI score ≥ 30 • Neck pain score ≥ 20	10% NI margin overall success rate on 5 criteria at 24 mo: 1. NDI score improvement from baseline of ≥ 15 points 2. Maintenance of or improvement in neurologic status 3. No serious implant-associated or implant/surgical procedure–associated adverse events 4. No lure-related additional surgery 5. Functional spinal unit height maintained[a]	• Neck pain • Arm pain (measured via VAS) • SF-36 score • Radiologic outcomes (angular and translational motion, disc height, flexion/extension on angular motion, fusion, HO, radiolucency) • Gait analysis (evaluated by Nurick grade) • Foramen compression • Work status • Patient satisfaction • Surgeon satisfaction	• Radiographic assessmernts by independent core lab	• Avoid heavy lifting • Avoid repetitive bending, high-impact exercise, and athletic activity for 60 d • NSAIDs recommended for 2 wk • Use of electrical bone growth stimulators prohibited for 24 mo

C-ADR Device, FDA Approval Date, Post-approval Request	Recruitment Period, No. of Sites, N Randomized (C-ADR/Fusion), % Followed Up at 2 Yr	Inclusion Criteria	Primary Endpoint/ Noninferiority Margin	Secondary Endpoints	Review Committees	Post-operative Management
Prestige-LP[37] for one-level disease						
July 24, 2014 • 10-yr extended post-approval study with expected 80% follow-up • Collection of metal ions in cohorts • Extensive clinical and radiographic information on adjacent levels • Annual survey of physicians for adverse events • Analysis of explants by third-party vendor	• January 13, 2005–November 8, 2005 • 20 sites • 280 enrolled + 54 additional patients (30 enrolled in metal ion cohort and 24 continued-access patients); safety implant cohort (n = 333) included trial subjects (n = 280) and continued-access subjects (n = 53) • C-ADR: 94%; fusion: 76%	• ≥ 18 yr of age • Clinically and radiographically confirmed single-level disease, C3–4 to C6–7 • Radiculopathy (failed to benefit from 6 wk of conservative therapy) or myelopathy • NDI score ≥ 30 • Neck pain score ≥ 20	10% NI margin overall success rate on 5 criteria at 24 mo: 1. NDI score improvement from baseline of ≥ 15 points 2. Maintenance of or improvement in neurologic status 3. Functional spinal unit height maintained[a] 4. No serious implant- or surgery-associated adverse events 5. No failure-related additional surgeries	• Neck pain • Arm pain (measured via VAS) • SF-36 score • Radiologic outcomes (angular and translational motion, disc height, flexion/extension on angular motion, fusion, HO, radiolucency) • Gait analysis (evaluated by Nurick grade) • Foramen compression • Work status • Patient satisfaction • Surgeon satisfaction	• Clinical adjudication • Committee of 3 independent spine surgeons • Radiographic assessments by independent core lab • Explant analysis by third-party vendor	• Avoid heavy lifting, repetitive bending, high-impact exercise, and athletic activity for 60 d • NSAIDs recommended for 2 wk • Bracing, including use of soft collar, left to discretion of physicians • Use of electrical bone growth stimulators prohibited for 24 mo[b] (if used were classed as surgery supplemental fixation

C-ADR Device, FDA Approval Date, Post-approval Request	Recruitment Period, No. of Sites, N Randomized (C-ADR/Fusion), % Followed Up at 2 Yr	Inclusion Criteria	Primary Endpoint/ Noninferiority Margin	Secondary Endpoints	Review Committees	Post-operative Management
Prestige-LP[42] for two-level disease						
July 16, 2007 • 7-yr extended post-approval study including those in continued-access arm • 5-yr enhanced surveillance study to more fully characterize adverse events when device was used in broader patient population	• January 13, 2005–November 8, 2005 • 32 sites • 541 enrolled (276/265) • A priori interim analyses of 250 implanted cases that reached 24-mo follow-up	• ≥ 18 yr of age • Clinically and radiographically confirmed 2-level disease, C3–4 to C6–7 • Radiculopathy (failed to benefit from 6 wk of conservative therapy) or myelopathy • NDI score ≥ 30 • Neck pain score ≥ 8	10% NI margin overall success rate on 5 criteria at 24 mo: 1. NDI score improvement from baseline of ≥ 15 points 2. Maintenance of or improvement in neurologic status 3. No serious implant- or surgery-associated adverse events 4. No failure-related additional surgery 5. Functional spinal unit height maintained[a]	• Neck pain • Arm pain (measured via VAS) • SF-36 score • Radiologic outcomes (angular and translational motion, disc height, flexion/extension on angular motion, fusion, HO, radiolucency) • Gait analysis (evaluated by Nurick grade) • Foramen compression • Work status • Patient satisfaction • Surgeon satisfaction	• Radiographic assessments by independent core lab	• Avoid heavy lifting • Avoid repetitive bending, high-impact exercise, and athletic activity for 60 d • NSAIDs recommended for 2 wk • Use of electrical bone growth stimulators prohibited for 24 mo

C-ADR Device, FDA Approval Date, Post-approval Request	Recruitment Period, No. of Sites, N Randomized (C-ADR/Fusion), % Followed Up at 2 Yr	Inclusion Criteria	Primary Endpoint/ Noninferiority Margin	Secondary Endpoints	Review Committees	Post-operative Management
ProDisc-C[22]						
December 17, 2007 • 7-yr post-approval study with RCT subjects and 99 continued-access subjects • 5-yr enhanced surveillance study to characterize adverse events in intended population under general conditions of use	• 13 sites • 209 enrolled (103/106) • C-ADR: 98%; fusion: 95%	• 18–60 yr of age • Clinically and radiographically confirmed single-level disease, C3–4 to C6–7 • Radiculopathy (failed to benefit from 6 wk of conservative therapy) or myelopathy • NDI score ≥ 15 (30%) considered moderate disability	10% and 15% NI margin overall success rate on 4 criteria at 24 mo: 1. NDI score improvement from baseline of ≥ 15 points or 20% 2. Maintenance of or improvement in neurologic status (motor function, sensory function, reflexes) 3. No removal, revision, reoperation, or additional fixation required to modify implant 4. No adverse events related to treatment, disc, implantation, surgery, or implant or graft material	• SF-36 score • Radiographic assessments (device migration, device subsidence, disc height (decrease > 3 mm),[c] HO, radiolucency, ROM) • Work status • Medication use • Patient satisfaction	• Radiographic endpoints were evaluated independently by core lab and reviewed by independent radiologist	• Surgeons advised to prescribe appropriate rehabilitation programs • Use of hard or soft cervical collars left to discretion of physicians • Limitations placed on prolonged or strenuous activity for weeks or months. depending on patient's progress • NSAIDs were not used post-operatively

C-ADR Device, FDA Approval Date, Post-approval Request	Recruitment Period, No. of Sites, N Randomized (C-ADR/Fusion), % Followed Up at 2 Yr	Inclusion Criteria	Primary Endpoint/ Noninferiority Margin	Secondary Endpoints	Review Committees	Post-operative Management
Secure-C[40]						
September 28, 2012 • 7-year post-approval study of pre-market RCT subjects with 85% follow-up • 10-year enhanced surveillance study to characterize adverse events in intended population under general conditions of use in United States and elsewhere • Annual survey of physicians for information related to HO, device malfunction, device removal, or other serious device-related complications	• July 7, 2005–April 25, 2008 • 18 sites • 288 enrolled (148/140) + 88 additional patients (non-randomized C-ADR lead-in cases) • C-ADR: 46%; fusion: 31%	• 18–60 yr of age • Clinically and radiographically confirmed single-level disease, C3–4 to C6–7 • Radiculopathy (failed 6 wk of conservative therapy) or myelopathy • NDI score ≥ 15 (30%) considered moderate disability • Able to participate in post-operative management program	10% and 15% NI margin overall success rate on 4 criteria at 24 mo: 1. ≥ 25% improvement in NDI score from baseline 2. No device failure requiring revision, removal, reoperation, or supplemental fixation 3. No major complications, defined as major vessel injury, neurologic damage, or nerve injury 4. For fusion patients, radiographic fusion as defined by presence of bridging trabecular bone, without evidence of pseudarthrosis (defined radio-graphically as no apparent bridging trabecular bone, ROM > 3 mm in translation, and > 2º in rotation) FDA overall success: 1. NDI score improvement from baseline of ≥ 15 points 2. No secondary surgery at index site (revision, removal, reoperation, supplemental fixation) 3. No potentially device-related adverse events 4. Maintenance of or improvement in all components of neurologic status 5. No intraoperative C-ADR changes in treatment	• SF-36 score improvement of 15% • Radiographic assessments (disc height,[d] HO, lordosis, ROM) • Radiolucency • Operative measures (time, blood loss) • Work status • Medication use • Patient satisfaction	• Clinical events committee consisting of 2 practising spine surgeons • Radiographic endpoints evaluated by core central lab and reviewed by independent radiologist	• Use of external orthosis for 3 wk • Physical therapy program consisting of active ROM exercises • Restrictions on athletic activities, repetitive bending, and lifting for 6 mo • NSAIDs were not used pre- or post-operatively

Abbreviations: ASD, adjacent-segment disease; C-ADR, cervical artificial disc replacement; CEC, clinical events committee; FDA, U.S. Food and Drug Administration; FOSS, Functional Outcome Swallowing Scale; HO, heterotopic ossification; NDI, Neck Disability Index; NI, noninferiority; NSAID, nonsteroidal anti-inflammatory drug; ROM, range of motion; SF-12, 12-Item Short-Form Health Survey; SF-36, 36-Item Short-Form Health Survey; VAS, Visual Analogue Scale.

[a]Functional spinal unit height loss was defined as decrease in height of ≥ 2 mm.

[b]If electrical bone growth stimulators were used, they were classified as supplemental fixation.

[c]Decrease in disc height was considered to be > 3 mm.

[d]Decrease in disc height was considered to be > 2 mm.

Short-Term Treatment Outcomes

Two-Year Overall Treatment Success

Ten FDA regulatory trials evaluated C-ADR versus fusion for cervical degenerative disc disease. Overall treatment success (OTS) at a 2-year follow-up was the primary outcome defined by the FDA and the investigators as a composite outcome based on several subcriteria involving safety, technical success, secondary surgeries, and reduction of symptom-related disability. Definitions and thresholds for these subcriteria differed across studies and between the FDA and trial investigators. In each trial, the primary outcome was defined as a successful treatment outcome when patients were found successful on each subcriterion.

Subcriteria in the various trials included various combinations of NDI scores (improved ≥ 15 points or ≥ 20% over baseline); neurologic status (maintained or improved); safety (free from serious adverse events related to the device or surgical procedure); and technical failure (free from subsequent surgery at the index site [site of first surgery] performed for technical failures). Radiographic subcriteria involving spinal height, fusion success, or development of heterotopic ossification (bridging bone across the intervertebral space, limiting motion) in the C-ADR arm were included in some trials but not others. In the FDA RCT of ProDisc-C,[57] a second primary outcome involving additional criteria of patient satisfaction and reduction in use of strong narcotics (Schedule 2 drugs) was used to evaluate noninferiority and superiority.

The overall GRADE assessment of results on 2-year OTS was moderate, rated down for risk of bias (Appendix 4, Table A5).

Table 4 summarizes the 2-year mean proportion of OTS for the C-ADR and fusion groups in the RCTs. Figure 2 is a graphical display of the difference in mean proportion of OTS for C-ADR versus fusion for each C-ADR device and the 10% noninferiority threshold. Most people in each study arm achieved OTS, ranging from 74% to 90% for the different C-ADR devices and from 61% to 74% for fusion.

The noninferiority status of C-ADR compared with fusion for OTS was demonstrated in all studies by one of two statistical approaches, Bayesian or frequentist. In the FDA RCT of ProDisc-C,[22] OTS for C-ADR was further assessed with an alternative subcriterion that included patient satisfaction with surgery and reduced use of strong narcotics (Schedule 2 drugs) and muscle relaxants. Radiographic subcriteria were not always included in OTS estimates because of limitations in imaging investigations and loss of information. However, the FDA trial of Prestige-LP[37] that included radiographic information showed a lower proportion of OTS in both surgery groups (lower for ProDisc-C than for fusion). Statistical superiority to fusion was not consistently demonstrated across C-ADR devices.

For RCTs evaluating two C-ADR devices for two-level C-ADR versus fusion, both the Mobi-C[41] and the Prestige-LP[58] C-ADR devices achieved statistically significant noninferiority to fusion.

Table 4: Statistical Noninferiority and Superiority Status of Two-Year Overall Treatment Success of C-ADR Versus Fusion for Cervical Degenerative Disc Disease

C-ADR Device Author No. of Sites	Two-Year Overall Treatment Success[a] C-ADR P_C (n/N) (95% CI)	Fusion P_F (n/N) (95% CI)	$Pc - P_F$ (95% CI)	10% Noninferiority Margin[b]	Superiority[c]
One-level disease					
Bryan Heller et al,[20] 2009 30 sites	82.6% (190/229) (77.1%–87.3%)	72.7% (141/194) (65.8%–78.8%)	9.9% (2.0%–17.9%)	$P < .001$	$P = .010$
Kineflex Coric et al,[36] 2011 21 sites	84.9% (101/119)	71.3% (82/115)	13.6% (3.1%–24.1%)	$P = .05$	$P = .009$
Mobi-C Hisey et al,[38] 2014 23 sites	73.7% (115/156)	65.3% (49/75)	8.2% (−2.44% to NR)	$P = .0021$	$P = .2162$
PCM Phillips et al,[56] 2013 24 sites	75.1% (142/189) (68.5%–80.8%)	64.9% (98/151) (57.3%–72.5%)	10.2% (2.0%–18.5%)	$P < .0001$	$P = .020$
Prestige-LP[e] Gornet et al,[59] 2015 20 sites (WOSH)	79.3% (215/271) (73.6%–83.8%)	66.8% (147/220) (61.3%–74.1%)	10.1% (2.7%–19.6%)	P_B[d] ≈ 1.00	P_B ≈ .995
(WSH)	70.4% (159/226) (62.3%–74.9%)	63.2% (108/171) (57.9%–73.1%)	3.1% (−7.0% to 13.5%)	$P_B = .995$	$P_B = .736$
Prestige-ST Mummaneni et al,[21] 2007 2 sites (WOSH)	79.3% (177/223)	67.8% (134/198)	11.5% (NR)	$P = .0001$	$P = .0053$
ProDisc-C Murrey et al,[22] 2009 13 sites	77.2% (78/101)	74.3% (75/101)	2.9% (−8.8% to 14.8%)	$P = .0017$	NS
(OTS, satisfaction, narcotics)	73.5%	60.5%	13% (−0.3% to 26.1%)	P_B ≈ 1.00	NS
Secure-C Vaccaro et al,[40] 2013 18 sites (Investigator, OTS, NDI ≥ 25%)	90.1% (127/141)	71.1% (81/114)	19% (8.2%–27%)	P_B ≈ 1.00	P_B ≈ 1.00
(FDA OTS, NDI ≥ 15 points, no surgery change)	83.8% (109/130)	73.2% (82/112)	10.6% (0.6%–20.2%)	P_B ≈ 1.00	P_B ≈ .98
Two-level disease					
Mobi-C Davis et al,[41] 2013 24 sites	69.7% (154/221)	37.4% (37/99)	32.3% (22.8%–43.6%)	$P = .0001$	$P = .0001$
Prestige-LP Gornet et al,[42] 2017 30 sites	81.4% (162/199)	69.4% (111/160)	12.0% (2.3%–20.1%)	P_B ≈ 1.00	$P_B = .993$

Abbreviations: C-ADR, cervical artificial disc replacement; CI, confidence interval; FDA, U.S. Food and Drug Administration; NDI, Neck Disability Index; NR, not reported; NS, not significant; OTS, overall treatment success; P_B, Bayesian probability; P_C, proportion of C-ADR patients; P_F, proportion of fusion patients; WOSH, without spinal unit height change; WSH, with spinal unit height change.

[a]Overall treatment success is based on achieving success in several subcriteria including neck disability, technical success, effectiveness, and safety.

[b]Noninferiority of treatment success for C-ADR versus fusion is met when mean difference in proportions does not exceed an a priori defined 10% difference (noninferiority margin).

[c]Superiority is met if lower limit of 95% CI of difference in treatment proportions does not include zero.

[d]For Bayesian estimation of OTS, posterior probability that OTS of C-ADR was not inferior to fusion by more than 10% was > 95%.

[e]The Prestige-LP[37] one-level study was a prospective, controlled clinical trial with propensity-adjusted fusion subjects selected from the previous FDA RCT of Prestige-ST[21]; it employed Bayesian logistic regression methods.

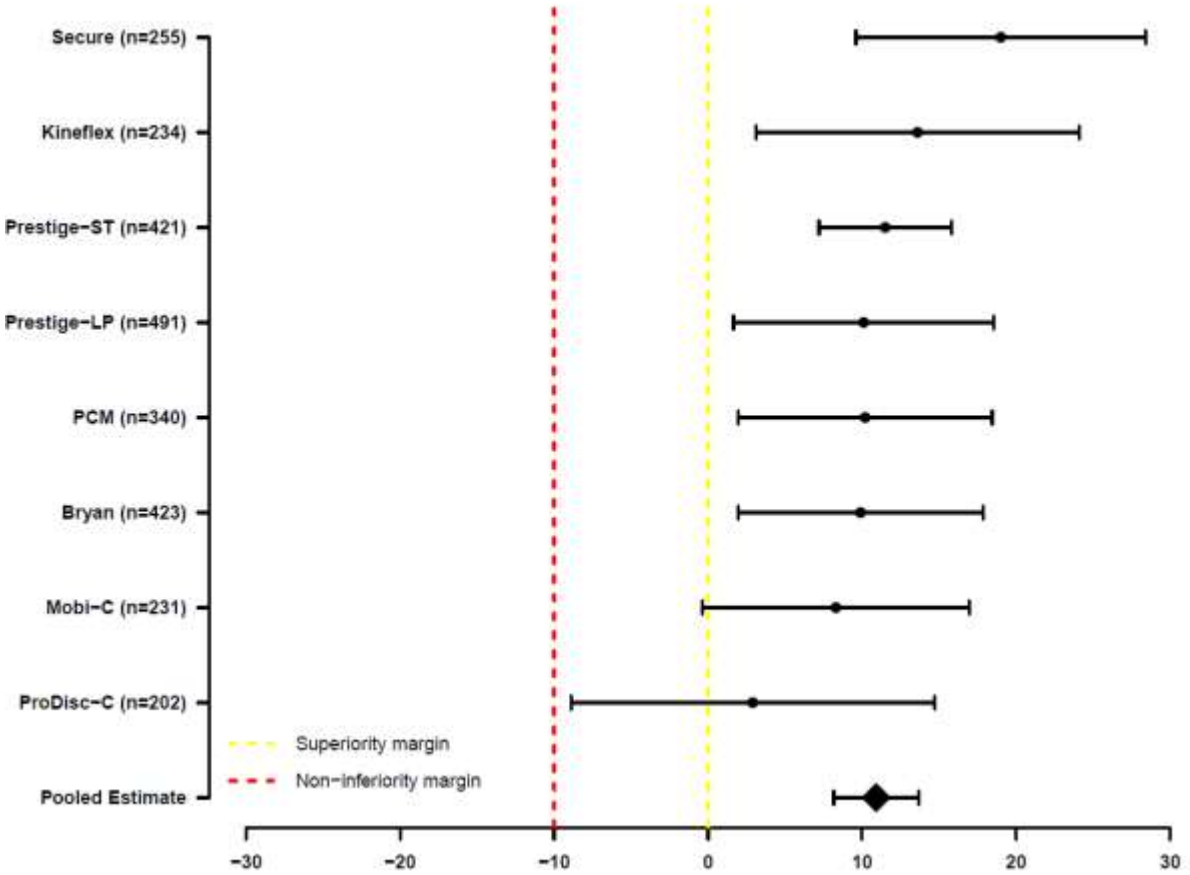

Figure 2: Graphical Display of Two-Year Mean OTS for One-level Disease—Proportion of C-ADR Versus Fusion OTS and 10% Noninferiority Margin

Sources: Heller et al,[20] 2009; Coric et al,[36] 2011; Hisey et al,[38] 2014; Phillips et al,[56] 2013; Gornet et al,[59] 2015; Mummaneni et al,[21] 2007; Murrey et al,[22] 2009; Vaccaro et al,[40] 2013.

Two studies compared longer-term differences in the OTS subcriteria of C-ADR versus fusion for one-level and two-level indications.[60,61] Clinical outcomes at 4 years were compared for indication-matched patients undergoing C-ADR in the FDA RCT of one-level Mobi-C versus the FDA RCT of two-level Mobi-C.[60] Baseline reduction in NDI (37.5 vs. 36.5), treatment satisfaction (88.6% vs. 85.0%), and major complications (4.3% vs. 4.0%) did not differ between patients undergoing one-level or two-level C-ADR.[60] There were also no differences between groups for secondary surgeries at the index level (3.0% vs. 4.0%) through a 4-year follow-up.[60] Rates of clinically relevant heterotopic ossification (Grades 3 and 4) were also similar between the groups: 23.8% (7.9% Grade 4) for one-level and 25.7% (10.2% Grade 4) for two-level C-ADR surgeries.[60]

Longer-term clinical outcomes at 5 years were also compared for patients undergoing fusion, who were also indication-matched patients, in the FDA RCTs of one-level Mobi-C and 2-level Mobi-C.[61] Patients undergoing fusion in both trials had significant reductions in mean NDI over baseline, but mean NDI values were significantly better for patients having one-level fusion than those having two-level fusion at all follow-up points ($P < .05$). Patients having one-level fusion were more satisfied with their treatment than those having two-level fusion (83.9% vs. 75.0%), but they were equally likely to recommend the surgery to friends with the same indication

(78.6% vs. 76.3%). Secondary surgeries increased in both groups from 2 years to 5 years: 6.2% to 11.1% for one-level patients and 11.4% to 16.2% for two-level patients. Fusion failure rates were statistically significantly lower ($P < .05$) for one-level versus two-level surgeries at 6 months (39.1% vs. 55.3%), 1 year (17.4% vs. 36.2%), and 2 years (10.7% vs. 20.2%) and remained lower at 3 years (9.7% vs. 13.3%), 4 years (6.6% vs. 15.0%), and 5 years (6.7% vs. 13.9%).

Neck Disability Index

A greater proportion of patients in the C-ADR group than in the fusion group achieving a clinically relevant reduction of more than 15 points in their NDI score over baseline was a driver (higher rate for C-ADR) in the OTS of one-level involvement for three C-ADR devices: 4.3% more patients with Secure, 7% more patients with Bryan[38], and 7% more patients with Prestige-LP.[37] For trials involving two-level degeneration, a greater proportion of patients having a statistically significant reduction in their NDI score was a subcriterion driver for the OTS of both Prestige-LP[42] (8.7% more patients) and Mobi-C[41] (16.4% more patients).

Patients with myelopathy undergoing C-ADR or fusion were evaluated in two studies.[62,63] Cheng et al[62] reported an RCT evaluating patients with cervical myelopathy undergoing one-level or multi-level (n = 17/41) treatment with C-ADR using the Bryan device or fusion. Neck disability and HRQOL scores statistically significantly improved in both surgery groups at a 3-year follow-up. Differences between surgery groups were not significantly different for mean NDI scores but were statistically significantly better for C-ADR than fusion for mean SF-36 physical component summary (PCS) scores after a 2-year follow-up. The median return-to-work time, however, was statistically significantly shorter for the C-ADR group than for the fusion group (20 vs. 84 days, $P < .01$). Neurologic assessments were not reported.

Riew et al[63] compared the effectiveness of C-ADR versus fusion for treating people with myelopathy participating in two RCTs of C-ADR devices (Bryan, Prestige-ST) at a single site with a 2-year follow-up. Of the 1,007 patients randomized and treated at one cervical level, 199 (19.7%) had evidence of cervical myelopathy. Patients in all groups had significant improvements in their NDI scores, but more patients in the C-ADR group using the Bryan device experienced statistically significantly better reductions in NDI score than in the fusion group at the 2-year follow-up.

Neurologic Status

Neurologic status was defined and evaluated differently across trials. In some trials, any deterioration in neurologic status was the outcome, whereas in other trials, neurologic testing involved various motor, reflex, and sensory parameters, and various degrees of change for neurologic criteria were rated as neurologic decline. Neurologic success, defined as neurologic status reported as either maintained or improved at a 2-year follow-up, was a subcriterion driver (higher rate for C-ADR) for four C-ADR devices. The proportion of patients achieving neurologic success was higher than fusion for four C-ADR devices: 4.0% more for Bryan[20], 5.4% more for PCM[56], 8.5% more for Prestige-ST[21], and 9.7% more for Prestige-LP.[37]

Neurologic status was reported separately as either maintained or improved in two trials.[20,21] Neurologic status that was not maintained could be interpreted as a complication (symptom worsening or a side effect); neurologic status that improved could reflect improvement in the original presenting symptoms. In the FDA RCT of Bryan,[20] maintenance of (30% for C-ADR vs. 35% for fusion) and improvement in (66% for C-ADR vs. 61% for fusion) neurologic status was

similar between surgery groups. The relationship of baseline neurologic status to any changes in neurologic status after surgery was not evaluated.

In the FDA RCT of Prestige-ST,[21] neurologic function was maintained or improved at 2 years and was similar in both surgery groups: 89.8% (95% CI 77.8%–96.6%) for C-ADR versus 81.1% (64.9%–92.0%) for fusion.

Gait assessment was also performed in these two trials.[20,21] No patients in any study group experienced gait deterioration, and gait improvement was greater with C-ADR than with fusion in both trials. In the FDA RCT of Prestige-ST,[21] gait improved in 47.9% (95% CI 33.3%–62.8%) of the C-ADR group and 37.8% (95% CI 22.5%–55.2%) of the fusion group. In the FDA RCT of Bryan,[20] gait improved in 46.2% (95% CI 30.1%–62.8%) of the C-ADR group and 26.7% (95% CI 12.3%–45.9%) of the fusion group. The authors noted that the neurologic function and gait outcomes in this study apply to myelopathy due to anterior cord compression and not to posterior cord compression (behind the vertebral body).

Second Index-Level Surgeries

Rates of secondary surgeries at the index site needed to address technical failures were generally low for both surgery groups, but index-level secondary surgeries were subcriterion drivers (lower rate for C-ADR) of higher OTS for four C-ADR devices. Fewer patients in the C-ADR group needed a second index-level surgery: 5% fewer for Mobi-C,[38] 5.4% fewer for Secure-C, 6.6% fewer for ProDisc-C[40], and 17% fewer for Prestige-ST.[21] In the FDA RCT of ProDisc-C,[22] an alternative subcriterion of OTS (use of opioids and anti-inflammatories, classified as Schedule 2 drugs) resulted in 8.4% fewer patients in the C-ADR group using medications at the 2-year follow-up.

For two-level procedures, second surgeries at the index site were subcriterion drivers of higher OTS for both C-ADR devices compared with fusion (Table 5). There were 5.6% fewer second index-level surgeries for Prestige-LP[37] and 8.3% fewer for the Mobi-C[38] device at a 2-year follow-up.

Table 5: Second Index-Level Surgeries for Technical Failures at Two Years

C-ADR Device Author, Year	No. of Patients (C-ADR/Fusion)	Percentage of Patients Not Needing a Second Index-Level Surgery[a]		
		% C-ADR	% Fusion	% C-ADR − % Fusion
One-level disease				
Bryan Heller et al,[20] 2009	229/194	97.5	96.4	1.1%
Kineflex Coric et al,[36]2014	119/108	94.2	93.4	0.8%
Mobi-C Hisey et al,[38] 2014	164/81	98.8	93.8	5%
PCM Phillips et al,[56] 2013	189/149	94.2	93.4	0.8%
Prestige-ST Mummaneni et al,[21] 2007	223/198	96.0	79.0	17%
Prestige-LP Gornet et al,[59] 2015	280 /265	95.0	92.1	2.9%
ProDisc-C Murrey et al,[22] 2009	103/106	98.1	91.5	6.6%
Secure Vaccaro et al,[40] 2013	145 /133	97.9	92.5	5.4%
Two-level disease				
Mobi-C Davis et al,[41] 2013	225/105	96.9	88.6	8.3%
Prestige-LP Gornet et al,[42] 2017	209/188	97.6	92.0	5.6%

Abbreviation: C-ADR, cervical artificial disc replacement.
[a]Second index-level surgeries for technical failures included removal, revision, re-operation, and supplemental fixation.

Radiologic Outcomes

Radiographic evaluations involving measures of treatment failure for both C-ADR and fusion in the FDA RCTs were conducted at central laboratories and by independent radiologists. Radiologic criteria for technical success were variably defined and reported in the trials.

For the FDA trials, motion preservation at the index site for C-ADR devices was evaluated as the degree of cervical flexion/extension range of motion (ROM) on plain radiographs and defined as maintaining a ROM of 4° or higher. For C-ADR, the formation of heterotopic ossification prevents motion, resulting in a naturally fused site and is considered a failed disc implant. This condition does not typically require additional surgery unless symptoms develop. Evaluation of heterotopic ossification was usually based on the radiographic classification system by the McAfee[64] index, which relates the degree of bridging bone to the loss of vertebral motion at that level. Clinically relevant values of the index are Grade 3 (some loss in motion) and Grade 4 (complete bridging and loss of motion).

The mean angles of cervical flexion/extension ROM, percentage achieving more than 4° of flexion/extension ROM, and incidence of heterotopic ossification 2 years after C-ADR are detailed in Table 6. Mean angles for all C-ADR devices were above the 4° threshold for both one-level and two-level treated cervical segments. The percentage of patients achieving the threshold for angular ROM varied among devices from 83% (Kineflex)[36] to 97% (Mobi-C).[38] Grade 4 heterotopic ossification, ranging from 1% to 5% for one-level treated disease, was reported for all C-ADR devices at a 2-year follow-up. For two-level treated disease, rates of heterotopic ossification were similar to rates for one-level treated disease.

In two reports, Coric et al[65,66] compared radiologic success in terms of angular ROM for three C-ADR devices (Bryan, Kineflex-C, Discover) in three separate FDA regulatory RCTs at their site. Both Bryan and Kineflex-C are semiconstrained designs, whereas Discover is an unconstrained design. In the first report,[65] involving 98 patients (57 with C-ADR), the change in mean angular ROM from before surgery to a 2-year follow-up varied by device. Coric et al[65] reported that cervical angles increased 5.1º (7.9º–13.0º) with Kineflex-C, increased 0.9º (7.6º–8.5º) with Bryan, and decreased 1.3º (8.2º–6.9º) with Discover.

At a 4-year follow-up, Kineflex-C continued to preserve angular ROM; angular motion for Discover was not reported.[66] The percentage of patients maintaining more than 2º of angular motion was 89.5% with Kineflex-C and 77.3% with Bryan.[66] At all follow-up points, the mean angular ROM was less than 1º for the fusion group.[66]

Table 6: Cervical Range of Motion at Two Years

C-ADR Device Author, Year	No. Patients	Cervical Flexion/ Extension ROM		Cervical Flexion/ Extension ROM at 2 yr	Heterotopic Ossification at 2 yr[a]
		Baseline Degree Mean ± SD	2-Yr Degree Mean ± SD	% > 4º	% With Grade 4
One-level disease					
Bryan Heller et al,[20] 2009	229	6.4º ± 3.4	7.9 ± 4.7	NR	0
Discover[b] Skeppholm et al,[67] 2015	28	--	5.1º ± 3.8	NR	5.0
Kineflex-C Coric et al,[66] 2014	119	8.2º	9.8º	83.0	1.0
Mobi-C Hisey et al,[38] 2014	156	8.2º ± 4.5	10.8º ± 6.5	97.0	3.0
PCM Phillips et al,[56] 2013	189	8.0º ± 4.5	6.2º ± 4.0	98.9	1.1
Prestige-ST Mummaneni et al,[21] 2007	223	7.6º	7.6º	99.2	0.8
Prestige-LP Gornet et al,[59] 2015	271	8.5º ± 6.1	7.5º	94.0	5.9
ProDisc-C Murrey et al,[22] 2009	101	8.5º	9.4º ± 6.1	84.4	2.9
Secure Vaccaro et al,[40] 2013	151	8.5º ± 4.8	9.7º	84.6	9.0
Two-level disease					
Mobi-C Davis et al[41] 2013	221				
Superior level		9.1º ± 4.9	10.1º ± 5.9	96.4	3.7
Inferior level		7.4º ± 4.3	8.3º ± 5.3	—	2.8
Prestige-LP Gorne et alt[42] 2017	199				
Superior level		6.8º ± 4.2	6.9º ± 3.9	69.5	2.0
Inferior level		6.9º ± 4.0	5.6º ± 3.9	64.6	3.0

Abbreviations: C-ADR, cervical artificial disc replacement; NR, not reported; ROM, range of motion; SD, standard deviation.
[a]Heterotopic ossification at Grade 4 level based on McAfee classification is bridging bone between vertebral endplates and no ROM.
[b]In the Skeppholm et al study of Discover, C-ADR radiographic measures were assessed with three-dimensional computed tomography, and patients had one-level (n = 18) and two-level (n = 10) disc placement.
[c]Radiographic success for motion was defined as 2º or more flexion/extension ROM, and success had to be evident at both levels.

For fusion, radiologic success is defined as absence of motion at the index site, the opposite of success for C-ADR. Fusion radiologic success was defined as either less than 2º or less than 4º of cervical flexion/extension ROM at the treated cervical site. Mean angles of ROM were not reported for the fusion study group, but the percentage of patients achieving the radiologic threshold for fusion success was high (> 90%) in all trials (Table 7). In FDA RCTs involving two-level disease, radiologic success was reported differently across trials. Fusion success for Mobi-C[41] was reported as both levels achieving the motion threshold (79.8%), whereas success for

Prestige-LP[42] was reported separately for each level and was higher at the superior than at the inferior treated level (95.4% vs. 85.6%).

Table 7: Fusion Radiologic Success at Two Years

Fusion Arm of C-ADR RCT: C-ADR Device Author, Year	No. Fusion Patients	Cervical Flexion/Extension ROM		
		Baseline Degrees Mean ± SD	2-yr Degrees Mean ± SD	Flexion Extension at 2 yr % < 4º
One-level disease				
Bryan Heller et al,[20] 2009	194	6.4º ± 3.4	NR	95%
Discover Skeppholm et al,[67] 2015		—	NR	
Kineflex Coric et al,[66] 2014	115	8.2º	NR	97%
Mobi-C Hisey et al,[38] 2014	81	7.5º	NR	89.3%
PCM Phillips et al,[56] 2013	150	8.0º ± 4.5	NR	92%
Prestige-ST Mummaneni et al,[21] 2007	198	7.6º	NR	97.5%
Prestige-LP Gornetet al,[59] 2015	219	8.5º ± 6.1	NR	
ProDisc-C Murrey et al,[22] 2009	106	8.5º	NR	90.2%
Secure Vaccaro et al,[40] 2013	140	--	NR	89.1
Two-level disease				
Mobi-C Davis et al,[41] 2013	105			
Superior level		9.3º ± 4.9	NR	79.8%
Inferior level		7.1º ± .3.9	NR	
Prestige-LP Gornet et al,[42] 2017	159			
Superior level		6.8º ± 4.2	NR	95.4%
Inferior level		6.9º ± 4.0	NR	85.6%

Abbreviations: C-ADR, cervical artificial disc replacement; NR, not reported; RCT, randomized controlled trial; ROM, range of motion; SD, standard deviation.

Short-Term Safety

Major Adverse Events

The avoidance of major adverse events was a key subcriterion in the assessment of OTS. Definitions and judgment of what constituted serious adverse events, however, varied across the C-ADR trials. Most trials had a Clinical Events Committee adjudicate on the seriousness of events and the likelihood that they were related to the device or to the surgery. Most patients in

both surgery groups avoided serious adverse events in the 2-year follow-up reports (Table 8). The likelihood of avoiding serious adverse events was higher, but not statistically significantly so, for all C-ADR devices than for fusion. The difference in the rate of serious adverse events between surgery groups was a driver (lower for the C-ADR devices) in OTS, with 6.4% fewer patients in the C-ADR surgery groups using the Kineflex device and 7% fewer patients using the Mobi-C device than fusion having major adverse events at 2-year follow-up.

Wound- and Hardware-Related Adverse Events

Rates of surgery-related adverse events (wound-related and hardware-related) were extracted from the systematic review by Anderson et al,[33] who reported on adverse events in detail from the FDA Summary of Effectiveness and Safety Reports for C-ADR versus fusion RCTs (Table 8).

Wound-related adverse events (infections and hematomas) included all wounds reported. Anderson et al[33] reported that wounds were mainly superficial and did not require incision or drainage in either group, and no implant infections required removal. Wound rates were generally less than 10% and were lower for fusion than C-ADR. The FDA trial of Prestige-LP was an outlier, with much higher wound rates for both the C-ADR (21%) and fusion (15%) groups. Wound rates for two-level cervical treatments were similar for the Mobi-C (4.9%) and fusion (5.7%) groups.

Anderson et al[33] also reported on hardware-related adverse events, which were defined as technically related events, such as those occurring during insertion, malpositioning with subsequent subsidence (sinking), or migration. These rates were low (< 3%) in one-level fusion surgery. Hardware-related adverse events in two trials were higher for C-ADR (6.4% for Prestige-LP; 11.7% for PCM) than for fusion. Hardware-related adverse event rates for other C-ADR devices were low (< 2%). The adverse event rate increased with the Mobi-C device from one- to two-level disease for fusion (3.7%–9.5%) but not for C-ADR (3.7%–0.9%).

Table 8: Adverse Events—C-ADR Versus Fusion for Cervical Degenerative Disc Disease

C-ADR Device Author, Year	Free From 2-Yr Serious Device- or Procedure-Related Adverse Events (%)			Wound-Related Adverse Events[a] (%)			Hardware-Related Adverse Events[a] (%)		
	C-ADR	Fusion	Δ C-ADR–Fusion	C-ADR	Fusion	Δ C-ADR–Fusion	C-ADR	Fusion	Δ C-ADR–Fusion
One-level disease									
Bryan Heller et al,[20] 2009	98.3	96.8	1.5	7.0	4.5	2.5	0.83	0.45	0.38
Kineflex Coric et al,[36] 2014	95.0	88.6	6.4	NR	NR	NR	NR	NR	NR
Mobi-C Hisey et al,[38] 2014	92.1	85.2	7.0	4.3	4.9	−0.6	3.7	3.7	0
PCM Phillips et al,[39] 2013	100	98.7	1.3	2.6	6.1	−3.5	11.7	1.6	10.1
Prestige-ST Mummaneni et al,[21] 2007	93.8	95.8	−2.0	9.8	7.6	2.2	1.5	1.1	0.4
Prestige-LP Gornet et al,[59] 2015	87.9	84.5	3.4	21.0	15.0	6.0	6.4	1.9	4.5
ProDisc-C Murrey et al,[22] 2009	97.1	93.4	3.7	2.9	2.8	0.1	0.0	1.9	−1.9
Secure Vaccaro et al,[40] 2013	100	100	0.0	0.0	4.2	−4.2	0.0	0.68	−0.68
Two-level disease									
Mobi-C Davis et al,[41] 2013	87.6	72.4	15.2	4.9	5.7	−0.8	0.89	9.5	−8.6
Prestige-LP Gornet et al,[42] 2017	98.0	94.0	4.0	NR	NR	NR	NR	NR	NR

Abbreviations: Δ, difference; C-ADR, cervical artificial disc replacement; NA, not applicable.
[a]Events extracted from Anderson et al.[33]

Dysphagia

Dysphagia (difficulty swallowing) often occurs after anterior spine cervical surgery, either C-ADR or fusion, and is considered a serious complication if it does not resolve over time. Five RCTs compared dysphagia after C-ADR and fusion for one-level cervical degenerative disc disease (Table 9).[53,62,68-71] In all trials, dysphagia was assessed by non-instrumental assessment tools: the Bazaz dysphagia severity questionnaire for North American trials[68-70] or the Dysphagia Short Questionnaire (DSQ) for Scandinavian trials.[53,71]

In the FDA RCT of Bryan,[70] fewer patients in the C-ADR group than in the fusion group reported symptoms of dysphagia at 2 years, and at a 7-year follow-up, rates were statistically significantly lower for patients in the Bryan group (0/15 vs. 28% (5/18), *P* = .04). In a second trial[62] involving the Bryan device, dysphagia was assessed in patients with myelopathy. Although rates were again lower in the C-ADR group than in the fusion group (2.4% [n = 1] vs. 16.7% [n = 7]), neither the severity of dysphagia nor the method of assessment was reported, and the number of cases was low.[62]

In a report on ProDisc-C[69] evaluating dysphagia at 2 of the 14 FDA RCT sites, the incidence of dysphagia at 12 months was statistically significantly (*P* = .01) lower for C-ADR than for fusion (15.6% [n = 6] vs. 42.1% [n = 16]). In the full FDA multicentre trial,[22] dysphagia events reported for this subset of patients were under-reported: Only one event in the C-ADR group and two events in the fusion group were reported.

In the FDA RCT of PCM,[68] the incidence of dysphagia was reported to gradually decrease over follow-up in both surgery groups. At a 2-year follow-up, dysphagia incidence was statistically significantly lower for the C-ADR than for the fusion group (14.9% vs. 27.5%, *P* = .04), although no cases of severe dysphagia were reported.

Two Scandinavian trials[53,71] evaluating dysphagia in RCTs of the Discover C-ADR reported median scores for the DSQ at a 2-year follow-up. Although the Norwegian trial[53] showed no difference in median dysphagia scores at any follow-up points, fusion used a stand-alone cage rather than an anterior plate. In the Swedish study,[71] fusion used plates; patients in the C-ADR group had statistically significantly lower median DSQ scores (0, range 0–6) than those in the fusion group (1, range 0–6; *P* = .04) and a statistically significantly lower odds ratio (OR 0.019; 95% CI 0.001–0.53; *P* = .02) for severe dysphagia.

Table 9: Dysphagia Following C-ADR or Fusion for Cervical Degenerative Disc Disease

C-ADR Device Author, Year Country	Study Sample C-ADR/Fusion Dysphagia Measurement	Post-operative Follow-Up Period	Dysphagia		*P* Value
			C-ADR	Fusion	
Bryan Smucker et al,[70] 2016 United States	1 site N = 22/25 BSS[a]	2 yr	9% (2/22)	12% (3/25)	*P* = .56
		5 yr (range 5.5–8.5 yr)	None of 15 patients completing questionnaire rated BSS > grade 1[a]	5 of 18 patients completing questionnaire rated BSS > grade 1[a]: grade 1 (n = 3), grade 3 (n = 2)	*P* = .042
Bryan Cheng et al,[62] 2011 China	1 site N = 41/42 Method not reported	3 yr	2.4% (1/41)	16.7% (7/42)	*P* = .057
ProDisc-C Segebarth et al,[69] 2010 United States	2 of 14 sites N = 45/42 Bazaz-Yoo dysphagia questionnaire[b]	1 yr	15.6% (6/38) reported some degree of dysphagia (> grade 0): mild (n = 2), moderate (n = 2), severe (n = 2)[b]	42.1% (16/38) reported some degree of dysphagia (> grade 0): mild (n = 9), moderate (n = 6), severe (n = 1)[b]	*P* = .01
PCM McAfee et al,[68] 201 United States	5 of 20 sites N = 151/100 BSS[a]	6 wk	44.6% (62/139) reported some degree of dysphagia (> grade 0): mild (n = 36), moderate (n = 26)[a]	57.8% (48/83) reported some degree of dysphagia (> grade 0): mild (n = 17), moderate (n = 27), severe (n = 4)[a]	*P* = .007
		3 mo	20.3% (24/118) reported some degree of dysphagia (> grade 0): mild (n = 15), moderate (n = 9)[a]	38.6% (27/70) reported some degree of dysphagia (> grade 0): mild (n = 18), moderate (n = 8), severe (n = 1)[a]	*P* = .008
		1 yr	14.1% (13/92) reported some degree of dysphagia (> grade 0): mild (n = 9), moderate (n = 4)[a]	29.0% (11/38) reported some degree of dysphagia (> grade 0): mild (n = 6), moderate (n = 4), severe (n = 1)[a]	*P* = .04
		2 yr	14.9% (10/67) reported some degree of dysphagia (> grade 0): mild (n = 8), moderate (n = 2)[a]	27.5% (8/29) reported some degree of dysphagia (> grade 0): mild (n = 4), moderate (n = 4)[a]	*P* = .04
Discover Sundseth et al,[53] 2017 Norway	5 sites N = 73/70 DSQ[c]	3 mo	Median 0 (range 0–9.0)	Median 1.0 (range 0–7.0)	*P* = .17
		6 mo	Median 0 (range 0–9.0)	Median 0 (range 0–12.0)	*P* = .62
		1 yr	Median 0 (range 0–9.0)	Median 0 (range 0–8.0)	*P* = .48
		2 yr	Median 1.0 (range 0–7.0)	Median 0 (range 0–8.0)	*P* = .16
Discover Skeppholm et al,[71] 2013 Sweden	3 sites N = 76 (22 at 2-level)/ 60 (16 at 2-level) DSQ[c]	4 wk	Median 2.5 (range 0–11)	Median 3.0 (range 0–9)	*P* = .24
		1 yr	Median 0 (range 0–6)	Median 1 (range 0–7)	*P* = .20
		2 yr	Median 0 (range 0–6)	Median 1 (range 0–6)	*P* = .04

Abbreviations: BSS, Bazaz severity scale; C-ADR, cervical artificial disc replacement; DSQ, Dysphagia Short Questionnaire.

[a]Bazaz severity scale: grade 1 (no symptoms); grade 2, mild (some difficulty in swallowing pills and food); grade 3, moderate (very difficult to swallow pills and hard foods); grade 4, severe (can't swallow pills or soft foods); grade 5, unbearable (requires hospital stay or readmission).

[b]Bazaz-Yoo dysphagia grading score: grade 0 (no difficulty with liquids or solids); grade 1, mild (no difficulty with liquids, rare with solids); grade 2, moderate (no or rare difficulty with liquids, occasional difficulty with specific solids); grade 3, severe (no or rare difficulty with liquids, frequent difficulty with most solids).
[c]Dysphagia Short Questionnaire consists of 5 items (ability to swallow, incorrect swallowing, globus sensation, involuntary weight loss, and pneumonia) with values ranging from 0 to 18. Total scores < 4 were classed as less severe and ≥ 4 as more severe.

Neck Disability Index in Non-regulatory Randomized Controlled Trials

Other RCTs of C-ADR versus fusion for cervical degenerative disc disease were non-IDE regulatory trials for five C-ADR devices (Bryan, Mobi-C, Prestige-LP, ProDisc-C, and Discover). All trials were conducted in Europe, Asia or Scandinavia: Bryan (China,[44,55,62] Netherlands[46]), Mobi-C (China[47,54]), Prestige-LP (Spain,[45] India[50]), ProDisc-C Germany[48,49,72,73]) and the Discover (China,[43,74] Croatia,[51] Sweden,[52,67,71,75] and Norway.[53,76])The Discover C-ADR device has only European regulatory approval.

Unlike the primary outcome of OTS employed in the FDA regulatory trials, the primary outcome for these trials was the NDI. Eight of these trials reported the mean NDI score, rather than the responder proportion, as the main outcome (Table 10). In all trials, patients in both surgery groups experienced statistically significant and clinically relevant improvements in neck disability at 2 years over their baseline levels. There were no significant differences between the surgery groups in NDI improvement at 2 years.

Table 10: Neck Disability Index Scores—C-ADR Versus Fusion in Nonregulatory Randomized Controlled Trials

C-ADR Device	Author, Year Country	No. of Patients Receiving C-ADR or Fusion	C-ADR NDI	Fusion NDI
Bryan	Zhang et al,[55] 2012 China	N = 60/60		
		Baseline,[a] mean ± SD	51.6 ± 7.2	54.5 ± 8.5
		2 yr, mean ± SD	14.9 ± 2.9	15.3 ± 3.8
		Mean change over baseline,[b] mean ± SD	Δ 36.9 ± 7.4	Δ 39.0 ± 5.9
	Cheng et al,[44] 2009[a] China	N = 31/34		
		Baseline, mean	50	51
		2 yr, mean	11	19
		Mean change over baseline[b]	Δ 39	Δ 32
Discover	Skeppholm et al,[52,67] 2015 Sweden	N = 83/70		
		Baseline, mean ± SD	64.6 ± 16.2	61.4 ± 14.2
		2 yr, mean ± SD	39.1 ± 20.2	40.1 ± 18.5
		Mean change over baseline[b]	Δ 25.5	Δ 21.3
	Sundseth et al,[53] 2017 Norway	N = 73/70		
		Baseline, mean (95% CI)	45.7 (42.9–48.6)	51.2 (48.0–54.4)
		2 yr, mean (95% CI)	25.0 (20.1–29.9)	21.2 (16.7–25.6)
		Mean change over baseline[b]	Δ 20.7	Δ 30.0
		≥ 10 points over baseline (%)	70	78
	Chen et al,[43] 2013 China	N = 16/16		
		Baseline, mean ± SD	47.8 ± 16.3	45.2 ± 13.7
		6 mo, mean ± SD	28.6 ± 6.5	21.4 ± 7.4
		2 yr, mean ± SD	16.5 ± 6.2	18.6 ± 6.7
	Rozankovic et al,[51] 2016 Croatia	N = 51/50		
		Baseline,[a] mean ± SD	50.9 ± 11.5	51.2 ± 8.6
		3 mo, mean ± SD	13.0 ± 5.2	19.8 ± 4.4
		2 yr, mean ± SD	11.6 ± 4.4	19.7 ± 6.0
		Mean change 2 yr over baseline[b]	Δ 39.3	Δ 31.5
Mobi-C	Zhang et al,[54] 2014 China	N = 55/56		
		Baseline,[a] mean	37.4	37.8
		1 yr, mean	19.0	19.2
		2 yr mean	19.0	19.3
		5 yr, mean ± SD	19.7 ± 8.1	18.5 ± 7.9
		Mean change 2 yr over baseline[b]	Δ 18.4	Δ 18.5
Prestige-LP	Pandey et al,[50] 2017 India	N = 17/17		
		Baseline,[a] mean	58.5	59.1
		6 wk, mean	29.9	39.6
		3 mo mean	21.5	33.9
		6 mo, mean	17.3	31.0
		1 yr, mean ± SD	13.6 ± 1.8	23.8 ± 2.4
		Mean change 1 yr over baseline[b]	Δ 44.9	Δ 35.3

Abbreviations: Δ, difference; C-ADR, cervical artificial disc replacement; CI, confidence interval; NDI, Neck Disability Index; SD, standard deviation.
[a]Cheng et al study involved patients with myelopathy.
[b]Crude NDI change over baseline was calculated by this author.

Perioperative Outcomes

One-Level Cervical Degenerative Disc Disease

Perioperative outcomes include operative time, blood loss, and length of stay for C-ADR versus fusion for one-level cervical disease (Table 11). Operative times for fusion groups were relatively constant, ranging from 1.2 to 1.4 hours, and those for C-ADR groups varied more and were longer. The 2.2-hour operative time for the Bryan device was longest.[20] Although operative times were generally longer for C-ADR, the longer time was unlikely to be of practical significance to operating room schedules given the low volume of these procedures. Estimated blood loss was generally not statistically significantly different. Two studies had opposite results; estimated blood loss in the FDA RCT of ProDisc-C[22] was statistically significantly lower (by 20 mL) for the C-ADR group, and in the FDA RCT of Secure,[40] estimated blood loss was statistically significantly higher (by 10 mL) for the C-ADR group. The length of hospital stay, from 1 to 2 days, was not significantly different between surgery groups.

Two-Level Cervical Degenerative Disc Disease

Operative outcomes including operative time, blood loss, and length of stay for C-ADR versus fusion for two-level cervical degenerative disc disease are also outlined in Table 11. Results are generally similar to results in one-level groups. Operative times are longer for C-ADR implants, but there are no statistically significant or clinically relevant differences in terms of blood loss or hospital length of stay between groups.

The overall GRADE assessment of results for perioperative outcomes was high (Appendix 4, Table A5).

Table 11: Perioperative Outcomes of C-ADR Versus Fusion for Cervical Degenerative Disc Disease

C-ADR Device No. Randomized (C-ADR/Fusion)	Total Operative Time Mean ± SD			Estimated Mean Blood Loss mL ± SD			Length of Hospital Stay Days ± SD		
	C-ADR	Fusion	*P* Value	C-ADR	Fusion	*P* Value	C-ADR	Fusion	*P* Value
One-level disease									
Bryan Heller et al,[20] 2009 N = 240/221	2.2 hr	1.4 hr	$P < .05$	91.5	59.6	NR	1.1	1.0	NS
Bryan Zhang et al,[55] 2012 N = 60/60	1.54 ± 0.49 hr	1.18 ± 0.27 hr	$P < .001$	40[a] (20, 75)	50[a] (30,70)	NS	3.32 ± 0.79	3.2 ± 1.02	NS
Kineflex-C Coric et al,[36] 2011 N = 136/133	80.2 ± 28.9 min	74.7 ± 26.9 min	NS	40.6 ± 30.5	41.1 ± 32.4	NS	2.1 ± 0.43	2.1 ± 0.51	NS
Mobi-C Hisey et al, [38] 2014 N = 164/81	1.5 ± 0.64 hr	1.3 ± 0.63 hr	NS	47.7 ± 46.8	48.1 ± 55.2	NS	2.1 ± 0.52	2.1 ± 0.47	NS
PCM Phillips et al,[56] 2013 N = 224/192	100.8 ± 42.0 min	85.7 ± 40.5 min	$P < .001$	65.6 ± 48.3	58.6 ± 46.1	NS	1.2 ± 0.6	1.4 ± 0.7	.024
Prestige-ST Mummaneni et al,[21] 2007 N = 276/265	1.6 hr	1.4 hr	$P < .001$	60.1	57.5	NS	1.1	1.0	.041
Prestige-LP[a] Gornet et al,[77] 2015 N = 280/256	1.49 hr	1.38 hr	$P = .015$	51	57.1	NS	0.9	0.95	NS
ProDisc-C Murrey et al,[22] 2009 N = 106/103	98.7 ± 47 min	107.2 ± 35.7 min	$P = .008$	63.5 ± 50.3	83.5 ± 64.9	$P_W = .009$	1.3 ± 0.83	1.4 ± 1.18	NS
Secure-C Vaccaro et al,[40] 2013 N = 151/140	87.7 ± 33.02 min	72.1 ± 25.4 min	$P < .0001$	55.2 ± 44.22	45.6 ± 33.21	$P_B = .025$	1.0 ± 0.6	0.9 ± 0.46	NS
Two-level disease									
Mobi-C Davis et al,[41] 2013 N = 225/105	2.2 ± 0.8 hr	1.8 ± 0.9 hr	$P = .0002$	67.2 ± 90.0	70.3 ± 78.8	NS	2.2 ± 0.5	2.4 ± 2.1	NS

C-ADR Device No. Randomized (C-ADR/Fusion)	Total Operative Time Mean ± SD			Estimated Mean Blood Loss mL ± SD			Length of Hospital Stay Days ± SD		
	C-ADR	Fusion	*P* Value	C-ADR	Fusion	*P* Value	C-ADR	Fusion	*P* Value
Prestige-LP Gornet et al,[42] 2017 N = 209/188	2.1 ± 0.8 hr	1.7 ± 0.7 hr	$P = .001$	67.2 ± 64.1	55.7 ± 46.3	$P_B = .019$	1.2 ± 0.5	1.3 ± 1.0	NS

Abbreviations: C-ADR, cervical artificial disc replacement; NR, not reported; NS, not significant; P_B, Bayesian probability; P_W, Wald probability; SD, standard deviation; SSED, Summary of Safety and Effectiveness Data.

Return to Work

One-Level Cervical Degenerative Disc Disease

Return-to-work times after surgeries are outlined in Table 12. Trials were conducted in the United States and Europe. Characteristics and work status of patients before their surgeries were reported inconsistently. Approximately half the study groups were female, and patients were generally in their mid-40s. The proportion of patients with Workers' Compensation status or involvement in spinal litigation, when reported, was approximately 10%. Return to work was estimated as the mean or median time to return at various follow-up points up to 2 years. For all devices, the estimated return to work was shorter for the C-ADR than for the fusion group, ranging from a median of 13 to 20 days: Bryan by 13 days (*P* = .015),[20] Discover by 2 weeks (*P* = .17),[53] Prestige-ST by 16 days (*P* = .022),[21] and Prestige-LP by 20 days (*P* =.02).[59]

In all trials, the 2-year return-to-work or employment rates were similar between the two surgery groups (C-ADR vs. fusion) for Bryan,[20] Prestige-ST,[21] Prestige-LP,[59] and ProDisc-C.[22] With ProDisc-C,[22] the percentage of patients reported to be involved in moderate to heavy physical work before surgery had declined similarly at 2 years in both surgery groups: from 57% to 48% for patients who received ProDisc-C and from 52% to 45% for patients who had undergone fusion. In the European trials, 2-year employment rates among patients receiving Discover were higher than among patients undergoing fusion in Sweden (91% vs. 85%).[52] In Norway, rates in the C-ADR group were lower than in the fusion group (60% vs. 72%, *P* = .17).[53]

Two-Level Cervical Degenerative Disc Disease

Return-to-work times after two-level surgery were also shorter for C-ADR than for fusion. Patients who received with the Mobi-C device[41] had a mean time shorter by 21 days, and patients who received the Prestige-LP[59] device had a median time shorter by 6 days. The 2-year employment rates were reported to be similar for the two surgery groups.

One study[78] involved a subgroup analysis of Workers' Compensation patients from the two surgery arms of two RCTs, one on Prestige-ST and the other on Bryan. Pre-operative employment rates were low in both surgery groups (36% for C-ADR vs. 33% for fusion), and litigation rates were 23% for C-ADR and 20% for fusion.[78] Female participation was higher in the C-ADR group than in the fusion group: 51% vs. 30%.[78] Although return-to-work times were shorter for the C-ADR than for the fusion group by 24 days (75 vs. 99 days), median return-to-work times varied greatly in both groups.[78] Median confidence intervals ranged from 26 to 121 days for the C-ADR group and from 87 to 191 days for the fusion group.[78] At a 2-year follow-up, return-to-work rates were also higher for the C-ADR group than for the fusion group (63% vs. 53%), although the difference was not statistically significant.[78]

The overall GRADE assessment of results on return to work was moderate, downgraded for imprecision (Appendix 4, Table A5).

Table 12: Return-to-Work Status After C-ADR or Fusion for Cervical Degenerative Disc Disease

C-ADR Device Author, Year Country	No. of Patients Age (Mean ± SD) % Female		Pre-operative Employment Status		Return-to-Work Outcome		Conclusions
	C-ADR	Fusion	C-ADR	Fusion	C-ADR	Fusion	
One-level disease							
Bryan Heller et al,[20] 2009 United States	• N = 242 • 44.4 ± 7.9 yr • 54.5% F	• N = 221 • 44.7 ± 8.6 yr • 48.9% F	• Employed 64.5% • Workers' Compensation 6.2% • Unresolved spinal litigation 2.5%	• Employed 65.0% • Workers' Compensation 5% • Unresolved spinal litigation 2.7%	• Median RTW time 48 d • 2-yr RTW 76.8%	• Median RTW time 61 d • 2-yr RTW 73.6%	• Median RTW was significantly quicker in C-ADR group (by 13 d, 48 d vs. 61 d) (*P* = .015) • 2-yr RTW rate was similar in surgery groups
Discover Skeppholm et al,[52] 2015 Sweden	• N = 81 • 46.7 ± 6.7 yr • 50.6% F	• N = 70 • 47.0 ± 6.9 yr • 52.8% F	• Unemployed 10% (n = 8) • Sick leave FT[a] 38% (n = 31) • Sick leave PT[a] 20% (n =16) • Other reason 7% (n = 6)	• Unemployed 14% (n = 10) • Sick leave FT[a] 36% (n = 25) • Sick leave PT[a] 17% (n =12) • Other reason 4% (n = 3)	Employed FT at: • 4 wk 46% • 3 mo 57% • 1 yr 71% • 2 yr 91%	Employed FT at: • 4 wk 37% • 3 mo 45% • 1 yr 62% • 2 yr 85%	• Employment rate was significantly higher in C-ADR group than in fusion group at all follow-up points • At 2-yr follow-up, employment rate was still higher in C-ADR group (by 6%, 91% vs. 85%)
Sundseth et al,[53] 2017 Norway	• N = 68 • 44.7 ± 7.2 yr • 52.9% F	• N = 68 • 43.4 ± 6.8 yr • 54.4% F	• Employed 20.6% (n = 14) • Median duration of sick leave[a] 21 wk (IQR 6–39 wk)	• Employed 25.8% (n = 17) • Median duration of sick leave[a] 24 wk (IQR 1–55 wk)	• Median duration of post-operative sick leave 10 wk (IQR 6–27 wk) • 2-yr RTW at 2 yr 59.7%	• Median duration of post-operative sick leave 12 wk (IQR 6–30 wk) • RTW at 2 yr 71.7%	• Median duration of post-operative sick leave was shorter in C-ADR group (by 2 wk, 10 vs. 12 wk), but difference was not significant (*P* = .17) • 2-yr RTW rate was lower in C-ADR group (by 12%, 59.7% vs. 71.7%), but difference was not significant (*P* = .16)
Mobi-C Hisey et al,[38] 2014 United States	• N = 164 • 43.3 ± 9.2 yr • 52.4% F	• N = 81 • 44.0 ± 8.2 yr • 55.6% F	• Employed 65.9% • NA 11.6%	• Employed 56.8% • NA 16%	• RTW mean time 30.1 d ± 24.6 d • RTW median time 21 d	• RTW mean time 36.8 d ± 40.3 d • RTW median time 22 d	• Mean RTW time was shorter in C-ADR group (by 6.7 d, 30.1 vs. 36.8 d), but difference was not significant

C-ADR Device Author, Year Country	No. of Patients Age (Mean ± SD) % Female		Pre-operative Employment Status		Return-to-Work Outcome		Conclusions
	C-ADR	Fusion	C-ADR	Fusion	C-ADR	Fusion	
PCM Phillips et al,[56] 2013 United States	• N = 218 • 45.3 ± 9.0 yr • 48.2% F	• N = 185 • 43.7 ± 8.3 yr • 48.1% F	• Workers' Compensation 11.9%	• Workers' Compensation 11.4%	NR	NR	
Prestige-ST Mummaneni, et al[21] 2007 United States	• N = 276 • 43.3 yr, range 25–72 yr • 53.5% F	• N = 265 • 43.9 yr, range 22–73 yr • 54% F	• Employed 66% • Workers' Compensation 11.6% • In litigation 10.9%	• Employed 63% • Workers' Compensation 13.2% • In litigation 12.1%	• RTW median time 45 d • 2-yr employment rate 75.4%	• RTW median time 61 d • 2-yr employment rate 74.7%	• Median RTW time was significantly shorter (by 16 d, 45 vs. 61 d) after C-ADR (*P* = .022), but median RTW time was not significantly different by log-rank time to event analysis • 2-yr employment rates were similar in surgery groups, and employment rates increased over baseline in both groups
Prestige-LP Gornet et al,[59] 2015 United States	• N = 280 • 44.5 ± 8.8 yr 44 yr (IQR 39–49 yr) • 53.9% F	• N = 265 • 43.9 ± 8.8 yr • 44 yr (IQR 38–49 yr) • 54% F	• Employed 67.1% (n = 188) • Workers' Compensation 11.4% (n = 32) • Unresolved spinal litigation 12.1% (n = 34)	• Employed 62.6% (n = 166) • Workers' Compensation 13.2% (n = 35) • Unresolved spinal litigation 10.9% (n = 32)	• RTW median time 40 d • 2-yr employment rate 73.4%	• RTW median time 60 d • 2-yr employment rate 75.9%	• Median RTW time was significantly shorter (by 20 d, 40 vs. 60 d) in C-ADR group (*P* = .02) • 2-yr employment rate was similar between surgery groups
ProDisc-C Murrey et al,[22] 2009 United States	• N = 106 • 42.1 ± 8.4 yr • 55.3% F	• N = 103 • 43.5 ± 7.1 yr • 46.2% F	• Employed 82.5% • Physical labour status (moderate to heavy work) 57.1%	• Employed 84.9% • Physical labour status (moderate to heavy work) 52.2%	• 2-yr employment rate 82.8% • 2-yr physical labour status 48.1%	• 2-yr employment rate 80% • 2-yr physical labour status 44.7%	• 2-year employment rate was similar between surgery groups • Percentage of patients involved in moderate or heavy physical work at 2 yr declined in both surgery groups
Secure Vaccaro et al,[40] 2013 United States	• N = 151 • 43.4 ± 7.5 yr • 47.2% F	• N = 140 • 44.4 ± 7.9 yr • 46.4% F	• NR	• NR	• RTW mean time 44 ± 74.5 d	• RTW mean time 50 ± 72.2 d	• Mean RTW time was shorter (by 6 d, 44 vs. 50 d) in C-ADR group

C-ADR Device Author, Year Country	No. of Patients Age (Mean ± SD) % Female		Pre-operative Employment Status		Return-to-Work Outcome		Conclusions
	C-ADR	Fusion	C-ADR	Fusion	C-ADR	Fusion	
Mixed C-ADR Devices (Prestige and Bryan) Steinmetz et al,[78] 2008 United States[b]	• N = 47 • 44.3 ± 6.5 yr • 51% F • (Workers' Compensation subgroup from 2 RCTs)	• N = 46 • 43.9 ± 8.3 yr • 30.4 % F • (Workers' Compensation subgroup from 2 RCTs)	• Employed 36.2% • In litigation 23.4% (n = 11)	• Employed 32.6% • In litigation 19.6% (n = 9)	• Median RTW time (all patients) 101 d (LCI 69 d, UCL 168 d) • Median RTW time (employed before surgery) 75 d (95% CI 26 – 121 d) • 2-yr RTW rate 63.2%	• Median RTW time (all patients) 222 d (LCL 125 d, UCL NSE) • Median RTW time (employed before surgery) 99 d (95% CI 87 – 191 d) 2-yr RTW rate 52.8%	• Median RTW time was shorter in C-ADR group (by 24 d, 75 vs. 99 d), but difference was not significant • 2-yr RTW rate was higher in C-ADR group (by 10.4%, 63.2% vs. 52.8%), but difference was not significant
Two-level disease							
Mobi-C Davis et al,[41] 2013 United States	• N = 225 • 45.3 ± 8.1 yr • 49.8% F	• N = 105 • 46.2 ± 8.0 yr • 46.2% F	• Employed 62.7% (n = 141) • Employment NA 15.1% (n = 34) • Workers' Compensation 4.9% (n = 11)	• Employed 61.0% (n = 64) • Employment NA 18.1% (n = 19) • Workers' Compensation 6.7% (n = 7)	• Mean RTW time (for working patients) 46 ± 101 d	• Mean RTW time (for working patients) 67 ± 113 d	• Mean RTW time was shorter in C-ADR group (by 21 d, 46 vs. 67 d), but difference was not statistically significant
Prestige-LP Gornet et al,[42] 2017 United States	• N = 209 • 47.1 ± 8.3 yr • Range 22–75 yr • 56% F	• N = 188 • 47.3 ± 7.7 yr • Range 25–69 yr • 52.1% F	• Employed 69.9% (n = 146) • Workers' Compensation 12.4% (n = 26) • Unresolved spinal litigation (n = 0)	• Employed 60.1% (n = 113) • Workers' Compensation 10.1% (n = 19) • Unresolved spinal litigation 0.5% (n = 1)	• Median RTW time 49 d • 2-yr RTW rate 72.9%	• Median RTW time 55 d • 2-yr RTW rate 71.1%	• Median RTW time was shorter in C-ADR group (by 6 d, 49 vs. 55 d), but difference was not statistically significant (P_{CPH} = .48) • 2-yr employment rates were similar between surgery groups

Abbreviations: C-ADR, cervical artificial disc replacement; CI, confidence interval; F, female; FT, full time; IQR, interquartile range; NA, not applicable; NE, not evaluated; NR, not reported; NSE, not statistically estimable; P_{CPH}, probability based on Cox proportional hazards ratio; PT, part time; RCT, randomized controlled trial; RTW, return to work.

[a]Sick leave included indications for neck pain or other medical conditions.

[b]Steinmetz et al report combined Workers' Compensation populations from both FDA RCTs of Prestige and Bryan. Use of post-operative orthosis differed by surgery groups: 35% in C-ADR and 62% in fusion group.

Health-Related Quality of Life

Table 13 outlines the health-related quality of life (HRQOL) of patients undergoing C-ADR or fusion for one- or two-level cervical degenerative disc disease with a 2-year prospective follow-up. Trials in the United States used the 36-item Short-Form Health Survey (SF-36) or the 12-item Short-Form Health Survey (SF-12), and the European trials used the SF-36 or the European Quality of Life instrument for measuring quality of life in five dimensions (EQ-5D) as the HRQOL measure.

One-Level Cervical Degenerative Disc Disease

In all C-ADR device trials, patients in both surgery groups significantly improved their HRQOL scores assessed as either changes over baseline in mean SF-36 or SF-12 score or in the percentage whose scores improved by 15% or more. Improvements were noted at 3 months for both the physical component summary (PCS) subscores and the mental component summary (MCS) subscores, and improvements remained stable throughout the 2-year follow-up period. Improvements in PCS subscores, however, were higher than in MCS subscores. Although mean PCS subscores improved more at early follow-up for C-ADR groups than for fusion groups, differences between groups were not statistically significant at a 2-year follow-up.

Two European trials[52,53] involving the Discover C-ADR device for one-level cervical degenerative disc disease assessed HRQOL status with the EQ-5D. Surgery groups in both trials achieved statistically significant improvements in mean EQ-5D scores over baseline. Improvements for both groups at a 2-year follow-up were also greater than a minimally clinically important difference of 0.24 for EQ-5D scores.

How cervical surgery affects HRQOL was also evaluated through the SF-36 in trials involving two patient subpopulations: those with myelopathy[63] and those with Workers' Compensation.[78] For the myelopathy subpopulation, statistically significant improvements in both the SF-36-PCS and SF-36-MCS subscores were achieved over baseline at a 2-year follow-up for both surgery groups.[63] In the study evaluating a subpopulation of patients with Workers' Compensation, both surgery groups achieved significant improvement over their baseline scores.[78] Patients in the C-ADR arm, however, achieved higher gains in both the SF-36-PCS and SF-36-MCS scores than patients in the fusion group (no statistical comparison was undertaken).[78]

Two-Level Cervical Degenerative Disc Disease

Two trials reported on HRQOL values after two-level surgery. In the FDA RCT of Mobi-C[41] for two-level cervical degenerative disc disease, HRQOL was assessed by the SF-12. Both surgery groups achieved statistically significant improvement ($P < .001$) in mean physical summary subscores over baseline. Improvements in mean SF-36-PCS subscores were statistically significantly greater ($P < .05$) in the C-ADR group than in the fusion group at all follow-up points. In the FDA RCT of Prestige-LP,[42] HRQOL was assessed by the SF-36, and the percentage achieving a clinically relevant change of a 15% improvement in scores over baseline was reported. A similar percentage achieved the clinically relevant improvement in scores: 69% for C-ADR and 72% for fusion surgery.

Three studies[79-81] employed subsections of the NDI, a disease-specific measure, to evaluate how C-ADR and fusion affect patients' HRQOL. Two studies[80,81] reported on headache relief after cervical surgery using the headache subscores of the NDI, and one study[79] used the driving subsection of the NDI to compare the effect of surgery on driving disability.

Headache scores on the NDI were rated 0 (no headaches at all), 1 (slight headaches that occur infrequently), 2 (moderate headaches that occur infrequently), 3 (moderate headaches that occur frequently), 4 (severe headaches that occur frequently), or 5 (headaches almost all the time). Headaches were not characterized further in terms of quality, location, or duration.

In the Schrot et al RCT[81] on Mobi-C versus fusion, almost all patients (82%) pre-operatively reported having headaches; 52% of patients reported NDI headache scores of 3 or more. Mean post-operative headache scores in both groups decreased over baseline: from 2.5 to 1.1 in the C-ADR group and from 2.4 to 1.2 in the fusion group. The lower headache score for the C-ADR group was not statistically significantly (P = .15) different from the fusion group.

In the second study evaluating headache relief, Liu et al[80] compared headache relief in one-level or two-level surgery with Mobi-C versus fusion. Preoperatively, 50% of the C-ADR group and 53% of the fusion group in the one-level study reported headache scores of 3 or more. At each follow-up, patients in both surgery groups experienced statistically significant (P < .0001) reduction in mean headache scores. At a 5-year follow-up, fewer patients reported headache scores of 3 or more in both surgery groups: 15.4% for C-ADR versus 10.5% for fusion. Improvements remained stable throughout the 5-year follow-up period, and between-group differences were not statistically significant.

For those undergoing two-level surgery, patients in the C-ADR arm experienced statistically significantly (P < .05) greater reductions in mean headache scores than patients in the fusion arm throughout follow-up.[80] At a 5-year follow-up, the C-ADR group had a greater reduction in mean headache score of 1.37± 1.68 compared with the fusion arm's reduction of 0.986 ± 1.37 (P = .08). The proportion of patients with headache scores of 3 or more also decreased from baseline: from 55.1% to 18.8% in the C-ADR arm and from 51.4% to 25.7% in the fusion arm.

In the study by Kelly et al,[79] the driving subsection of the NDI questionnaire was used to measure driving disability for patients in the RCT comparing Bryan versus fusion.[20] The driving subsection involved a 6-point severity scale: 0 (can drive without any neck pain), 1 (can drive as long as I want with slight neck pain), 2 (can drive as long as I want with moderate neck pan), 3 (can't drive my car as long as I want because of moderate neck pain), 4 (hardly drive at all because of severe neck pain), 5 (can't drive at all). For the study, driving disability was dichotomized as none or little (Grades 0, 1, or 2) or moderate to severe (Grades 3, 4, and 5).

Before surgery, 49.6% (67 patients) reported moderate or severe driving disability, and mean driving disability scores were similar for patients in the C-ADR (2.6 ± 1.0) and fusion (2.5 ± 1.1) groups.[79] Throughout the 2-year follow-up, both groups experienced a statistically significant (P < .001) improvement in their driving disability scores. At 6 weeks (0.9 ± 1.8 vs. 1.4 ± 2.8; P = .044) and 3 months (0.6 ± 0.9 vs. 1.0 ± 1.1; P = .023), mean driving disability scores were statistically significantly better for the C-ADR than the fusion group. At a 2-year follow-up, the proportion of patients reporting moderate or severe driving disability was similar between surgery groups: 8.5% in the C-ADR group and 8.2% in the fusion group.

The overall GRADE assessment of results on HRQOL was high (Appendix 4, Table A5).

Table 13: Health-Related Quality of Life With C-ADR Versus Fusion for Cervical Degenerative Disc Disease

C-ADR Device Author, Year Country N (C-ADR/ Fusion)	HRQOL Measure[a]			Conclusions
	Follow-Up Point	C-ADR	Fusion	
One-level disease				
Bryan Heller et al,[20] 2009 United States 30 sites, 242/221	**SF-36-PCS, mean**			• Both groups improved their SF-36-PCS subscores over baseline • Mean SF-36-PCS values were significantly higher for C-ADR group at early follow-up, but difference was not statistically significant at 2-yr follow-up (P_{ANCOVA} = .15)
	Baseline	32.6	31.8	PCS between groups
	3 mo	46.3	43.9	3 mo, P = .017
	6 mo	47.5	45.1	6 mo, P = .019
	1 yr	48.4	45.5	1 yr, P = .010
	2 yr	47.9	46.3	2 yr, P = .150
	2-yr mean change from baseline[b]	Δ 15.3	Δ 14.5	
	SF-36-MCS, mean			• Both groups significantly improved their SF-36-MS subscores over baseline • Mean SF-36-MCS values were significantly higher for C-ADR group at early follow-up, but values were not statistically different at 24 mo (P_{ANCOVA} = .270)
	Baseline	42.3	44.6	MCS between groups
	3 mo	52.6	50.8	3 mo, P = .002
	6 mo	53.0	50.8	6 mo, P < .001
	1 yr	52.5	51.6	1 yr, P = .048
	2 yr	51.7	51.7	2 yr, P = .270
	2-yr mean change from baseline[b]	Δ 9.4	Δ 7.0[a]	
Discover Skeppholm et al,[52] 2015 Sweden 3 sites, 81/70	**EQ-5D, mean ± SD (95% CI)**			• Both groups achieved statistically significant and clinically relevant (> MCID 0.24) improvements over baseline in EQ-5D scores at 2 yr • Improvements in 2-yr mean and median EQ-5D values were not different between surgery groups (P_{MW} = .50)
	Baseline	0.36 ± 0.32 (0.29–0.43)	0.47 ± 0.30 (0.40–0.54)	
	2 yr	0.70 ± 0.30 (0.63–0.77)	0.71 ± 0.26 (0.65–0.77)	
	EQ-5D, median (range)			
	Baseline	0.25 (−0.18 to 0.8)	0.69 (−0.24 to 0.8)	
	2 yr	0.79 (−0.29 to 1.000	0.76 (−0.17 to 1.00)	

C-ADR Device Author, Year Country N (C-ADR/ Fusion)	HRQOL Measure[a]			Conclusions	
	Follow-Up Point	C-ADR	Fusion		
Sundseth et al,[53] 2017 Norway 5 sites, 73/70	**SF-36-PCS, mean (95% CI)**			Mean between-group difference (95% CI)	• Mean SF-36-PCS scores increased significantly over baseline in both surgery groups at all follow-up points • At 2 yr, mean increased over baseline by 11.8 points in C-ADR and by 12 points in fusion arms • Mean between-group SF-36-PCS subscores were not significantly different at any follow-up point
	Baseline	34.6 (32.9–36.2)	34.9 (32.9–36.8)		
	3 mo	44.6 (42.5–46.8)	43.9 (41.3–46.6)	3 mo 1.1 (−2.2 to 4.3) *P* = .53	
	6 mo	46.9 (44.0–49.7)	47.9 (45.3–50.4)	6 mo −0.5 (−4.0 to 3.1) *P* = .79	
	1 yr	46.1 (43.7–48.4)	48.2 (45.5–50.9)	12 mo −1.4 (−4.8 to 1.9) *P* = .40	
	2 yr	46.4 (43.7–49.3)	46.9 (44.5–49.1)	24 mo 0.1 (−3.2 to 3.3) *P* = .9	
	2-yr mean change Δ from baseline[b]	Δ 11.8	Δ 12.0		
	SF-36-MCS, mean (95% CI)			Mean between-group difference (95% CI)	• Mean SF-36-MCS scores increased significantly over baseline in both surgery groups at all follow-up points • At 2 yr, mean increased over baseline by 4.9 points in C-ADR and by 6.1 points in fusion arms • Between-group differences in mean SF-36-MCS subscores were not significantly different at any follow-up point
	Baseline	47.4 (44.5–50.1)	44.2 (41.6–46.5)		
	3 mo	52.0 (49.8–54.3)	50.7 (47.7–53.2)	3 mo 0.3 (−3.0 to 3.7) *P* = .85	
	6 mo	51.0 (48.2–53.7)	51.7 (49.2–54.8)	6 mo −1.8 (−5.5 to 1.9) *P* =.34	
	1 yr	52.2 (49.5–54.8)	53.3 (51.0–55.6)	1 yr −1.9 (−5.4 to 1.5) *P* = .28	
	2 yr	52.3 (49.3–54.9)	50.3 (47.2–53.3)	2 yr 0.6 (−2.8 to 4.0) *P* = .72	
	2-yr mean change Δ from baseline[b]	Δ 4.9	Δ 6.1		

C-ADR Device Author, Year Country N (C-ADR/ Fusion)	HRQOL Measure[a]			Conclusions	
	Follow-Up Point	C-ADR	Fusion		
	EQ-5D, mean (95% CI)			Mean between-group difference (95% CI)	• Mean EQ-5D scores increased significantly over baseline at 3 mo and remained stable over follow-up in both surgery groups • At 2 yr, mean in scores increased over baseline by 0.35 for C-ADR and by 0.44 for fusion • Between-group differences in mean EQ-5D scores were not significantly different at any follow-up point
	Baseline	0.37 (0.29–0.45)	0.28 (0.20–0.35)		
	3 mo	0.73 (0.67–0.79)	0.67 (0.60–0.74)	3 mo 0.04 (−0.05 to 0.13) *P* = .36	
	6 mo	0.80 (0.74–0.85)	0.81 (0.75–0.86)	6 mo −0.03 (−0.13 to 0.07) *P* =.54	
	1 yr	0.70 (0.63–0.78)	0.72 (0.65–0.79)	1 yr −0.03 (−0.12 to 0.06) *P* = .50	
	2 yr	0.72 (0.64–0.80)	0.72 (0.67–0.79)	2 yr −0.02 (−0.11 to 0.07) *P* = .72	
	2-yr mean change Δ from baseline[b]	Δ 0.35	Δ 0.44		
Mobi-C Hisey et al,[38] 2014 United States 23 sites, 164/81	**SF-12-PCS**			• Mean change in SF-12-PCS score over baseline was higher (15.7 vs. 13.0 points), but not significantly, in C-ADR group at 2-yr follow-up • Mean improvement in SF-12-PCS scores over baseline was statistically significant (*P* < .05) and clinically relevant in both surgery groups at 2-yr follow-up	
	Baseline, mean ± SD	32.5 ± 5.91	33.8 ± 6.36		
	2 yr, mean	48.3	46.5		
	2-yr mean change Δ over baseline	Δ 15.7	Δ 13.0		
	SF-12-MCS			• Mean change in SF-12-MCS score over baseline was higher (8.5 points vs. 7.2 points) but not statistically significant in C-ADR group at 2-yr follow-up • Mean improvements in SF-12-MCS scores over baseline were statistically significant (P < .05) in both surgery groups at 2-yr follow-up	
	Baseline, mean ± SD	42.1 ± 13.1	42.1 ± 13.1		
	2 yr, mean	51.0	49.2		
	2-yr mean change Δ over baseline	Δ 8.5	Δ 7.2		

C-ADR Device Author, Year Country N (C-ADR/ Fusion)	HRQOL Measure[a]			Conclusions
	Follow-Up Point	C-ADR	Fusion	
PCM Phillips et al,[56] 2013 United States 24 sites, 224/192	**SF-36-PCS** Percentage at 2 yr improved ≥ 15% over baseline	71.1% (133/187)	64.9% (98/151)	• Most patients achieved clinically relevant improvement in SF-36-PCS scores at 2-yr follow-up • More patients in C-ADR group experienced clinically relevant improvement in SF-36-PCS scores over baseline at 2 yr by 6.2% (71.1% vs. 64.9%) • Improvement in SF-36-PCS scores (*P* = .241) was not statistically different between surgery groups at 2-yr follow-up
	SF-36-MCS Percentage at 2 yr improved ≥ 15% over baseline	46.5% (87/187)	49.7% (75/151)	• Improvement in SF-36-MCS scores (*P* = .585) was not statistically different between surgery groups at 2-yr follow-up • Fewer patients achieved clinically relevant improvement in SF-36-MCS scores than in SF-36-PCS scores at 2-yr follow-up
Prestige-ST Mummaneni et al,[21] 2007 United States 32 sites, 276/265	**SF-12-PCS**, mean (3, 6, 12, 24 mo)	(Data graphed)	(Data graphed)	• Between-group mean differences for SF-12-PCS subscores were not significantly different at any follow-up point • Both surgery groups achieved significant improvement in SF-12-PCS scores over baseline at all follow-up points • Improvements were stable over 2-yr follow-up in both groups
	1-yr mean change Δ over baseline	Δ 12.8	Δ 11.2	
	2-yr mean change Δ over baseline	Δ 13.1	Δ 7.7	
	SF-12-MCS, mean (3 mo, 6 mo, 1 yr, 2 yr)	(Data graphed)	(Data graphed)	• Between-group mean differences for SF-12-MCS subscores were not significantly different at any follow-up point • Both surgery groups achieved significant improvement over baseline SF-12-PCS and SF-12-MCS at all follow-up points • Improvement was stable over 2-yr follow-up in both groups
	1-yr mean change Δ over baseline	Δ 7.7	Δ 6.1	
	2-yr mean change Δ over baseline	Δ 7.4	Δ 7.5	
Prestige-LP Gornet et al,[59] 2015 United States 20 sites, 280/265	**SF-36-PCS**			• At 2 yr, proportion of people achieving clinically relevant improvement (≥ 15% in SF-36-PCS subscores) was lower in C-ADR group (82.8% vs. 87.5%) • Bayesian 95% CI of mean difference at 2 yr over baseline was −4.7% (−11.5% to 2.2%) with a P_B of noninferiority of 93.8% • Improvement in SF-36-PCS scores over baseline was greater in C-ADR group at early follow-up (by 15% at 6 mo, by 14% at 12 mo) but was similar at 2-yr follow-up (82.8% vs. 87.5%) • Similar proportions of patients had lower SF-36-PCS scores (> 15% change) at 6 mo (4.1% C-ADR vs. 6.8% fusion) and at 24 mo (5.3% C-ADR vs. 6.0% fusion)
	Baseline, mean ± SD	32.2 ± 7.4	32.0 ± 7.5	
	Baseline, median (IQR)	32.4 (27.4–36.7)	31.5 (27.1–36.6)	
	S-F36-PCS Percentage improved ≥ 15% over baseline			
	6 mo	80.1% (213/280)	65.0% (143/265)	
	1 yr	82.2% (221/280)	68.2% (150/265)	
	2 yr	75% (198/280)	71.3(154/265)	

C-ADR Device Author, Year Country N (C-ADR/ Fusion)	HRQOL Measure[a]			Conclusions
	Follow-Up Point	C-ADR	Fusion	
	SF-36-PCS Percentage maintained (−15% to 15%)			
	6 mo	15.8% (42/280)	28.2% (62/265)	
	1 yr	14.1% (38/280)	25.5% (56/265)	
	2 yr	19.7% (52/280)	22.7% (49/265)	
	SF-36-MCS			• Improvement in SF-36-MCS subscores over baseline was similar between surgery groups at all follow-up points • Proportion with improved SF-36-MCS subscores in C-ADR group was statistically superior to that in fusion group by Bayesian analyses • Bayesian 95% CI of mean difference at 2 yr between C-ADR and fusion in proportion improving: MCS Δ 10.5% (2%–*19%)* noninferiority $P_B \approx 100\%$, $P_B = 99.3\%$
	Baseline, mean ± SD	44.5 ± 11.5	42.7 ± 12.4	
	Baseline, median (IQR)	46.5 (4.9–53.6)	42.1 (33.1–53.0)	
	SF-36-MCS Percentage improved ≥ 15% over baseline			
	6 mo	44.0% (117/280)	47.3% (104/265)	
	1 yr	47.2% (127/280)	45.5% (100/265)	
	2 yr	1 yr 49.6% (131/280)	47.7% (103/265)	
	SF36-MCS Percentage maintained (−15% to 15%)			
	6 mo	46.6% (124/280)	42.3% (93/265)	
	1 yr	43.9% (118/280)	42.3% (93/265)	
	2 yr	41.3% (109/280)	38.4% (83/265)	
ProDisc-C Murrey et al,[22] 2009 United States 13 sites, 106/103	**SF-36-PCS** At 2 yr, percentage improved ≥ 15% over baseline[d]	51.5% (51 /99)	34.4% (31 /90)	• More patients in C-ADR group experienced a clinically relevant improvement in SF-36-PCS subscores over baseline at 2 yr (by 17.1%, 51.5% vs. 34.4). and between-group differences were significant (*P* = .017)
	SF-36-MCS At 2 yr, percentage improved ≥ 15% over baseline[d]	36.4% (36/99)	42.2% (38/90)	• Fewer patients in C-ADR group experienced a clinically relevant improvement in SF-36-MCS subscores over baseline at 2 yr (by 15.8%, 36.4% vs. 42.2%) • Difference between surgery groups was not statistically significant

C-ADR Device Author, Year Country N (C-ADR/ Fusion)	HRQOL Measure[a]			Conclusions
	Follow-Up Point	C-ADR	Fusion	
Secure Vaccaro et al,[40] 2013 United States 18 sites, 151/140	**SF-36-PCS**			• Percentage of patients achieving clinically relevant improvement in PCS subscores was similar between groups
	Baseline, mean ± SD	33.9 ± 7.4	32.0 ± 6.5	
	At 2 yr, percentage improved ≥ 15% over baseline	79.0%	78.1%	
	SF-36-MCS			• Percentage of patients achieving clinically relevant improvement in SF-36-MCS subscores was higher for C-ADR group, but differences between groups were not significant • MCS Δ (−2.5 to 21.2)
	Baseline, mean ± SD	44.0 ± 13.6	44.4 ± 12.0	
	At 2 yr, percentage improved ≥ 15% over	50.7%	42.1%	
Mixed-device RCTs for one-level disease				
Bryan and, Prestige-ST Riew et al,[63] 2008 United States Myelopathy, 106/93	**SF-36-PCS and SF-36-MCS,** mean baseline and 2-yr follow-up	(Data graphed)	(Data graphed)	• Patients with myelopathy in both surgery groups showed significant improvement over baseline at 2 yr in mean SF-36-PCS and MCS scores Differences at 2-yr follow-up between groups for both PCS and MCS subscores were reported to be not statistically significant
Bryan and Prestige-ST Steinmetz et al,[78] 2008 U.S. Workers' Compensation, 47/46	**SF-36-PCS**			• Workers' Compensation patients in both groups showed significant improvement over baseline at 2 yr in mean SF-36-PCS subscores • Patients in C-ADR group achieved greater improvement in SF-36-MCS subscores, but differences between study subgroups were not significant (P_{ANOVA} = .219)
	Baseline, mean ± SD	31.1 ± 6.1	30.1 ± 6.6	
	2-yr, mean ± SD	42.5 ± 12.8	38.1 ± 12.8	
	2-yr mean change Δ over baseline	Δ 11.4	Δ 8.0	
	SF-36-MCS			• Significant improvements in SF-36-MCS subscores over baseline were achieved in both surgery groups • Differences between groups were not significant (P_{ANOVA} = .219)
	Baseline, mean ± SD	39.1 ± 14.1	38.4 ± 11.7	
	2 yr, mean ± SD	46.2 ± 13.6	44.6 ± 14.0	
	2-yr mean change Δ over baseline	Δ 7.1	Δ 6.2	

C-ADR Device Author, Year Country N (C-ADR/ Fusion)	HRQOL Measure[a]			Conclusions
	Follow-Up Point	C-ADR	Fusion	
Two-level disease				
Mobi-C Davis et al,[41] 2013 United States 24 sites, 225/105	**SF-12-PCS**			• Mean changes in SF-36-PCS subscores over baseline were significantly (*P* < .05) higher in C-ADR than in fusion group at 2-yr follow-up • Improvement over baseline in mean SF-36-PCA subscores at 2 yr was statistically significant at all follow-up points (*P*w < .0001) in both groups at all follow-up points to 2 yr
	Mean (3, 6, 12, 18 and 24 mo)	(Data graphed)	(Data graphed)	
	2-yr mean change Δ over baseline	Δ 13.5	Δ 10.5	
	SF-12-MCS			• Improvement over baseline in mean SF-36-MCS subscores was statistically significant (*P*w < .0001) in both surgery groups at all follow-up points to 2 yr • Between-group differences were not significant
	Mean (3, 6, 12, 18, and 24 mo)	(Data graphed)	(Data graphed)	
	2-yr mean change Δ over baseline	Δ 9.5	Δ 7.2	
Prestige-LP Gornet et al,[42] 2017 30 sites, 226/230	**SF-36-PCS**			• Most patients achieved clinically relevant improvement in SF-36-PCS subscores over baseline at 2-yr follow-up
	Baseline, mean ± SD	31.8 ± 7.8	30.8 ± 7.4	
	Percentage at 2-yr follow-up improved (> 15%) over baseline	90.4% (178/197)	87.8% (137/156)	
	SF-36-MCS			• Both groups achieved clinically relevant improvement in SF-36-PCS subscores over baseline and at 2-yr follow-up. Results for C-ADR were statistically noninferior and not superior to fusion • Patients in both surgery groups achieved clinically relevant improvement in SF-36-MCS subscores over baseline
	Baseline, mean ± SD	43.9 ± 11.8	43.8 ± 12.2	
	Percentage maintained (−15% to 15%) or improved (> 15%) score over baseline at 2-yr follow-up	69.0% (136/197)	72.4% (113/156)	

Abbreviations: Δ, difference; C-ADR, cervical artificial disc replacement; CI, confidence interval; EQ-5D-3L, European Quality of Life questionnaire in five dimensions, 3-level version; HPD, highest posterior density; HRQOL, health-related quality of life; IQR, interquartile range; NE, not evaluated; NR, not reported; P_{ANCOVA}, probability based on ANCOVA with pre-operative score as covariate; P_B, Bayesian probability; P_{MW}, Mann-Whitney probability; Pw, Wilcoxon probability; RCT, randomized controlled trial; SD, standard deviation; SF12-MCS, 12-item Short-Form Health Survey–Mental Component Summary; SF12-PCS, 12-item Short-form Health Survey–Physical Component Summary; SF36-MCS, 36-item Short-Form Health Survey–Mental Component Summary; SF36-PCS, 36-item Short-Form Health Survey–Physical Component Summary; SSED, Summary of Safety and Effectiveness Data.

[a]Scores in Short-Form Health Survey range from 0 to 100; higher scores indicate better health. Scores in EQ-5D range from −0.59 to 1; higher scores indicate better health.

[b]Crude mean changes over baseline were calculated by this author.

[c]SF36 improvement rates for groups were abstracted from FDA 2-yr SSED report.

[d]SF36 information in report by Murrey et al involved classifying any change in score as improvement. Data were abstracted from FDA SSED report in which SF36 was scored as percentage, achieving ≥ 15% improvement over baseline.

Treatment Satisfaction

Patients' Treatment Satisfaction

One-Level Cervical Degenerative Disc Disease
The satisfaction of patients and surgeons with eight C-ADR devices and with fusion was evaluated for one-level degenerative disc disease (Table 14). In the trials, treatment satisfaction was evaluated at various post-operative follow-up points and as varying degrees of satisfaction: willingness to recommend surgery to others; willingness to repeat the surgery; global perception of treatment effectiveness; degree helped by surgery; or degree recovered after surgery.

In comparing responses at a 2-year follow-up across trials, more patients were "very satisfied" with C-ADR than with fusion: Prestige-LP (73% vs. 59%),[59] Kineflex (75% vs. 67%),[36] and Mobi-C (89% vs. 84%). Patients were as likely to recommend (definitely or probably) C-ADR with the Mobi-C device (87% vs. 84%)[38] or with the PCM device (91.9% vs. 87.5%)[56] as they were to recommend fusion. Patients, however, were more likely to report willingness (definitely or mostly) to repeat C-ADR than fusion for Prestige-LP (80% vs. 70.3%),[59] ProDisc-C (85.6% vs. 80.9%),[82] and Secure (92% vs. 81%).[40] In terms of treatment expectations, patients were more likely to report that C-ADR, rather than fusion, helped them as much as they had expected (definitely or mostly): Secure (91% vs. 79%).[40] In the FDA trial of Prestige-LP, the proportion of patients reporting that surgery "definitely helped as much as I had thought" was much lower for both C-ADR (68.9%) and fusion (60.7%).[59]

Two-Level Cervical Degenerative Disc Disease
Patients' satisfaction with two C-ADR devices and fusion was evaluated for two-level disease. In the FDA RCT of Mobi-C, patients' satisfaction remained similarly high for both surgery groups at early follow-up. Patients were more likely to report satisfaction with C-ADR than with fusion at the 5-year (96.4% vs. 89.5%)[83] and 7-year (86% vs. 73.9%)[84] follow-up points. They were also more likely to recommend C-ADR than fusion to a friend at all follow-up points. In the FDA trial of Prestige-LP,[42] patients reported having similarly high satisfaction rates (> 90%) with their surgery as assessed by three questions: rating the degree of satisfaction, whether surgery helped as much as they thought it would, and willingness to repeat the surgery.

Patient-reported satisfaction measures have been validated to some extent by their agreement with other patient-reported outcome measures. A retrospective analysis of the data from the FDA RCT of Mobi-C reported patient satisfaction measures to be strongly related to patient-reported outcome measures, such as the NDI.[85] Reductions in mean NDI scores at 2 years were statistically significantly ($P < .0001$) related to the degree of patient satisfaction: very satisfied (NDI 15.1 ± 16.7), somewhat satisfied (NDI 38.7 ± 20.1), and somewhat dissatisfied (NDI 41.6 ±.13.7). Patient dissatisfaction was also found to be related to subsequent surgeries at 2 years. Rates of subsequent surgery for the very satisfied, somewhat satisfied, and somewhat dissatisfied patients were 3.6%, 12.9%, and 26.7%, respectively. Overall, those who underwent secondary surgeries were statistically significantly less likely to be "very satisfied" with their treatment than those not having secondary surgeries (56% vs. 85.8%, $P = .0004$).

Surgeons' Treatment Satisfaction

One-Level Cervical Degenerative Disc Disease
Surgeons' judgment of treatment outcomes for one-level involvement was also reported in trials of three C-ADR devices (PCM, Prestige-ST, Prestige-LP). Surgeons' judgment was based on Odom's criteria[86]: 1 (excellent) all pre-operative symptoms were relieved and abnormal findings improved; 2 (good) minimal persistence of pre-operative symptoms and abnormal findings were unchanged or improved; 3 (fair) some pre-operative symptoms were definitely relieved and other symptoms were unchanged or slightly improved; 4 (poor) symptoms and signs were unchanged or exacerbated.

In the FDA RCT of PCM, surgeons' judgment of treatment success (excellent or good outcomes at a 2-year follow-up) was high for both C-ADR (91.5%) and fusion (86.3%).[56] However, in this trial, patients' satisfaction with treatment (very or moderately satisfied) was lower (84% for PCM and 79% for fusion) than surgeons' judgments. In the FDA RCT of Prestige-ST, surgeons were more likely to judge treatment outcomes as "excellent" after C-ADR (70.9%) than after fusion (56.2%).[21] Patients, however, were much less likely than surgeons to rate their recovery as a "complete recovery" after either surgery, although they were more likely to report complete recovery after C-ADR (45.7%) than after fusion (39.4%). In the FDA RCT of Prestige-LP, surgeons were more likely to judge treatment success at 2 years as excellent for the C-ADR than for the fusion group (71.6% vs. 56.8%).[59]

Two-Level Cervical Degenerative Disc Disease
Surgeons' judgment of treatment outcomes for two-level involvement was reported for one C-ADR device, Prestige-LP. Their judgment that surgical outcomes were good or excellent was statistically significantly higher for patients in the C-ADR than in the fusion group (96.9% vs. 84.3%).[42]

The overall GRADE assessment of results on treatment satisfaction was high (Appendix 4, Table A5).

Table 14: Satisfaction With C-ADR Versus Fusion for Cervical Degenerative Disc Disease

C-ADR Device	Author, Year Follow-Up Point Satisfaction Measures	Treatment Satisfaction Ratings	C-ADR	Fusion	Conclusions
One-level disease					
Bryan	Hackeret al,[87] 2005 1 yr Patients' global perception of treatment success	Excellent	77%	63%	• Patients in both surgical groups perceive their treatment success as excellent or good (91%, 88%) 1 yr after surgery • Global perceptions of success with Bryan device were higher, but not significantly higher, than with fusion
		Good	14%	25%	
		Fair	9%	8%	
		Poor	0	4%	
Kineflex-C	Coric et al,[36] 2011 2 yr Patients satisfied with treatment	Very satisfied	75%	67%	• Patients in both surgical groups are satisfied or very satisfied (88%, 87%) with treatments 2 yr after surgery, and levels do not differ between groups
		Satisfied	13%	20%	
		Somewhat satisfied or dissatisfied	12%	13%	
Mobi-C	Hisey et al,[38] 2014 2 yr Hisey et al,[88] 2015 4 yr Hisey et al,[88] 2016 5 yr Radcliff et al,[84] 2017 7 yr Patients' treatment satisfaction	Very satisfied at follow-up			• Patients in both surgical groups have high levels of satisfaction (89%, 84%) with their treatment 2 yr after surgery, and levels in both groups remain high (>80%) at 4- and 5-yr follow-up with levels consistently higher, but not significantly higher, for Mobi-C device • At 7 yr, satisfaction was significantly (*P* =.028) higher in C-ADR group
		2 yr	89.0%	84.3%	
		4 yr	88.6%	83.6%	
		5 yr	92.0%	83.9%	
		7 yr	90.9%	77.8%	
	Recommend surgery	Definitely recommend at follow-up			
		2 yr	87.0%	84.3%	
		4 yr	87.8%	81.8%	
		5 yr	97.1%	91.1%	
		7 yr (definitely or probably)	96.2%	88.9%	

C-ADR Device	Author, Year Follow-Up Point Satisfaction Measures	Treatment Satisfaction Ratings	C-ADR	Fusion	Conclusions
PCM	Phillips et al,[56] 2013 2 yr Philips et al,[89] 2015 5 yr Patients' treatment satisfaction	Mean treatment VAS			• Patients in both surgical groups rated their treatment satisfaction highly (≥ 60 mm is completely satisfied) at both follow-up points, but rating at 5-yr follow-up was significantly (*P* = .005) higher for PCM device
		2 yr	82.8/100 mm	81.4/100 mm	
		5 yr	86.9/100 mm	78.3/100 mm	
	Treatment satisfaction	Very or moderately satisfied at follow-up			• Satisfaction, evaluated as moderately or very satisfied, was high (> 80%) in both surgical groups, although satisfaction was higher in PCM group at 5-yr follow-up
		2 yr	84.4%	79.4%	
		5 yr	88.8%	78.7%	
	Recommend surgery	Definitely or probably recommend at follow-up			• Willingness to recommend surgery to a friend was high in both surgical groups, but was higher for PCM group at 5-yr follow-up
		2 yr	91.9%	87.5%	
		5 yr	94.4%	85.0%	
	Surgeons' judgment of treatment outcomes[b]	Excellent or good outcomes at 2 yr	91.5%	86.3%	• Surgeons' judgment that treatment outcomes were excellent or good was high for both surgical groups 2 yr after surgery
Prestige-ST	Mummaneni et al,[21] 2007 2 yr[c] Patients' global perception of health	Perception of health at 2 yr			• Although patients' global perception of health as complete recovery or much improved was high (84%, 82%) for both surgical groups at 2-yr follow-up, proportion reporting complete recovery was higher for Prestige-ST group
		Complete recovery	45.7%	39.4%	
		Much improved	38.0%	43.0%	
	Surgeons' judgment of treatment outcomes	Surgeons' judgment at 2 yr			• Surgeons' judgment that post-operative treatment outcomes were excellent at 2 yr was also higher (70.9% vs. 56.2%) for Prestige-ST
		Excellent outcome	70.9%	56.2%	
		Good outcome	23.6%	35.2%	

C-ADR Device	Author, Year Follow-Up Point Satisfaction Measures	Treatment Satisfaction Ratings	C-ADR	Fusion	Conclusions
Prestige-LP	Gornet et al,[59] 2015 2 yr Gornet et al,[90]2016 7 yr Patients' global perception of effectiveness of surgery	Patients' perception of treatment			• Patients' global perceived effectiveness of surgery was high in both surgical groups and was similar to estimates for Prestige-ST study • Post-operative reports of global effectiveness (complete or much improved) was higher for Prestige-LP group 2 yr and 7 yr after surgery
		Complete recovery at 2 yr	47.0%	40.2%	
		Much recovered at 2 yr	39.6%	40.6%	
		Completely or much recovered at 7 yr	86.1%	77.4%	
	I am satisfied with my surgery (at 2 yr, at 7 yr)	Definitely true at 2 yr, 7 yr	73.3%, 79.3%	59.4%, 60.8%	• Level of dissatisfaction (mostly and definitely) was lower for Prestige-LP group both 2 yr (3.4% vs. 7.9%) and 7 yr (2.9% vs. 8.3%) after surgery
		Mostly true at 2 yr, 7 yr	16.7%, 11.5%	28.3%, 24.9%	
		Do not know at 2 yr, 7 yr	6.7%, 6.3%	5.0%, 6.1%	
		Mostly false at 2 yr, 7 yr	1.9%, 1.0%	2.3%, 5%	
		Definitely false at 2 yr, 7 yr	1.5%, 1.9%	5.0%, 3.3%	
	I was helped as much as I thought I would be (at 2 yr, at 7 yr)	Definitely true at 2 yr, 7 yr	68.9%, 74.5%	60.7%, 59.9%	• Patients' expectations of their treatment outcomes (mostly or definitely not met) were less often met in fusion group both at 2 yr (4.8% vs. 12.8%) and at 7 yr (5.7% vs. 9.6%) after surgery
		Mostly true at 2 yr, 7 yr	18.5%, 14.9%	21.5%, 23.7%	
		Do not know at 2 yr, 7 yr	7.8%, 4.8%	5.0%, 6.8%	
		Mostly false at 2 yr, 7 yr	3.7%, 3.8%	6.4%, 4.0%	
		Definitely false at 2 yr, 7 yr	1.1%, 1.9%	6.4%, 5.6%	
	All things considered, I would have surgery again (2 yr, 7 yr)	Definitely true at 2 yr, 7 yr	80.0%, 80.1%	70.3%, 72.3%	• Patients in Prestige-LP group feel more strongly about having surgery again if they needed it than those in fusion group at 2-yr (80% vs. 70.3%) and 7-yr follow-up (80.1% vs. 72.3%)
		Mostly true at 2 yr, 7 yr	10.7%, 7.7%	13.2%, 12.4%	
		Do not know at 2 yr, 7 yr	6.3%, 7.7%	11.0%, 10.7%	
		Mostly false at 2 yr, 7 yr	1.1%, 1.4%	2.3%, 2.3%	
		Definitely false at 2 yr, 7 yr	1.9%, 1.9%	3.2%, 2.3%	
	Surgeons' judgment of treatment outcomes	Excellent outcome at 2 yr	71.6%	56.8%	• Surgeons were more likely to judge surgical outcomes as excellent for Prestige-LP than for fusion group (71.6% vs. 56.8%) at 2-yr follow-up
		Good outcome at 2 yr	22.9%	31.4%	

C-ADR Device	Author, Year Follow-Up Point Satisfaction Measures	Treatment Satisfaction Ratings	C-ADR	Fusion	Conclusions
ProDisc-C	Murrey et al,[22] 2009 2 yr Zigler et al,[91] 2013 5 yr Janssen et al,[92] 2015 7 yr Treatment satisfaction VAS rating Proportion very or completely satisfied (VAS ≥ 60 mm)	Treatment satisfaction rating and proportion very or completely satisfied			• Patients in both surgical groups rated their satisfaction with their treatment high. Mean VAS scores were all ≥ 60 mm, indicating complete satisfaction in both groups at all follow-up points. Scores between groups were not significantly different
		Mean VAS at 2 yr	83.4 mm ± 24.8	80.0 mm ± 28.4	
		Proportion VAS ≥ 60 mm at 2 yr	86.3%	83.0%	
		Mean VAS at 5 yr	86.6 mm	82.7 mm	
		Mean VAS at 7 yr	85.8 mm	81.8 mm	
	Delamarter et al,[82] 2010 2 yr, 4 yr Willing to repeat surgery	Yes, willing to repeat surgery			• Almost all (> 80%) patients in both surgical groups 2 yr and 4 yr after surgery would recommend their surgery with ProDisc-C or fusion to a friend
		Repeat surgery at 2 yr	85.6%	80.9%	
		Repeat surgery, at 4 yr	88.9%	81.0%	
Secure	Vaccaro et al,[40] 2013 2 yr Satisfaction with treatment success	Satisfied, definitely or mostly true	95.7%	85.2%	• More patients in Secure implant group were satisfied with their surgery than in fusion group (95.7% vs. 85.2%)
	Surgery helped as much as I had thought	Helped, definitely or mostly true	91%	79%	• More patients in Secure implant group than in fusion group thought that surgery helped them as much as they thought it would at 2-yr follow-up (91% vs. 79%)
	Willing to repeat surgery	Repeat surgery, definitely or mostly true	92%	81%	• More patients in Secure implant group than in the fusion group (92% vs. 81%) were willing to recommend their surgery to a friend

C-ADR Device	Author, Year Follow-Up Point Satisfaction Measures	Treatment Satisfaction Ratings	C-ADR	Fusion	Conclusions
Two-level disease					
Mobi-C	Davis et al,[41] 2013 2 yr Davis et al,[93] 2015 4 yr Radcliff et al,[83] 2016 5 yr Radcliff et al,[84] 2017 7 yr	Satisfaction with surgical treatment			• At 2-yr follow-up, patients in both groups reported similar levels of satisfaction with their treatment, but in longer follow-up, more patients in Mobi-C group than in fusion group reported satisfaction with their treatment at 4 yr (96.4% vs. 86.3%, *P* = .011) and at 5 yr (96.4% 89.5%) • At 7-yr follow-up, more patients reported being satisfied with their C-ADR surgery (*P* =.039)
		2 yr	95.8%	92.0%	
		4 yr	96.4%	89.0%	
		5 yr	96.4%	89.5%	
		7 yr	86.0%	73.9%	
	Recommend same surgery to a friend with same indications	Recommend, definitely or probably true			• Patients in Mobi-C group were more likely to recommend their surgery than those in fusion group at all follow-up points. • At 7-yr follow-up, more patients in C-ADR group were willing to recommend their surgery (*P* = .025)
		2 yr	95.8%	88.5%	
		4 yr	95.9%	86.3%	
		5 yr	94.8%	84.2%	
		7 yr	96.8%	88.4%	

C-ADR Device	Author, Year Follow-Up Point Satisfaction Measures	Treatment Satisfaction Ratings	C-ADR	Fusion	Conclusions
Prestige-LP	Gornet et al,[42] 2017 2 yr Lanman et al,[94] 2017 5 yr, 7 yr I am satisfied with my surgery	Satisfied, definitely or mostly true			• Patients in both surgery groups reported similar very high levels of satisfaction with their surgery at all follow-up points
		2 yr	94.5%	89.3%	
		5 yr	94.6%	93.3%	
		7 yr	94.8%	92.6%	
	Helped me as much as I thought I would be	Helped, definitely or mostly true			• Patients in both surgery groups reported that their expectations for surgery were met at all follow-up points
		2 yr	94.5%	94.5%	
		5 yr	92.7%	88.1%	
		7 yr	NR	NR	
	All things considered I would have surgery again	Would repeat surgery, definitely or mostly true			• Patients in both surgery groups were willing to undergo surgery again if required, a statement that remained high with longer-term follow-up
		2 yr	93.4%	88.7%	
		5 yr	92.7%	88.1%	
		7 yr	94.8%	89.4%	
	Surgeons' judgment of treatment outcomes 2- and 5-yr physician reports SSED	Excellent outcome at 2 yr, 5 yr	69.8%, 71.7%	56.0%, 48.5%	• In surgeons' judgment, excellent or good treatment outcomes (96.9% vs. 84.3%) are more likely to occur with patients in Prestige-LP group • Fair or poor outcomes were judged to occur more often in fusion group at 2-yr (15.5% vs. 3%) and 5-yr (13.2% vs. 4.2%) follow-up
		Good outcome at 2 yr, 5 yr	27.1%, 24.1%	28.3%, 38.2%	
		Fair outcome at 2 yr, 5 yr	2.0%, 3.0%	13.2%, 11.0%	
		Poor outcome at 2 yr, 5 yr	1.0%, 1.2%	2.5%, 2.2%	

Abbreviations: C-ADR, cervical artificial disc replacement; NR, not reported; SF-12; 12-item Short-Form Health Survey; SSED, Summary of Safety and Effectiveness Data; VAS, Visual Analogue Scale.
[a]Schroeder et al examined predictive relationship between Patient reported outcomes (PRO) (Neck Disability Index, neck pain, and SF-12) with patient satisfaction.
[b]Odom's criteria: 1 (excellent) all pre-operative symptoms relieved, abnormal findings improved; 2 (good) minimal persistence of pre-operative symptoms, abnormal findings unchanged or improved; 3 (fair) definite relief of some pre-operative symptoms, other symptoms unchanged or slightly improved; and 4 (poor) symptoms and signs unchanged or exacerbated,
[c]Satisfaction data were extracted from U.S. Food and Drug Administration SSED report for randomized controlled trial by Mummaneni et al.

Longer-Term Treatment Outcomes

The key difference between C-ADR and fusion is that C-ADR devices are intended to maintain motion at the treated cervical level, whereas fusion does not allow for motion at the treated level. An anticipated outcome taking this difference into account is that the more normal biomechanics with C-ADR impose less stress on adjacent spinal levels than fusion and might result in less adjacent-segment disease and fewer adjacent-level secondary surgeries.

Cervical Kinematics: Preservation of Cervical Range of Motion at the Treated Level

One-Level Cervical Degenerative Disc Disease

Radiologic assessment of spinal motion was performed in longitudinal follow-ups of C-ADR trials (Table 15). The main reported radiologic measure for cervical motion was cervical flexion and extension range of motion (ROM)at the treated site evaluated by plain radiographs. The radiologic threshold of 4º ROM at the treated cervical site was maintained at all follow-up points for all C-ADR devices. The mean cervical flexion/extension ROM angles at the longest follow-up for the C-ADR devices were 8.5º at 4 years for Bryan,[95]; 5.7º at 3 years for ProDisc-C[73]; 10.6º at 5 years for Kineflex[96]; 10.2º at 7 years for Mobi-C[84]; 5.2º at 5 years for PCM[89]; 6.8º at 7 years for Prestige-ST[97]; 6.8º at 7 years for Prestige-LP[77]; 8.1º at 7 years for ProDisc-C[98], and 9.7º at 2 years for Secure.[40] In all trials, radiologic thresholds of no motion (< 2º) in the fusion groups was maintained at all follow-up points.

Other radiographic techniques were employed to evaluate cervical ROM. Radiographic measurements based on three-dimensional computed tomography (CT) for the Discover C-ADR device also confirmed that a mean ROM angle of 5.7º for cervical flexion/extension at a 3-year follow-up satisfied radiologic success.[67] Extensive radiographic measurements were taken for the ProDisc-C. Measurements were based on radiosterometric analysis in which placement of tantalum markers in the vertebrae enabled more precise measurement. These measurements showed some declines in post-operative ROM followed by stabilization of angles after 1 year. The mean flexion/extension angles, however, were lower than ROM angles obtained with plain radiographs. Fusion success, defined by a radiographic threshold of mean angles of less than 1º, was again confirmed for the fusion group.

The effect of fusion on the ROM of cervical levels adjacent to the index level was evaluated with the PCM[99] and ProDisc-C[49] C-ADR devices. In the 1-year follow-up of the FDA RCT of PCM,[99] mean flexion/extension angles at the superior adjacent level increased more above the fusion-treated (9.6º–11.0º) than the C-ADR-treated (9.8º–10.8º) level. In the FDA RCT of ProDisc-C,[49] which used radiosterometic assessments in a 6-month-to-1-year follow-up, rotations in all planes increased at superior levels adjacent to the fused level but remained within 1º for levels adjacent to the C-ADR-treated level. Auerbach et al[100] assessed total ROM of the cervical spine and found that the overall loss of motion at the fused level (15.4º) was compensated for by an increase in every other adjacent level.

Two-Level Cervical Degenerative Disc Disease

In the FDA RCTs of Mobi-C[93] and Prestige-LP[94] for two-level disease, mean angles for both angular flexion/extension and lateral bending ROM were reported for both inferior and superior treated cervical segments (Table 15). Radiologic thresholds for ROM were maintained at a 7-year follow-up with both C-ADR devices at inferior and superior levels for flexion/extension

and lateral bending ROM. Fusion groups also maintained radiologic thresholds of no motion (< 2º) at the treated site at long-term follow-up.

The overall GRADE assessment of results on cervical kinematics was moderate, downgraded for risk of bias (Appendix 4, Table A5).

Table 15: Cervical Motion at Treated and Adjacent Sites After C-ADR or Fusion

C-ADR Device Author, Year	Study Design No. of Patients (C-ADR/Fusion)	Follow-Up Duration Measurement of Cervical Motion at Treated Level	Conclusions
One-level disease			
Bryan			
Powell et al,[101] 2010	IDE RCT 1 site N = 22/26	2 yr Static and dynamic digital radiographs assisted by Quantitative Motion Analysis software	• Mean flexion/extension angles in C-ADR group were 6º at 3 mo and 8º at 1 and 2 yr • Mean flexion/extension angles remained at 2º or lower in fusion group over follow-up
Sasso et al,[102] 2008	IDE RCT 1 site N = 9/13	2 yr Neutral and lateral radiographs assisted by Quantitative Motion Analysis software	• C-ADR group retained average ROM of 6.7º at 2 yr and fusion group of 2.0º at 3 mo that decreased to 0.6º at 2 yr
Sasso et al,[103] 2011	IDE RCT 1 site N = 22/26	2 yr Static and dynamic digital radiographs assisted by Quantitative Motion Analysis software	• Overall (C2–C7) cervical lordosis was not significantly different between groups (4.4º ± 12.8 vs. 5.8º ± 9.4) • Lordosis at treated level varied greatly in all surgery groups
Sasso et al,[104] 2008	IDE RCT 31 sites N = 242/221	2 yr Angular motion assessed by lateral flexion/extension motion on radiographs	• Angulation, mean flexion/extension angles at treated level for C-ADR group remained high in follow-up: from 6.4º before surgery to 6.3º at 3 mo, 7.2º at 6 mo, 7.8º at 1 yr, and 8.0º at 2 yr • For fusion group, mean angulation angles decreased and remained low at follow-up from 8.4º before surgery to 1.1º at 3 mo, 1.1º at 6 mo, 0.9º at 1 yr, and 0.9º at 2 yr
Sasso et al,[95] 2011	IDE RCT 31 sites N = 242/221	4 yr Angular motion assessed by lateral flexion/extension motion on radiographs	• Angulation, mean flexion/extension angles at treated level for C-ADR group remained high at 4-yr follow-up: 8.5º (95% CI 7.7º–9.2º). • In fusion group, angulation remained low at 1.1º
Discover			
Skeppholm et al,[67] 2015	Non-IDE RCT 1 site N = 28	3 yr 3-D CT volume spatial registration scans for angular flexion/extension motion	• Mean rotation with flexion/extension at 1-level treated group (n = 18) was 5.7º ± 4.1 and at 2-level treated group (n = 10) was 3.3º ± 2.7 at inferior level and 5.7º ± 3.9 at superior level
Kineflex			
Coric et al,[96] 2018	IDE RCT 21 sites N = 136/133	5 yr Angular motion assessed by lateral flexion/extension motion on radiographs	• Mean ROM as flexion/extension in C-ADR group was maintained at or above pre-operative values (8.2º): 8.9º at 6 mo, 9.5º at 1 yr, 9.7º at 2 yr, 10.8º at 3 yr, 11.2º at 4 yr, and 10.6º at 5 yr
Mobi-C			
Hisey et al,[105] 2015	IDE RCT 23 sites N = 164/81	4 yr Angular motion on flexion/extension	• In C-ADR group, mean flexion/extension ROM was maintained at all follow-up points above 8º annually through 5-yr follow-up • In fusion group, ROM was < 2º angular motion on flexion/extension at all follow-up points

C-ADR Device Author, Year	Study Design No. of Patients (C-ADR/Fusion)	Follow-Up Duration Measurement of Cervical Motion at Treated Level	Conclusions
Hisey et al,[88] 2016	--	5 yr Angular motion on flexion/extension and lateral bending	• In C-ADR group, both angular motion on flexion/extension and lateral bending were maintained at or above pre-operative values throughout follow-up (data graphed) • In fusion group, both flexion/extension and lateral bending was reduced to < 2º throughout follow-up
Radcliff et al,[84] 2017	--	7 yr Angular motion on flexion/extension and lateral bending	• In C-ADR group, ROM at treated level was maintained over 7-yr follow-up. At 7 yr, mean angles for flexion/extension were 10.2º ± 6.3 and for lateral bending were 5.1º ± 3.5 • In fusion group, average ROM was 1º or less for both flexion/extension and lateral bending
PCM			
Park et al,[99] 2011	IDE RCT 23 sites N = 272/182	1 yr Neutral flexion and extension radiographs	• In C-ADR group, mean flexion/extension motion at 8.0º ± 4.5 before surgery reduced to 6.2º ± 4.0, and for fusion group, angles decreased to 1.0º ± 1.1 at 12 mo • At superior adjacent level to C-ADR index level, mean angle did not increase over perioperative value from 9.8º ± 5.2 to 10.8º ± 5.5 (*P* = .043) • After fusion, mean angle at superior level adjacent to index level significantly increased over perioperative value from 9.6º ± 5.1 to 11.0º ± 5.5 (*P* =.003) • Mean flexion/extension angles at inferior adjacent levels did not change over baseline in either C-ADR or fusion group
Phillips et al,[56] 2013	IDE RCT 24 sites N = 224/192	2 yr Plain radiographs of neutral, lateral and anteroposterior flexion/extension and lateral bending	• At 2 yr, mean flexion/extension for cervical ROM at treated level was 5.7º ± 3.9 (range 0º–17.2º) for C-ADR and was 0.8º ± 0.8 for fusion
Phillips et al,[89] 2015	IDE RCT 24 sites N = 163/130	5 yr Plain radiographs of neutral lateral and anteroposterior flexion/extension and lateral bending	• In C-ADR group, mean flexion/extension motion was reduced over baseline of 8º to 5º and was maintained over follow-up and at 5 yr. Mean angle was 5.2º ± 3.8 (range 0º–16.1º) • In fusion group, mean motion decreased to less than 2º and was maintained at follow-up. At 5 yr, mean angle was 0.5º ± 0.5 (range 0º–4.1º)
Prestige-ST			
Mummaneni et al,[21] 2007	IDE RCT 32 sites N = 276/265	2 yr Neutral anteroposterior and lateral radiographs and dynamic flexion/extension lateral radiographs	• In C-ADR group, mean angular motion was 7.6º before surgery and was 7.6º at 2 yr
Burkus et al,[106] 2010	IDE RCT 32 sites N = 144/127	5 yr Neutral anteroposterior and lateral radiographs and dynamic flexion/extension lateral radiographs	• In C-ADR group, mean angular motion was maintained over 7.9º mean pre-operative angle during 5-yr follow-up: 7.6º at 6 mo, 7.6º at 1 yr, 7.7º at 2 yr, 7.3º at 3 yr, and 6.5º at 5 yr • In fusion group, mean angle at treated site decreased and remained less than 1º throughout follow-up

C-ADR Device Author, Year	Study Design No. of Patients (C-ADR/Fusion)	Follow-Up Duration Measurement of Cervical Motion at Treated Level	Conclusions
Burkus et al,[97] 2014	IDE RCT 31 sites N = 212/183	7 yr Neutral anteroposterior and lateral radiographs and dynamic flexion/extension lateral radiographs	• In C-ADR group, mean angle remained high and at 7 yrs was 6.8º • In fusion group, mean angle at 7 yr remained less than 1º
Prestige-LP			
Gornet et al,[77] 2015	IDE CCT 20 sites N = 280/265	7 yr Segmental motion at index and superior or inferior adjacent levels to the index treated site using Cobb method of angles	• In C-ADR group, mean angular motion was maintained over follow-up and at 7 yr was 6.8º • In C-ADR group, ratio of mean angular motion before surgery and at 7 yr at adjacent superior level to index-treated site was 8.5º/9.3º, and ratio at adjacent inferior level was 6.1º/6.2º • In fusion group, ratio of mean angular motion before surgery and at 7 yr at superior level adjacent to index treated site was 10.8º/10.7º, and ratio at adjacent inferior level was 7.8º/8.5º
ProDisc-C			
Nabhan et al,[49] 2011	Non-IDE RCT 1 site N = 10/10	1 yr Cervical motion was evaluated by RSA with tantalum markers placed into vertebrae adjacent to treated level to evaluate segment mobility at treated and adjacent levels	• Segmental motion at treated level in C-ADR group decreased in all 3 planes from 1 wk to 6 mo to 1 yr—extension (4.8º ± 0.85 to 3.6º ± 1.8 to 2.1º ± 1.1), right-sided axial rotation (5.9º ± 1.65 to 4.4º ± 2.37 to 3.1º ± 1.38) and right-sided bending (2.6º ± 0.95 to 1.3º ± 0.9 to 2.1º ± 0.5) • Rotations at superior level adjacent to C-ADR index treated level remained within 1º in planes at follow-up • For fusion group, angles were 1º or less in all planes at all follow-up points • Rotations at superior adjacent level of fused level increased from 6 mo to 1 yr in all planes—extension (14.8º ± 2.2 to 18.1º ± 3.0), right-sided axial rotation (13.0º ± 2.4 to 16.4º ± 2.9), and right-sided bending (8.3º ± 1.8 to 10.1º ± 2.7)
Nabhan et al,[48] 2007	Non-IDE RCT 1 site N = 16/17	6 mo Motion was evaluated by RSA with tantalum markers placed into vertebrae	• Rotations in all 3 planes (extension, right-sided axial, and right-sided bending) decreased over 6-mo follow-up in both groups • At 6 mo, angles in all 3 planes were significantly higher for C-ADR than for fusion group—extension (2.36º ±1.0 vs. 0.95º ± 0.80; *P* = .0001); right-sided axial (2.56º ± 0.7 vs. 0.79º ± 0.64; *P* > .01) and right-sided bending (2.24º ± 1.3 vs. 0.74º ± 0.44; *P* = .001)
Nabhan et al,[72] 2007	Non-IDE RCT 1 site N = 24/24	1 yr Translation motion was evaluated by RSA with tantalum markers placed into vertebrae	• Mean segmental translation values declined in both groups to 3 wk (more for C-ADR group than for fusion) • Segmental translation after 3 wk remained relatively stable in both groups, although at higher values for C-ADR than for fusion group, in all 3 directions • At 1 yr, C-ADR mean translation distances in 3 axes were higher than fusion—mediolateral (0.39 mm ± 0.17 vs. 0.06 mm ± 0.05), craniocaudal (0.26 mm ± 0.13 vs. 0.06 mm ± 0.06), and anterioposterior (0.66 mm ± 0.42 vs. 0.07 mm ± 0.05) at all follow-up points

C-ADR Device Author, Year	Study Design No. of Patients (C-ADR/Fusion)	Follow-Up Duration Measurement of Cervical Motion at Treated Level	Conclusions
Nabhan et al,[73] 2007	Non-IDE RCT 1 site N = 25/24	3 yr Motion was evaluated by RSA with tantalum markers placed into vertebrae	• In C-ADR group, segmental rotations decreased, mainly in extension at 1 yr over post-operative measures, but then all angles remained stable at 2 and 3 yr: extension (5.4º ± 0.46 to 3.2º ± 0.64, 3.3º ± 0.65 to 3.2º ± 0.68); right-sided axial (3.8º ± 0.8 to 3.2º ± 0.6, 3.3º ± 0.7 to 3.2º ± 0.5); right-sided bending remained (2.9º ± 0.8 to 2.9º ± 0.7, 3.1º ± 0.6 to 3.0º ± 0.56) • In fusion group, all segmental rotational angles decreased over follow-up, were significantly lower than in C-ADR group at follow-up, and were all less than 1º at 3 yr
Park et al, 2010[107]	IDE RCT N = 164 Levels C6/C7 (n = 44) C5/C6 (n = 96) C4/C5 (n = 18) C3/C4 (n = 6)	2 yr Radiographs in different views – neutral, lateral, flexion/extension and bending with image analysis software (Quantitative Motion Analysis, Medical Metrics, Inc)	• In C-ADR, pre-operative level C4/C5 (12.5º) had significantly (*P* <.001) more flexion/extension ROM than other cervical levels: C5/C6 (8.6º), C6/C7 (6.3º), C3/C4 (6.3º) and significantly (*P* = .015) more lateral bending ROM at level C4/C5 (7.6º) than at levels C3/C4 (6.5º), C5/C6 (5.8º), or C6/C7 (4.7º) • In almost all cases there was no difference in angles before or after surgery, except for level C4/C5 (loss of 2.5º flexion/extension ROM) and level C3/C4 (also 2.5º loss in lateral bending)
Kelly et al,[108] 2011	IDE RCT 13 sites N = 199	2 yr Flexion/extension radiographs and image analysis software (Quantitative Motion Analysis, Medical Metrics, Inc)	• In C-ADR group, pre-operative mean flexion/extension (7.6º ± 4.3) at index level was maintained above 8º through 2-yr follow-up. • In fusion group, mean angles dropped to 2º or below throughout follow-up • Univariate analysis showed significant increases in adjacent ROM angles for fusion group at superior and inferior levels for C4-C5 (+3.2º superior, +3.9º inferior). Increases did not persist in multivariate analysis
Anakwenze et al,[109] 2009	IDE RCT 13 sites N = 89/91	2 yr Neutral radiographs evaluating cervical lordosis with C6 distal endpoint with image analysis software (Quantitative Motion Analysis, Medical Metrics, Inc)	• Both C-ADR and fusion improved lordosis—no between-group differences in total cervical lordosis or superior adjacent-level lordosis • Greater lordosis restoration at index level with fusion, and greater loss in lordosis at inferior adjacent level with C-ADR
Auerbach et al[100] 2011	IDE RCT 13 sites N = 93/94	2 yr Flexion/extension on plain radiographs with image analysis software (Quantitative Motion Analysis, Medical Metrics, Inc)	• Total cervical ROM had greater improvement at 2 yr with C-ADR (+5.9º) than with fusion (−0.8º). Loss of 15.4º at fusion-treated site accounted for biggest contribution to overall loss in ROM • Loss at fusion-treated level was compensated for by increases at every other adjacent level, particularly at vertebrae 2 levels away from treated level
Peng et al,[110] 2009	IDE RCT N = 166	2 yr Neutral lateral flexion/extension and lateral bending x-rays with image analysis software (Quantitative Motion Analysis, Medical Metrics, Inc)	• Mean flexion/extension ROM (8.4º ± 0.7 to 9.6º ± 0.84) and lateral bending ROM (5.6º ± 0.5 to 5.7º ± 0.5) was maintained after surgery • Patients with less than 4-mm pre-operative disc height had an increased flexion/extension ROM, whereas those with more than 4-mm disc height had no change in flexion/extension • Two post-operative disc heights were associated with ROM: 1) those with more than 5 mm had significantly higher post-operative flexion/extension ROM than those with less than 5 mm (10.1º ± 1.0 vs. 8.3º ± 1.4; *P* = .014) and 2) those with more than 7-mm disc height had significantly lower post-operative lateral bending ROM (4.1º ± 1.3 vs. 5.7º ± 0.5; *P* = .04)

C-ADR Device Author, Year	Study Design No. of Patients (C-ADR/Fusion)	Follow-Up Duration Measurement of Cervical Motion at Treated Level	Conclusions
Delmarter et al,[82] 2010	IDE RCT 13 sites N = 103/106	4 yr Anteroposterior and lateral standing radiographs to assess flexion/extension ROM	• In C-ADR group, mean flexion/extension angles were 9.4º ± 5.96 at 2 yr and 9.1º ± 6.06 at 4 yr • In fusion group, 91% had ≤ 2º ROM at 2 yr and 96% had ≤ 2º ROM at 4 yr
Zigler et al,[91] 2013	IDE RCT 13 sites N = 103/106	5 yr Anteroposterior and lateral standing radiographs to assess flexion/extension ROM	• In C-ADR group, mean flexion/extension angles were 8.5º before surgery, 9.4º at 2 yr, and 8.1º at 5 yr
Janssen et al,[98] 2014	IDE RCT 13 sites	7 yr Anteroposterior and lateral standing radiographs to assess flexion/extension ROM	• In C-ADR group, mean flexion/extension angles' ROM at index level was 8.1º ± 5.91 vs. 0.66º ± 0.58 in fusion group ($P < .0001$)
Secure			
Vaccaro et al,[40] 2013	IDE RCT 18 sites 151/140	2 yr Anteroposterior and lateral standing radiographs to assess flexion/extension ROM	• For C-ADR group, mean flexion/extension ROM was 9.7º at 2 yr and mean flexion/extension translation was 1.2 mm
Two-level disease			
Mobi-C			
Davis et al,[93] 2015	IDE RCT 24 sites N = 225/105	4 yr Angular motion on flexion/extension and lateral bending radiographs	• In C-ADR group, ROM was maintained over 4-yr follow-up, and mean angles at 4 yr were at the superior level for flexion/extension (10.0º ± 6.0) over baseline (9.1º ± 4.9) and lateral bending (5.5º ± 3.6) over baseline 5.8º ± 3.4) • At inferior level, flexion/extension and lateral angles were also maintained: inferior level for flexion/extension (8.2º ± 5.3) over baseline (7.4º ± 4.3) and lateral bending (5.1º ± 3.4) over baseline (4.9º ± 3.3)
Radcliff et al,[84] 2017	IDE RCT 24 sites N = 225/105	7 yr Angular motion on flexion/extension and lateral bending radiographs	• In C-ADR group, ROM was maintained over 7-yr follow-up and mean angles at 7 yr were at superior level for flexion/extension (9.3º ± 5.8) and lateral bending (4.8º ± 3.4) and also at inferior level for flexion/extension (7.4º ± 5.2) and lateral bending (4.9º ± 3.4) • In fusion group, all angles for flexion/extension and lateral bending were less than 1º at both superior and inferior levels
Prestige-LP			
Lanman et al,[94] 2017	IDE RCT 30 sites N = 209/188	7 yr Angular motion on lateral flexion/extension radiographs	• In C-ADR group, mean angular ROM for target superior and inferior levels were maintained over 7 yr follow-up [data were graphed] • Mean angles for fusion group were not reported

Abbreviations: 3D, three-dimensional; C-ADR, cervical artificial disc replacement; CCT, controlled clinical trail; CI, confidence interval, CT, computed tomography; IDE, Investigational Device Exemption; RCT, randomized controlled trial; ROM, range of motion; RSA, radiostereometric analysis.

Cervical Kinematics: Other Radiologic Measures

Radiographic evaluations also involved other measures of treatment failure over longer terms for both C-ADR and fusion. Results of C-ADR and fusion failure at 2 years are reported earlier and included here for comparison. C-ADR failure and fusion failure have different time courses; C-ADR devices are essentially stable at implantation and heterotopic ossification can gradually develop, limiting motion of the disc at the treated level. Fusion is essentially not fixed at insertion, as it takes time to fill in or for the bone to fuse. In this case, pseudofusion will only decrease in time, unless it results in further instability and causes symptoms that require secondary surgery at the initially treated site.

C-ADR Radiographic Failure

For C-ADR, the formation of bridging bone between the endplates, heterotopic ossification, and loss of motion at the index site are considered to indicate a fused state and thus a failed disc implantation. This condition does not generally require additional surgery unless it causes symptoms. Evaluation of heterotopic ossification was usually based on the radiographic classification system by McAfee et al[64] of the degree of bridging bone and loss of vertebral motion at that level. The system consists of four grades: Grade 3 (some loss in motion) and Grade 4 (complete bridging and loss of motion) are clinically relevant. Rates of heterotopic ossification reported in follow-up for C-ADR arms of RCTs are summarized in Table 16. For one-level disease, rates of heterotopic ossification varied among the devices at 2-year follow-up from 1% to 5.9%. Heterotopic ossification gradually increased during follow-up; 5-year rates ranged from 2.5% to 10.2% and 7-year rates from 10% to 13%. For two-level involvement, rates were similar, with a rate of 11.9% reported at a 7-year follow-up.[94] The overall GRADE assessment of results on cervical kinematics was moderate, downgraded for risk of bias (Appendix 4, Table A5).

Table 16: Heterotopic Ossification at C-ADR-Treated Index Level on Follow-Up

C-ADR Device	Heterotopic Ossification–Grade 4 or Complete Bridging Bone, %			
	First Post-operative Examination	Second Post-operative Examination	Third Post-operative Examination	Fourth Post-operative Examination
One-level disease				
Kineflex	2 yr 1.0%[36]	5 yr 2.9%[96]	--	--
Mobi-C	2 yr 3.0%[38]	4 yr 7.9%[105]	7 yr 11.1%[84]	--
PCM	2 yr 1.1%[56]	5 yr 6.0%[89]	--	--
Prestige-ST	2 yr 0.8%[21]	3 yr 2.2%[106]	5 yr 6.2%[106]	7 yr 10%[97]
Prestige-LP	2 yr 5.9%[59]	3 yr 9.5%[77]	5 yr 10.2%[77]	7 yr 13%[77]
ProDisc-C	2 yr 2.9%[22]	4 yr 4.9%[82]	5 yr 8.3%[91]	7 yr 11%[92]
Two-level disease[a]				
Mobi-C	2 yr 3.7% (s) 2 yr 2.8% (i) 2 yr 4.9% (e)[41]	4 yr 10.2% (e)[93]	5 yr 10% (e)[83]	7 yr 6.5% (s) 7 yr 4.7% (i) 7 yr 11.1% (e)[84]
Prestige-LP	2 yr 2.0% (s) 2 yr 3.0% (i) 2 yr NR (e)[42]	3 yr 4.9% (s) 3 yr 4.9% (i) 3 yr 7.6% (e)[94]	5 yr 8.5% (s) 5 yr 8.5% (i) 5 yr 11.5% (e)[94]	7 yr 8.6% (s) 7 yr 7.3% (i) 7 yr 11.9% (e)[94]

Abbreviations: C-ADR, cervical artificial disc replacement; e, either inferior or superior level; i, inferior level; NR, not reported; s, superior level.
[a]Heterotopic ossification rate for two-level involvement included ossification at either or both levels.

Fusion Radiographic Failure

One-Level Cervical Degenerative Disc Disease

For fusion, pseudofusion resulting in instability and increasing motion at the index site is considered a failure. This failure is the main reason for secondary surgery at the index level in patients undergoing fusion. Pseudofusion in clinical trials has been defined by various radiologic characteristics: degree of bridging bone across intervertebral space, presence of graft radiolucent lines, or degree of segmental angular motion. Failure rates for pseudofusion would be expected to decline over time as bone gradually fills in the intervertebral space. Fusion failure rates were highly variable for one-level degenerative disc disease at the 2-year follow-up ranging from 2.5% to 20% (Table 17). In all cases, failure rates declined over time and at the 7-year follow-up was 4.5%.

Two-Level Cervical Degenerative Disc Disease

In the two-level fusion-treated patient groups, fusion failure at 2 years was high for both Mobi-C (20%)[41] and Prestige-LP (18%).[42] In both trials, fusion failure decreased over time and at 7 years was 9.1% for Mobi-C and 8% for Prestige-LP. Fusion failure rates also varied by radiology or investigator report. At the 5-year follow-up in the FDA RCTof Mobi-C,[84] the radiographically defined fusion failure rate was 9.5%, and the investigator-reported fusion failure rate was 14%.

Table 17: Fusion Failure at Index Level on Follow-Up

Fusion Arm in C-ADR Trial	Fusion Failure at Longitudinal Follow-Up			
	First Post-operative Examination	Second Post-operative Examination	Third Post-operative Examination	Fourth Post-operative Examination
One-level disease				
Mobi-C	2 yr, 20%[38]	4 yr, 5.6%[105]	7 yr, 4.5%[84]	--
PCM	2 yr, 8%[56]	5 yr, 5.6%[89]	--	--
Prestige-ST	2 yr, 2.5%[21]	5 yr, 1.5%[106]	--	--
ProDisc-C	2 yr, 9.8%[22]	4 yr, 4.5%[82]	5 yr, 7.6%[91]	--
Two-level disease				
Mobi-C	2 yr, 20%[41]	4 yr, 14.8%[93]	7 yr, 9.1%[84]	--
Prestige-LP	2 yr, 18%[42]	3 yr, 16.7%[94]	5 yr, 6%[94]	7 yr, 8%[94]

Abbreviation: C-ADR, cervical artificial disc replacement.

Adjacent-Segment Degenerative Disease

Longer-term follow-up reports (> 2 years) in RCTs also included a comparative evaluation of C-ADR versus fusion for adjacent-segment deterioration and the need for surgeries at the adjacent site. Radiologic assessments made for adjacent-level degeneration for one- and two-level involvement are outlined in Table 18.

Radiographic measurements of adjacent-segment deterioration (rASD) in clinical trials were variable and included various criteria: disc degeneration; facet degeneration; herniated nucleus pulposus or herniated disc; stenosis and/or instability. Different measurements and grading systems for severity were also applied depending on the radiographic methods: plain radiography, CT or MRI. Most studies either did not report radiographic findings of adjacent-level disc deterioration or did not report their definitions of rASD.

The overall GRADE assessment of results on adjacent-segment degeneration was low, rated down for risk of bias and imprecision (Appendix 4, Table A5).

One-Level Degenerative Disc Disease

The occurrence of rASD in longer-term follow-up (> 5 years) was reported for four C-ADR device trials, three involving FDA regulatory trials. For the Bryan C-ADR device, a second trial was conducted in the Netherlands by Donk et al[46] with a 9-year mean follow-up; this was the only trial to involve MRI investigations. In that trial, adjacent-segment deterioration was evaluated as MRI signal loss (degree of disc desiccation) of the intervertebral discs and no cases of symptomatic adjacent-segment degeneration were reported for 50 patients undergoing C-ADR and five of the 47 patients (10.6%) undergoing fusion.

In the FDA RCT of Kineflex-C,[96] for which severe adjacent-segment degeneration was reported, occurrence at 5-year follow-up was significantly lower in the C-ADR group than in the fusion group at the superior adjacent level (17% vs. 32%; $P < .01$) to the treated site, but not the inferior adjacent level (25% vs. 29%; $P > .05$).

In the FDA RCTs of Mobi-C and PCM, adjacent-segment degeneration was defined as any change in a degeneration score in follow-up, measured by the Kellgren–Lawrence classification for Mobi-C[84] and the Waraevens score for PCM.[89] In both trials, rASD progression was statistically significantly lower for the C-ADR group than for the fusion group only at the superior adjacent level to the index site in the 5-year follow-up for PCM (33% vs. 51%; *P* = .006),[89] and the 7-year follow-up for Mobi-C (40% vs. 65%; *P* < .010).[84]

Two other studies[111,112] evaluated the development of adjacent-segment deterioration and its potential risk factors in patients undergoing C-ADR or fusion in multiple C-ADR device FDA RCTs at their institutions.

The study by Jawahar et al[111] involved three FDA RCTs on C-ADR devices (Kineflex-C, Mobi-C, and Advent), with 93 patients treated for one- or two-level cervical degenerative disc disease followed for 3 years. Of these, 59 patients (24 with neurologic deficits) had disc replacement, and 34 patients (13 with neurologic deficits) had fusion. The criteria for adjacent-segment degeneration were defined by Hillibrand et al[16] as disc space narrowing, decreased disc height, herniated disc, or spinal nerve or cord compression.

The 3-year symptom-free survival rate was similar for the two surgery groups: 67.6 % ± 0.7% for C-ADR patients and 68.5% ± 1.1% for fusion patients.[111] None of the factors evaluated—patient age, sex, or smoking habit; affected levels (one vs. two); or osteopenia—were risk factors for symptom-free survival. Given clinical and radiologic parameters, cervical adjacent-segment degeneration was present in 18% of C-ADR and 15% of fusion patients. No differences between surgery groups were found for the actuarial survival rates (free of adjacent-segment disease) of 84.1% ± 0.6 for the C-ADR patients versus 78.3% ± 1.0 for the fusion patients. Previous symptom-related potential risk factors were also not associated with cervical adjacent-segment degeneration, although concurrent lumbar degeneration was a risk factor for both C-ADR and fusion groups. The presence of lumbar degeneration (compared with no lumbar degeneration) significantly decreased rates of survival free of adjacent-segment degeneration from 84.1% ± 0.6 to 52.3% ± 1.5% in the C-ADR group and from 78.3% ± 1.0 to 54.3% ± 1.9% in the fusion group.

Study results from Jawahar et al[111] were repeated in a larger study by Nunley et al[112] involving 170 patients treated at two sites in three FDA IDE RCTs. The same clinical and radiologic criteria were employed for cervical adjacent-segment degeneration. The 4-year survival time free of adjacent-segment degeneration was again similar for the two surgery groups: 78.3% ± 0.85 for the C-ADR group and 76.7% ± 0.56 for the fusion group. As in the earlier analysis, none of the factors, other than coexisting lumbar degeneration, were found to be risk factors for development of adjacent-segment degeneration. The percentage of patients in the two groups who had had lumbar degeneration was similar: 29% of disc implant patients and 31% of fusion patients. In the overall cohort, the 4-year survival rate free of adjacent-segment dengeration was statistically significantly higher for patients without lumbar degeneration than for those with lumbar degeneration (74.5% ± 0.6 vs. 55.5% ± 0.12; *P* = .023).

Two-Level Degenerative Disc Disease

Reports for two-level C-ADR devices with a 7-year follow-up are available for Mobi-C[84] and Prestige-LP,[94] although the occurrence of radiographically detected adjacent-segment degeneration at 7 years was reported only for Mobi-C. Degeneration was reported to be statistically significantly (*P* < .0001) lower for the C-ADR than for the fusion group at segments both inferior (30.3% vs. 66.7%) and superior (37.5% vs. 80.8%) to treated sites. Although disc

degeneration was greater for the fusion group, degeneration also steadily increased for the C-ADR group between the 4-year and 7-year follow-ups at both the superior (27.6%–37.5%) and inferior (16.4%–30.3%) adjacent levels (no time trend test reported).

Table 18: Adjacent-Segment Disease at Longer-Term Follow-Up of C-ADR

C-ADR Device Author, Year	Study Design	Radiologic ASD Outcome Measure	Conclusion
One-level disease			
Bryan			
Garrido et al,[113] 2010	IDE RCT 1 of 31 sites	4-yr adjacent-level HO (Grade 0–3)	• At 4 yr, adjacent-level HO (any grade) was less common in C-ADR (52%) than in fusion (84%) group • Median rank scores were significantly lower in C-ADR group (16 vs. 26.5; *P* = .004)
Donk et al,[46] 2017	Non-IDE RCT 1 site	9-yr MRI-defined rate of intervertebral disc desiccation	• At 9-yr follow-up of small study, no cases of symptomatic ASD occurred in C-ADR vs. 5 in fusion group
Kineflex-C			
Coric et al,[96] 2018	IDE RCT 21 sites	5-yr rASD changes from baseline at levels inferior and superior to index treated level were based on modified Waraevens score[a]	• At 5 yr, rASD was more prevalent at inferior than superior level in both surgery groups • C-ADR group had significantly lower (*P* < .01) rate of severe rASD than fusion at superior (17% vs. 32%) but not inferior (25% vs. 29%) levels
Mobi-C			
Hisey et al,[105] 2015	IDE RCT 23 sites	4-yr rASD measured by Kellgren-Lawrence scale[b]	• At 4 yr, rASD rate was significantly lower (*P* < .025) for C-ADR than for fusion group at both superior (34% vs. 53%) and inferior (30% vs. 50%) levels
Radcliff et al,[83] 2016		5-year rASD measured by Kellgren-Lawrence scale[b]	• At 5 yr, rASD rate was significantly lower (P < .0001) for C-ADR than for fusion at both superior (32.6% vs. 70.8%) and inferior (22% vs. 55%) levels
Radcliff et al,[84] 2017		7-yr rASD measured by Kellgren-Lawrence scale[b]	• At 7 yr, rASD rate was lower for C-ADR than for fusion group at both superior (40% vs. 65%) and inferior (44% vs. 63%) levels, but was significantly lower (*P* < .01) only at superior level
PCM			
Phillips et al,[89] 2015	IDE RCT 24 sites	7-yr rASD degeneration measured by Waraevens score[a]	• At 5 yr, rASD rate was significantly lower (*P* = .006) for C-ADR only at superior level (33% vs. 51%), not at inferior level (49% vs. 52%)
Two-level disease			
Mobi-C			
Davis et al,[93] 2015	IDE RCT 24 sites	4-yr rASD degeneration measured by Kellgren-Lawrence scale[b]	• At 4 yr, rASD rate was significantly lower (*P* < .0001) for C-ADR than for fusion at both superior (27% vs. 64%) and inferior (16% vs. 56%) levels

C-ADR Device Author, Year	Study Design	Radiologic ASD Outcome Measure	Conclusion
Radcliff et al,[83] 2016		5-yr rASD degeneration measured by Kellgren-Lawrence scale[b]	• At 5 yr, rASD rate was significantly lower ($P < .0001$) for C-ADR than for fusion at both superior (32.6% vs. 70.8%) and inferior (22% vs. 55%) levels
Radcliff et al,[84] 2017		7-yr rASD degeneration measured by Kellgren-Lawrence scale[b]	• At 7 yr, rASD rate was significantly lower ($P < .0015$) for C-ADR than for fusion at both superior (38% vs. 81%) and inferior (30% vs. 67%) levels
Prestige-LP			
Lanman et al,[94] 2017	IDE RCT	7-yr rASD NR	• 7-yr r-ASD NR

Abbreviations: ASD, adjacent-segment disease; C-ADR, cervical artificial disc replacement; HO, heterotopic ossification; IDE, Investigational Device Exemption; MRI, magnetic resonance imaging; NR, not reported; r-ASD, radiologic adjacent-segment degeneration; RCT, randomized controlled trial.

[a]Waraevens score measures changes in disc height, presence and size of osteophytes, and vertebral endplate sclerosis. Degeneration is judged to be no change at 0 points; mild (33% disc space narrowing, mild osteophytes, no endplate sclerosis) at 1 to 3 points; moderate (33%–66% disc space narrowing, moderate osteophytes, mild to moderate endplate sclerosis) at 4 to 6 points; or severe (> 66% disc space narrowing, severe osteophytes, moderate or greater endplate sclerosis) at 7 to 9 points.

[b]Kellgren-Lawrence scale measures disc degeneration as change in score from baseline; lateral views are used to divide disc degeneration into 4 grades: Grade 1, minimal anterior osteophytes; Grade 2, definite anterior osteophytosis with possible narrowing of disc space and some sclerosis of vertebral plates; Grade 3, moderate narrowing of disc space with definite sclerosis of vertebral plates and osteophytosis; and Grade 4, severe narrowing of disc space with sclerosis of vertebral plates and multiple large osteophytes.[114]

Secondary Adjacent-Level Surgeries

One-Level Cervical Degenerative Disc Disease

Trials of six C-ADR devices for one-level degeneration and secondary cervical surgeries at the index and adjacent site with longer-term (> 2 years) follow-up are detailed in Table 19. Index-level surgery rates were statistically significantly lower at a 7-year follow-up in the C-ADR than in the fusion group for three of the implants (Prestige-ST,[97] Prestige-LP,[77] ProDisc-C[92]). In the FDA RCT of Mobi-C, index-surgery rates for the fusion group were adjusted by subtracting index-level surgeries performed to enable fusion procedures at two levels. With this adjustment, index-surgery rates were no longer statistically significantly different between surgery groups at the 5-year,[115] or 7-year,[84] follow-up. In the FDA RCT of PCM,[89] surgery rates were combined for index and adjacent levels, and the overall 7-year surgery rates were not statistically significantly different (P = .123) between the surgery groups.

Adjacent-level surgery rates for the C-ADR devices, all with appropriate longitudinal survival or time to event analysis, varied. For the Mobi-C,[84] and Prestige-ST,[97] devices, cumulative rates for 7-year adjacent-level surgeries were statistically significantly lower for C-ADR than for fusion. However, 7-year adjacent-level surgeries for patients who had received the ProDisc-C device were not statistically significantly (P = .084) lower for the C-ADR than for the fusion group.[92] Adjacent-level surgeries for Prestige-LP were also not statistically significantly different between surgery groups.[77] However, this trial was a prospective controlled cohort study employing propensity-matched fusion controls from the prior FDA RCT of Prestige-ST, limiting comparisons.

Two-Level Cervical Degenerative Disc Disease

Trials of C-ADR devices for two-level involvement and secondary surgeries at the index-and adjacent-level sites for longer-term (> 2 years) are detailed in Table 19. Index-level surgery rates were significantly lower at a 7-year follow-up in the C-ADR than in the fusion group for Prestige-LP[94] but not for Mobi-C.[84] In the FDA RCT of Mobi-C, index surgery rates for the fusion group had again been adjusted by subtracting index-level surgeries performed to enable fusion procedures at two levels. Adjacent level-surgery rates for C-ADR devices, also with appropriate longitudinal survival or time to event analysis, were statistically significantly lower for both Mobi-C[84] and Prestige-LP[94] than for their fusion comparators.

In a report by Blumenthal et al,[116] longer-term surgeries were compared for C-ADR (n = 84) and fusion (n = 52) patients followed up for more than 2 years and participating in one of six FDA RCTs at one site. The overall surgery rate was 8.3% for the C-ADR group and 21.2% for the fusion group, but index-level (3.6% [n = 2] vs. 7.7% [n = 4]) and adjacent-level (4.8% [n = 4] vs. 13.5% [n = 7]) surgery rates were similar between the groups. At a 2-year follow-up, no surgeries were performed in the C-ADR groups, whereas five surgeries (4 for pseudarthrosis) were performed for the fusion groups. The mean length of time to all surgeries in time-to-event survival analysis was statistically significantly (P < .01) longer for patients in the C-ADR group than in the fusion group (54.6 months vs. 31.1 months) and also longer to adjacent-level surgery (52.8 months vs. 37.7 months), though this result was not statistically significant.

The overall GRADE assessment of results for secondary adjacent-level surgery was low, rated down for inconsistency and imprecision (Appendix 4, Table A5).

Table 19: Secondary Surgeries at Longer-Term Follow-Up of C-ADR Versus Fusion

C-ADR Device Author, Year No. C-ADR/ Fusion	Follow-up Duration, Surgery Level	Index- and Adjacent-Level Surgery		Δ C-ADR vs. Fusion
		C-ADR, % (No. of Patients)	Fusion, % (No. of Patients)	
One-level disease				
Bryan				
Sasso et al,[95] 2011 N = 242/221	4-yr index level	3.7% (n = 9)	4.5% (n = 10)	Δ 0.8% NS
	4-yr adjacent level	4.1% (n = 10)	4.1% (n = 9)	Δ 0 NS
Mobi-C				
Jackson et al,[115] 2016 N = 179/81	5-yr index level Adjusted index level[a]	3.4% (n = 6)	12.3% (n = 10) 8.6% (n = 7)	Δ 8.9%, *P* = .0097 Δ 5.2% *P* = .1194
	5-yr adjacent level	2.2% (n = 4)	11.1% (n = 9)	Δ 8.9%, *P* =.0043
Radcliff et al,[84] 2017 N = 179/81	7-yr index level Adjusted index level[a]	3.0 % (n = 5)	12.3 % (n = 10) 6.2 % (n = 5)	Δ 9.4% Δ 3.2% $P_{KM\ log\text{-}rank}$ = .219
	7-yr adjacent level	3.7% (n = 6)	13.6% (n = 11)	Δ 9.9% *P* = .007 $P_{KM\ log\text{-}rank}$ =.002
PCM				
Phillips et al,[89] 2015 N = 224/192	7-yr overall (index- and adjacent-level) surgery	8.5% (n = 18, 1 for ASD)	13% (n = 24, 19 for ASD)	Δ 4.5%, *P* = .237 $P_{KM\ log\text{-}rank}$ =.123
Prestige-ST				
Burkus et al,[106] 2010 N = 276/265	5-yr index level	4.0% (n = 11)	10.9% (n = 29)	Δ 6.9% *P* < .001
	5-yr adjacent level	2.9% (n = 8, 11 surgeries)	4.9% (n = 13, 16 surgeries)	Δ 2% *P* =.376
Burkus et al,[97] 2014 N = 276/265	7-yr index level	4.8% (n = 11)	13.7% (n = 29)	Δ 8.9% $P_{KM\ log\text{-}rank}$ < .001
	7-yr adjacent level	4.6% (n = 11)	11.9% (n = 24)	Δ 7.3% $P_{KM\ log\text{-}rank}$ < .008
Prestige-LP[c]				
Gornet et al,[77] 2015 N = 280/265	7-yr index level	6.4% (n = 18, 20 surgeries),16 procedures at 2 yr	10.9% (n = 29, 31 surgeries) 26 procedures at 2 yr	Δ 4.5% $P_{KM\ log\text{-}rank}$ < .004 For supplemental fixation only
	7-yr adjacent level	9.6% (n = 27)	8.3% (n = 22)	Δ −1.3% NS
ProDisc-C				
Delamarter, et al,[117] 2013 N = 103/106	5-yr index level	0.9% (n = 1)	7.5% (n = 8)	Δ 6.6%
	5-yr adjacent level[b]	1.9% (n = 2)	5.7% (n = 6)	Δ 3.8%
	5-yr probability of no secondary surgery (index or adjacent)	97.1%	85.5%	Δ 11.6% P_{wald} = .008 (for either surgery site

C-ADR Device Author, Year No. C-ADR/ Fusion	Follow-up Duration, Surgery Level	Index- and Adjacent-Level Surgery C-ADR, % (No. of Patients)	Index- and Adjacent-Level Surgery Fusion, % (No. of Patients)	Δ C-ADR vs. Fusion
Janssen et al,[92] 2015	7-yr index level	6.0% (n = 6, 6 procedures	15% (n = 16, 19 procedures)	Δ 6% $P_{KM\ log\text{-}rank}$ = .022
	7-yr adjacent level[b]	5.8% (n = 6, 6 procedures)	12.3% (n = 13, 22 procedures)	Δ 6.5% $P_{KM\ log\text{-}rank}$ = .084
Two-level disease				
Mobi-C				
Radcliff et al,[83] 2016 N = 234/105	5-yr index level	4.3% (n = 9)	16.2% (n =17)	Δ 13.2% $P_{KM\ log\text{-}rank}$ = .0002
	5-yr adjacent level	3.1% (n = 7, 5 for ASD, 2 involving herniated disc)	11.4% (n = 12, 11 for ASD, 2 involving herniated disc)	Δ 8.0% *P* = .0059
Radcliff et al,[84] 2017 N = 234/105	7-yr index level Adjusted index level[a]	4.4% (n = 10)	16.2% (n =17) 10.5% (n = 11)	Δ 11.8% Δ 6.1% $P_{KM\ log\text{-}rank}$ = .062
	7-yr adjacent level	4.4% (n = 10)	11.4% (n = 12)	Δ 7.0% $P_{KM\ log\text{-}rank}$ = .009
Prestige-LP				
Lanman et al,[94] 2017 N = 209/188	7- yr index level	4.2% (n = 8, 10 procedures)	14.7% (n = 22, 27 procedures)	Δ 10.5% log HR (95% CI) = −1.29 (−212 to −0.46)
	7-yr adjacent level	6.5% (n = 12, 12 procedures)	12.5% (n = 17, 22 procedures)	Δ 6.0% NS log HR (95% CI) = −0.59 (−1.35 to 0.156)

Abbreviations: Δ, difference; ASD, adjacent-segment disease; C-ADR, cervical artificial disc replacement; CI, confidence interval; HR; hazard ratio; KM, Kaplan-Meier; NS, not significant.

[a]Corrected index surgery rate involves subtracting surgeries performed at the index level, usually plate removal, in order to facilitate surgeries at adjacent level(s).

[b]Adjacent-level surgery included both stand-alone adjacent and adjacent along with index procedure.

[c]The Prestige-LP study[37] was a prospective cohort study employing fusion controls from an earlier Prestige-ST trial.[21]

Longer-Term Safety

Adverse events in longer-term follow-up were reported for different time intervals, such as events occurring cumulatively, combining early and late events, or those occurring from defined post-operative time points beyond 2 years (Table 20). The trials also varied widely in their reporting of adverse events in terms of the severity of the event or the likelihood that the event was device or surgery related.

One-Level Cervical Degenerative Disc Disease

In the three trials reporting adverse events occurring beyond 2 years, differences between the surgery groups were not statistically significant for any adverse event between 2 to 4 years for Bryan[95]; major adverse events between 2 to 7 years for PCM[89]; any adverse event between 2 and 7 years for Prestige-ST.[97] In the four trials reporting cumulative adverse events, differences between surgery groups were also not statistically significantly different for device-

or surgery-related adverse events up to 5 years for Kineflex-C[96]; any major adverse event or any device-or surgery-related adverse event up to 4 years and any adverse event up to 7 years for Mobi-C[84]; any adverse event up to 7 years for Prestige-LP[77] and any device- or surgery-related adverse event up to 5 years and any device-related adverse event up to 7 years for ProDisc-C.[91]

Very late complications, those occurring beyond 4 years for C-ADR devices, were reported by Hacker[118] in two separate FDA RCTs, for Bryan and the Prestige-LP. Overall, 10.6% (5/47) returned for neck- or extremity-related symptoms. Although female patients were equally represented (53%) in the trials, all with these late complications were female. Conditions causing symptoms involved device subluxation (movement) with ventral cord effacement (1 patient); marked vertebral bone loss and deformity (1 patient); haloing (bone loss) around the device (2 patients); and loss of vertebral height, deformity, and heterotopic ossification compressing a nerve root in the neural foramen (1 patient).

Two-Level Cervical Degenerative Disc Disease

For Mobi-C, any clinical event committee–confirmed major adverse event was not statistically significant between the surgery groups at 4,[93] 5,[83] or 7[84] years. Three levels of adverse events were reported for Prestige-LP[94], and device-related serious adverse events up to 7 years were statistically significantly lower for the C-ADR group.

The overall GRADE assessment of results for long-term safety was moderate, rated down for risk of bias (Appendix 4, Table A5).

Table 20: Longer-Term Safety of C-ADR Versus Fusion for Cervical Degenerative Disc Disease

C-ADR Device Author, Year	Observation Period	C-ADR	Fusion	Comments[a]
One-level disease				
Bryan				
Sasso et al,[95] 2011 N = 242/221	Between 2 and 4 yr, all AEs. (any clinical adverse sign, symptom, syndrome, or illness occurring or worsening by physical examination, clinical evaluation, subject interview, or medical charts)	18.2% (n = 44, 63 AEs)	16.7% (n = 36, 64 AEs)	• Most AEs were medically related, but not related to index surgery or device • Differences between surgery groups were not significant
Kineflex-C				
Coric et al,[96] 2018 N = 136/133	5 yr, any AEs Device related (definitely, probably, possibly) Procedure related (definitely, possibly)	Device related 38.2% (n = 52, 38 possibly) Procedure related 5.9% (n = 8)	Device related 35.3% (n = 47, 37 possibly) Procedure related 4.5% (n = 6)	• Device- or procedure-related AE rates did not differ between surgery groups • Serious AEs[a] were not reported, and AEs occurring after surgery and longer term were combined
Mobi-C				
HIisey et al,[105] 2015 N = 164/81	4 yr, overall major AEs Major AE (related to device or surgery)	Overall major AEs 9.8% Major AEs 4.3% Between 2- and 4-yr AEs 5.6% (n = 10, 2 were major AEs)	Overall major AEs 9.9% Major AEs 3.7% Between 2- and 4-yr serious AEs 2.5% (n = 2, none were major AEs)	• Differences between surgery groups in overall major AEs were not significant • After 2 yr, post-operative serious AEs were uncommon
Radcliff et al,[84] 2017 N = 179/81	7 yr, any AE. Defined as any clinical adverse sign, symptom, or syndrome or illness occurring or worsening and assessed by CEC	6.1% (10/164)	3.7% (3/81)	• Overall AE rate was higher in C-ADR group than in fusion group, but difference between surgery groups was not significant

C-ADR Device Author, Year	Observation Period	C-ADR	Fusion	Comments[a]
PCM				
Phillips et al,[89] 2015 N = 224/192	Between 2 and 7 yr, new serious AEs, judged by CEC Device-related serious AEs	21% (45/214) 0.5% (N = 1)	17.4% (33/190) 1.1% (N = 2)	• Most AEs were systemic or medical in nature • Device-related serious AEs were infrequent after 2-yr post-operative follow-up and did not differ between surgery groups
Prestige-ST				
Burkus et al,[97] 2014 N = 276/265	7 yr, cumulative overall AEs	Any AE after 2 yr 8.7% (n = 24) Spinal event (undefined) 8% (n = 22) Urogenital AEs (undefined) 8.7% (n = 24) Dysphonia/dysphagia (undefined) 1.8% (n = 5)	Any AE after 2 yr 4.9% (n = 13) Spinal event (undefined) 12.1% (n = 32) Urogenital AEs (undefined) 5.7% (n = 15) Dysphonia/dysphagia (undefined) 1.5% (n = 4)	• Overall AE rate after 2 yr was higher, but not significantly higher, in C-ADR group • Spinal AEs (20.9% vs. 38.9%, *P* < .001) and urogenital AEs (20.1% vs. 12.2%, *P* = .02) were significantly different between surgery groups when 2-yr early follow-up events were added
Prestige-LP				
Gornet et al, [77] 2015[b] N = 280/265	7 yr, cumulative overall AEs	Any AE up to 7 yr 96.8% (95% HPD 94.3%–98.5%) Dysphonia/dysphagia (undefined) up to 2 yr, n = 27 Between 2 and 7 yr, n = 6	Any AE up to 7 yr 87.7% (95% HPD 83.2%–91.7%) Dysphonia/dysphagia up to 2 yr, n = 23 Between 2 and 7 yr, n = 9	• 7-yr difference (9%, 95% HPD 4.4%–14.1%) in overall AEs is statistically lower in fusion group • Dysphonia/dysphagia occurred more often within 2 yr in both surgery groups, and rates were low and not significantly different between groups at 7-yr follow-up

C-ADR Device Author, Year	Observation Period	C-ADR	Fusion	Comments[a]
ProDisc-C				
Zigler et al,[91] 2013 N = 103/106	5 yr, device-related AE rate 5 yr, surgery-related AE rate (dysphagia, edema, or gastrointestinal or genitourinary symptoms)	1% (n = 1) 11.7% (n = 12)	2.8% (n = 3) 20.8% (n = 22)	• Rate of device-related AEs was low and not significantly different between surgery groups • Rate of surgery-related AEs was also lower for C-ADR group, but not significantly lower (*P* = .09)
Janssen et al,[92] 2015	7 yr, device-related AEs	28% (n = 30, 48 AEs)	27% (n = 28, 41 AEs)	• Most common AE was neck pain (isolated or with shoulder or arm pain) • Differences between surgery groups were not clinically relevant for any category of AE
Two-level disease				
Mobi-C				
Davis et al,[93] 2015 N = 225/105	4 yr, CEC confirmed major AE rate	4.0%	7.6%	• 4-yr major AE rate was not significantly different between surgery groups
Radcliff et al,[83] 2016 N = 225/105	5 yr, CEC confirmed major AE rate	4.4%	8.6%	• 5-yr major AE rate was not significantly different between surgery groups
Radcliff et al,[84] 2017 N = 225/105	7 yr, CEC confirmed major AE rate	5.3% (n = 12)	8.6% (n = 9)	• 7-yr major AE rate was not significantly different between surgery groups
Prestige-LP				
Lanman et al,[94] 2017 N = 209/188	7 yr, ≥1, any AE Any serious AE Possible device-related AEs Serious, possible device-related AEs[c]	99.1% 56.7% 26.6% 3.2%	98.2% 68.2% 27.7% 7.2%	• Most AEs occurred in first 2 yr after surgery • Any serious AE (HR −0.38, 95% BCI −0.65 to −0.11) and possible serious device-related AE (HR −1.19, 95% BCI, −2.29 to −0.15) categories were statistically lower in C-ADR group

Abbreviations: AE, adverse event; BCI, Bayesian credible interval; C-ADR, cervical artificial disc replacement; CEC, Clinical Events Committee; HR, hazard ratio, HPD, highest posterior density.

[a]Serious adverse events defined as World Health Organization Grade 3 or 4.

[b]Controlled clinical prospective trial employing indication-matched fusion controls from prior Prestige-ST randomized controlled trial.

Disc Durability: Wear Characteristics and Biological Reactivity

C-ADR devices are made of various polymers and metal alloys and do not have the long-term history of polyethylene and cobalt-chromium alloys as orthopedic bearing material.[119] We do not yet know how particulate debris, metal ions, corrosion products, and biological responses to these byproducts could vary among these materials.

All C-ADR devices with regulatory approval submitted extensive information on in vitro testing of wear characteristics in their FDA Summary Safety and Effectiveness Data reports. Most employed testing cycles of 10 to 20 million repetitions under various degrees of stress in various positions. Testing scenarios, however, usually involve perfect placement of the disc, which is not always the case in vivo. In vitro testing also involves the disc in isolation without interactions with other components in the spine and does not consider device sizing. Although wear characteristics of implant devices can be evaluated with in vitro tests, biological responses to any wear debris cannot.

Reports on two metal-on-metal implants, Kineflex-C and Prestige-LP, included information on in vivo serum metal ion levels. The Prestige-LP had prospective follow-up with blood sampling for metal ions on a 30-subject cohort in the FDA trial. Serum metal ion concentrations were analyzed specifically for titanium, vanadium, and aluminum. In the FDA RCT of two-level Prestige-LP, an independent laboratory performed the metal ion analysis and reported that none of the patients had symptoms related to metal ion sensitivity. Serum levels, however were not reported.

In the FDA RCT of Kineflex-C, information was collected on serum metal ions (cobalt and chromium) prospectively during follow-up visits of 27 patients and reported after the 5-year follow-up.[96] Metal ion analysis was performed by an independent laboratory that compared values regulatory agencies recommend monitoring for hip metal-on-metal implants. Both mean cobalt (0.21 µg/L) and chromium (0.31 µg/L) serum levels remained low throughout the 5-year follow-up and were significantly below the levels of 7 µg/L for both cobalt and chromium specified by the Medicines and Healthcare Products Regulatory Agency. Serum levels were also lower than a lower suggested threshold level of 4 µg/L for cobalt and 4.6 µg/L for chromium.[120]

Two systematic reviews[35,121] reported on the in vivo wear characteristics and biological responses to implanted artificial cervical discs.

Systematic Review by Lehman et al

Lehman et al[121] reported on studies involving wear of discs in biomechanical simulations and in vivo explants. The authors classified the quality of the evidence as low. Four reports on wear testing of C-ADR devices (Bryan, Active-C, ProDisc-C, Prestige) in biomechanical simulations all used more than 10 million cycles under an applied force of 150 N for various flexion, extension, lateral, and axial rotations. There was minimal loss in mass, height, or volume, and no device failures were reported. However, for one device cracks at screw holes and heads appeared after 5 million cycles.

Ex vivo reports involved 10 case studies on histologic and pathologic findings for C-ADR retrievals involving four devices. All but one of the devices (explanted at 39 months) had been explanted within 2 years, most because of pain or ongoing symptoms. Infection was uncommonly reported. Only one implant became infected; one other developed osteolysis (bone

loss), which was interpreted as a reactive metal hypersensitivity. Metallic debris in periprosthetic tissues was noted in five cases, and in four cases reactions were judged to be hypersensitivity to metal. Implant wear noted in two cases was reported to be less than in vitro biomechanical simulations showed. Inflammation was common, affecting seven cases.

Systematic Review by Veruva et al

A systematic review by Veruva et al[35] on device wear involved seven reports (five case reports and two clinical series). The authors judged the studies to be of good quality based on the MINORS scale (Methodological Index for Non-randomized Studies). [122]
Both polymeric and metallic debris was reported in the series. Inflammatory responses involving both innate and adaptive responses were noted for three metal devices: titanium (Bryan), cobalt-chromium-molybdenum (Co-Cr-Mo) alloy (ProDisc-C, Kineflex), and stainless steel (Prestige-ST). This systematic review mentioned implant retrievals from two larger series for ProDisc-C[123] and for Prestige-ST and Bryan.[119]

One international implant-retrieval program[123] for ProDisc-C from 2005 to 2011 involved 24 spine surgeons. Thirty implants from 29 patients of mean age 45.1 years (range 31–57 years) survived around 1 year (range 2 days to 3.5 years) before explantation. Surgeons reported pain was the most common indication for implant removal; reactivity to metal was suspected in one case 5 months after implantation. The most common finding was metal-on-metal (Co-Cr-Mo) endplate impingement in 80% of the explanted devices. Evidence of wear was also indicated by third-body particles of titanium displaced from the porous coated surface of the implant in seven (23%) implants. Wear pattern was not associated with any device depth or heights. Most (70%) cases developed bone ongrowth in areas on the titanium–plasma sprayed endplates, but two of the six devices explanted for atraumatic loosening showed no bone ongrowth. Atraumatic loosening of implants was associated with various conditions: hypermobility due to bone spurs, osteolysis, collapse of endplate, osteoporotic bone, and implant migration.

Kurtz et al[119] reported on explants of Prestige-ST (stainless steel) and Bryan (titanium alloy), both through Medtronic's IDE studies or their quality system. Twenty Prestige-ST devices were explanted from 20 patients, on average 2 years (range 0.3–7.0 years) after implantation. Evidence of wear, anterior impingement, was noted in 69% (11/16) of implants. Localized screw hole fretting and fretting near bone screw heads were typically observed along with locking screw fractures (three cases) and bone screw fractures (one case). Focal microscopic metallic debris was commonly observed within fibrous tissues, particularly at the device interface. The immune response in host tissue adjacent to the implant was characterized by an innate chronic inflammatory response.

The Bryan explant review consisted of 35 implants from 30 patients explanted after 3.2 years (range 0.3–7 years). Endplate impingement was noted in 30% (9/30) of cases; rim impingement resulted in titanium debris with third-body wear of the polycarbonate urethane nucleus causing substantial height loss (1.6 and 2.6 mm). Degradation of the polyurethane sheath was observed in some cases (27%, 4/15) and did not appear to be related to implant time. Biodegradation was thought to be caused by release of reactive oxygen by macrophages and foreign body giant cells. No implant removals in this series were attributed to degradation of the polyether urethane sheath surrounding the nucleus. Follow-up in the series was too short to evaluate inflammatory responses to any debris.

Discussion

The safety and effectiveness of a number of C-ADR devices were investigated in several multicentre regulatory randomized controlled trials with external oversight of key outcome assessments. Primary outcomes in the trials involved the same composite primary outcomes of 2-year overall treatment success, although the composite subcriteria were selected, defined, and evaluated differently across trials. The subcriteria for overall treatment success specified success in each of several subcriteria: no major adverse events, no secondary surgery at the index site for technical failure, maintenance or improvement of neurologic status, and disability reduction with clinically defined thresholds. Both types of spinal surgery had high levels of overall treatment success for most patients, and all studies evaluating C-ADR devices easily satisfied criteria for statistical noninferiority to fusion (i.e., C-ADR was found not to be worse than fusion) for the treatment of cervical degenerative disc disease.

Major adverse events related to C-ADR devices versus fusion were evaluated in all trials with independent verification of events and their severity by clinical event committees. Both the C-ADR and fusion groups had low rates (< 5%) of serious adverse events. Dysphagia, a common complication of cervical surgery in which an anterior approach is used, is thought to be a greater concern with fusion than C-ADR because the fusion plates project into the anterior space. Several studies that specifically evaluated dysphasia found that dysphagia declined over time in both surgery groups and was less common with C-ADR. However, the limited diagnostic ability of non-instrumental survey methods and the under-assessment in follow-up likely means that reported estimates for this condition are unreliable.

Various definitions and methods were used to evaluate maintaining or improving neurologic status, but in several trials, improvement was greater for C-ADR than fusion. However, as the first steps in either C-ADR or fusion involve disc removal and decompression of the neural elemements (nerves and spinal cord) through an anterior approach, it is unclear how C-ADR could have a greater effect on neurologic success.

Pain-related disability was evaluated with a reliable disease-specific instrument, the Neck Disability Index. Improvements were both statistically significant and clinically relevant in both surgery groups.

Much of the difference in overall treatment success between surgery groups is attributed to differences in the number of second surgeries performed at the index site (initial surgery site). Second surgeries were performed for different reasons in various surgery groups. As C-ADR devices are stable (i.e., fixed in place) at the time of implantation, a second surgery should be needed only if the device was incorrectly sized or had been unsuccessfully placed. Fusion, on the other hand, requires time for bone to form between the vertebral endplates to achieve stability (i.e., fusion). Judgment is required to assess whether fusion is in progress or whether pseudarthrosis (i.e., failed fusion) is present. Generally, a second surgery is offered and undertaken if radiologically detected pseudarthrosis is accompanied by unrelieved pain. The rate for index-level second surgery within 2 years of the index surgery varied across studies but was statistically significantly higher for fusion than for C-ADR in all trials. For two-level cervical degenerative disc disease, the level of fusion failure was much higher, suggesting that C-ADR offers a technical advantage over fusion if a two-level surgery is being considered.

C-ADR and fusion were also compared across several other secondary outcomes. Perioperative outcomes were largely comparable, but duration of surgery varied among C-ADR devices and was usually longer than for fusion. Differences in duration might be attributable to

several factors, including endplate preparation (for fusion) and whether lead-in or first cases were included in results. Other perioperative outcomes, such as estimated blood loss and length of stay, did not differ between surgery groups.

Recovery or return to work after spinal surgery was reported in all trials, with return-to-work times being statistically significantly shorter for patients undergoing C-ADR than for those undergoing fusion in all trials. By 2 years, however, return-to-work and employment rates were similar between surgery groups. However, there were some limitations to the reports. Baseline or pre-operative unemployment or retirement status was generally not reported, and measures of return to work would thus not appropriately reflect recovery for patients unemployed or retired prior to surgery. Return to usual activities, which would better reflect recovery for retired or unemployed patients, was not evaluated in any studies. One trial did focus on a population of special interest—a Workers' Compensation population in the United States—and recovery rates were again statistically significantly better for those who had undergone C-ADR than for those who had undergone fusion.

Health-related quality of life (HRQOL) was also extensively evaluated longitudinally after surgery with validated instruments in most trials. Results were consistent in that patients in both surgery groups had statistically significant and clinically relevant improvements in mean HRQOL scores—improvements that were maintained at a 2-year follow-up. In the months following surgery, improvements in HRQOL mean scores were statistically significantly better for patients in C-ADR groups than in fusion groups, although differences at a 2-year follow-up were no longer statistically significant. Generally, mean score improvements in physical subdomains in both surgery groups were greater than in those for the mental subdomains, which aligns with the objective of surgery to improve physical function.

Treatment satisfaction was investigated in most trials, and high levels of satisfaction were consistently reported by patients in both surgery groups despite the use of different satisfaction measures and response formats. Satisfaction levels reported were always higher, although not statistically significantly, among patients in C-ADR groups versus fusion groups. Surgeons were also more likely to judge outcomes for patients undergoing C-ADR as being excellent with complete recovery than for patients undergoing fusion. For two-level procedures surgeons' judgment that patients had excellent outcomes was even higher for C-ADR than for fusion. As results were reported separately for patients and surgeons, it was impossible to evaluate any concordance or agreement in their opinions or judgments of treatment success.

C-ADR Advantages

Cervical motion at the treated level was demonstrated in each C-ADR trial and was maintained in longer-term follow-up. The success of C-ADR devices was balanced with failures in other cases in which bridging bone that develops between vertebral endplates limits motion, essentially producing a fused disc. The approximate disc fusion rate of 10% at a 5-year follow-up in the trials is likely to be an underestimate for several reasons. The occurrence of disc fusion is a function of the adequacy of follow-up, measurement limitations, and length of follow-up. Radiographic failure of C-ADR has more limited clinical consequences than radiographic failure of fusion because a fused disc, unlike fusion failure, rarely requires further surgery unless the condition begins to cause symptoms.

The general assumption that maintaining motion limits the occurrence or progression of cervical degenerative disc disease at adjacent levels is the major anticipated advantage of C-ADR over fusion. Radiologic investigations into these degenerative changes were problematic for several

reasons. Difficulties in performing radiographic assessments were often reported. When examinations were performed, the diverse criteria used to identify degenerative changes in cervical discs and the use of different rating systems limited comparisons across trials. Although degenerative changes were reported to occur less frequently in patients undergoing C-ADR than fusion, the prevalence of degenerative changes in the C-ADR group was still high, likely representing an underestimate given short-term follow-up in relatively young patients. Subsequent surgery rates for symptomatic adjacent-level disease were similarly limited. With the short-term follow-up so far available, few surgeries have occured and the full burden of adjacent-level surgeries has likely not yet been realized.

Limitations

Several conditions limited the generalizability of trial results. Although patients with a range of ages were eligible for trials, those participating were mainly in their 40s, limiting conclusions about treatment effectiveness in older patients. Patients with symptoms of radiculopathy or myelopathy were eligible for trials, but few patients had myelopathy as a presenting symptom, and data for these conditions are limited. Investigations into C-ADR devices initially focused on effectiveness in one-level cervical degenerative disc disease and only later began to investigate effectiveness in two-level disease, resulting in less trial evidence and shorter follow-up for two-level disease. There is some concern about the durability of C-ADR devices, as there is for any implantable device, but the evidence we reviewed was insufficient to determine the long-term durability of C-ADR devices.

Conclusions

In carefully selected patients with cervical degenerative disc disease undergoing C-ADR or fusion, there is evidence that:

- C-ADR is an alternative to fusion for cervical degenerative disc disease given outcomes that are statistically noninferior to fusion: perioperative outcomes (GRADE high), health-related quality of life (GRADE high), patient satisfaction (GRADE high), and overall treatment success for one-level cervical degenerative disc disease (GRADE moderate)
- C-ADR might be preferable to fusion for cervical degenerative disc disease given outcomes that are statistically superior to fusion: quicker recovery and return to work (GRADE moderate), higher technical success and lower rate of re-operation at the index site (GRADE moderate), maintenance of more normal spinal segment kinetics (GRADE moderate), and higher overall treatment success for two-level cervical degenerative disc disease (GRADE moderate)

We are uncertain if adjacent-level surgery rates differ between C-ADR and fusion for one-level and two-level cervical degenerative disc disease (GRADE low). Evidence was also insufficient to determine the long-term durability of C-ADR devices.

ECONOMIC EVIDENCE

Research Questions

1. Based on the published literature, what is the cost-effectiveness of cervical artificial disc replacement (C-ADR) compared with cervical anterior discectomy and fusion (fusion) for people with one-level symptomatic cervical degenerative disc disease?

2. Based on the published literature, what is the cost-effectiveness of C-ADR compared with fusion for people with two-level symptomatic cervical degenerative disc disease?

Methods

Economic Literature Search

We performed an economic literature search on July 17, 2017, for studies published from inception until the search date. To retrieve relevant studies, we developed a search using the clinical search strategy with an economic filter applied.

Database auto-alerts were created in MEDLINE and Embase and monitored for the duration of the health technology assessment. We performed targeted grey literature searching of health technology assessment agency sites, clinical trial registries, and Tufts Cost-Effectiveness Analysis Registry. See Clinical Evidence, Literature Search, above, for further details on methods used, and Appendix 1 for literature search strategies, including all search terms.

Literature Screening

A single reviewer reviewed titles and abstracts, and, for those studies likely to meet the eligibility criteria, we obtained full-text articles and performed further assessment for eligibility.

Inclusion Criteria

- Studies comparing C-ADR with fusion in people with cervical disc degeneration
- English-language full-text publications
- Studies published between inception and the search date
- Cost-utility analyses, cost-effectiveness analyses, cost-benefit analyses, or cost minimization analyses

Exclusion Criteria

- Reviews
- Abstracts, letters, and editorials
- Unpublished studies

Outcomes of Interest

- Cost
- Quality-adjusted life-years (QALYs)
- Incremental cost and incremental effectiveness

- Incremental cost per QALY gained

Data Extraction

We extracted relevant data on the following:

- Source (i.e., name, location, year)
- Population and comparator
- Interventions
- Outcomes (i.e., health outcomes, costs, QALYs, and incremental cost-effectiveness ratio [ICER])

We contacted authors of the studies to provide clarification as needed.

Study Applicability and Limitations

We determined the usefulness of each identified study for decision-making by applying a modified applicability checklist for economic evaluations that was originally developed by the National Institute for Health and Care Excellence (NICE) in the United Kingdom. The original checklist is used to inform development of clinical guidelines by NICE.[124] We retained questions from the NICE checklist related to study applicability and modified the wording to remove references to guidelines and to make questions Ontario-specific. The number of studies judged to be directly applicable, partially applicable, or inapplicable to the research question is summarized.

Results

Literature Search

The literature search yielded 145 citations published from inception until July 17, 2017, after removing duplicates. We excluded a total of 128 articles based on information in the title and abstract. We then obtained full texts of 17 potentially relevant articles for further assessment. Three articles were excluded because they were utility studies.[125-127] In the end, 14 articles were included in the final review.[128-141] Figure 3 presents the flow diagram for the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA).

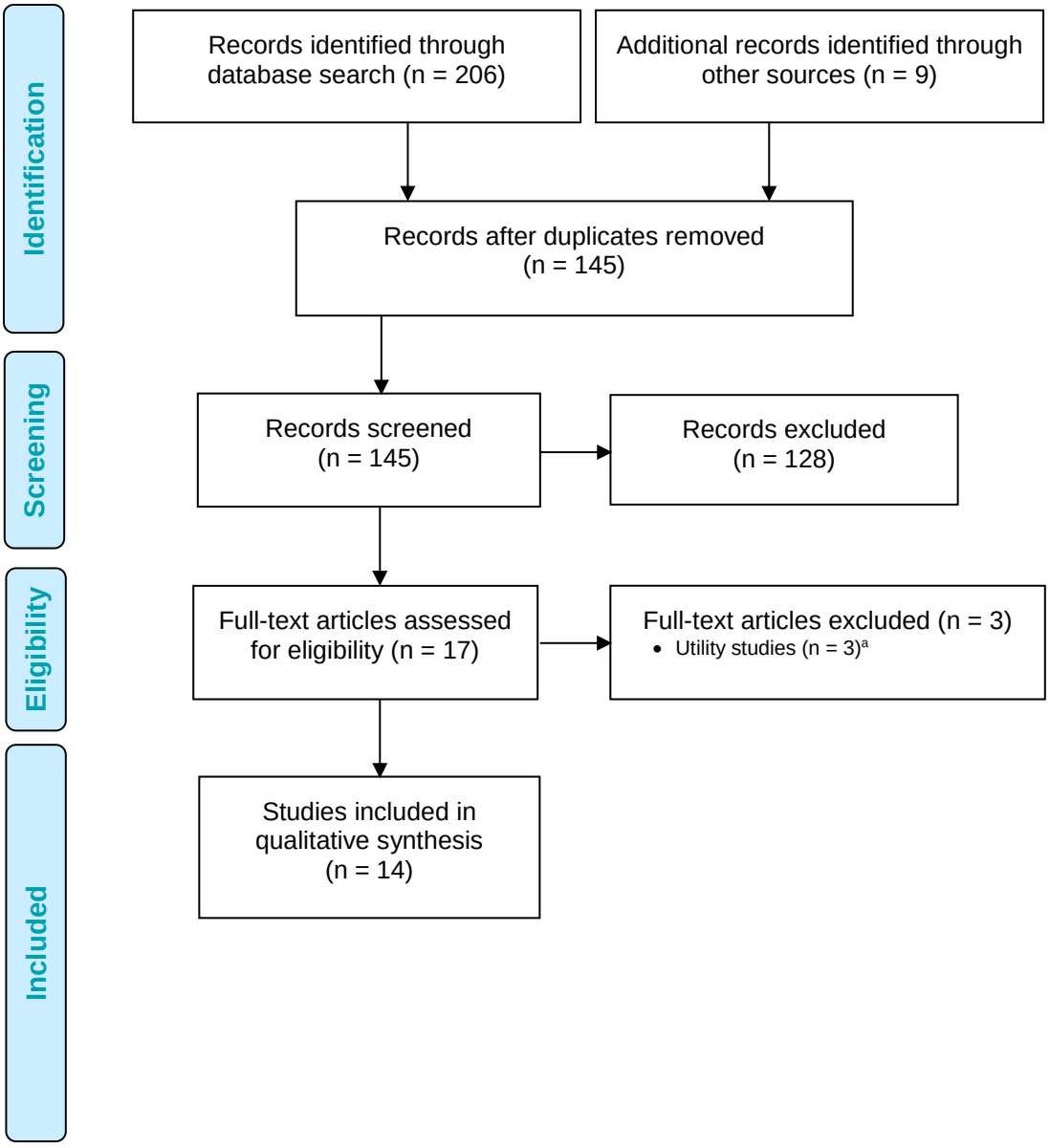

Figure 3: PRISMA Flow Diagram—Economic Search Strategy

Abbreviation: PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-analyses.
[a]These utility studies were used to inform our model.
Source: Adapted from Moher et al.[32]

Review of Included Economic Studies

In this section we use the term “subsequent surgery” to refer to any subsequent corrective surgery at the index level or any subsequent surgery to treat adjacent-segment disease. In future sections of this report, we will further use specific terms, defined by the U.S. Food and Drug Administration (FDA), for subsequent corrective surgeries at index level, including “revision,” “removal,” “supplemental fixation,” and “reoperation.” Unfortunately, in many publications, the term “reoperation” has been used to refer to any subsequent surgery. In such cases, we use “subsequent surgery” instead to eliminate possible confusion.

One-level Symptomatic Degenerative Disc Disease

Table 21 summarizes the results of the 10 included studies for one-level cervical degenerative disc disease. Five studies applied models[129,131,132,134,137]; three were economic evaluations alongside randomized controlled trials (RCTs),[128,130,135] and two were retrospective reviews of administrative data.[133,136] Of the five studies that used models, two used RCTs that reported results for specific artificial discs.[132,137]

Menzin et al[128] (United States, 2010) performed a cost-benefit assessment of C-ADR that compared the Prestige-ST device with fusion to evaluate the difference between incremental medical costs and gains in work productivity in people with one-level cervical degenerative disc disease and radiculopathy or myelopathy using a societal perspective. They used 2-year follow-up data from an RCT performed for investigational device exemption (IDE) (Mummaneini et al,[21] 2007) and published sources of cost and wage data. Their defined direct cost was the costs of the index surgery, secondary procedures, and medical devices. The mean initial procedure cost for C-ADR was $111 higher than for fusion, and the mean secondary procedure cost was $542 lower. The total direct cost per patient was therefore $431 lower with C-ADR relative to fusion. They estimated that C-ADR patients worked 38 days longer, on average, than fusion patients, yielding an average gain in work productivity of $6,547. They found that, compared with fusion, C-ADR was associated with an average total savings of $6,978 per patient over 2 years.

Table 21: Economic Literature Review for One-Level Symptomatic Cervical Degenerative Disc Disease

Author, Year, Country	Study Design and Perspective	Population	Intervention and Comparator	Results		
				Health Outcomes	Costs	Cost-Effectiveness
Menzin et al,[128] 2010, United States	• Cost-benefit analysis • RCT-based model • 2-yr follow-up • Perspective of U.S. society at large	• Patients with 1-level symptomatic DDD and radiculopathy or myelopathy • N = 541 patients: 276 C-ADR, 265 fusion • Sex (% female): 53.6 C-ADR, 54.0 fusion • Mean age (yr): 43.3 C-ACR, 43.9 fusion	• Intervention: C-ADR (Prestige-ST) • Comparator: fusion	• No. of secondary procedures: 11 (3.98%) C-ADR, 26 (9.81%) fusion • Mean No. of working days after surgery: 523 C-ADR, 485 fusion	• 2007 USD • Discount rate: NR • Index surgery: $11,147 C-ADR; $11,036 fusion • Mean net benefit = saving from work days − (cost of index surgery + cost of secondary surgery): $43,979 C-ADR; $37,175 fusion	Net economic benefit: C-ADR was associated with average savings of $6,978 over fusion
Qureshi et al,[129] 2013, United States	• Cost-utility analysis • Decision tree • Lifetime horizon • Perspective of U.S. private insurers	• Patients with 1-level symptomatic DDD and radiculopathy • Age: 45 yr	• Intervention: C-ADR • Comparator: fusion	• QALYs: 3.94 C-ADR, 1.92 fusion	• 2010 USD • Discount rate: 3% • Total cost: $11,987 C-ADR; $16,823 fusion	C-ADR dominated fusion with higher QALYs and lower cost if both prostheses survived for 20 yr
Warren et al,[130] 2013, United States	• Cost-utility analysis • Retrospective review of single institution data from a prospective multicentre trial • 2-yr follow-up • Perspective of U.S. private insurers	• Patients with 1-level symptomatic DDD with radiculopathy • N = 28 patients: 18 C-ADR, 10 fusion • Sex (% female): 50 C-ADR, 60 fusion • Mean age (yr): 41.9 C-ADR, 40.3 fusion	• Intervention: C-ADR (ProDisc-C) • Comparator: fusion	• QALYs, NDI: 0.27 C-ADR; 0.37 fusion • QALYs, SF-36: 0.32 C-ADR; 0.47 fusion	• 2010 USD • Discount rate: 3% • Based on Medicare reimbursement, Mean cost: $13,171 C-ADR; $16,162 fusion • Based on institutional financial modeling • Mean cost: $18,440 C-ADR; $19,811 fusion	• Based on Medicare reimbursement, C-ADR is less costly and less effective than fusion • Based on institutional financial modeling, C-ADR is less costly and less effective than fusion
McAnany et al,[132] 2014, United States	• Cost-utility analysis • Markov state-transition model • 5-yr time horizon • Perspective of U.S. private insurers	• Patients with acute 1-level disc herniation and associated myelopathy or radiculopathy • Age: 40 yr	• Intervention: C-ADR • Comparator: fusion	• QALYs: 2.84 C-ADR, 2.81 fusion	• 2010 USD • Discount rate: 3% • Total cost: $102,274 C-ADR; $119,814 fusion	C-ADR dominated fusion with higher QALYs and lower cost

Author, Year, Country	Study Design and Perspective	Population	Intervention and Comparator	Results		
				Health Outcomes	Costs	Cost-Effectiveness
Radcliff et al,[135] 2016, United States	• Cost-utility analysis • RCT-based model • 7-yr follow-up • Perspective of U.S. private insurers	• Patients with 1-level symptomatic DDD and radiculopathy or myelopathy • N = 209 patients: 103 C-ADR, 106 fusion • Sex (% female): 55.3 C-ADR, 53.8 fusion • Mean age (yr): 42.1 C-ADR, 43.5 fusion	• Intervention: C-ADR (ProDisc-C) • Comparator: fusion	QALYs: 4.52 C-ADR, 4.36 fusion	• 2014 USD • Discount rate: 3% • Mean cost: $29,697 C-ADR; $42,486 fusion • Inclusion of productivity loss • Mean cost: $30,471 C-ADR, $45,596 fusion	C-ADR dominated fusion with higher QALYs and lower cost
Lewis et al,[131] 2014, United States	• Cost-utility analysis • Decision tree • 5-yr time horizon • Perspective of U.S. private insurers	• Patients with 1-level symptomatic DDD with radiculopathy • Meta-analysis of 156 case series from 443 publications • N = 16,992 patients	• Interventions: C-ADR, fusion with autograft, fusion with allograft, fusion with spacer, or ACD (no fusion)	• QALYs: 4.843 C-ADR, 4.714 fusion with autograft, 4.781 fusion with allograft, 4.787 fusion with spacer, 4.885 ACD (no fusion)	• 2014 USD • Discount rate: NR • Total cost: $20,154 C-ADR; $20,511 fusion with autograft; $19,973 fusion with allograft; $19,539 fusion with spacer; $16,558 ACD (no fusion)	• ICERs not reported • ACD (no fusion) dominated all others • Fusion with autograft is dominated by the others • Fusion with spacer dominated fusion with allograft • Our computed ICER ($/QALY): 10,982 (C-ADR vs. fusion with spacer)
Radcliff et al,[133] 2015, United States	• Cost-effectiveness analysis • Retrospective review of prospectively collected administrative data • 4-yr (maximum) follow-up • Perspective of U.S. private insurers	• Patients with 1-level symptomatic DDD N = 6,962 patients: 327 C-ADR; 6,635 fusion • Sex: NR • Mean age (yr): 43.97 C-ADR, 46.57 fusion	• Intervention: C-ADR • Comparator: fusion	• Subsequent surgery incidence (within 36 mo): 5.7% C-ADR, 10.5% fusion • Post-operative complication rates (treated non-operatively): lower for C-ADR	• Reference year: NR • Discount rate: NR • Index surgery: $20,722 C-ADR; $22,379 fusion • Total cost (at final follow up): $34,979 C-ADR; $39,820 fusion	C-ADR is less costly and more effective than fusion

Author, Year, Country	Study Design and Perspective	Population	Intervention and Comparator	Results		
				Health Outcomes	Costs	Cost-Effectiveness
Ghori et al,[134] 2016, United States	• Cost analysis • Markov state-transition model • Lifetime horizon • Perspective of U.S. society at large	• Patients with 1-level symptomatic DDD • Age: 45–65 yr	• Intervention: C-ADR • Comparator: fusion		• 2012 USD • Discount rate: 3% • Perioperative cost (each procedure): $10,859 C-ADR; $12,361 fusion) • Lost productivity cost (each procedure): $4,621 C-ADR; $6,066 fusion • Total cost (for 45-year-old cohort): $24,119 C-ADR; $42,486 fusion	• C-ADR remains less expensive throughout modeled age range of 45–65 yr • Sensitivity analysis demonstrated that C-ADR remains less expensive than fusion if its annual subsequent surgery rate remains below 10.5% annually
Wiedenhofer et al,[136] 2017, Germany	• Cost analysis • Retrospective review of prospectively collected administrative data (January 2003–June 2008) • 4-yr (maximum) follow-up • Perspective of Germany's state health insurance	• Patients with 1- or 2-level symptomatic DDD • N = 467 patients: 199 C-ADR, 268 fusion • Sex (% female): 58 C-ADR, 46 fusion Mean age (yr): 45 C-ADR, 51 fusion	• Intervention: C-ADR • Comparator: fusion		• Discount rate: NR • Hospitalization cost (1-yr follow-up): €6,517 C-ADR; €6,504 fusion Difference in hospitalization cost (fusion − C-ADR): €469 (2-yr follow-up); €3,012 (3-yr follow-up); €3,397 (4-yr follow-up)	C-ADR is less costly than fusion
McAnany et al,[137] 2018, United States	• Cost-utility analysis • Markov state-transition model • 7-yr time horizon • Perspective of U.S. private insurers	• Patients with 1-level symptomatic DDD and radiculopathy or myelopathy • Age: 40 yr	• Intervention: C-ADR (Prestige-LP) • Comparator: fusion	QALYs: 4.53 C-ADR, 3.85 fusion	• 2014 USD • Discount rate: 3% • Total cost: $172,989 C-ADR); $143,714 fusion	ICER ($/QALY): 43,522 (C-ADR vs. fusion)

Abbreviations: ACD, anterior cervical discectomy (without fusion); C-ADR, cervical artificial disc replacement; DDD, degenerative disc disease; ICER, incremental cost-effectiveness ratio; NDI, Neck Disability Index; NR, not reported; QALY, quality-adjusted life-year; RCT, randomized controlled trial; SF-36, 36-item Short-Form Health Survey; USD, U.S. dollars.

Qureshi et al[129] (United States, 2013) conducted a cost-effectiveness analysis to compare C-ADR and fusion for patients aged 45 years who had cervical degenerative disc disease with associated radiculopathy using a decision tree from the perspective of private insurers. This model was based purely on health state utility values taken from the literature, and effectiveness was expressed in terms of QALYs. Cost data were calculated by the authors using data from the Nationwide Inpatient Sample and from Medicare reimbursement records. Uncertainty for costs and QALYs was assessed using sensitivity analyses. In the reference case, over a lifetime horizon and under the assumption that C-ADR prostheses would last for 20 years, C-ADR was associated with higher QALYs gained at a lower cost (3.94 QALYs for C-ADR versus 1.92 QALYs for fusion; $11,987 for C-ADR versus $16,823 for fusion). Sensitivity analysis indicated C-ADR devices needed to last at least 9.75 years to be considered a more cost-effective strategy than fusion devices.

Warren et al[130] (United States, 2013) conducted a cost-utility analysis comparing C-ADR with fusion from the perspective of private insurers. The authors reviewed data from a single institution with 2-year follow-up from a prospective multicentre study of ProDisc-C in which an RCT (Murrey et al[22], 2009) compared the efficacy of C-ADR with fusion for treating symptomatic cervical degenerative disc disease. Both Medicare reimbursement costs and actual hospital costs were reviewed and analyzed to calculate the treatment cost per patient. The QALYs gained were calculated at 1 and 2 years after surgery based on the Neck Disability Index (NDI) and 36-question Short-Form Health Survey (SF-36) outcomes score. The ICER was calculated to determine relative cost-effectiveness. Study results showed that, at 2 years on average and using SF-36, a patient gained 0.47 QALYs with fusion and 0.32 QALYs with C-ADR. The Medicare reimbursement costs associated with fusion were estimated at $16,162 and with C-ADR at $13,171. The total hospital cost associated with fusion was estimated at $19,811 and with C-ADR at $18,440. When comparing fusion with C-ADR, this translated into an ICER of $19,940 per QALY when considering Medicare reimbursement and $9,140 per QALY when considering hospital cost. According to the willingness-to-pay threshold of $50,000 per QALY, fusion would be more cost-effective than C-ADR. The main limitation of this study is the small size (28) of the chosen subpopulation of the original trial, which caused considerable variation in both utilities and costs. However, no sensitivity analysis was conducted to associate a level of certainty with the findings around cost-effectiveness. Qureshi et al,[125] using whole data from the same trial, reported different values for health-related utilities that favoured C-ADR over fusion.

McAnany et al[132] (United States, 2014) developed a Markov state-transition model to explore the 5-year cost-effectiveness of C-ADR versus fusion from the perspective of private insurers. Utilities were derived from responses to the SF-36 at baseline and at 5 years from the treatment arms of the FDA RCT of ProDisc-C (Murrey et al,[22] 2009; Zigler et al,[91] 2013). The study showed that, at 5 years, C-ADR generated a total cost of $102,274 and fusion generated a total cost of $119,814. Artificial disc replacement generated 2.84 QALYs, but fusion generated 2.81 QALYs. Because C-ADR was less costly and more effective than fusion, it is a cost-saving strategy.

Radcliff et al[135] (United States, 2016) evaluated the 7-year cost-effectiveness of C-ADR versus fusion for patients with one-level degenerative disc disease from the perspective of private insurers. Health care resource use and QALYs were prospectively obtained from the randomized, multicentre study and post-approval study of ProDisc-C total disc replacement for IDE (Murrey et al,[22] 2009; Janssen et al,[92] 2015). The results were presented through probabilistic sampling using 10,000 Monte Carlo simulations. In the base-case analysis, C-ADR resulted in cost savings of $12,789 (95% confidence interval [CI]: $5,362–$20,856) per patient and QALY gains of 0.16 (95% CI 0.073–0.39) per patient compared with fusion over 7 years.

C-ADR was more effective and less costly in 90.8% of simulations. In conclusion, C-ADR was found to be more effective and less costly over a 7-year time horizon for patients with one-level symptomatic degenerative disc disease. The results were proven to be robust through scenario analyses: C-ADR dominated fusion with higher QALYs and lower cost.

Lewis et al[131] (United States, 2014) built a decision tree model to determine the relative cost-effectiveness of fusion (with autograft, allograft, or spacers), anterior cervical discectomy without fusion, and C-ADR for the treatment of one-level cervical disc disease with radiculopathy, over the 5-year time horizon from the perspective of private insurers. They performed a literature search and yielded 156 case series describing nearly 17,000 cases. Using pooled meta-analyses, they estimated the incidence of various outcomes including index-level and adjacent-level subsequent surgeries. The ICERs were not reported. However, analysis of cost and QALYs showed that anterior cervical discectomy without fusion dominated the other strategies, and fusion with autograft was dominated by all others. Also, fusion with spacer dominated fusion with allograft. Using the values reported by Lewis et al,[131] we computed the ICER for C-ADR versus fusion with spacer to be $10,982 per QALY.

Radcliff et al[133] (United States, 2015) performed a 4-year cost-effectiveness analysis using a retrospective, matched cohort analysis of a prospectively collected database of costs and outcomes for patients enrolled in a Blue Cross plan, from the perspective of private insurers. This study aimed to determine subsequent surgery rates, adverse event rates, and both direct and follow-up costs of C-ADR compared with fusion in a real-world population of patients with one-level symptomatic cervical degenerative disc disease. There were 6,635 fusion patients and 327 C-ADR patients. By 36 months after surgery, the subsequent surgery rate was significantly greater in the fusion group (10.5%) than in the C-ADR group (5.7%) (hazard ratio, $P = .0214$). The index surgery and 90-day total costs were significantly lower in the C-ADR group. At final follow-up, there was a statistically significant reduction in total costs paid by insurer for C-ADR patients (C-ADR $34,979 vs. fusion $39,820). Researchers concluded that C-ADR was less costly and more effective than fusion.

Ghori et al[134] (United States, 2016) used a Markov model to compare the relative long-term societal costs of C-ADR and fusion by considering upfront surgical costs, lost productivity, and risk of subsequent surgeries. They reported long-term costs to society for a 45-year-old patient undergoing fusion to be $31,178; long-term costs for C-ADR were $24,119. Long-term costs for C-ADR remained lower throughout the modeled age range of 45 to 65 years of age. Sensitivity analysis demonstrated that C-ADR remains less expensive than fusion if the annual subsequent surgery rate remains below 10.5%. Given the subsequent surgery rates of 2.5% for C-ADR reported in their references, the authors concluded that C-ADR was the preferred treatment for cervical radiculopathy from an economic perspective.

Wiedenhofer et al[136] (Germany, 2017) retrospectively reviewed patient-related data from Germany's state health care insurance from January 2003 to June 2008 and performed a costing analysis from the perspective of the state health insurance. A total of 467 cases (199 C-ADR and 268 fusion) with one- or two-level surgery for cervical degenerative disc disease were included. Both groups obtained less pain medication post-operatively than pre-operatively, but between-group difference was not significant. Post-operative absenteeism from work was significantly higher in the C-ADR group; however, patients with C-ADR underwent less rehabilitation covered by the state. Both groups had the same amount of pre-operative and post-operative physiotherapy covered by the state. Researchers concluded that the collected data showed no differences between medical outcomes with C-ADR versus fusion. At the same time, C-ADR incurred significantly lower costs (€3,397 over 4-year follow-up). Therefore, both

medical and financial factors show C-ADR to be feasible for treatment of cervical degenerative disc disease.

McAnany et al[137] (United States, 2018) used a Markov transition model to determine 7-year cost-utility of C-ADR and fusion from the perspective of private insurers. They analyzed data from the SF-36 collected in the Prestige-LP Cervical Disc IDE study (Gornet et al,[59] 2015; Gornet et al,[90] 2016), and used the Short-Form Health Survey classification algorithm (SF-6D) to convert these data into health state utilities. They performed Monte Carlo simulations to assess the probabilistic sensitivity of their model: C-ADR generated a 7-year cost of $172,989, fusion of $143,714. The authors found C-ADR generated 4.53 QALYs; fusion generated 3.85 QALYs. The ICER of C-ADR versus fusion was $43,522 per QALY. Under the willingness-to-pay threshold of $50,000 per QALY, their probabilistic sensitivity analysis demonstrated C-ADR to be cost-effective 56% of the time. Given the Markov simulation and the Monte Carlo simulation, they concluded that C-ADR was the more cost-effective strategy over 7 years.

Two-Level Symptomatic Degenerative Disc Disease

Table 22 summarizes the results of the four included studies for two-level cervical degenerative disc disease. All four studies used models directly informed by reported results of RCTs for specific artificial disc devices.

Ament et al[138] (United States, 2014) conducted a cost-utility analysis to compare C-ADR with fusion in two-level symptomatic cervical degenerative disc disease from the perspectives of both society at large and private insurers. The authors constructed a Markov transition model to evaluate QALYs for both surgery groups. Clinical data were taken from a 2-year follow up IDE trial for the Mobi-C device (Davis et al,[41] 2013). Costs were calculated by extracting diagnosis-related group codes and then applying 2012 Medicare reimbursement rates. Data from SF-12 questionnaires were transformed into utilities values using the SF-6D. Results showed that C-ADR had an average of 1.58 QALYs after 2 years and fusion of 1.50 QALYs. A greater average cost associated with C-ADR of $2,139 translated into an ICER of $24,594 per QALY at 2 years. Sensitivity analysis showed that ICER values stayed below the threshold of $50,000 per QALY in most scenarios. However, when disability-related productivity loss was not accounted for, ICER increased to $100,257 per QALY, indicating that at least part of the benefit of C-ADR was realized outside the health sector.

Ament et al[139] (United States, 2016), in their second study, conducted a similar cost-utility analysis to compare C-ADR with fusion for treatment of two-level symptomatic cervical degenerative disc disease, but used clinical data from a 5-year follow-up of IDE trial for the Mobi-C device (Davis et al,[41] 2013; Radcliff et al,[83] 2016). The authors applied the Markov model developed for the first study (Ament et al,[138] 2014) to compare the two treatments. Costs were calculated by extracting diagnosis-related group codes and then applying 2014 Medicare reimbursement rates. The analysis was conducted from the perspectives of both the health system and society at large and applied a 3% discount rate. Univariate and multivariate sensitivity analyses were conducted to test the robustness of the model. Results showed that C-ADR was associated with 3.574 QALYs and fusion with 3.376 QALYs. From the perspective of the health system, C-ADR cost $1,687 more than fusion over 5 years. From the perspective of society at large, fusion had $34,377 more productivity lost than C-ADR. Therefore, from the societal perspective, C-ADR was less costly than fusion. From the health system perspective, the ICER for C-ADR was $8,518 per QALY. Sensitivity analyses showed that the ICER for C-ADR remained below the willingness-to-pay threshold of $50,000 per QALY in all scenarios.

The ICERs were more favourable for C-ADR in this study using 5-year clinical data than in the previous study using 2-year clinical data.

Overley et al[140] (United States, 2018) developed a Markov transition model to determine the 5-year cost-effectiveness of two-level C-ADR versus fusion from the perspective of private insurers. Data to populate the model were taken from the two-level Prestige-LP cervical disc IDE study (Gornet et al,[42] 2017). Data from the SF-36 were converted into utilities using the SF-6D. Probabilistic sensitivity analyses were performed to test the robustness of the model. In the reference case analysis, a 40-year-old patient with symptoms of degenerative disc disease incurred costs with C-ADR of $130,417 and with fusion of $116,717. The QALYs generated in the C-ADR and fusion arms were 3.45 and 3.23, respectively. The ICER comparing C-ADR with fusion was $62,337 per QALY gained. Probabilistic sensitivity analyses showed that, at the willingness-to-pay threshold of $100,000 per QALY, there would be a 61.5% chance that C-ADR was more cost-effective than fusion. Costs assigned to complications in their work (which were taken to be adverse events) were much higher for C-ADR than for fusion ($8,068.12 for C-ADR versus $3,961 for fusion). Researchers did not report which kind of adverse events they had considered, but the rather high yearly rates they used did not match rates for costly adverse events commonly reported for C-ADR or fusion. This discrepancy seems to have contributed to the high overall 5-year cost and hence the large ICERs reported.

In a subsequent publication, Merrill et al[141] (United States, 2018) developed a Markov transition model to determine the 7-year cost-effectiveness of two-level C-ADR versus fusion from the perspective of private insurers, using data from a Prestige-LP IDE study (Gornet et al,[42] 2017; Lanman et al,[94] 2017). Data from the SF-36 were converted into health utility scores using the SF-6D. In the reference case analysis, the two-level C-ADR had a 7-year cost of $176,654 and generated 4.65 QALYs. The two-level fusion had a 7-year cost of $158,373 and generated 4.44 QALYs. This translated into an ICER of $89,021 per QALY when comparing C-ADR with fusion. Probabilistic sensitivity analyses showed that, in 46% of the runs, C-ADR would be more cost-effective than fusion under the willingness-to-pay threshold of $50,000 per QALY. As in the study by Overley et al,[140] assumptions about complications (adverse events) seemed to affect the overall 7-year costs and the reported ICERs.

Table 22: Economic Literature Review for Two-Level Symptomatic Cervical Degenerative Disc Disease

Author, Year, Country	Study Design and Perspective	Population	Intervention and Comparator	Results		
				Health Outcomes	Costs	Cost-Effectiveness
Ament et al,[138] 2014, United States	• Cost-utility analysis • Markov state-transition model • Perspectives of U.S. society (reference case) and U.S. private insurers (scenario) • 2-yr time horizon	• Patients with symptomatic DDD and radiculopathy or myeloradiculopathy at 2 contiguous levels from C3 to C7 • Age: 44 yr	• Intervention: C-ADR (Mobi-C) • Comparator: fusion	QALYs: 1.59 C-ADR, 1.50 fusion	• 2012 USD • Discount rate: 3% • Reference case: Total cost: $43,060 C-ADR; $40,920 fusion • Perspective of private insurers: Total cost: $29,689 C-ADR; $21,067 fusion	• Reference case, ICER ($/QALY): 24,594 C-ADR vs. fusion • Perspective of private insurers, ICER ($/QALY): 100,257 (C-ADR vs. fusion)
Ament et al,[139] 2016, United States	• Cost-utility analysis • Markov state-transition model • Perspectives of U.S. society (reference case) and U.S. private insurers (scenario) • 5-yr time horizon	• Patients with symptomatic DDD and radiculopathy or myeloradiculopathy at 2 contiguous levels from C3 to C7 • Age: 44 yr	• Intervention: C-ADR (Mobi-C) • Comparator: fusion	QALYs: 3.57 C-ADR, 3.38 fusion	• 2014 USD • Discount rate: 3% • Reference case: Total cost: $80,906 C-ADR; $113,596 fusion • Perspective of private insurers: Total cost: $23,459 C-ADR; $21,772 fusion	• Reference case, C-ADR dominated fusion with higher QALYs and lower cost • Perspective of private insurers, ICER ($/QALY): 8,518 (C-ADR vs. fusion)
Overley et al,[140] 2018, United States	• Cost-utility analysis • Markov state-transition model • Perspective of U.S. private insurers • 5-yr time horizon	• Patients with symptomatic DDD and radiculopathy or myelopathy at 2 contiguous levels from C3 to C7 • Age: 40 yr	• Intervention: C-ADR (Prestige-LP) • Comparator: fusion	QALYs: 3.45 C-ADR, 3.23 fusion	• 2014 USD • Discount rate: 3% • Total cost: $130,417 C-ADR; $116,717 fusion	ICER ($/QALY): 62,337 (C-ADR vs. fusion)
Merrill et al,[141] 2018, United States	• Cost-utility analysis • Markov state-transition model • Perspective of U.S. private insurers • 7-yr time horizon	• Patients with symptomatic DDD and radiculopathy or myelopathy at 2 contiguous levels from C3 to C7 • Age: 40 yr	• Intervention: C-ADR (Prestige-LP) • Comparator: fusion	QALYs: 4.65 C-ADR, 4.44 fusion	• 2014 USD • Discount rate: 3% • Total cost: $176,654 C-ADR; $158,373 fusion	ICER ($/QALY): 89,021 (C-ADR vs. fusion)

Abbreviations: C-ADR, cervical artificial disc replacement; DDD, degenerative disc disease; QALY, quality-adjusted life-year; USD, U.S. dollars.

Applicability of Included Studies

The applicability checklist for economic evaluations was applied to the included articles (Appendix 6). All were deemed partially applicable to the research question. None of the studies were relevant for Ontario. However, we benefited from the patient pathways discussed in these studies in building our model. We also used some of their reported utilities of health states in our analysis.

Discussion

For one-level symptomatic degenerative disc disease, C-ADR appeared to reduce costs more than fusion (while having superior health effects) in all but three studies: a study by Warren et al[130] in which fusion was more costly but also more effective than C-ADR, a study by Lewis et al[131] in which C-ADR had a favourable cost-effectiveness profile compared with fusion with spacer (as dominant fusion option), and a study by McAnany et al[137] in which C-ADR had a favourable ICER compared with fusion. Most researchers used the perspective of private insurers, but some used the perspective of society at large. There was a notable difference in reported costs for different items, including surgery (initial or subsequent), treatment of adverse events and complications, and lost productivity. In most cases, results depended heavily on specific values picked for these cost items, and also depended on complication rates and the rates for subsequent surgical interventions. Some studies used rates from follow-up of RCTs, and others extracted these rates from retrospective searches of administrative databases. Some studies combined data from different publications. Results from studies using RCT data are specific to the device under consideration and protocols followed during the follow-up period. Further, costs for items used are not comparable or transferable to other settings. For studies using administrative databases, it is hard to judge whether the C-ADR and fusion populations were comparable. The issue of cost comparability and transferability was also a limitation.

For two-level symptomatic degeneration disc disease, one study[139] showed that, depending on the perspective, C-ADR had either a favourable ICER in relation to fusion or a greater cost saving and superior health effect over a 5-year time horizon. Another study[138] showed that, depending on the perspective, C-ADR was more cost-effective than fusion for willingness-to-pay thresholds of $50,000 (societal) or $100,000 (private insurer) over a 2-year time horizon. The two other studies[140,141] showed that C-ADR was costlier but more effective than fusion over 5- and 7-year time horizons. It was impossible to determine the cost-effectiveness profile because results depended on the willingness-to-pay threshold. At the willingness-to-pay threshold of $100,000 per QALY, C-ADR would be more cost-effective than fusion. All two-level studies used rates from follow-up of RCTs.

For both one- and two-level cases, studies used credible values for utilities after index surgery, usually coming from SF-36 or SF-12 questionnaires. However, the utilities used for adverse events, complications, and for subsequent surgical interventions seemed arbitrary and lacked firm measurements. These questionable choices for utilities were not subject to clear sensitivity or scenario analysis, which was another limitation.

A similar issue is raised when we look at the costs used by these studies, where the cost of index surgery has been chosen to resemble real-world values but the costs for complications and subsequent surgeries were not based on a clear argument. In some studies, the chosen values were not even reported.

Conclusions

All studies we reviewed were conducted in the United States (13) or Germany (1). Results from these studies are not easily transferable to Ontario mainly because costs for items in Ontario are very different from costs in the jurisdictions reported in these studies. Further, none of the studies reported enough results, in the form of sensitivity analysis or scenario exploration, to transfer to the Ontario setting.

PRIMARY ECONOMIC EVALUATION

Published economic evaluations identified in the literature review addressed cervical artificial disc replacement (C-ADR) versus fusion, but none of these published studies took a Canadian perspective. Further, studies did not use nor capture all available clinical data, and their scenario and sensitivity analyses did not reflect situations or parameters of special interest in Ontario. Owing to these limitations, we conducted a primary economic evaluation.

Research Questions

Within the context of the Ontario Ministry of Health and Long-Term Care, we asked two questions:

1. What is the cost-effectiveness of C-ADR versus fusion for the treatment of adults with one-level symptomatic cervical degenerative disc disease in Ontario?

2. What is the cost-effectiveness of C-ADR versus fusion for the treatment of adults with two-level symptomatic cervical degenerative disc disease in Ontario?

Methods

Information presented in this report follows the reporting standards set out by the Consolidated Health Economic Evaluation Reporting Standards statement.[142]

Type of Analysis

For both research questions we conducted a cost-utility analysis to measure the costs and quality-adjusted life-years (QALYs) of adopting C-ADR versus fusion. We conducted a reference case analysis and sensitivity analyses. Our reference case analysis adhered to the Canadian Agency for Drugs and Technologies in Health (CADTH) guidelines when appropriate and represents the analysis with the most likely set of input parameters and model assumptions. Our sensitivity analyses explored how the results are affected by varying input parameters and model assumptions. We conducted both probabilistic and deterministic sensitivity analysis. We identified further scenarios after consultation with experts and the Ministry of Health and Long-Term Care and verified the robustness of our conclusions under alternate assumptions.

Target Population

The study population is adults presenting with symptomatic one-level or two-level cervical degenerative disc disease who are eligible for both C-ADR and fusion and have been unresponsive to conservative treatment.

The age range reported in clinical studies and previous economic evaluations is 18 years for the lower limit and 65, 69, or 72 years for the upper limit. The average age reported in these studies is around 44 years. After consulting with experts, we selected a cohort of 44-year-old patients for our reference case analysis. We also used an approximate distribution for the ages (18 to 72 years with a mode of 44 years) and simulated an age-distributed cohort in our probabilistic analysis.

Perspective

We conducted this analysis from the perspective of the Ontario Ministry of Health and Long-Term Care.

Interventions

We evaluated C-ADR compared with fusion (Table 23). Each intervention included presurgical preparations, inpatient surgery, treatment of postsurgical complications, and diagnosis and treatment of long-term complications. The three alternative methods for fusion are autograft (bone from another part of a person's body), allograft (bone from another person or cadaver), and a synthetic cage with spacer (placed between two vertebrae to maintain space, preserve spinal alignment, and promote fusion). In this study, we do not distinguish between these alternatives. However, our sensitivity analysis aimed to capture the resulting differences in cost, QALY, and incremental cost-effectiveness ratio (ICER) for important parameters (i.e., costs, utilities, and complications) that could vary between these alternatives. Similarly, C-ADR devices differ in cost and reported effectiveness. In this analysis we focused on C-ADR devices licensed by Health Canada, and we performed sensitivity and scenario analyses by varying parameters across C-ADR devices. We also considered C-ADR devices whose manufacturers have applied for a licence in Canada but not yet received approval. Thus, we considered the following C-ADR devices in our economic analysis: M6-C (one-level), ProDisc-C (one-level), Prestige-LP (one- or two-level), and Mobi-C (one- or two-level). (See Tables A6 and A7 of Appendix 7 and the Clinical Evidence section for more information).

Table 23: Disease Interventions and Comparators Evaluated in the Primary Economic Model

Intervention	Comparator	Patient Population	Outcomes
1-level C-ADR	1-level fusion	Adults with symptomatic 1-level disease	Cost, QALYs, ICER
2-level C-ADR	2-level fusion	Adults with symptomatic 2-level disease	Cost, QALYs, ICER

Abbreviations: C-ADR, cervical anterior discectomy and artificial disc replacement; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year.

Discounting and Time Horizon

We applied an annual discount rate of 1.5% to both costs and QALYs.[143] We also tested rates of 0%, 3%, and 5% in our sensitivity analyses. We used a 7-year time horizon in all analyses. This time horizon reflected the most recent available long-term follow-up data from clinical studies. In scenario analyses we used a 20-year time horizon, making assumptions for the times (beyond 7 years) for which we have no comparative data for C-ADR versus fusion.

Main Assumptions

The major assumptions for this model are:

- The index level always refers to the spinal level at which the index (first) surgery is performed.
- Surgery performed at two adjacent levels is called a two-level surgery. However, we do not consider two adjacent levels operated on at different times as two-level surgery. We also do not include the possibility that two nonadjacent levels would be operated on at

once. In two-level surgery, the two levels are collectively called the index level in our diagrams and models.
- We used similar patient pathways for one- and two-level cervical degenerative disc disease but with different rates of transition between states and with different costs and health outcomes.
- If C-ADR is chosen, then all subsequent surgeries should give priority to C-ADR over fusion. This means that C-ADR is chosen whenever possible. However, in many cases, the subsequent complication (at the index or adjacent level) is ineligible for C-ADR, and fusion is the only option. Hence, in the model, we varied the proportion eligible for C-ADR (10–51% for adjacent level, based on expert consultation and published literature; see the Budget Impact Analysis for further details).
- If fusion is chosen for the index surgery, then all subsequent surgeries should give priority to fusion over C-ADR. This, in practice, means that no C-ADR procedures will be performed as a subsequent surgery for patients who have had fusion.
- The time elapsed after index surgery, which will affect the probability and rates of complications, is globally measured from the start of simulation.
- We do not distinguish between brands for fusion.
- Complications for index-level and adjacent-level surgeries will not appear at the same time. This should be valid for most situations.
- Disease-specific mortalities are negligible for both types of surgery. We consider only age-specific general mortality.

Model Structure and Structure of Analysis

We developed a Markov model (Figure 4) to determine the incremental cost per QALY gained. The model is based on the patient pathways that we identified from clinical studies and we confirmed by clinical experts. In brief, a patient with symptomatic degenerative disc disease who is eligible for both C-ADR and fusion will be simulated to go through each surgery to determine the costs and QALYs for the treatment associated with that intervention.

The Markov model consists of four health states based on health-related quality of life and on the history of treatments received. There are costs associated with transitions to and costs for being in states. Cycle length was 1 month, meaning patients transitioned to a different health state only once a month. The states have same names and descriptions for one-level and two-level models, and for C-ADR and fusion alternatives on the left and right arm of the diagram, respectively. The four health states are described below.

- **Index-Level, Initial**—This is the state a patient enters after their initial surgery (i.e., their index surgery). This state is expected to last for a long time (ideally lifetime). However, some patients in this state will have complications either at or adjacent to the index level. If a complication is subject to surgery, the patient will follow one of two paths depending on whether the subsequent surgery will be at or adjacent to the index level.
- **Index-Level, Corrective**—This is the state a patient enters after undergoing a subsequent surgical intervention at the index level. These surgeries are classified as “revision,” “removal,” “supplemental fixation,” or “reoperation.” We use the term “corrective surgery” for any of these procedures.
- **Adjacent-Segment**—This is the state a patient enters after surgery at a segment adjacent to the index level for the treatment of adjacent-segment disease. This state includes cases in which the index level is involved; for example, when a plate is replaced with a longer one to extend fusion to an adjacent segment.

- **Dead—**At any point during the model time horizon, patients have a probability of death from background mortality.

In each state there is a chance that adverse events might occur. These (usually) temporary complications are modeled as events within a cycle with associated disutilities and costs applied.

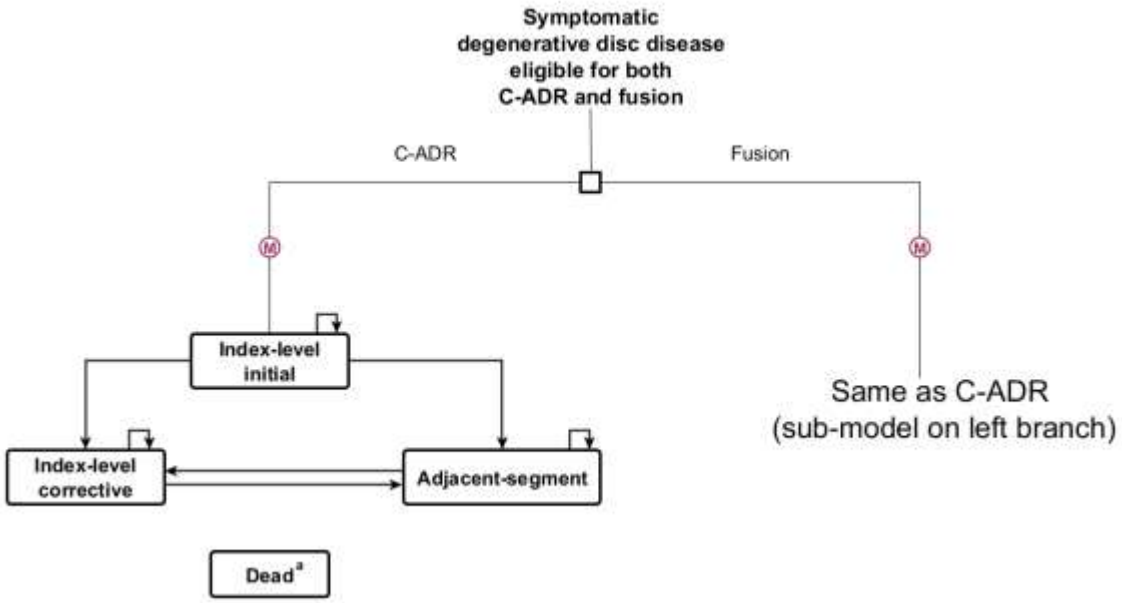

Figure 4: Decision Tree and Markov Models for One-Level and Two-Level C-ADR Versus Fusion

Abbreviation: C-ADR, cervical anterior discectomy and artificial disc replacement.
[a]There are transitions from every state of the Markov model to the "Dead" state, which represents general mortality. Here, we consider transitions only for condition-specific mortality (if any) to the "Dead" state.

Clinical Outcome and Utility Parameters

We used several input parameters to populate the model. These included:

- Variables used to model the natural history of the disease
- Variables used to modify the natural history model to account for treatment effects of C-ADR

Natural History

The main natural history parameters of this model are the rate of adverse events, complications at index level, and development of degenerative disc disease at adjacent levels. In our analysis, fusion is the current practice and our comparator. Therefore, the short-term and long-term rates of adverse events and complications are important and are estimated from the clinical published literature.

Intervention Impact on Natural History

Results from clinical studies suggest that C-ADR results in fewer adverse events and in fewer subsequent surgeries at index and adjacent levels. We report values for 2-year and 7-year

follow-up of RCTs for devices considered in our analysis. We use 2-year incidence during the first 2 years and the adjusted difference of 7-year and 2-year incidence after 2 years. This means that, if the simulation time extends beyond 7 years (in our scenarios), we assume the same complication rates for after 7 years.

- **Adverse Events**—Different definitions for adverse events result in a variety of values reported for the incidence. Some studies report incidence of specific events, such as dysphagia, dysphonia, or infection. However, in most cases these events are also grouped, for example, as serious, device related, or surgery related (Tables 24 and 25). We picked the latter definition because 2-year and 7-year incidences were similar across the different devices and covered most important adverse events. We assume adverse events have constant risk from time 0 to 2 years, and again from 2 to 7 years, and are associated with temporary disutilities.
- **Index-Level Corrective Surgeries**—Collective rates for these surgeries are reported in Tables 26 and 27 for 2-year and 7-year follow-up visits. Relative distributions for each class (revision, removal, supplemental fixation, and reoperation), along with their definitions, are reported in Appendix 8 and Table A11.
- **Adjacent-Segment Surgeries**—Extracted rates for the adjacent-segment surgeries are reported in Tables 26 and 27 for 2-year and 7-year follow-up.

Table 24: Probability (Cumulative Incidence) of Adverse Events in 2-Year Follow-Up

C-ADR Device Author, Year	Levels	Adverse Events (%)			Definition According to Publication
		Disc[a]	Fusion[a]	Δ (Disc – Fusion)[a]	
Prestige-LP Gornet et al,[59] 2015	1-level	**12.1**	**15.5**	**−3.4**	Device- and device- or surgical procedure–related[b]
		5.0	4.9	0.1	Serious device- and device- or surgical procedure–related
		9.2	8.3	0.9	Dysphagia or dysphonia
ProDisc-C Murrey et al,[22] 2009	1-level	2.9	6.6	−3.7	Adverse events resulting in overall study failure
		1.9	**6.6**	**−4.7**	Device related[b]
		0.0	0.9	−0.9	Dysphagia
Mobi-C Hisey et al,[38] 2014	1-level	3.9	7.4	−3.5	Related to study device
		10.6	18.5	−7.9	Dysphagia
		1.7	3.7	−2.0	Dysphonia
Prestige-LP Gornet et al,[42] 2017	2-level	**15.8**	**20.7**	**−4.9**	Related to implant and surgical procedure[b]
		1.9	5.8	−3.9	Grade 3 or 4 adverse event related to implant or to implant and surgical procedure
		6.7	11.2	−4.5	Dysphagia or dysphonia
Mobi-C Davis et al,[41] 2013	2-level	**3.6**	**6.7**	**−3.1**	Determined by CEC to be major complication of treatment (or adverse events resulting in overall study failure)[b]
		16.7	34.3	−17.6	Related to study device
		3.4	14.3	−10.9	Serious, and definitely or possibly related to device
		3.8	7.6	−3.8	Dysphagia
		0.4	1.0	−0.6	Dysphonia

Abbreviations: Δ, difference; C-ADR, cervical artificial disc replacement; CEC, Clinical Events Committee.
[a]Values in bold are those chosen for our analysis.
[b]Definition chosen for our analysis.

Table 25: Probability (Cumulative Incidence) of Adverse Events in 7-Year Follow-Up

C-ADR Device Author, Year	Levels	Adverse Events (%)			Definition According to Publication
		Disc[a]	Fusion[a]	Δ (Disc − Fusion)[a]	
Prestige-LP Gornet et al,[90] 2016	1-level	**17.5**[a]	**16.6**	**0.9**	Device- or device- and surgical procedure–related[b]
		6.1	5.6	0.5	Serious device- or device- and surgical procedure–related
		11.1	9.8	1.3	Dysphagia or dysphonia
ProDisc-C Janssen et al,[92] 2015	1-level	**27.2**	**28.3**	**−1.1**	Device-related[b]
		0.0	1.9	−1.9	Dysphagia
Mobi-C Radcliff et al,[84] 2017	1-level	6.1	3.7	2.4	Determined by CEC to be a major complication of treatment (or adverse events resulting in overall study failure)
Prestige-LP Lanman et al,[94] 2017	2-level	**26.6**	**27.7**	**−1.1**	Possibly device-related[b]
		3.2	7.2	−4.0	Serious, possibly device-related (or serious, Grade 3 or 4 implant or implant- and surgical procedure–related)
		1.3	0.0	1.3	Dysphagia or dysphonia
Mobi-C Radcliff et al,[84] 2017	2-level	**5.3**	**8.6**	**−3.3**	Determined by CEC to be major complication of treatment (or adverse events resulting in overall study failure)[b]

Abbreviations: Δ, difference; C-ADR, cervical artificial disc replacement; CEC, Clinical Events Committee.
[a]Values in bold are those chosen for our analysis.
[b]Definition chosen for our analysis.

Table 26: Probabilities (Cumulative Incidence) of Index-Level Corrective Surgeries and Adjacent-Segment Surgeries in 2-Year Follow-Up

C-ADR Device Author, Year	Levels	Index-Level Corrections (%)			Adjacent-Segment Surgery (%)		
		Disc	Fusion	Δ (Disc − Fusion)	Disc	Fusion	Δ (Disc − Fusion)
Prestige-LP Gornet et al,[59] 2015	1-level	5.0	7.9	−2.9	2.5	4.2	−1.7
ProDisc-C Murrey et al,[22] 2009	1-level	1.9	8.5	−6.6	0	0.9	−0.9
Mobi-C Hisey et al,[38] 2014, FDA SSED	1-level	1.2	6.2	−5.0	0.6	4.9	−4.3
Prestige-LP Gornet et al,[42] 2017	2-level	2.4	8.0	−5.6	2.4	3.2	−0.8
Mobi-C Davis et al,[41] 2013, FDA SSED	2-level	3.1	11.4	−8.3	0.9	4.8	−3.9

Abbreviations: Δ, difference; C-ADR, cervical artificial disc replacement; FDA, Food and Drug Administration; SSED, Summary of Safety and Effectiveness Data.

Table 27: Probabilities (Cumulative Incidence) of Index-Level Corrective Surgeries and Adjacent-Segment Surgeries in 7-Year Follow-Up

C-ADR Device Author, Year	Levels	Index-Level Corrections (%)			Adjacent-Segment Surgery (%)		
		Disc	Fusion	Δ (Disc − Fusion)	Disc	Fusion	Δ (Disc − Fusion)
Prestige-LP Gornet et al,[90] 2016	1-level	6.4	10.9	−4.5	9.6	8.3	1.3
ProDisc-C Janssen et al,[92] 2015	1-level	6.0	15.0	−9.0	5.8	12.3	−6.5
Mobi-C Radcliff et al,[84] 2017	1-level	3.0	6.2	−3.2	3.7	13.6	−9.9
Prestige-LP Lanman et al,[94] 2017	2-level	4.2	14.7	−10.5	6.5	12.5	−6.0
Mobi-C Radcliff et al,[84] 2017	2-level	4.4	10.5	−6.1	4.4	11.4	−7.0

Abbreviation: Δ, difference; C-ADR, cervical artificial disc replacement.

Intervention Utilities

We reviewed the literature to retrieve data for the following health state utilities:

- Pre-operation utilities
- Post-operation utilities (time-dependent) for the intervention and the comparator
- Utilities and disutilities associated with adverse events, index-level complications and surgery, and adjacent-segment disease and surgery

Our review of previous economic evaluations indicated that, in most cases, pre-operation and post-operation utilities are derived from SF-36 and SF-12 questionnaires collected during

clinical trials and their follow-ups (≤ 7 years) (Table 28). Although not reported here, some measurements of utilities use a Neck Disability Index (NDI) or experts' opinions. For one-level surgery, we have utilities for Prestige-LP (7 years) and ProDisc-C (7 years). For two-level surgery, we have utilities for Prestige-LP (7 years) and Mobi-C (5 years). We use linear interpolation of utilities for time points falling within the reported period but use the last reported utility (at 5 or 7 years) for later times (beyond 5 or 7 years), if needed.

Table 28: Utilities (Health-Related Quality of Life) for Artificial Disc and Fusion Surgeries at 5- or 7-Year Follow-Up

C-ADR Device Author, Year, Method	Level, Years of Follow-Up	Time After Surgery	C-ADR Mean	C-ADR Standard Error	Fusion Mean	Fusion Standard Error
Prestige-LP McAnany et al,[137] 2018, SF-36 to SF-6D	1-level, 7 yr	Pre-operative	0.55	NR	0.54	NR
		mo 12	0.73		0.68	
		24	0.72		0.69	
		36	0.73		0.69	
		60	0.72		0.70	
		84	0.72		0.69	
ProDisc-C Radcliff et al,[135] 2016; Qureshi et al,[125] 2014, SF-36 to SF-6D	1-level, 7 yr	Pre-operative	0.54	NR	0.54	NR
		wk 6	0.54		0.54	
		mo 3	0.65		0.62	
		6	0.71		0.68	
		12	0.71		0.69	
		24	0.72		0.69	
		36	0.71		0.68	
		48	0.73		0.71	
		60	0.72		0.70	
		72	0.73		0.69	
		84	0.70		0.68	
Prestige-LP Merrill et al,[141] 2018, SF-36 to SF-6D	2-level, 7 yr	Pre-operative	0.55	NR	0.53	NR
		mo 12	0.72		0.69	
		24	0.73		0.71	
		36	0.74		0.72	
		60	0.74		0.70	
		84	0.73		0.70	
Mobi-C Ament et al,[126] 2015, SF-12 to SF-6D	2-level, 5 yr	Pre-operative	0.556	0.113	0.545	0.111
		mo 6	0.767	0.156	0.720	0.171
		12	0.766	0.164	0.722	0.176
		18	0.765	0.160	0.718	0.183
		24	0.779	0.153	0.719	0.166
		36	0.781	0.150	0.716	0.184
		48	0.777	0.156	0.722	0.177
		60	0.776	0.145	0.711	0.173

Abbreviations: C-ADR, cervical artificial disc replacement; NR, not reported; SF-6D, six-dimension classification of results from SF-36 and SF-12; SF-12, 12-item Short-Form Health Survey; SF-36, 36-item Short-Form Health Survey.

For adverse events in each state, a disutility is applied to each occurrence, proportional to the duration of the adverse event (on average 3 months according to experts). Table 29 lists average utilities for adverse events after cervical spine surgery. These values are relative to the

perfect health value of 1. For our disutility calculations, we applied these values as multiplicative factors to reduce the health state utilities of our model. We used the average of three alternative methods for fusion in our reference analysis. In our sensitivity analysis we explore different values for these parameters.

Table 29: Multiplicative Factor to Calculate Disutilities for Patients Experiencing Adverse Events

	Fusion							
	Autograft		Allograft		Spacer		C-ADR	
Source	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Lewis et al,[131] 2014	0.870	0.023	0.827	0.044	0.838	0.049	0.842	0.043

Abbreviations: C-ADR, cervical artificial disc replacement; SD, standard deviation.

For the period after occurrence/diagnosis and while waiting for surgery, we apply a disutility proportional to the duration of wait (on average 3 months according to experts). These patients are assumed to have a utility less than the average expected post-operative value, but usually higher than pre-operative values. In our analysis we chose the mid-point, but let it vary between the two limits in our sensitivity analysis.

Similarly, for corrective surgeries at index level and surgeries at adjacent segments, we used an intermediate utility value (70% of the normal gain in the utility from pre-operation to post-operation, subject to sensitivity analysis).

Cost Parameters

We searched the clinical literature for operations and treatments with costs that would potentially be paid directly by the Ontario health system. We cross-referenced results with costs reported in the economic literature and consulted experts to finalize costs specific to Ontario.

We divided the costs into three categories:

- Costs associated with the index surgery including hospital costs, surgeon/anesthetist/ assistant fees, implant (hardware) cost, and follow-up visits and required imaging costs
- Costs resulting from treatment of adverse events
- Costs associated with long-term complications (including corrective surgeries at the index level and surgeries at adjacent segments)

C-ADR Device Costs

We contacted the manufacturers and distributors of the M6-C, ProDisc-C, Mobi-C, and Prestige-LP C-ADR devices to obtain cost information (Table 30). We also inquired about prices for fusion instruments (Table 31). Compared with costs from U.S. studies (Tables A9 and A10), Canadian prices are lower for both disc prostheses and fusion instruments. However, prices for artificial discs are consistently higher than prices for instruments used in fusion for both one-level and two-level surgeries.

Table 30: Unit Prices and Sales for Artificial Discs

	Sales in Ontario (n)		
Unit Price, $ (2018)	2015	2016	2017
2,450–5,500[a]	35	38	25

[a]Prices provided by manufacturers.

Table 31: Price Details for Instruments Used in Fusion Surgery in Ontario

2018 Cost per Fusion Surgery	
One-Level, $	Two-Level, $
1,350–2,500	2,100–3,924

Surgeon, Anesthetist, and Surgical Assistant Costs

We used the Ontarios Schedule of Benefits[144] to confirm these fees after confirming the procedure codes with experts. Additional codes are used for multi-level surgeries. For subsequent surgeries, some parts of surgeon fees are increased by 30%. For corrective surgeries, fees vary depending on the type of surgery (revision, removal, supplemental fixation, or reoperation). These fees are summarized in Table 32 (see Tables A11–A14 for more details).

Table 32: 2018 Surgeon, Anesthetist, and Surgical Assistant Fees for C-ADR Versus Fusion

C-ADR		Fusion	
One-Level, $	Two-Level, $	One-Level, $	Two-Level, $
Index surgery			
2,427.61	3,192.61	1,941.46	2,400.46
Revision surgery			
3,039.61	3,804.61	2,216.86	2,675.86
Removal surgery (removal followed by fusion)[a]			
NA	NA	2,216.86	2,675.86
Corrective surgery: supplemental fixation[b]			
NA	NA	1,599.47	1,752.47
Corrective surgery: reoperation[c]			
NA	NA	1,594.59	1,849.59

Abbreviations: C-ADR, cervical artificial disc replacement; NA, not applicable.
[a]Secondary surgery is almost always fusion, irrespective of type of previous surgery.
[b]According to the U.S. Food and Drug Association's Summary of Safety and Effectiveness Data, supplemental fixation is posterior arthrodesis in most cases, irrespective of type of previous surgery.
[c]According to the U.S. Food and Drug Association's Summary of Safety and Effectiveness Data, reoperation is a posterior decompression, in most cases, irrespective of type of previous surgery.
Source: Data provided by Ontario Health Insurance Plan.

Hospital Cost

We used the Ontario Case Costing (OCC) database for hospital costs.[145] For index surgery, the 6-year average (2011–2016) cost for C-ADR was $11,255 compared with $9,030 for fusion (both including implants). The average hospital cost for supplemental fixation was $14,379 (including required hardware) and $9,577 for reoperation surgery (Tables A15 and A16).

We estimated the costs of C-ADR devices using the approximated distribution for one-level and multi-level surgeries derived from OHIP billing codes and Alberta Health Evidence Reviews[146] (2009), and using sales data for C-ADR devices from manufacturers for the years 2015 to 2017. These adjusted hospital costs (excluding hardware) are listed in Table 33 (see also Tables A17–A20).

Table 33: Approximation of Hospital-Only Costs for C-ADR and Fusion in Ontario

Entity	C-ADR	Fusion
Upper Bound, $	7,392	7,105
Lower Bound, $[a]	6,652	6,214

Abbreviation: C-ADR, cervical artificial disc replacement.
[a]Lower-bound values are based on data from OHIP billing codes and Alberta Health Evidence Reviews.[146]

Specialist Visit Cost

Both surgeries required one presurgical visit. Fusion patients usually had three postsurgical visits, and C-ADR patients typically had two postsurgical visits within the first year after surgery. The presurgical cost for both surgeries was $102.10. The postsurgical costs were $151.20 for the fusion group and $100.80 for the C-ADR group (Tables A21 and A22).

Diagnostic Imaging Cost

Both surgeries used one set of presurgical images (x-ray and magnetic resonance imaging [MRI]). The fusion group usually has three postsurgical x-ray films and the C-ADR group has two in the first year after surgery. The presurgical cost for both groups was $963. The postsurgical costs were $540.20 for the fusion group and $360 for the C-ADR group (Tables A23–A25).

Adverse Events Cost

We searched available economic evaluations and converted values to 2018 Canadian dollars. We estimated $129 for an episode of dysphagia and $1,480 for surgical treatment of serious infection. We used these values for dysphagia and infection as lower and upper bounds, respectively, for the cost of an adverse event, subject to sensitivity analysis (Table A29).

Analysis

We calculated the reference case of this analysis by average values for transitions, utilities, and costs. We set distributions for some of these variables within the model. We ran the model for a cohort of patients and computed appropriate statistics.

For one-level analysis we used parameters associated with Prestige-LP and ProDisc-C, and for two-level analysis we used Prestige-LP and Mobi-C. Because of differences in settings of the corresponding randomized controlled trials (RCTs) and methodologic variations in collecting and reporting data, it seemed inappropriate to average the parameters for different devices. We therefore report results for one device with more conservative estimates as reference and results for the other as an alternative.

For sensitivity and scenario analyses, we assessed variability and uncertainty in the model through one-way and probabilistic sensitivity analyses. For one-way sensitivity analyses, we varied specific model variables (device cost, hospital cost, complication rates, disutilities) within plausible ranges and examined the effect on the results. Results of the one-way sensitivity analyses are presented in a tornado diagram. To determine the effect of simultaneously varying numerous variables using the assigned distributions, we conducted a probabilistic sensitivity analysis by running simulations of the model. Results of the probabilistic sensitivity analysis are presented on a cost-effectiveness plane as well as a cost-effectiveness acceptability curve. Variables and ranges are presented in Table 34. The base cases and ranges for these parameters were discussed in previous sections, except for the last parameter (which determines the proportion of adjacent-segment disease eligible for C-ADR). We have determined a minimum of 10% based on expert consultation and a maximum of 51% based on a study by Quirno et al.[147] We discuss the details of derivation of the upper limit for this parameter in our Budget Impact section.

Table 34: Parameters Varied in One-Way and Probabilistic Sensitivity Analysis

Variable	Range	Distribution[a]	Reference
Cost of disc prosthesis	$2,450–$5,500	PERT	Manufacturers/Distributors
Cost of fusion instruments	$1,350–$2,500	PERT	Manufacturers/Distributors
Hospital cost (various)	Base value ± 20%	PERT	OCC[145]
Cost of adverse events	$129–$1,480	PERT	Table A29
Probability of adverse events	Base value ± 50%	Beta	Tables 24, 25
Probability of index-level corrective surgery	Base value ± 50%	Beta	Tables 26, 27
Probability of adjacent-segment surgery	Base value ± 50%	Beta	Tables 26, 27
Postsurgical utility (PSA only)	Base value ± 20%	Beta	Table 28
Disutility factor of adverse events	0 to base value	Beta	Table 29
Disutility factor of index-level complications	0–1	Beta	Expert consultation
Disutility factor of ASD	0–1	Beta	Expert consultation
Disutility factor of index-level corrective surgery	0–1	Beta	Expert consultation
Disutility factor of adjacent-segment surgery	0–1	Beta	Expert consultation
Cohort age	18–72 yr	PERT	RCTs, Expert consultation
Proportion of ASD patients eligible for C-ADR	0.1–0.51	Beta	Expert consultation, Quirno et al[147]

Abbreviations: ASD, adjacent-segment disease; C-ADR, cervical artificial disc replacement; OCC; Ontario Case Costing; PSA; probabilistic sensitivity analysis; RCT, randomized controlled trial.

[a]PERT (also known as Beta-PERT) distribution allows to parametrize a generalized beta distribution based on expert opinion regarding a pessimistic estimate (minimum value), a most likely estimate (mode), and an optimistic estimate (maximum value).

Generalizability

The findings of this economic analysis cannot be generalized to all patients with degenerative disc disease and to other artificial disc implants. They may, however, be used to guide decision-making about the specific patient populations and specific disc prostheses addressed in the trials investigated by Health Quality Ontario.

Expert Consultation

From January to March 2018, we solicited expert advice on the use of C-ADR and fusion. Consultation included physicians in the specialty areas of orthopedic spine surgery. The role of the expert advisors was to provide important context on the use of C-ADR and fusion, including expertise on health conditions, patients, diffusion of the technology, or clinical issues in Ontario. However, the statements, conclusions, and views expressed in this report do not necessarily represent the views of the consulted experts.

Results

Reference Case Analysis

The reference case results for our analysis are presented in Table 35 for one-level and two-level surgeries. In both cases, C-ADR provided greater health gains for an incremental cost than fusion.

Table 35: Reference Case Analysis for One-Level and Two-Level Surgeries[a]

Strategy	Average Total Costs, $	Incremental Cost,[b] $	Average Total Effects, QALYs	Incremental Effect,[c] QALYs	ICER, $/QALY
One-level surgery					
Fusion	14,483.52		4.4448		
C-ADR	16,768.07	2,284.56	4.6416	0.1968	11,607.05
Two-level surgery					
Fusion	17,332.60		4.4996		
C-ADR	20,923.26	3,590.66	4.7135	0.2139	16,782.44

Abbreviations: C-ADR, cervical artificial disc replacement; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year.
[a]Results might appear incorrect because of rounding.
[b]Incremental cost = average cost of C-ADR – average cost of fusion.
[c]Incremental effect = average effect of C-ADR – average effect of fusion.

Sensitivity Analysis

One-Way Sensitivity Analysis

Figures 5 and 6 show the tornado diagrams for one-level and two-level cases. The most influential parameter in both cases is the price of the C-ADR device. The ICER varied between $3,522/QALY and $19,692/QALY for one-level surgery, and between $2,310/QALY and $31,254/QALY for two-level surgery, when the price of the device varied from $2,450 to $5,500. Other influential parameters are rate of index complications, rate of adjacent-segment disease, and cost of fusion instruments. The ICERs in all cases show robust acceptable changes and always fall well below the willingness-to-pay threshold.

In both one-level and two-level cases, measurements for utilities at post-operative times were done independently and show consistently higher values for C-ADR than for fusion (see Table 28). To measure the sensitivity of the cost-effectiveness results to systematic variations in these time-dependent utilities, we calculated the maximum amount of variation that can be applied (simultaneously) to all post-operative values without the ICER exceeding a willingness-to-pay of $50,000/QALY. Our simulations show that, in one-level surgery, utilities for C-ADR could be 3.5% lower, or the utilities for fusion could be 3.8% higher, and the ICER will remain below a willingness-to-pay of $50,000/QALY. Similarly, for two-level surgery, the utilities for C-ADR could be 3.1% lower, or the utilities for fusion could be 3.6% higher, and the ICER will remain below a willingness-to-pay of $50,000/QALY.

Variations beyond these bounds would cause values to draw closer and eventually invert, which would result in large ICERs or would cause C-ADR to be dominated by fusion surgery. (A dominated intervention is one found to be less costly and more effective than another intervention.) An alternative view for modeling deviations of utilities around the reported mean values is to pick random variations at each post-operative time point independently. We adopted this approach for our probabilistic sensitivity analysis, with results reported in the next section.

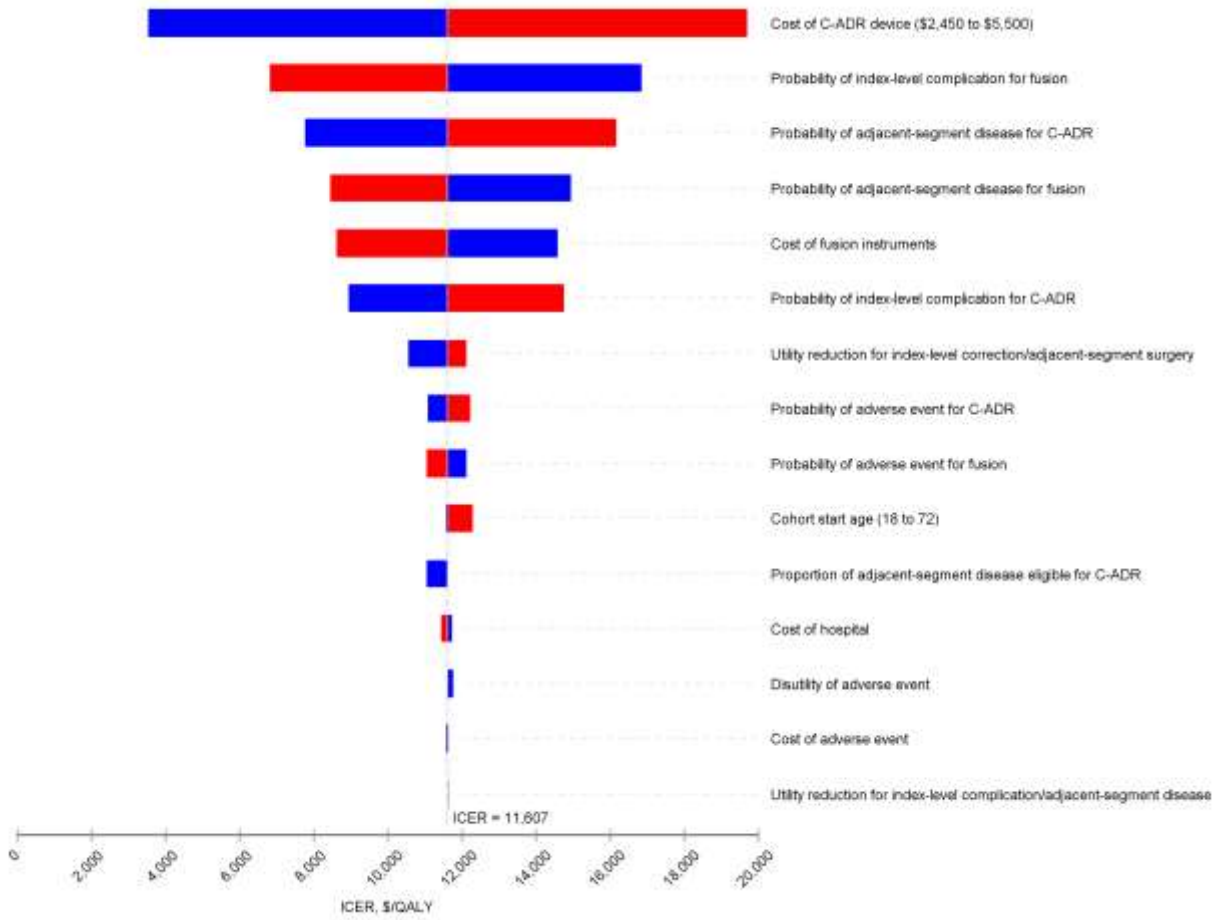

Figure 5: Tornado Diagram for One-Level Surgery

Abbreviations: C-ADR, cervical artificial disc replacement; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year.

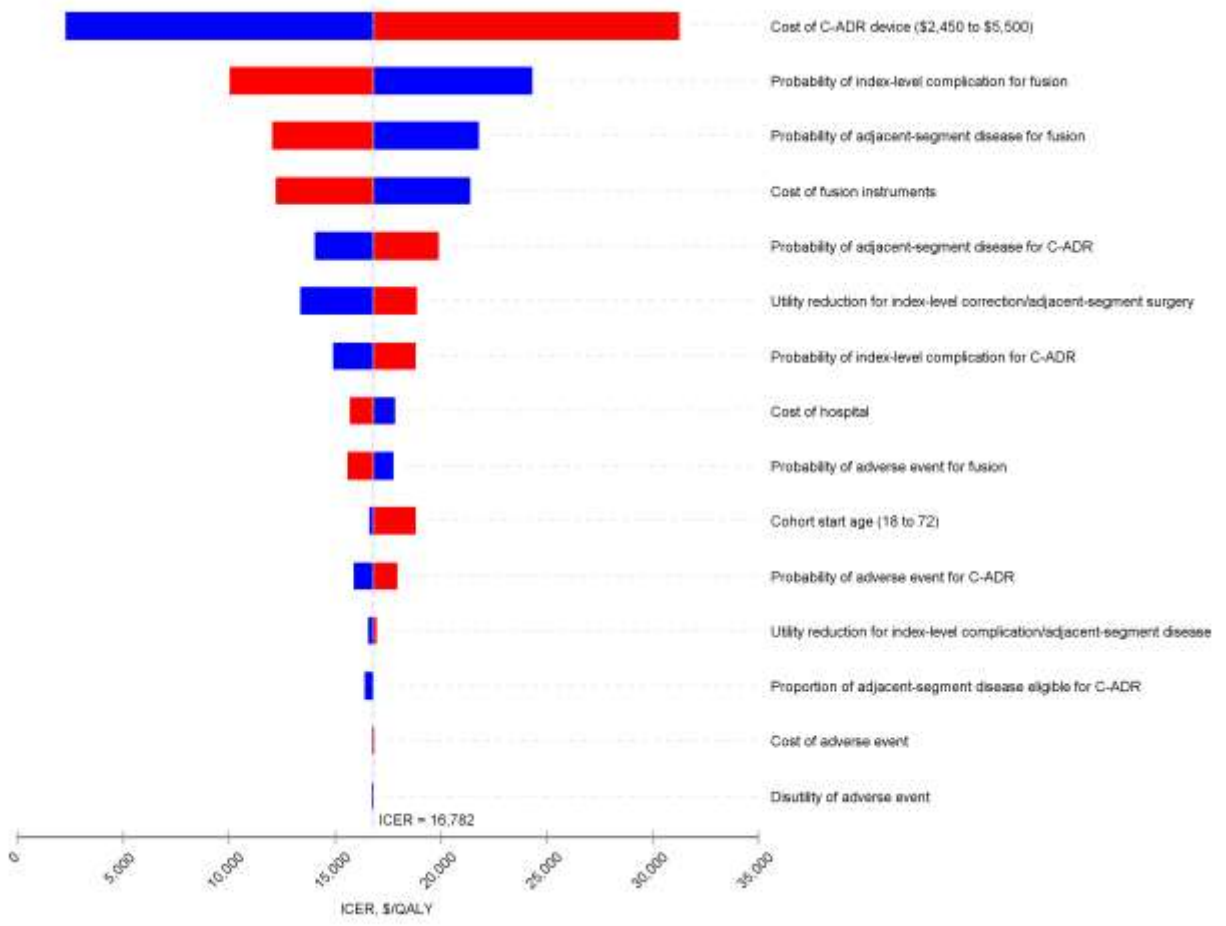

Figure 6: Tornado Diagram for Two-Level Surgery

Abbreviations: C-ADR, cervical artificial disc replacement; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year.

Probabilistic Sensitivity Analysis

Figures 7 and 8 show the incremental cost-effectiveness planes, where each point represents one ICER from one Monte Carlo simulation. We ran 5,000 simulations for each case.

For one-level degenerative disc disease: Using a willingness-to-pay of $50,000 per QALY, 98.08% of the simulations were considered cost-effective (below the willingness-to-pay line in quadrant 1), 0.24% of the simulations were considered not cost-effective (above the willingness-to-pay line in quadrant 1), and 1.68% of the simulations were superior (less costly and more effective in quadrant 2).

For two-level degenerative disc disease: Using a willingness-to-pay of $50,000 per QALY, 98.68% of the simulations were considered cost-effective (below the willingness-to-pay line in quadrant 1), 0.54% of the simulations were considered not cost-effective (above the willingness-to-pay line in quadrant 1), and 0.78% of the simulations were superior (less costly and more effective in quadrant 2).

Figures 9 and 10 show cost-effectiveness acceptability curves for one-level and two-level cases. These curves visually represent the probability of being cost-effective over a range of willingness-to-pay thresholds up to $100,000 per QALY.

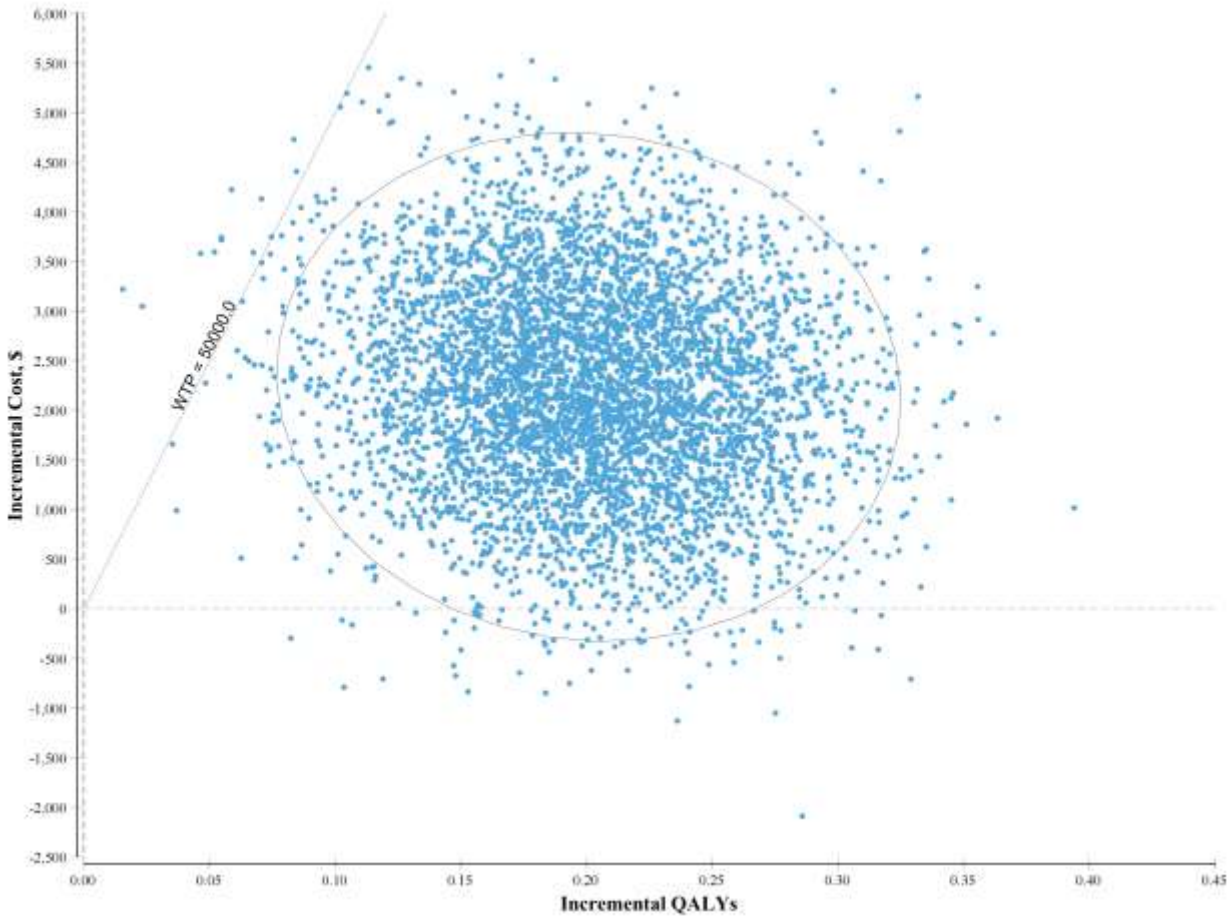

Figure 7: Incremental Cost-Effectiveness Plane for C-ADR Versus Fusion With a Willingness-to-Pay of $50,000 per QALY for One-Level Cervical Degenerative Disc Disease

Abbreviations: C-ADR, cervical artificial disc replacement; QALY, quality-adjusted life-year; WTP willingness to pay.

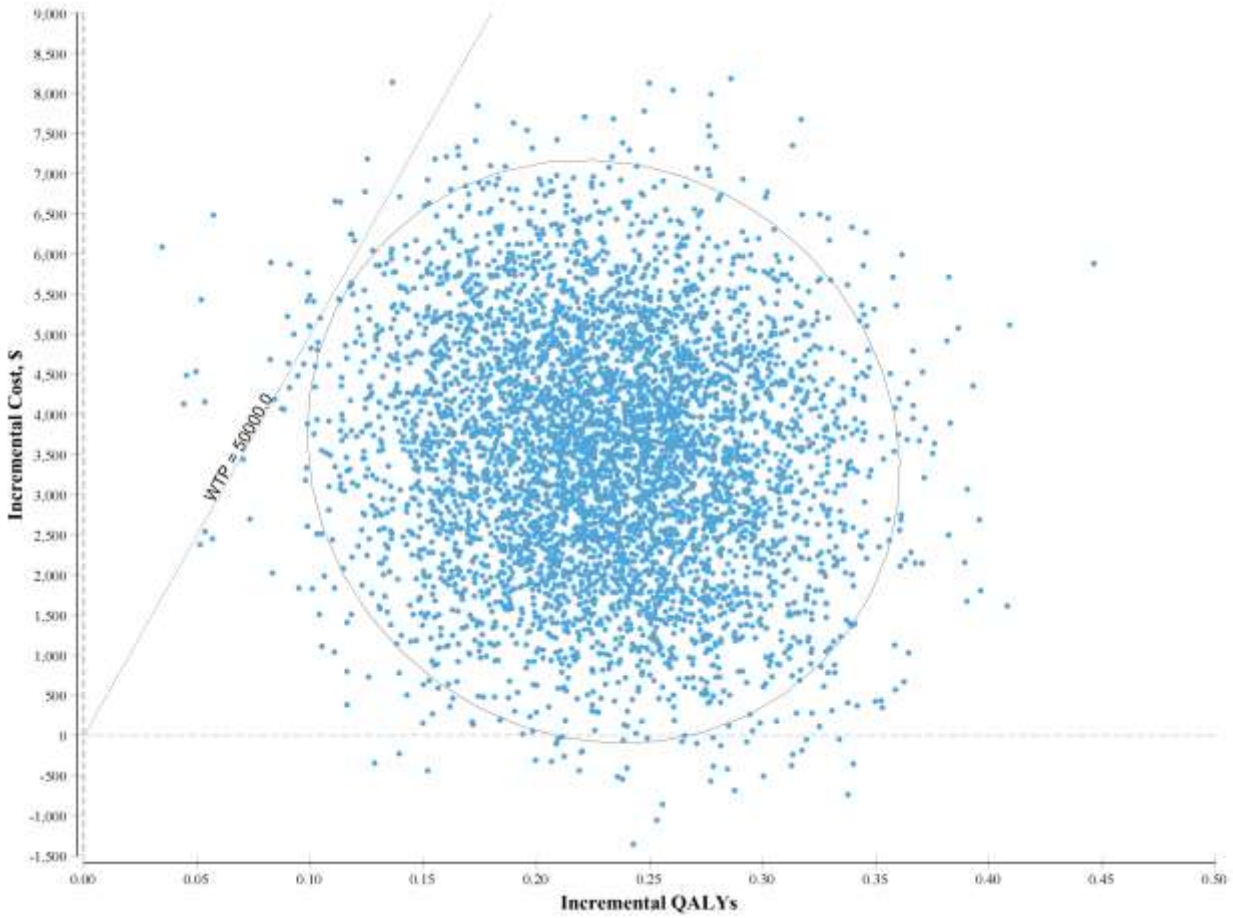

Figure 8: Incremental Cost-Effectiveness Plane for C-ADR Versus Fusion With a Willingness-to-Pay of $50,000 per QALY for Two-Level Cervical Degenerative Disc Disease

Abbreviations: C-ADR, cervical artificial disc replacement; QALY, quality-adjusted life-year; WTP, willingness to pay.

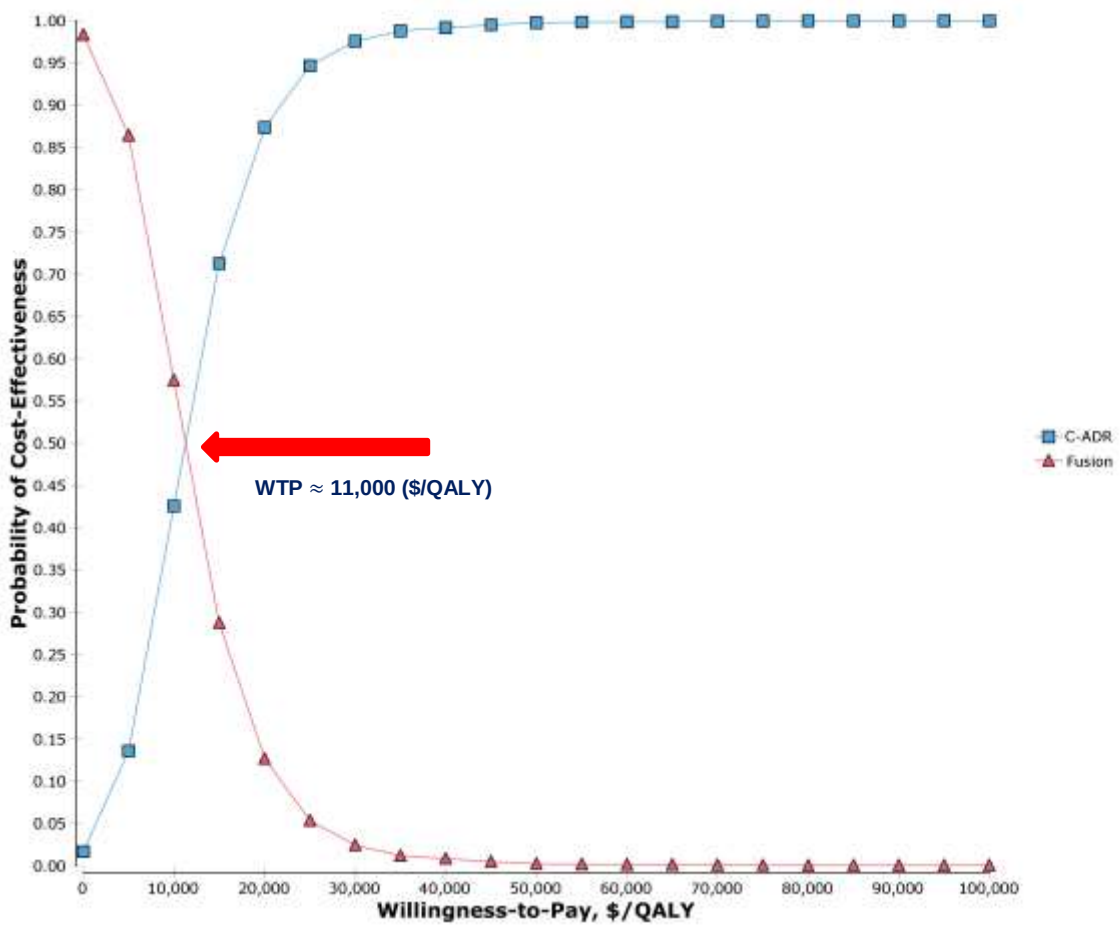

Figure 9: Cost-Effectiveness Acceptability Curve for C-ADR Versus Fusion for One-Level Cervical Degenerative Disc Disease

Abbreviations: C-ADR, cervical artificial disc replacement; QALY, quality-adjusted life-year; WTP willingness to pay.

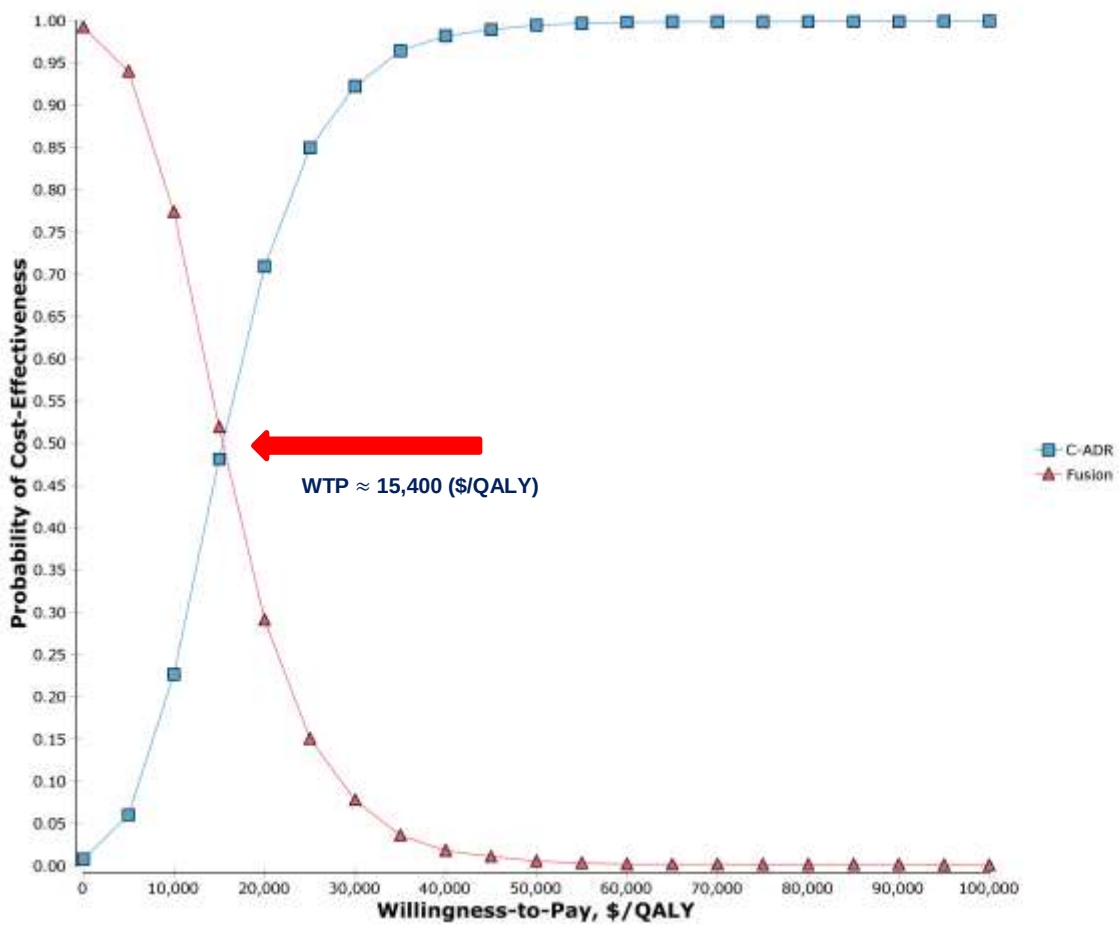

Figure 10: Cost-Effectiveness Acceptability Curve for C-ADR Versus Fusion for Two-Level Cervical Degenerative Disc Disease

Abbreviations: C-ADR, cervical artificial disc replacement; QALY, quality-adjusted life-year; WTP, willingness-to-pay.

Scenario Analysis

Results show robustness under different scenarios. Table 36 presents results for a longer follow-up times of 20 years and for alternative C-ADR devices. We also analyzed the scenario from the perspective of hospital payers by including only the hospital cost and the cost of the device for index surgery and for treatment of complications (Table A30).

For both one-level and two-level procedures, the ICER decreased when a 20-year, rather than 7-year, follow-up was chosen: $6,336.97/QALY versus $11,607.05/QALY for one-level surgery, and $804.20/QALY versus $16,782.44/QALY for two-level surgery.

Similarly, when data for alternative C-ADR devices were used to populate the model, the ICER decreased to $3,541.60/QALY for one-level surgery and $10,356.84/QALY for two-level surgery.

The ICERs for both one- and two-level surgery decreased to $9,866.85/QALY and $16,348.41/QALY, respectively, when the perspective of hospital payers was considered.

Table 36: Scenarios With Longer Follow-Up Times and for Various C-ADR Devices

Scenario	Average Total Costs, $	Incremental Cost,[b] $	Average Total Effects, QALYs	Incremental Effect,[c] QALYs	ICER, $/QALY
One-level surgery					
20-year follow-up					
Fusion	16,582.99		11.3898		
C-ADR	19,329.09	2,746.10	11.8231	0.4333	6,336.97
Alternative disc					
Fusion	15,577.29		4.4711		
C-ADR	16,266.88	689.59	4.6658	0.1947	3,541.60
Two-level surgery					
20-year follow-up					
Fusion	22,368.43		11.4449		
C-ADR	22,852.72	484.28	12.0471	0.6022	804.20
Alternative disc					
Fusion	16,406.39		4.6383		
C-ADR	20,464.76	4,058.37	5.0301	0.3918	10,356.84

Abbreviations: C-ADR, cervical artificial disc replacement; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year.
[a]Results might appear incorrect because of rounding.
[b]Incremental cost = average cost of C-ADR – average cost of fusion.
[c]Incremental effect = average effect of C-ADR – average effect of fusion.

Discussion

Our primary economic analysis is informed by a combination of sources. The utilities and rate of complications (adverse events and subsequent surgeries) are taken from RCTs. For costs, we used Ontario-specific data for most parameters, including device costs, hospital costs, and surgeon fees. Our estimated costs of subsequent surgeries were also derived from Ontario databases. We showed that it was important to use Ontario-specific data because of substantial differences in cost estimates from various sources.

We performed various sensitivity and scenario analyses to test the robustness of our results, changing parameter values and assumptions. Our ICERs for both one-level and two-level cases remained below a willingness-to-pay of $50,000/QALY.

We used data from two C-ADR devices for one-level and two-level cases. In both cases, the ICER decreased when the alternative device was chosen.

Our results for one-level surgery are in line with the trial-based and model-based economic evaluations identified in the economic review (9 of 10), and also with the only health technology assessment[148] that included a primary economic evaluation. All these studies found cost savings or cost-effectiveness in favour of C-ADR over fusion. However, the studies reporting cost savings were conducted from the perspective of society at large or assigned higher initial and recurring costs to fusion and its complications.

For two-level surgery, three of the four studies[138,140,141] identified in the economic review reported large ICERs ($69,337–$100,257 USD/QALY, 2012 or 2014) when the perspective of private insurers was chosen. However, the study by Ament et al[139] reports a small ICER ($8,518 USD/QALY, 2014). The large ICERs for those three studies result from relatively high costs associated with medical expenses for two-level cases and from the use of unrealistically high rates for costly adverse events. In our analysis, the increase in cost for two-level surgery is moderate and similar to what is reported in the study by Ament et al.[139]

Although we performed extensive sensitivity analyses to explore the variations of ICERs under changing parameter values, our results might not be generalizable to settings in which costs are very different from those we used.

A limitation of our analysis is the use of simplifying assumptions in the cohort for complication pathways (adverse events, index corrective surgery, and adjacent-segment surgery). In real-world clinical situations, these events do not necessarily occur independently, and patients sometimes experience a variety of complicated pathways.

Conclusions

The results of our primary economic analysis show that C-ADR represents good value for money compared with fusion for both one-level and two-level cervical degenerative disc disease.

BUDGET IMPACT ANALYSIS

We conducted a budget impact analysis from the perspective of the Ontario Ministry of Health and Long-Term Care to estimate the cost burden over the next 5 years of funding cervical artificial disc replacement (C-ADR). All costs are reported in 2018 Canadian dollars.

Research Questions

Within the context of the Ontario Ministry of Health and Long-Term Care, we asked two questions:

1. What is the potential 5-year budget impact of publicly funding C-ADR in Ontario for people with one-level cervical degenerative disc disease?

2. What is the potential 5-year budget impact of publicly funding C-ADR in Ontario for people with two-level cervical degenerative disc disease?

Methods

Analytic Framework

We estimated the budget impact of publicly funding C-ADR using the cost difference between two scenarios: current clinical practice without dedicated public funding for C-ADR (the current scenario) and the anticipated clinical practice with publicly funded C-ADR (the new scenario). The model schematic is shown in Figure 11.

We conducted a reference case analysis and sensitivity analyses. Our reference case analysis represents the analysis with the most likely set of input parameters and model assumptions. Our sensitivity analyses explored how results are affected by varying input parameters and model assumptions.

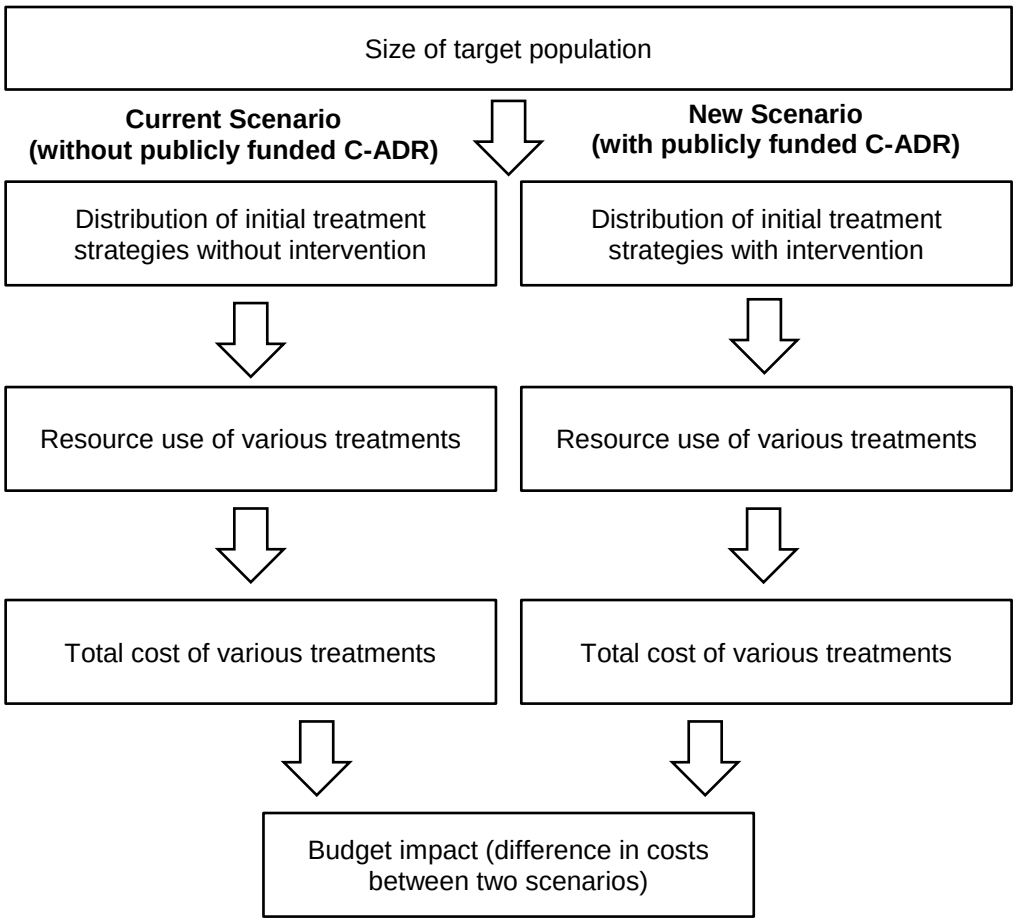

Figure 11: Model Schematic of Budget Impact

Abbreviation: C-ADR, cervical artificial disc replacement.

Key Assumptions

- The disease-specific rate of death is negligible
- All patients will receive treatment for their complications for the whole period (no emigration)

Target Population

1. The target population for Research Question 1 is adults with symptomatic one-level cervical degenerative disc disease unresponsive to conservative treatment who are eligible for both C-ADR and anterior cervical discectomy with fusion (fusion).

2. The target population for Research Question 2 is adults with symptomatic two-level cervical degenerative disc disease unresponsive to conservative treatment who are eligible for both C-ADR and fusion.

In many clinical trials, patient enrolment is restricted to minimize confounders; for example, by making sure patients have not undergone any similar surgeries before. Following Quirno et al,[147]

we relaxed the inclusion criteria by removing any constraint that is not considered an absolute contraindication in real clinical situations.

The population to be included in our budget impact analysis is all patients eligible for both fusion and C-ADR during the 5-year time horizon. The yearly size of the target population is estimated from population data and epidemiologic inputs (the prevalence and incidence) of patients in need of surgery for symptoms of degenerative disc disease. We also include predicted changes in the target population over the 5-year time horizon. We used the number of claims or cases in administrative databases for the size of the yearly incidence.

Current Intervention Mix

Table 37 lists the yearly number of C-ADR and fusion surgeries performed in Ontario from 2012 to 2017. These are yearly total numbers for one-level, two-level, and possibly multi-level (more than two levels) cases. The sum of the two columns is the total yearly number of patients who were eligible for fusion in that year. The ratio of these yearly numbers to the total adult population produces yearly incidence of fusion-eligible adult patients who had undergone surgery.

Table 37: Yearly Number of C-ADR and Fusion Surgeries in Ontario

Year	C-ADR[a]		Fusion[a]		Total Eligible for Fusion[a]	Ontario Population (≥ 18 yr)	Total Eligible for Fusion per 100,000 Population[a]
	n	%	n	%			
2012–2013	20	3.4	564	96.6	584	10,699,089	5.46
2013–2014	14	2.3	597	97.7	611	10,859,964	5.63
2014–2015	14	2.6	534	97.4	548	11,006,243	4.98
2015–2016	18	3.0	587	97.0	605	11,125,317	5.44
2016–2017	22	3.4	616	96.6	638	11,297,714	5.65

Abbreviation: C-ADR, cervical artificial disc replacement.
[a]Yearly total numbers for one-level, two-level, and possibly multi-level (more than two levels).
Source: Canadian Institute for Health Information [149] and Statistics Canada[150] (for populations).

We fit a linear model to the yearly incidence in the last column and predicted the relative incidence for the next 5 years from 2018 to 2022 (Table 38). We then used the predicted relative incidence and population projection data to predict the number of total eligible fusion patients for the same period, shown in the last column of Table 38.

Table 38: Predicted Yearly Number of Total Surgeries Eligible for Fusion in Ontario

Year	Total Eligible for Fusion per 100,000 Population[a] (Predicted)	Ontario Population (≥ 18 yr, Projected)	Total Eligible for Fusion[a] Predicted)
2018	5.51	11,699,192	644
2019	5.52	11,866,000	656
2020	5.54	12,015,646	666
2021	5.56	12,149,333	676
2022	5.58	12,285,584	686

Abbreviation: C-ADR, cervical artificial disc replacement.
[a]Yearly total numbers for one-level, two-level, and possibly multi-level (more than two levels).
Population Source: Ontario Ministry of Finance projections, spring 2017.[151]

To find the share of patients eligible for one-level and two-level surgery, we used data extracted from Alberta Health Evidence Reviews,[146] shown in Table 39. These data are similar to those in the study by Quirno et al,[147] who estimated that 45.2% of fusion-eligible patients are ineligible for one-level surgery because they have cervical degenerative disc disease at more than one level.

Table 39: Number and Relative Distribution of Fusion Surgeries Performed in Alberta, 2004–2007

	One-Level	Two-Level	Three-Level	Total (N)
Fusion surgeries (n)	639	257	239	1,135
Relative weight, n/N × 100, %	56%	23%	21%	

Source: Alberta Health Evidence Reviews.[146]

Based on these ratios and the total from Table 38, we calculated how many patients would be eligible for one-level and two-level fusion (Table 40).

Table 40: Predicted Yearly Number of Patients Eligible for One-Level and Two-Level Fusion in Ontario

	Eligible for Fusion (Predicted)		Eligible for C-ADR (Predicted)	
Year	One-Level	Two-Level	One-Level	Two-Level
2018	361	148	184	71
2019	367	151	187	72
2020	373	153	190	73
2021	379	155	193	74
2022	384	158	196	76

Abbreviation: C-ADR, cervical artificial disc replacement.

Quirno et al[147] showed that approximately 28% of patients eligible for fusion (any number of levels) are eligible for one-level C-ADR. Noting that 54.8% (100% − 45.2%) of all fusion patients are eligible for one-level fusion, we calculated that 51% (28% ÷ 54.8%) of patients eligible for one-level fusion would also be eligible for one-level C-ADR. For any-level C-ADR, Quirno et al[147] estimated 39% of all fusion patients would be eligible. Subtracting the 28% who were one-level C-ADR patients, we calculated 11% of all fusion patients would be eligible for two-level C-ADR. Using distribution from Alberta Health Evidence Reviews[146] (Table 39), we calculate that 47.8% (11% ÷ 23%) of all two-level fusion patients are also eligible for two-level C-ADR surgery. Table 40 shows the numbers predicted by this analysis.

Annual incidence of cervical radiculopathy is reported to be 83.2 per 100,000 population (≥ 15 years of age).[3] Of these patients, 8% to 35% are eventually treated surgically.[3,152] Therefore, the annual surgical incidence because of radiculopathy can be estimated to be 6.66 and 29.12 per 100,000 population (83.2 × 8% = 6.66 per 100,000; 83.2 × 35% = 29.12 per 100,000). As we can see from Tables 37 and 38, current and predicted rates for surgical incidence in Ontario are close to the lower bound of 6.66. However, because the upper bound is almost five times higher than the lower bound, it is possible that the health system currently has an unmet demand for cervical spine surgery. We report the predicted number of patients eligible for receiving fusion surgery in Ontario using these lower and upper bounds in Table A33.

Uptake of the New Intervention and Future Intervention Mix

If there were dedicated public funding for C-ADR, then uptake will likely gradually increase. The effect might be different for patients with one-level and two-level degenerative disc disease. The perceived relative clinical and lifestyle benefits, involving patients' understanding of the new technology, surgeons' belief in its effectiveness, and availability of suitable products, will affect future uptake rates.

We consulted clinical experts, manufacturers and distributors, and the Ontario Ministry of Health and Long-Term Care to estimate yearly post-funding uptake. We considered two scenarios: (1) standard uptake increase, and (2) quick uptake increase. Standard uptake corresponds to the typical diffusion speed of newly funded technologies (taken to be 25% increase per year, except for the first year). In the quick uptake scenario, it is supposed that the uptake will reach 100% within 2 years of funding and corresponds to a situation when the resources needed for C-ADR are set to meet the demand in a short time.

Table 41 shows the two scenarios, and Table 42 lists the yearly predicted numbers for one-level and two-level cases for both surgery types. We assume that current funding levels leading to the very limited uptake in past years will continue. We use 8% for future uptake that continues the trend in uptake shown in the first column of Table 37 (2.3%–3.4% of all cases eligible for fusion would mean approximately 8% of those cases would also be eligible for C-ADR).

Table 41: Anticipated Yearly Percentage of Patients Receiving C-ADR and Fusion in Ontario Among Patients Eligible for Both Surgeries, 2018–2022

			New Scenario			
	Current Scenario		Standard Uptake Increase		Quick Uptake Increase	
Year	C-ADR, %	Fusion, %	C-ADR, %	Fusion, %	C-ADR, %	Fusion, %
2018	8	92	10	90	10	90
2019	8	92	25	75	50	50
2020	8	92	50	50	100	0
2021	8	92	75	25	100	0
2022	8	92	100	0	100	0

Abbreviation: C-ADR, cervical artificial disc replacement.

Table 42: Predicted Yearly Number of Patients Receiving One-Level and Two-Level C-ADR and Fusion Surgery in Ontario Among Patients Eligible for Both Surgeries

	Current Scenario (n)				New Scenario (n)							
					Standard Uptake Increase				Quick Uptake Increase			
	One-Level		Two-Level		One-Level		Two-Level		One-Level		Two-Level	
Year	C-ADR	Fusion	C-ADR	Fusion	C-ADR	Fusion	C-ADR	Fusion	C-ADR	Fusion	C-ADR	Fusion
2018	15	169	6	65	18	166	7	64	18	166	7	64
2019	15	172	6	66	47	140	18	54	94	93	36	36
2020	15	175	6	67	95	95	36	37	190	0	73	0
2021	15	178	6	68	145	48	56	18	193	0	74	0
2022	16	180	6	70	196	0	76	0	196	0	76	0

Abbreviation: C-ADR, cervical artificial disc replacement.

Resources and Costs

Resources needed by hospitals and training required by surgeons for C-ADR, as well as any extra surgical tools needed, are provided by manufacturers and distributors free of charge. Therefore, except for negligible administrative expenses, costs for items are the same as those used in our primary economic evaluation.

We ran a simulation of our model with 0% discount rate to find yearly costs per patient needed for our analysis (Table 43).

Table 43: Undiscounted Yearly Costs of One-Level and Two-Level C-ADR and Fusion Surgeries in Ontario

	Undiscounted Yearly Cost, $[a]			
	One-Level		Two-Level	
Year After Index Surgery	C-ADR	Fusion	C-ADR	Fusion
1	15,084.81	12,670.51	19,675.92	14,216.10
2	549.53	909.97	398.24	908.14
3	251.95	203.45	188.53	488.01
4	251.10	203.08	188.03	487.12
5	250.22	202.68	187.50	486.15

Abbreviation: C-ADR, cervical artificial disc replacement.
[a]In 2018 Canadian dollars.

We base the budget impact on the estimated number of surgeries done with C-ADR and fusion in Ontario each year in the next 5 years (2018–2022). Cost details are provided separately for each year.

Analysis

In the reference case analysis, we calculate the budget required for funding C-ADR in Ontario. We also calculate the net budget impact as the difference between the costs of C-ADR and fusion treatment, including the cost of index surgery and the potential savings of fewer complications (adverse events, index-level corrective surgery, and subsequent adjacent-segment surgery).

Results

Table 44 shows the yearly budget impact and net budget impact of the two scenarios of increasing uptake, compared with the current practice of very limited uptake for C-ADR, in one-level and two-level cases.

We also calculated net budget impact using incidence rates from the published literature (Table A34) and from the perspective of hospital payers (Table A35).

One-Level C-ADR Budget Impact

For both scenarios the net budget impact increases with time, but with different trends. The yearly values for years 2018 and 2022 are the same in both scenarios, as expected from similar

uptakes, and the values for years 2019, 2020, and 2021 are higher for the scenario with quick uptake. The total 5-year net budget impact is $399,473 higher in the scenario of quick uptake.

Two-Level C-ADR Budget Impact

Trends similar to those of the one-level case are seen. However, values are smaller than one-level cases because fewer patients have two-level involvement. The total 5-year net budget impact is $339,489 higher in the scenario for quick uptake.

Table 44: Total and Net Budget Impact for Adoption of C-ADR Versus Fusion for One- and Two-Level Cervical Degenerative Disc Disease

	Total Cost, $[a]					
Scenario	**2018**	**2019**	**2020**	**2021**	**2022**	**5-Yr Total**
One-level surgery						
Current scenario: limited funding, 8% uptake	2,367,588	2,570,357	2,650,482	2,731,749	2,816,346	13,136,523
New scenario 1: standard uptake increase	2,374,831	2,636,081	2,818,671	3,011,297	3,211,969	14,052,849
Net budget impact	**7,243**	**65,724**	**168,188**	**279,548**	**395,623**	**916,326**
New scenario 2: quick uptake increase	2,374,831	2,732,612	3,018,394	3,114,515	3,211,969	14,452,322
Net budget impact	**7,243**	**162,255**	**367,912**	**382,766**	**395,623**	**1,315,799**
Two-level surgery						
Current scenario: limited funding, 8% uptake	1,042,102	1,118,645	1,167,597	1,217,461	1,284,815	5,830,619
New scenario 1: standard uptake increase	1,047,562	1,178,044	1,307,110	1,435,028	1,568,504	6,536,247
Net budget impact	**5,460**	**59,399**	**139,513**	**217,567**	**283,689**	**705,628**
New scenario 2: quick uptake increase	1,047,562	1,267,142	1,479,176	1,513,353	1,568,504	6,875,736
Net budget impact	**5,460**	**148,497**	**311,579**	**295,891**	**283,689**	**1,045,117**

[a]In 2018 Canadian dollars.

Discussion

The uptake for C-ADR in Ontario is currently about 8% of eligible cases. One of the main reasons for this low rate appears to be the limited funding available to hospitals for the cost of the C-ADR devices. The costs associated with hospital stays, operations, and surgeons' fees do not differ substantially between surgical procedures and are not limiting access, according to clinical experts. The costs associated with complications are lower for C-ADR than fusion, which leads to downstream savings that could partially recover the higher cost of the index surgery.

The technical aspects of increasing uptake are not very demanding, according to clinical experts. Surgeons need special training, which is offered by manufacturers and distributors free of charge, as well as any extra surgical tools needed.

There are several strengths to this analysis. First, we used Ontario-specific costs for most items. We also explored various scenarios for uptake increase. We used published literature and

expert opinion to inform estimates of people eligible for C-ADR among people who would receive fusion. We were able to consider the savings from fewer complications in our analysis because our primary economic evaluation included them.

Our results cannot be generalized to settings where costs are considerably different from those we used.

We were unable to find the true rates for patients with cervical degenerative disc disease who need surgery in Ontario. We instead used surgical incidence rates, representing the current capacity of the health system for surgical treatment determined by several factors. However, we reported the results for a range reported in published literature. The lower bound of this range is close to our surgical incidence rates for Ontario, and the upper bound is approximately five times higher, increasing budget impact by five times if this upper bound is used in calculations.

Conclusions

The current uptake for one-level and two-level C-ADR is about 8% of the total eligible. If dedicated public funding for C-ADR were to become available, we estimate a total net budget impact of $916,326 for one-level surgeries and $705,628 for two-level surgeries over the next five years.

PATIENT PREFERENCES AND VALUES

Objective

This analysis aimed to explore the underlying values, needs, impacts, and preferences of people with cervical degenerative disc disease. Our treatment of focus was cervical artificial disc replacement (C-ADR).

Background

Patient, caregiver, and public engagement provides a unique source of information about people's experiences of a health condition and the health technologies or interventions used to manage or treat that health condition. It explores the impact of the condition and its treatment on the patient, on the patient's family and other caregivers, and on the patient's personal environment. It also provides insights into how a health condition is managed by the province's health system.

Information shared from lived experience can also identify gaps or limitations in published research (e.g., typical outcome measures sometimes do not reflect what is important to those with lived experience).[153-155] Lived experience can add information and perspectives on implications of health technologies or interventions for ethical and social values.

Because the needs, priorities, preferences, and values of those with lived experience in Ontario are often inadequately explored in published literature, we contact and speak directly with people who live with a given health condition, including those who have experience with the intervention we are exploring.

For this project, we spoke with 12 people with cervical degenerative disc disease who had undergone C-ADR or fusion. We discussed how cervical degenerative disc disease affects their life and their experience with various treatments. Gaining an understanding of the day-to-day experience of living with cervical degenerative disc disease, including people's experiences with C-ADR or fusion, helps us assess the potential value of C-ADR from the perspective of patients and caregivers.

Methods

Engagement Plan

The engagement plan for this health technology assessment focused on consultation with patients and caregivers to examine the experiences of people with cervical degenerative disc disease, including their experience with C-ADR or fusion.

We used a qualitative interview, as this method of engagement allows us to explore the meaning of central themes in the experiences of people with cervical degenerative disc disease. Our main task in interviewing is to understand what people tell us and to understand the meaning of their experiences.[156] The sensitive nature of exploring people's experiences of a health condition and their quality of life supports our choice of method of engagement.

Participant Outreach

We used purposive sampling,[157-160] which involves actively reaching out to patients, families, and caregivers with direct experience of the health condition and health technology or intervention being reviewed. We approached a variety of partner organizations, health clinics, spine care support associations, and surgical specialists to spread the word about this engagement activity and to contact patients, families, and caregivers with experience of cervical degenerative disc disease and C-ADR.

Inclusion Criteria

We sought to speak with people with cervical degenerative disc disease and their families. These people were not required to have had direct experience with C-ADR to participate.

We sought broad geographic, cultural, and socioeconomic representation to elicit possible equity issues in accessing treatment for cervical degenerative disc disease, including C-ADR.

Exclusion Criteria

We did not set exclusion criteria.

Participants

We recruited participants from across the province and conducted interviews with 12 people, including 11 people with cervical degenerative disc disease and one family member. Of the people with cervical degenerative disc disease, five had undergone C-ADR, three had undergone fusion, and three were waiting for surgery to be scheduled.

All participants had direct experience with managing cervical degenerative disc disease and were familiar with its impact and various treatments. Because no participants had undergone both fusion and C-ADR, they were unable to compare experiences or speak conclusively about benefits or drawbacks of each surgery compared with the other.

Approach

At the beginning of the interview, we explained the role of Health Quality Ontario, the purpose of this health technology assessment, the risks of participation, and how participants' personal health information would be protected. We provided this information to participants both verbally and in a letter of information, if desired (Appendix 12). We then obtained each participant's verbal consent before starting the interview. With participants' consent, we audio-recorded the interviews and then transcribed the recordings.

Interviews were conducted by phone and lasted 20 to 40 minutes. The interview was loosely structured and consisted of a series of open-ended questions. Questions were based on a list developed by Health Technology Assessment International's Interest Group on Patient and Citizen Involvement in Health Technology Assessment.[161] Questions focused on the development of degenerative disc disease, its progression, and the impact on participants' and families' quality of life. Interview questions then covered participants' experiences with treatments and participants' perceptions of the benefits and limitations of C-ADR and fusion. See Appendix 13 for our interview guide.

Data Extraction and Analysis

We used a modified version of a grounded-theory method to analyze interview transcripts and written results. The grounded-theory approach allowed us to organize and compare information across participants. This method consists of a repetitive process of obtaining, documenting, and analyzing responses while simultaneously collecting, analyzing, and comparing information.[162,163] We used the qualitative data analysis software program Nvivo (QSR International, Doncaster, Victoria, Australia) to identify and interpret patterns in interview data. The patterns we identified then allowed us to highlight the impact of cervical degenerative disc disease and various therapies on those we interviewed.

Results

Lived Experience of Cervical Degenerative Disc Disease

For the people we interviewed, cervical degenerative disc disease appeared in a variety of ways. For most participants, the disease appeared as mild pain in the neck or back with slow progression over several years. Typically, symptoms of mild pain were the first sign of any medical issue, though a few participants reported that a traumatic event or injury precipitated the pain. This pain would often progress, extend to other parts of the body, and intensify over time. Participants added that often other symptoms, such as numbness or tingling, appeared before they sought medical care. Symptoms of degenerative disc disease were not reported as constant but could increase or decrease in severity depending on circumstances and activities.

> *I started having a lot of neck pain when I was 19. And muscle spasms … would come and go, to the point where I'd get probably two to three episodes a month with muscle spasms.*
>
> *I have cervical; it's only between discs 5 and 6, and I've had it for 20 years.*
>
> *[Ever since] my middle 20s, I'd say. I was in my mid-20s, and I had pain in the neck. And I mean, like any pain, you just take two Tylenol and go about your day. That's what I did. Every time I had pain, I would just take two Tylenol.*

Neck pain or stiffness often progressed over time, and other symptoms could include numbness, burning, and pain in other areas. These symptoms often restricted a person's ability to perform certain activities. Participants reported that simple daily tasks, such as raising arms to perform chores or even sleeping could be affected by these symptoms of cervical degenerative disc disease.

> *[M]y arm was so painful, it throbbed—with pain, not just like, muscle or anything. It was a real pain, and I ended up going to the hospital. So, I do have a numbness on the under part of my arm, from my underarm almost to my elbow. …*
>
> *Like right now I can't do anything over my head. Like I have to be above whatever I'm doing. And if I'm carrying, let's say, two jugs of water in, my neck will just—extreme pain and down the arm and a burning sensation.*
>
> *It affects my sleep. I'll wake up with [numbness] and tingling and just a stiff neck. So I find I have to move. But I have to be careful when I move because often that will trigger … in the beginning, it triggered neck spasms. And then I'd be out of commission for a*

> *good week. It would be brutal pain and I couldn't move my neck. But now with the years, I know what the triggers are, so I've adapted.*

More strenuous activities, such as walking, driving, or exercising, could also be impaired as cervical degenerative disc disease progressed. As symptoms increased in intensity and spread, participants reported that the effect became larger. Participants also reported adjusting their daily routines or activities to try to cope with pain and other symptoms.

> *[After] 3 years with what I have, just this massive pain and numbness in my arms, … it was starting to affect my legs, too. So it was starting to attack my physical walking where I couldn't walk anymore.*

> *So I was doing all sorts of things, like I've had a stand-up desk at the office for years and I just got rid of the chair completely; I never sat down.*

> *Driving wasn't too bad as long as I was in the right position; I had to kind of pull the seat really forward, so I was jamming my shoulders back into the seat while I drove because that was the position that gave me some relief.*

Most of those interviewed mentioned having to adjust their work schedules or career responsibilities in response to pain, stiffness, or numbness. In several instances, participants reported that they were forced to change their careers because of these symptoms or even to withdraw from the workforce entirely.

> *I was like constantly fatigued. I had … a physical job and had to leave that job and start to do a home daycare. And [I] started to do some night courses, because I knew I wasn't going to be able to do physical work with those symptoms.*

> *I practised that for about 3 years. But then, because of the persistent neck pain and the numbing and the tingling, it was just something that I couldn't do anymore. So, I basically had to find another type of job that didn't require me [to look] down all the time or fine hand motor movement. So it did affect my career. …*

Such restrictions or changes in lifestyle resulted in various levels of frustration and anxiety for participants. Those interviewed expressed this frustration and exasperation at their change in baseline level of activity and felt fearful of the future. Some expressed a desire to continue these activities despite the pain and discomfort but often found they could no longer perform certain activities. For those who had been fairly active before their pain had progressed, this activity restriction could be quite pronounced. Some of those interviewed reported on the emotional impact of this activity restriction. Beyond the frustration associated with activity restriction, participants spoke of levels of depression and dismay at the progression of pain and other restrictive symptoms.

> *It has been getting worse, and I think I've just kind of succumbed to the idea that that's just the way I am, the way that my bones are.*

> *I always had a lot of stress and anxiety because it was like … I'm a pretty independent person. But it was limiting my life.*

> *You know, it really had me down in the dumps. It really had me worried that, you know, just what am I going to be living with for the rest of my life, if I'm now not even at 50?*

These changes in activity level and the emotional burden of dealing with consistent levels of pain and other symptoms was reported to affect families and caregivers. Several participants reported detrimental effects of the condition on their relationships with family members and their own mental state.

> *Yeah, so again, from the outside I was somewhat functional; people at the office noticed that I was … in pain. People—you know, my wife and kids—noticed that … I was having good days, bad days, and I guess that … affected my moods, too.*
>
> *[T]hat was before I had kids. So, this is before, you know around the time where you're trying to build your career, contemplating parenthood. And, that level of pain, it gets in the way. And it does change how you look into the future.*
>
> *I was limited in anything, doing in the house. My wife had to take over all of the responsibilities. And you know, as a man, you don't like that. That's not what you want in life. You want to be the strong man and help your family.*

Treatments for Degenerative Disc Disease

Participants reported that, when the pain first developed, it was often treated with simple pain medications. Occasionally other treatments, such as acupuncture, physiotherapy, chiropractic care, or exercise were used to alleviate pain. In general, these treatments were perceived as having minimal benefit, because they were not treating the underlying cause of the pain, stiffness, or numbness.

> *Yeah, yeah. It was pretty much pain meds. I tried massage, physio, acupuncture, osteopath, chiropractic, you know, everything except leaving chickens on the front doorstep! I was ready for it. But if someone had suggested that would have helped, I would have done that too, right.*
>
> *And I went to physio, chiro, massage, not really knowing what the heck was going on.*
>
> *It started off with alternatives. I tried acupuncture and it was … it was good for the muscle spasms. But it was very short term and it's the same with, you know, physio and massage. It definitely helps, but it's short term, unfortunately.*

As the pain progressed, or other symptoms developed, more medical interventions were sought. Participants reported getting imaging—x-rays, computed tomography, or magnetic resonance imaging—as well as seeing a series of doctors and specialists. Occasionally, participants reported seeking out pain clinics to help manage their pain, which sometimes prescribed nerve blocks or opioid medication. While opioids were reported to have beneficial effect on the pain, participants reported side effects as well.

> *So, naproxen, anti-inflammatories … I've done nerve blocks for it, too.*
>
> *I found the opioids were interfering terribly with my job. They took care of the pain all right, but I just didn't … I lost … I wasn't sharp anymore. Not to say that I'm all that sharp, but however sharp I was, I wasn't as sharp as that anymore, and that was over a 2-year period.*

> *See, I did not want to be on pain pills for the rest of my life. It's not a life to have. I mean, not only do they affect you physically, … the amount of pain pills I was on, but emotionally, it kills you inside, because you can't do anything.*

Accessing health care services for these treatments could be a challenge. Some participants spoke of extended wait times and frustrations in accessing medical specialists or pain clinics. These wait times could increase frustration because the nature of the pain could vary while waiting to see a health care practitioner.

> *Well, that's the whole thing! That's the whole thing; when it flares up, that's when I go see my [general practitioner]. He refers me to an orthopedic surgeon or a neurosurgeon. And by the time I get in to see him, I haven't done [activities] because of the pain. And so the flare-up has decreased considerably. You know, there might not even really be anything there anymore.*

> *But they don't seem to understand what [my physician] is trying to refer for. The bureaucracy to get through to get to these things is mind-boggling.*

Fusion Versus Cervical Artificial Disc Replacement

Participants reported that, as the symptoms of cervical degenerative disc disease progressed and other options were exhausted, fusion became a more prominent option. Often, surgical treatment was proposed by their physicians to address the underlying issue in the cervical spine. Among participants who had had surgery several years ago, fusion was more likely. More recent surgeries could be either fusion or C-ADR. Two participants reported speaking to their surgeons regarding this change in care.

> *And when I saw one orthopedic surgeon, he said, "If you had come to me 5 years earlier, we would have fused your neck. But now, we don't fuse unless we really, really have to." He goes, "We let the body fuse the disc naturally." So basically, because you lose the movement in the disc, eh? So it's not moving all the time. So sooner or later, it's going to fuse together.*

People who had undergone fusion reported a variety of expectations regarding this surgery. People knew that fusion could restrict neck movement, which was a cause for hesitation. However, they reported that the continuous pain and discomfort from the underlying cervical issue—often lasting many years—meant that fusion was seen as the only way to provide effective and sustained relief of these symptoms.

> *I think once [we discussed magnetic resonance imaging results] and they could prove to me that this was the problem and they could fix it, then it was a no-brainer.*

> *I was never worried about being paralyzed or anything. I was pretty debilitated in what I could do at that point, so, I didn't think it could be any worse.*

> *So, you know, I'd like to have … to be pain-free for the next 10 years, maybe 15, on average, you know. By the time I get to 70, I'm not going to be doing that much. The time I'm going to be doing the most is now. So it makes sense to fuse them now, have no pain, and be able to do more. And then when I turn 70, 75, or that. I won't be doing as much, so I won't—do you know what I mean?*

Impact from this surgery varied; some participants reported generally positive outcomes while others considered the surgery less successful. Restriction in movement of the neck and shoulders was mentioned as the most common negative result, though one person reported more serious results from fusion.

> *I was lying in bed, and I said, "Can I try and lift my arm?" And, he said, "Yeah, try." And, my left arm went up over my head. It was like immediate … immediate relief.*

> *My range of motion is very, very limited and there is a grinding, like bone on bone. When I turn my head, I can hear it clicking and grinding. I'm doing it right now. If I could put a microphone up to my ear and have the ear buds at the back of my head and have it actually transfer through the phone, [you'd hear it]."*

> *They ended up doing a C1, C2 fusion, taking bone from my right hip to put in with two wires. I take it the surgery wasn't 110%. … It took a lot longer than it should've. But anyways, 2 years later the whole thing crumbled and fell apart because apparently my body hadn't accepted the bones, its own bones.*

Two people mentioned that this restriction in movement required extensive physiotherapy after the surgery and affected certain activities of daily living afterward.

> *You know, you learn to turn differently. I went for therapy, I went to chronic pain management, I've been to physio, you know. You learn how to adapt and do things differently. I can't ski anymore. I can't dive anymore.*

> *Oh, yeah. I had a brace. They tried to do a halo. Initially they were going to do a halo, but they never went with the halo. But they did put me in one of those braces that … covers your chest and … holds your chin up. Like there's a plate your chin sits on, and you can adjust it in the chest piece, like to [adjust] how high your head sits up. And, yeah, I did physio for a long time.*

Participants who had had surgery more recently were more likely to have been aware of C-ADR and to have had this surgery than fusion. Several participants reported researching C-ADR themselves when they realized it could be used for their cervical degenerative disc disease. These patients reported advocating for C-ADR over fusion.

> *The disc replacement is the better of the two. I did the research myself. I've talked to other people, other doctors. The fusion is not the better of the two. … I've seen it fail time after time with patients. I'm not saying anything [my physician] does would fail, but I'm just saying that … I've heard of people with fusion where it breaks after a while, and they're back to square one.*

> *And yeah, I guess the science and the statistics take years to settle, but … I'm an engineer, and I took a look at what they were proposing … on the fusion side and [looked] at what they were proposing on the artificial disc [side], and it's a bit of a no-brainer, really.*

Expectations for the surgery were generally high. While respondents expressed some fears about surgery in such a sensitive area, participants emphasized the benefit of not having their neck vertebrae fused and thus retaining mobility after surgery. The potential restriction on neck

mobility after surgery was reported as one of the reasons participants preferred C-ADR over fusion.

> *It was worrisome. I was terrified because of, … you know, anything to do with your neck. It's the spine area and I was … I saw [my physician], and he explained it all, saying he could make it better. And, I just said, "I think I'm going to think about it." I was terrified.*
>
> *And also, the fact that I'll keep my mobility and so everything seemed on the plus side.*
>
> *[It] was a shitty decision, because I had to put my surgery off again. But it was a decision that I know I made it right. There was no way that I was going to lose rotation in the neck at the age of 47, or 46 at the time, or 45. I wasn't going to lose rotation in my neck.*

Participants who had undergone C-ADR perceived positive results. They most commonly perceived a reduction in pain and stiffness. Those interviewed also mentioned that there was little need for rehabilitation or any sort of support device around their necks after the surgery, which was seen as a benefit.

> *So I finally got the surgery done back in [the winter]. And at first, it was touch and go with the pain. The pain was a lot. But then, after a month, the pain subsided, and I am now going down in the amount of pain pills that I'm taking, way down.*
>
> *By all accounts, I think I'm doing things absolutely right compared to fusion, of course. I was chatting with someone who had fusion the other day; he was in a collar for 7 weeks, but with the disc replacement, you walk out of the hospital the next day and can resume normal life.*
>
> *It was pretty incredible. Like I said, I couldn't believe the results, for one thing. And the fact that it made a lot of difference in … what was going on. A lot of the dizziness was gone.*

Participants reported they were able to resume activities of daily living that had previously been restricted by symptoms of their cervical degenerative disc disease. This ability was of great emotional benefit to the participants, according to those we interviewed.

> *And I woke up. And fortunately, I was kind of juggling around, but I was happy with that because hey, my arms were moving, my feet were moving, everything was working.*
>
> *It just changed my life. You know, I went from a decrepit person [who] couldn't lift a laundry basket to where I was vacuuming and mopping floors, just this morning.*

Participants who had either C-ADR or fusion reported some barriers to accessing these treatments. Typically, many appointments and referrals were needed, which could take months or even years to complete before surgery. Several people expressed frustration at this process and the perceived lack of options from the health care system.

> *They said, "Well, you know, there's no [C-ADR surgeon] in Ottawa." I said, "That's fine; I'll go anywhere." At that stage I was contemplating putting another mortgage on the house and paying the equivalent of $50,000 Canadian to get them to do it in the [United States].*

> *As a matter of fact, I went to my family doctor the other day and she said, "There's a 180-month waiting [period for the surgeon] now."*
>
> *I don't know. Like, … you get to a point where you're tired of appointments and, you know, getting shuffled through the system from one specialist to another specialist and nobody wants to touch you. Yeah, you just … get like, "Screw it."*

One specific barrier mentioned was the perceived limit on the number of C-ADR surgeries offered in Ontario. Several participants reported being told that there were cost barriers to additional C-ADR surgeries being performed in Ontario.

> *From there, it took 3 years to get the surgery, due to certain complications that were going on. One was the fact that, when I was supposed to go for surgery 3 years ago, the hospital's fiscal year was renewed, and the disc replacement wasn't feasible anymore. They were only doing spine fusion. And I didn't want spine fusion.*
>
> *But when the surgeon said to me … that he only does about three or four of these a week and he'd like to do more, … I asked him why, and he said it's because he can't get the approval to do it and it comes down to the cost of the artificial disc.*

Discussion

We engaged people with lived experience of cervical degenerative disc disease and surgical treatments for the condition, including C-ADR. We attempted to interview participants from across Ontario, but responses generally came from the London, greater Toronto, and Ottawa regions. Perhaps for this reason, we did not hear about any specific geographic or equity issues related to accessing surgery.

Eleven participants had direct experience with cervical degenerative disc disease, and one had experience as a family member of a person with cervical degenerative disc disease. All interviewees were familiar with different treatment options to manage symptoms. Therefore, those we interviewed were able to compare the perceived benefits and harms of different treatments and describe the impact of cervical degenerative disc disease on their activities of daily living, their ability to work, and their emotional well-being. In addition, participants were able to discuss their decision-making when it came to weighing surgical options and choosing a type of surgery.

People who had undergone C-ADR spoke positively of this surgery and its effect on symptoms. Participants perceived that the ability to retain mobility in the neck after surgery was a benefit of C-ADR over fusion. While we were unable to interview people with experience of both C-ADR and fusion, several people familiar with both options were able to reflect on their decision-making in choosing one type of surgery over the other.

No matter the type of surgery received, participants who had undergone surgery reported on the challenge of accessing surgery. While cost was not reported as a barrier to accessing surgery, the length of time involved was reported as frustrating by most of those interviewed. The time to receive surgery was perceived as a large barrier, and the limited number of surgeries performed in Ontario was a barrier particularly for C-ADR.

Conclusions

People with cervical degenerative disc disease reported on the negative effect that symptoms of pain and numbness can have on their quality of life. They reported using a variety of treatments with mild success, with surgery being perceived as the most beneficial and permanent solution for cervical degenerative disc disease. Those who had undergone C-ADR spoke positively of its effect on their quality of life and ability to maintain movement in their neck after surgery. The limited availability of C-ADR in Ontario was viewed as a barrier to receiving this treatment.

CONCLUSIONS OF THE HEALTH TECHNOLOGY ASSESSMENT

In carefully selected patients with cervical degenerative disc disease undergoing cervical artificial disc replacement (C-ADR) or fusion for cervical degenerative disc disease, there is evidence that:

- C-ADR is an alternative to fusion for cervical degenerative disc disease given outcomes that are statistically noninferior to fusion: perioperative outcomes (GRADE high), health-related quality of life (GRADE high), patient satisfaction (GRADE high), and overall treatment success for one-level cervical degenerative disc disease (GRADE moderate)

- C-ADR might be preferable to fusion for cervical degenerative disc disease given outcomes that are statistically superior to fusion: quicker recovery and return to work (GRADE moderate), higher technical success and lower reoperation rates at the index site (GRADE moderate); maintenance of more normal spinal segment kinetics (GRADE moderate), and higher overall treatment success for two-level cervical degenerative disc disease (GRADE moderate)

We are uncertain if adjacent-level surgery rates differ between C-ADR and fusion for one-level and two-level cervical degenerative disc disease (GRADE low). Evidence was also insufficient to determine the long-term durability of C-ADR devices.

All studies we reviewed in the economic literature search were conducted in the United States (13) or Germany (1). Results from these studies are not transferable to Ontario mainly because costs for items in Ontario are very different from reported costs. Further, none of the studies reported results, in the form of sensitivity analysis or scenario exploration, to transfer to the Ontario setting or to establish the cost-effectiveness of various brands of C-ADR device.

The results of our primary economic analysis show that C-ADR is likely to be a cost-effective intervention compared with fusion for both one-level and two-level cervical degenerative disc disease.

The current level of uptake for one-level and two-level C-ADR is about 8% of the total eligible. If public funding becomes available, uptake will increase, which will result in an estimated total net budget impact of $916,326 (501 procedures) for one-level surgeries and $705,628 (193 procedures) for two-level surgeries.

People with cervical degenerative disc disease reported on the negative effect that symptoms of pain and numbness can have on their quality of life. They used a variety of treatments with mild success, but surgery was perceived as the most effective and permanent solution. Those who had undergone C-ADR spoke positively of its impact on their quality of life and ability to move their neck following surgery. The limited availability of C-ADR in Ontario was viewed as a barrier to receiving this treatment.

ABBREVIATIONS

AMSTAR	A Measurement Tool to Assess Systematic Reviews
ASD	Adjacent segment disease
C-ADR	Cervical artificial disc replacement
CI	Confidence interval
DSQ	Dysphagia Short Questionnaire
EQ-5D	European Quality of Life Questionnaire in Five Dimensions
FDA	U.S. Food and Drug Administration
GRADE	Grading of Recommendations Assessment, Development, and Evaluation
HRQOL	Health-related quality of life
ICER	Incremental cost-effectiveness ratio
IDE	Investigational device exemption
MCS	Mental Component Summary (of the Short-Form Health Survey)
NDI	Neck Disability Index
NICE	National Institute for Health and Care Excellence
OR	Odds ratio
OTS	Overall treatment success
PCS	Physical Component Summary (of the Short-Form Health Survey)
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-analyses
QALY	Quality-adjusted life-year
RCT	Randomized controlled trial
SF-6D	Six-dimension classification of results from the SF-12 and SF-36
SF-12	12-Item Short-Form Health Survey
SF-36	36-Item Short-Form Health Survey

GLOSSARY

Adverse event	Any unexpected problem that happens during treatment, regardless of the cause or severity.
Allograft	The transplant of bone from another person or a cadaver.
Autograft	The transplant of bone from one part of the body to another in the same person.
Cervical spine	The upper section of spine in the neck region; it consists of seven vertebrae.
Cost–utility analysis	A type of analysis that estimates the value for money of an intervention by weighing the cost of the intervention against the improvements in length of life and quality of life. The result is expressed as a dollar amount per quality-adjusted life-year (QALY).
Degenerative disc disease	A deterioration of the intervertebral discs in the spine that in some cases results from loss of moisture and increased brittleness.
Generalizability	The degree to which study results may apply or be relevant to populations or groups that did not participate in the study.
Health-related quality of life	A measure of the impact of a health technology or intervention on a patient's health, including dimensions such as physiology, function, social life, cognition, emotions, sleep and rest, energy and vitality, health perception, and general life satisfaction.
Heterotopic ossification	The abnormal growth of bone in nonskeletal tissues, such as muscle or tendon. The condition typically occurs following spinal cord injury and results in jagged, painful joints.
Incremental cost-effectiveness ratio (ICER)	Determines a unit of benefit for an intervention by dividing the incremental cost by the effectiveness. The incremental cost is the difference between the cost of the treatment under study and an alternative treatment. The effectiveness is usually measured as additional years of life or as quality-adjusted life years (QALYs).
Index surgery	A first surgery.
Markov model	A type of modelling that measures the health state of a patient over the course of treatment. A patient may stay in one health state or move from one health state to another, depending on the effect of the treatment or the progression of the disease.
Minimally clinically important difference	Measurement scores that reflect changes in a clinical intervention that are meaningful to patients.
Natural history	The course of a disease from when it begins until it resolves, in the absence of treatment.

Noninferiority trial	A study designed to test that a treatment is not inferior to a comparison treatment; i.e., that a treatment is "not worse than" or is "at least as good as" another treatment.
One-level cervical degenerative disc disease	Cervical degenerative disc disease affecting a single disc.
Quality-adjusted life-year (QALY)	A measurement that takes into account both the number of years gained by a patient from a procedure and the quality of those extra years (considering such factors as ability to function and freedom from pain). The QALY is commonly used as an outcome measure in cost–utility analyses.
Scenario analysis	An analysis exploring a range of possible outcomes for an action by projecting the effects of different future events.
Sensitivity analysis	Every evaluation contains some degree of uncertainty. Study results can vary depending on the values taken by key parameters. Sensitivity analysis is a method that allows estimates for each parameter to be varied to show the impact on study results. There are various types of sensitivity analyses, including deterministic, probabilistic, and scenario.
Two-level degenerative disc disease	Cervical degenerative disc disease affecting two neighbouring discs.
Utility	The perceived benefit (value) placed on a treatment by a person or by society.

APPENDICES

Appendix 1: Literature Search Strategies

Clinical Evidence Search

Search date: July 11, 2017

Databases searched: All Ovid MEDLINE, Embase, Cochrane Database of Systematic Reviews, CRD Health Technology Assessment Database, Cochrane Central Register of Controlled Trials, NHS Economic Evaluation Database

Database: EBM Reviews - Cochrane Central Register of Controlled Trials <June 2017>, EBM Reviews - Cochrane Database of Systematic Reviews <2005 to July 6, 2017>, EBM Reviews - Health Technology Assessment <4th Quarter 2016>, EBM Reviews - NHS Economic Evaluation Database <1st Quarter 2016>, Embase <1980 to 2017 Week 28>, All Ovid MEDLINE(R) <1946 to Present>

Search Strategy

---

1 neck/ (76263)
2 exp cervical vertebrae/ (38023)
3 neck pain/ (24829)
4 (cervical or neck).ti,ab,kf. (823513)
5 or/1-4 (851625)
6 total disc replacement/ (1200)
7 ((disc*1 or disk*1) adj2 (replace* or prosthes#s or prosthetic* or arthroplast* or implant* or artificial*)).ti,ab,kf. (6876)
8 or/6-7 (7023)
9 5 and 8 (2673)
10 (neck adj2 (arthroplast* or ADR or TDR or TDA or CADR or CTDR or CDR or CDA or ACDA)).ti,ab,kf. (88)
11 (cervical adj3 (ADR or TDR or TDA or CADR or CTDR or CDR or CDA or ACDA)).ti,ab,kf. (673)
12 (cervical and arthroplast*).ti,ab,kf. (1955)
13 (discocerv or activC or activ C or ProDisc C or ProDiscC or Mobi C or Prestige ST or PrestigeST or Prestige LP or PrestigeLP or (NuVasive adj2 (PCM or disc*1 or disk*1)) or Kineflex C or KineflexC or cervicore or Cadisc C or CadiscC or Baguera C or BagueraC or Synergy Disc or (Axiomed adj2 (Freedom or disc*1 or disk*1)) or (Discover adj2 (DePuy or disc*1 or disk*1)) or ROTAIO or (Bryan adj2 (disc*1 or disk*1))).ti,ab,kf. (1044)
14 or/9-13 (3311)
15 Meta-Analysis/ or Meta-Analysis as Topic/ or exp Technology Assessment, Biomedical/ (271672)
16 Meta Analysis.pt. (83393)
17 (((systematic* or methodologic*) adj3 (review* or overview*)) or pooled analysis or published studies or published literature or hand search* or handsearch* or medline or pubmed or embase or cochrane or cinahl or data synthes* or data extraction* or HTA or HTAs or (technolog* adj (assessment* or overview* or appraisal*))).ti,ab,kf. (582230)
18 (meta analy* or metaanaly* or health technolog* assess*).mp. (397654)
19 Clinical Trials as Topic/ or Randomized Controlled Trials as Topic/ (459698)

20 (randomized controlled trial or controlled clinical trial).pt. (1068084)
21 trial.ti. (601318)
22 (randomi#ed or randomly or RCT$1 or placebo* or sham).ti,ab,kf. (2644535)
23 or/15-22 (3938556)
24 14 and 23 (1130)
25 exp Animals/ not Humans/ (14557058)
26 24 not 25 (835)
27 14 not 25 (2379)
28 26 use ppez,cleed (342)
29 27 use coch,cctr,clhta (326)
30 28 or 29 (668)
31 limit 30 to english language [Limit not valid in CDSR; records were retained] (556)
32 neck/ (76263)
33 exp cervical spine/ (32106)
34 neck pain/ (24829)
35 (cervical or neck).tw,kw. (828797)
36 or/32-35 (854228)
37 total dis*1 replacement.sh. (1223)
38 artificial dis*1 replacement.sh. (36)
39 dis*1 prosthesis.sh. (484)
40 ((disc*1 or disk*1) adj2 (replace* or prosthes#s or prosthetic* or arthroplast* or implant* or artificial*)).tw,kw,dv. (6955)
41 or/37-40 (7163)
42 36 and 41 (2753)
43 (cervical adj3 (replace* or prosthes#s or prosthetic* or arthroplast* or implant* or artificial)).hw. (480)
44 (neck adj2 (arthroplast* or ADR or TDR or TDA or CADR or CTDR or CDR or CDA or ACDA)).tw,kw,dv. (97)
45 (cervical adj3 (ADR or TDR or TDA or CADR or CTDR or CDR or CDA or ACDA)).tw,kw,dv. (671)
46 (cervical and arthroplast*).tw,kw,dv. (2020)
47 (discocerv or activC or activ C or ProDisc C or ProDiscC or Mobi C or Prestige ST or PrestigeST or Prestige LP or PrestigeLP or (NuVasive adj2 (PCM or disc*1 or disk*1)) or Kineflex C or KineflexC or cervicore or Cadisc C or CadiscC or Baguera C or BagueraC or Synergy Disc or (Axiomed adj2 (Freedom or disc*1 or disk*1)) or (Discover adj2 (DePuy or disc*1 or disk*1)) or ROTAIO or (Bryan adj2 (disc*1 or disk*1))).tw,kw,dv. (1147)
48 or/42-47 (3459)
49 Meta Analysis/ or "Meta Analysis (Topic)"/ or Biomedical Technology Assessment/ (266335)
50 (((systematic* or methodologic*) adj3 (review* or overview*)) or pooled analysis or published studies or published literature or hand search* or handsearch* or medline or pubmed or embase or cochrane or cinahl or data synthes* or data extraction* or HTA or HTAs or (technolog* adj (assessment* or overview* or appraisal*))).tw,kw. (605780)
51 (meta analy* or metaanaly* or health technolog* assess*).mp. (397654)
52 exp "controlled clinical trial (topic)"/ (136035)
53 randomized controlled trial/ or controlled clinical trial/ (1180250)
54 trial.ti. (601318)
55 (randomi#ed or randomly or RCT$1 or placebo* or sham).tw,kw. (2665966)
56 or/49-55 (3839756)
57 48 and 56 (1205)
58 (exp animal/ or nonhuman/) not exp human/ (10223450)

59 57 not 58 (1190)
60 limit 59 to english language [Limit not valid in CDSR; records were retained] (1086)
61 60 use emez (540)
62 31 or 61 (1096)
63 62 use ppez (336)
64 62 use coch (2)
65 62 use cctr (201)
66 62 use clhta (16)
67 62 use cleed (1)
68 62 use emez (540)
69 remove duplicates from 62 (637)

Economic Evidence Search

Search date: July 17, 2017

Databases searched: All Ovid MEDLINE, Embase, Cochrane Database of Systematic Reviews, CRD Health Technology Assessment Database, Cochrane Central Register of Controlled Trials, NHS Economic Evaluation Database

Database: EBM Reviews - Cochrane Central Register of Controlled Trials <June 2017>, EBM Reviews - Cochrane Database of Systematic Reviews <2005 to July 12, 2017>, EBM Reviews - Health Technology Assessment <4th Quarter 2016>, EBM Reviews - NHS Economic Evaluation Database <1st Quarter 2016>, Embase <1980 to 2017 Week 29>, All Ovid MEDLINE(R) <1946 to Present>

Search Strategy

––

1 neck/ (76405)
2 exp cervical vertebrae/ (38064)
3 neck pain/ (24893)
4 (cervical or neck).ti,ab,kf. (824426)
5 or/1-4 (852572)
6 total disc replacement/ (1207)
7 ((disc*1 or disk*1) adj2 (replace* or prosthes#s or prosthetic* or arthroplast* or implant* or artificial*)).ti,ab,kf. (6882)
8 or/6-7 (7029)
9 5 and 8 (2678)
10 (neck adj2 (arthroplast* or ADR or TDR or TDA or CADR or CTDR or CDR or CDA or ACDA)).ti,ab,kf. (88)
11 (cervical adj3 (ADR or TDR or TDA or CADR or CTDR or CDR or CDA or ACDA)).ti,ab,kf. (677)
12 (cervical and arthroplast*).ti,ab,kf. (1957)
13 (discocerv or activC or activ C or ProDisc C or ProDiscC or Mobi C or Prestige ST or PrestigeST or Prestige LP or PrestigeLP or (NuVasive adj2 (PCM or disc*1 or disk*1)) or Kineflex C or KineflexC or cervicore or Cadisc C or CadiscC or Baguera C or BagueraC or Synergy Disc or (Axiomed adj2 (Freedom or disc*1 or disk*1)) or (Discover adj2 (DePuy or disc*1 or disk*1)) or ROTAIO or (Bryan adj2 (disc*1 or disk*1))).ti,ab,kf. (1047)
14 or/9-13 (3318)
15 economics/ (253239)

16 economics, medical/ or economics, pharmaceutical/ or exp economics, hospital/ or economics, nursing/ or economics, dental/ (777133)
17 economics.fs. (403306)
18 (econom* or price or prices or pricing or priced or discount* or expenditure* or budget* or pharmacoeconomic* or pharmaco-economic*).ti,ab,kf. (761350)
19 exp "costs and cost analysis"/ (542686)
20 (cost or costs or costing or costly).ti. (234457)
21 cost effective*.ti,ab,kf. (272159)
22 (cost* adj2 (util* or efficacy* or benefit* or minimi* or analy* or saving* or estimate* or allocation or control or sharing or instrument* or technolog*)).ab. (176617)
23 models, economic/ (10751)
24 markov chains/ or monte carlo method/ (70579)
25 (decision adj1 (tree* or analy* or model*)).ti,ab,kf. (35072)
26 (markov or markow or monte carlo).ti,ab,kf. (111423)
27 quality-adjusted life years/ (33101)
28 (QOLY or QOLYs or HRQOL or HRQOLs or QALY or QALYs or QALE or QALEs).ti,ab,kf. (56797)
29 ((adjusted adj (quality or life)) or (willing* adj2 pay) or sensitivity analys*s).ti,ab,kf. (92015)
30 or/15-29 (2286558)
31 14 and 30 (223)
32 31 use ppez,coch,cctr,clhta (83)
33 14 use cleed (7)
34 32 or 33 (90)
35 limit 34 to english language [Limit not valid in CDSR; records were retained] (83)
36 neck/ (76405)
37 exp cervical spine/ (32163)
38 neck pain/ (24893)
39 (cervical or neck).tw,kw. (829729)
40 or/36-39 (855191)
41 total dis*1 replacement.sh. (1230)
42 artificial dis*1 replacement.sh. (36)
43 dis*1 prosthesis.sh. (489)
44 ((disc*1 or disk*1) adj2 (replace* or prosthes#s or prosthetic* or arthroplast* or implant* or artificial*)).tw,kw,dv. (6961)
45 or/41-44 (7170)
46 40 and 45 (2759)
47 (cervical adj3 (replace* or prosthes#s or prosthetic* or arthroplast* or implant* or artificial)).hw. (481)
48 (neck adj2 (arthroplast* or ADR or TDR or TDA or CADR or CTDR or CDR or CDA or ACDA)).tw,kw,dv. (97)
49 (cervical adj3 (ADR or TDR or TDA or CADR or CTDR or CDR or CDA or ACDA)).tw,kw,dv. (675)
50 (cervical and arthroplast*).tw,kw,dv. (2023)
51 (discocerv or activC or activ C or ProDisc C or ProDiscC or Mobi C or Prestige ST or PrestigeST or Prestige LP or PrestigeLP or (NuVasive adj2 (PCM or disc*1 or disk*1)) or Kineflex C or KineflexC or cervicore or Cadisc C or CadiscC or Baguera C or BagueraC or Synergy Disc or (Axiomed adj2 (Freedom or disc*1 or disk*1)) or (Discover adj2 (DePuy or disc*1 or disk*1)) or ROTAIO or (Bryan adj2 (disc*1 or disk*1))).tw,kw,dv. (1150)
52 or/46-51 (3466)
53 Economics/ (253239)
54 Health Economics/ or Pharmacoeconomics/ or Drug Cost/ or Drug Formulary/ (127222)

55 Economic Aspect/ or exp Economic Evaluation/ (417280)
56 (econom* or price or prices or pricing or priced or discount* or expenditure* or budget* or pharmacoeconomic* or pharmaco-economic*).tw,kw. (785615)
57 exp "Cost"/ (542686)
58 (cost or costs or costing or costly).ti. (234457)
59 cost effective*.tw,kw. (282935)
60 (cost* adj2 (util* or efficac* or benefit* or minimi* or analy* or saving* or estimate* or allocation or control or sharing or instrument* or technolog*)).ab. (177709)
61 Monte Carlo Method/ (57220)
62 (decision adj1 (tree* or analy* or model*)).tw,kw. (38779)
63 (markov or markow or monte carlo).tw,kw. (116347)
64 Quality-Adjusted Life Years/ (33101)
65 (QOLY or QOLYs or HRQOL or HRQOLs or QALY or QALYs or QALE or QALEs).tw,kw. (60543)
66 ((adjusted adj (quality or life)) or (willing* adj2 pay) or sensitivity analys*s).tw,kw. (111174)
67 or/53-66 (1934431)
68 52 and 67 (230)
69 limit 68 to english language [Limit not valid in CDSR; records were retained] (219)
70 69 use emez (123)
71 35 or 70 (206)
72 71 use ppez (58)
73 71 use coch (0)
74 71 use cctr (18)
75 71 use clhta (0)
76 71 use cleed (7)
77 71 use emez (123)
78 remove duplicates from 71 (142)

Grey Literature Search

Performed: May 30–June 5, 2017

Websites searched: HTA Database Canadian Repository, Alberta Health Technologies Decision Process reviews, Canadian Agency for Drugs and Technologies in Health (CADTH), Institut national d'excellence en santé et en services sociaux (INESSS), Institute of Health Economics (IHE), McGill University Health Centre Health Technology Assessment Unit, National Institute for Health and Care Excellence (NICE), Agency for Healthcare Research and Quality (AHRQ) Evidence-based Practice Centers, Australian Government Medical Services Advisory Committee, Centers for Medicare & Medicaid Services Technology Assessments, Institute for Clinical and Economic Review, Ireland Health Information and Quality Authority Health Technology Assessments, Washington State Health Care Authority Health Technology Reviews, ClinicalTrials.gov, Tufts Cost-Effectiveness Analysis Registry

Keywords used: Cervical, neck, disc, discs, disk, disks, arthroplasty, arthroplasties

Results: 9

Ongoing clinical trials: 29 (Clinicaltrials.gov)

Appendix 2: Randomized Controlled Trials Comparing C-ADR to Fusion Surgery

Table A1: Cervical Implant Devices Used in Randomized Controlled Trials of C-ADR Versus Fusion for Degenerative Disc Disease

Author, Year Location	Study Design Enrolment Period	No. in Randomized Groups, ADR/Fusion	Fusion Comparator Surgery	Target-Level Symptoms	Report Follow-Up	Outcomes			
						Operative Technical	Clinical and Functional	Radiologic	Patient reported outcomes
Bryan MoP Semiconstrained (19 reports)									
Heller et al,[20] 2009 United States	IDE NI RCT 30 sites May 2002 to October 2004	290/292	Allograft and Atlantis anterior cervical plating system	1-level radiculopathy or myelopathy	2 yr	AE, ROP	NDI, NS, OSR, VAS-N, VAS-A	ROM	SF-36-MCS, SF-36-PCS, RTW
Hacker,[87] 2005 United States	IDE NI RCT 1/31 sites	22/24			1 yr	BL, DYS; LOS, OT	NDI, VAS-A, VAS-N, OC		SAT
Anderson et al,[164] 2008 United States	IDE NI RCT 31 sites	241/221		1-level radiculopathy or myelopathy	2 yr	AE, ROP			
Kelly et al,[79] 2013 United States	IDE NI RCT 3/31 sites	66/69			2 yr				DA (NDI)
Sasso et al,[165] 2007 United States	IDE NI RCT 3 sites	56/59		1-level radiculopathy or myelopathy	2 yr	AS, ROP, DM	NDI, VAS-A, VAS-N	FSUH, ROM	SF-36-MCS, SF-36-PCS
Sasso et al,[166] 2007 United States	IDE NI RCT 3 sites	56/59			2 yr	AE, BL, DM, LOS, ROP	NDI, VAS-A, VAS-N	ROM	SF-36-MCS, SF-36-PCS
Sasso et al,[103] 2011 United States	IDE NI RCT 1/31 sites	22/26			2 yr		NDI	SA, FSUH, FSUA, ASD	
Powell et al,[101] 2010 United States	IDE NI RCT 1/31 sites	22/26			2 yr			ROM	
Sasso and Best,[102] 2008 United States		9/13			2 yr			KM	
Sasso et al,[104] 2008 United States	IDE NI RCT 31 sites	242/221			2 yr			KM, ROM, HO	
Sasso et al,[95] 2011 United States	IDE NI RCT 31 sites	181/138			4 yr	AE, ROP	NDI, NS, VAS-A, VAS-N, OSR	ROM	SF-36-PCS, RTW
Garrido et al,[167] 2011 United States	IDE NI RCT 1/31 sites	21/25			4 yr			HO	

Author, Year Location	Study Design Enrolment Period	No. in Randomized Groups, ADR/Fusion	Fusion Comparator Surgery	Target-Level Symptoms	Report Follow-Up	Outcomes			
						Operative Technical	Clinical and Functional	Radiologic	Patient reported outcomes
Garrido et al,[113] 2010 United States	IDE NI RCT 1/31 sites	21/26			4 yr	AE, BL, DM, OT	NDI, VAS-A, VAS-N		SF-36-MCS, SF-36-PCS
Smucker et al,[70] 2016 United States	IDE NI RCT 1/31 sites	43/46			5 yr	OT, DYS			
Sasso et al,[168] 2017 United States	IDE NI RCT 1/31 sites	24/23			10 yr	ROP	NDI, NS, VAS-A, VAS-N		
Cheng et al,[44] 2009 China	Non-IDE RCT 1 site December 2004 to April 2006	31/34	Iliac crest autograft and Orion anterior cervical plating	2-level radiculopathy or myelopathy	2 yr	AE, DYS	VAS-N, VAS-A, NDI, OC	ROM	SF-36-PCS
Cheng et al,[62] 2011 China	Non-IDE RCT 1 site December 2004 to September 2006	41/42		1-, 2-, or 3-level myelopathy	3 yr	BL, OT, ROP, AE, DYS	OC	ROM	RTW
Zhang et al,[55] 2012 China	Non-IDE RCT 3 sites May 2004 to May 2006	60/60	Allograft and anterior cervical plate	1-level radiculopathy or myelopathy	2 yr	BL, LOS, OT, ROP	NDI, VAS-A, VAS-N	HO, ROM, SA	
Donk et al,[46] 2017 Netherlands	Non-IDE RCT 1 site October 2003 to April 2010	50/47	Cervical interbody cage filled with autologous cancellous bone	1-level radiculopathy (arm pain)	9 yr			ASD	
Discover MoP Unconstrained (8 reports)									
Chen et al,[43] 2013 China	Non-IDE RCT 1 site November 2008 to October 2010	16/16	PEEK cage and SlimLoc anterior cervical plating	1-level radiculopathy or myelopathy kyphosis	2 yr		JOA, VAS, NDI	FSUH, ROM, SA	
Rozankovic et al,[51] 2016 Croatia	Non-IDE RCT 1 site October 2008 to June 2010	51/50	DuoCage Allograft	1-level radiculopathy or myelopathy NDI ≥ 30%	2 yr		NDI, VAS-A, VAS-N	HO	

Author, Year Location	Study Design Enrolment Period	No. in Randomized Groups, ADR/Fusion	Fusion Comparator Surgery	Target-Level Symptoms	Report Follow-Up	Outcomes			
						Operative Technical	Clinical and Functional	Radiologic	Patient reported outcomes
Skeppholm et al,[52] 2015 Sweden	Non-IDE RCT 3 sites April 2007 to May 2010	83/70	Iliac crest tricortical bone autograft and cervical anterior cervical plate by choice	1- or 2-level radiculopathy	2 yr	AE, DC, ROP	NDI, MU	ASD	EQ-5D, RTW
Skeppholm and Olerud,[71] 2013 Sweden		76/60			2 yr	DYS			
Skeppholm et al,[67] 2015 Sweden		28/NR			2 yr			ROM, SA	
Skeppholm et al,[75] 2017 Sweden		81/70			2 yr		HAD, VAS-A, VAS-N	KM, ROM	
Sundseth et al,[53] 2017 Norway	Non-IDE RCT 5 sites November 2008 to January 2013	73/70	Cervios PEEK interbody cage preloaded with chronOS	1-level radiculopathy NDI ≥ 30%	2 yr	AE, DYS, OT, ROP	NDI, NRS-A		RTW, SF-36-MCS, SF-36-PCS, EQ-5D and utility index
Sundseth et al,[76] 2016 Norway	Non-IDE RCT 1 of 5 sites	39/NR			2 yr	ROP	NDI, NRS-A, NRS-N	HO	SF-36-MCS, SF-36-PCS, EQ-5D
Kineflex-C MoM Semiconstrained (2 reports)									
Coric et al,[36] 2011 United States	IDE NI RCT 21 sites	136/133	Allograft and Slim-Loc anterior cervical screw/plate	1-level radiculopathy or myelopathy	2 yr	LOS, BL, AE, DYS, DYP, ROP	NDI, VAS NS	ROM, ASD	SAT, AT
Coric et al,[96] 2018		136/133			5 yr	AE, MISA, ROP	NDI, VAS-NAS, OTS	HO, ROM, ASD	
Mobi-C MoP Semiconstrained (16 reports)									
Hisey et al,[38] 2014 United States	IDE NI RCT 23 sites	164/81	Allograft and Slim-Loc or Atlantis anterior cervical plate system	1-level radiculopathy or myeloradiculopathy	2 yr	AE, BL, DYS, DYT, LOS, OT, ROP	NDI, NS, OSR, VAS-A, VAS-N	DC, ROM, HO, ASD	RTW, SAT, SF-12-PCS, SF-12-MCS

Author, Year Location	Study Design Enrolment Period	No. in Randomized Groups, ADR/Fusion	Fusion Comparator Surgery	Target-Level Symptoms	Report Follow-Up	Outcomes			
						Operative Technical	Clinical and Functional	Radiologic	Patient reported outcomes
Schrot et al,[81] 2014 United States	IDE NI RCT 23 sites	179/81		1-level radiculopathy or myelopathy	2 yr		CH		
Hisey et al,[88] 2015	IDE NI RCT 23 sites	164/81	Allograft and Slim-Loc or Atlantis anterior cervical plate system	1-level radiculopathy or myeloradiculopathy	4 yr		NDI, NS, OSR, VAS-A, VAS-N	ASD, ROM, HO	SAT, SF-12-PCS, SF-12-MCS
Hisey et al,[88] 2016 United States		128/55		1-level	5 yr	AE, ROP,	NDI, OSR, VAS-A, VAS-N,	ASD, HO, ROM	SF-12-PCS, SF-12-MCS
Radcliff et al,[84] 2017 United States		164/81		1-level	7 yr	AE, ROP	NDI, OSR, VAS-A, VAS-N,	ASD, HO, ROM	SAT, SF-12-PCS, SF-12-MCS
Two-level cervical degenerative disc disease									
Davis et al,[41] 2013 United States	IDE NI and superiority RCT 24 sites April 2006 to March 2008	225/105	Allograft and Slim-Loc anterior cervical plate	2-level radiculopathy or myelopathy	2 yr	LOS, BL, OT, ROP, AE, NS	OSR, NDI, VAS-N, VAS-A	DC, ROM, ASD	SAT, SF-12-PCS, SF-12-MCS
Davis et al,[93] 2015 United States		202/89		2-level	4 yr	AE, ROP	NDI, VAS-N, VAS-A, NS	ROM, HO, ASD	SAT, SF-12-PCS, SF-12-MCS, RTW
Radcliff et al,[83] 2016 United States		225/105		2-level	5 yr	AE, DYS, DYP, DC	NS, OSR	ASD, HO, ROM	
Radcliff et al,[84] 2017 United States		164/81		1- vs. 2-level	7 yr	AE, ROP	NDI, OSR, VAS-A, VAS-N	ASD, HO, ROM	SAT, SF-12-PCS, SF-12-MCS
Bae et al,[60] 2015 United States		1-level 169, 2-level 232		1- vs. 2-level C-ADR only	4 yr	AE, ROP	NDI, VAS-N, VAS-A, NS	ASD, HO, ROM	SAT, SF-12-PCS, SF-12-MCS
Jackson et al,[115] 2016 United States		1-level 179/81, 2-level 234/105		1- vs. 2-level	5 yr	ROP			
Zigler et al,[61] 2016 United States		1-level 81, 2-level 105		1- vs. 2-level fusion only	5 yr	BL, LOS	NDI, NS, VAS-A, VAS-N	ASD, ROM	SAT, SF-12-PCS, SF-12-MCS, RTW

Author, Year Location	Study Design Enrolment Period	No. in Randomized Groups, ADR/Fusion	Fusion Comparator Surgery	Target-Level Symptoms	Report Follow-Up	Outcomes			
						Operative Technical	Clinical and Functional	Radiologic	Patient reported outcomes
Liu et al,[80] 2016 United States		1-level 164/81, 2-level 225/105		1- vs. 2-level	5 yr		NDI (CH)	ROM, HO	
Schroeder et al,[85] 2017 United States		1-level 164/81, 2-level 225/105		1- or 2-level	5 yr	ROP	NDI, VAS-A, VAS-N		SAT, SF-36-MCS, SF-36-PCS
Zhang et al,[54] 2014 China	Non-IDE RCT 11 sites February 2008 to November 2009	55/56	Autologous iliac or clavicle bone graft and anterior cervical plate	1-level cervical spondylosis	4 yr	AE, DM, ROP, DYS, DYP	JOA, NDI, VAS	ROM, FSUA	
Hou et al,[47] 2016 China	Non-IDE RCT 11 sites January 2008 to July 2009	51/48	Autologous Iliac bone graft and intervertebral body cage	1-level	5 yr	BL, LOS, OT, ROP	JOA, NDI, VAS-A and VAS-N	ROM	
PCM MoP Semiconstrained (4 reports)									
Phillips et al,[56] 2013 United States	IDE NI and superiority RCT 24 sites January 2005 to December 2007	224/192	Tricorticol allograft and either CLSP or Slim-Loc anterior cervical plate	1-level radiculopathy or myelopathy NDI ≥30/100	2 yr	AE, ROP, DC	NS, OSR, VAS-A, VAS-N	ROM	SAT, SF-36-MCS, SF-36-PCS
Park et al,[99] 2011 United States	IDE RCT 23 sites NR	272/182			1 yr			KM, ROM, SA, FSUH, FSUA	
McAfee et al,[68] 2010 United States	IDE RCT 5 of 20 sites	151/100			2 yr	LOS, DYS, DYP			
Phillips et al,[89] 2015 United States	IDE NI and superiority RCT 24 sites	163/130			5 yr (final), 7 yr (interim)	AE, DYS, ROP	NDI, NS, VAS-A, VAS-N	HO, ROM, ASD	SF-36-MCS SF-36-PCS SAT
Prestige-ST MoM Semiconstrained (4 reports)									
Mummaneni et al,[21] 2007 United States	IDE NI RCT 32 sites October 2002 to August 2004	276/265	Allograft intra-disc spacer and Atlantis cervical anterior plate	1-level radiculopathy or myelopathy NDI ≥ 30	2 yr (interim)	AE, BL, DM, LOS, OT, ROP	OSR, VAS-N, VAS-A, NDI, NS	ROM	SF-36 MCS, SF-36 PCS, RTW, SAT
Burkus et al,[106] 2010 United States					2 yr (final), 5 yr (interim)	DYS, DYP, ROP	OSR, VAS-N, VAS-A, NDI	DC, ROM, HO, SA	RTW

Author, Year Location	Study Design Enrolment Period	No. in Randomized Groups, ADR/Fusion	Fusion Comparator Surgery	Target-Level Symptoms	Report Follow-Up	Outcomes			
						Operative Technical	Clinical and Functional	Radiologic	Patient reported outcomes
Burkus et al,[97] 2014 United States	FDA post-approval 5-yr extension RCT 31 sites	212/183			7 yr (final)	AE, DYS, DYP, ROP	OSR, VAS-N, VAS-A, NDI, NS	SA, FSUH, HO, ASD	SF-36-PCS, RTW
Prestige-LP MoM Semiconstrained (6 reports)									
Gornet et al,[59] 2015 United States	IDE NI CPMC study, 20 sites January 2005 to November 2005	280/265	Cortical ring allograft and anterior cervical plate	1-level radiculopathy or myelopathy	2 yr	AE, ROP	NDI, NRS-A, NRS-N, NS, OSR	FSUH, ROM	SAT, SF-36-MCS, SF-36-PCS RTW
Gornet et al,[77] 2016 United States		280/265		1-level radiculopathy or myelopathy	7 yr	AE, BL, LOS	NDI, NRS-A, NRS-N, NS, OSR	FSUH, ROM	SAT, SF-36-MCS, SF-36-PCS
Cincu et al,[45] 2014 Spain	Non-IDE RCT 1 site March 2004 to June 2005	25/28	Solis cage	1- and 2-level (n = 6) radiculopathy or myelopathy	7 yr	AE, LOS, ROP		ROM	RTW
Pandey et al,[50] 2017 India	Non-IDE RCT 1 site July 2012 to April 2014	17/17	Iliac crest autologous bone graft and anterior cervical locking plate	1-level radiculopathy, myelopathy	1 yr		JOA, NDI, OC, VAS-AN, NS	ASD, SA, ROM	
Two-level cervical degenerative disc disease									
Gornet et al,[42] 2017 United States	IDE NI and superiority RCT 30 sites June 2006 to November 2007 (surgeries performed)	226/230	Cortical ring allograft and Atlantis anterior cervical plate	2-level radiculopathy or myelopathy	2 yr	BL, DYP/DYS, LOS, OT, ROP	GS, NDI, NS, OSR, FE	FSUH, ROM, HO	SF-36-MCS, SF-36-PCS, RTW, SAT
Lanman et al,[94] 2017 United States					7 yr	AE, DYS, DYP, ROP	GS, NDI, NS, OSR, VAS-A, VAS-N	FSUH, HO, ROM	SAT, SF-36-MCS, SF-36-PCS

Author, Year Location	Study Design Enrolment Period	No. in Randomized Groups, ADR/Fusion	Fusion Comparator Surgery	Target-Level Symptoms	Report Follow-Up	Outcomes			
						Operative Technical	Clinical and Functional	Radiologic	Patient reported outcomes
ProDisc-C MoP Semiconstrained (17 reports)									
Murray et al,[22] 2009 United States	IDE NI and superiority RCT 13 sites August 2003 to October 2004 (surgeries performed)	106/103	Allograft bone spacer of surgeons' choice and cervical anterior plating	1-level radiculopathy or myelopathy NDI ≥15/50 (30%)	2 yr	AE, BL, LOS, OT, ROP	NS, VAS-N, VAS-A, OSR, NDI, MU	ROM, FUSH, HO	RTW, SAT, SF-36-MCS, SF-36-PCS
Anakwenze et al,[109] 2009 United States		89/91			2 yr			SA	
Auerbach et al,[100] 2011 United States		111/117			2 yr			ROM	
Murrey et al,[57] 2008 United States	IDE NI RCT 2 of 13 multicentre sites	44/43			2 yr	AE, OT, BL	VAS-N, VAS-A, NDI	ROM	SAT, SF-36-MCS
Delamarter et al,[82] 2010 United States	IDE NI and superiority RCT with CA group	103/136 (CA)/106			2 yr, 4 yr (interim)	BL, LOS, OT, ROP	NS, VAS-N, VAS-A, OSR, NDI, MU	ROM, HO	SAT, SF-36-MCS, SF-36-PCS
Park et al,[107] 2010 United States	IDE NI RCT 13 multicentre sites	164		1 level	2 yr			ROM	
Segebarth et al,[69] 2010 United States	IDE NI RCT 2 of 13 multicentre sites	45/42		1 level	1 yr	DYS			
Peng et al,[110] 2009 United States	IDE NI RCT 13 multicentre sites	166 (102 RCT and 64 CA)		1 level	2 yr		NDI, VAS-A, VAS-N	FSUH, ROM	
Kelly et al,[108] 2011 United States		100/99		1 level	2 yr			ROM	
Delamarter and Zigler,[117] 2013 United States		72/99 (73%) 61/96 (64%)		1 level	5 yr	ROP			
Zigler et al,[91] 2013 United States				1 level	5 yr	AE, ROP	NS, VAS-A, VAS-N, NDI	DC, HO, ROM	SAT, SF-36-MCS, SF-36-PCS

Author, Year Location	Study Design Enrolment Period	No. in Randomized Groups, ADR/Fusion	Fusion Comparator Surgery	Target-Level Symptoms	Report Follow-Up	Outcomes			
						Operative Technical	Clinical and Functional	Radiologic	Patient reported outcomes
Janssen et al,[92] 2015 United States		103/106		1 level	7 yr	AE, ROP, MU	NS, NDI	HO, ROM	SAT, SF-36-MCS, SF-36-PCS
Loumeau et al,[169] 2016 United States	IDE NI RCT 1 of 13 multicentre sites	22/22		1 level	7 yr	ROP	VAS-A, VAS-N, NDI	DC, ROM	SAT, SF-36-MCS, SF-36-PCS
Nabham et al,[48] 2007 Germany	Non-IDE RCT 1 site April 2004 to May 2005	16/17	Solis PEEK cage and anterior titanium alloy plate	1-level radiculopathy (symptomatic soft disc herniation)	6 mo		VAS-A, VAS-N	ROM	
Nabhan et al,[72] 2007 Germany	Non-IDE RCT 1 site April 2004 to May 2005	25/24			1 yr		VAS-A, VAS-N	ROM	
Nabhan et al,[73] 2007 Germany	Non-IDE RCT 1 site April 2004 to May 2005	25/24			3 yr		VAS-A, VAS-N	ROM	
Nabhan et al,[49] 2011 Germany	Non-IDE RCT 1 site January 2006 to August 2001 (surgeries performed)	10/10			1 yr			ROM	
Secure-C MoP Semiconstrained (1 report)									
Vaccaro et al,[40] 2013 United States	IDE NI and superiority RCT 18 sites	151/140	Structural allograft and Assure anterior cervical plate	1-level radiculopathy, myelopathy	2 yr	AE, BL, DC, DM, DYS, DYP, OT, LOS, ROP	NDI, NS, OSR, VAS-A, VAS-N	ROM	RTW, SAT, SF-36-MCS, SF-36-PCS
Mixed-Device RCTs (8 reports)									
Blumenthal et al,[116] 2013 United States	6 IDE RCTs (devices NR) 1 site	84/52	Reoperation rate	1- or 2-level radiculopathy or myelopathy	2 yr	ROP			

Author, Year Location	Study Design Enrolment Period	No. in Randomized Groups, ADR/Fusion	Fusion Comparator Surgery	Target-Level Symptoms	Report Follow-Up	Outcomes			
						Operative Technical	Clinical and Functional	Radiologic	Patient reported outcomes
Coric et al,[65] 2010 United States	3 IDE RCTs (Bryan, Discover, Kineflex-C) 1 site	57/41	Long-term follow-up outcomes	1- or 2-level radiculopathy or myelopathy	2 yr	AE, ROP	NDI, VAS-N	ROM, HO	
Coric et al,[66] 2013 United States	2 IDE RCTs (Bryan, Kineflex-C) 1 site	41/33	Long-term follow-up outcomes	1-level radiculopathy	4 yr	AE, ROP	NDI, VAS-N	ROM, HO	
Jawahar et al, [111] 2010 United States	3 IDE RCTs (Kineflex-C, Mobi-C, Advent) 1 site	59/34	Adjacent-segment degeneration	1- or 2-level radiculopathy, myelopathy	4 yr		NDI, VAS-N	ASD	
Hackeret al,[118] 2013 United States	2 IDE RCTs (Bryan, Prestige-LP) 1 site	47/47	Late complications (>4 yr)	1-level radiculopathy or myelopathy	4 yr	AE			
Nunley et al,[112] 2012 United States	3 IDE RCTs (devices ND) 2 sites	120/62	Symptomatic ASD	1- or 2-level radiculopathy or myelopathy	4 yr	ROP	NDI, VAS-A VAS-N	ASD	SF-12
Riew et al,[63] 2008 United States	2 IDE RCTs sites (Bryan, Prestige-ST) Multiple sites, subset of myelopathy	106/93	Efficacy for myelopathy	1-level myelopathy	2 yr	AE, ROP	NDI, NG, VAS-A, VAS-N		SAT, SF-36-MCS, SF-36-PCS
Steinmetz et al,[78] 2008 United States	2 IDE RCTs (Bryan, Prestige-ST) 63 sites	47/46	Mobilization in Workers' Compensation population	1-level radiculopathy, myelopathy NDI ≥30	2 yr		NDI, VAS-A, VAS-N		RTW, SF-36-MCS, SF-36-PCS

Abbreviations: ADR, artificial disc replacement; AE, adverse event; ASD, adjacent-segment disease; AS, adjacent segment; AT, activity level; BL, blood loss; CA, continued access; C-ADR, cervical artificial disc replacement; CH, cervicogenic headache; CPMC, controlled prospective multicenter; DA, driving ability; DC, device condition; DM, discharge management; DYP, dysphonia; DYS, dysphagia; EQ-5D, European Quality of Life questionnaire for measuring quality of life in 5 dimensions; FDA, U.S. Food and Drug Administration; FE, foraminal encroachment; FSUA, functional spinal unit angle; FSUH, functional spinal unit height; GS, gait success; HAD, hospital anxiety and depression scale; HO, heterotrophic ossification; IDE, investigational device exemption; JOA, Japanese Orthopedic Association score; KM, kinematics; LOS, length of stay; MISA, metal ion serum analysis; MoM, metal on metal; MoP, metal on plastic; MU, medication use; NG, Nurich grade; ND, not defined; NDI, neck disability index; NDI (CH), Neck Disability Index, cervicogenic headache; NG, Nurich grade NI, noninferiority; NR, not reported; NRS-A, numerical rating scale arm pain; NRS-N, numerical rating scale neck pain; NS, neurological status; OC, Odom's criteria; OSR, overall success rate; OT, operative time; PEEK, polyetheretherketone; RCT, randomized controlled trial; ROM, range of motion; ROP, reoperation; RTW, return to work; SA, sagittal alignment; SAT, satisfaction; SF36-MCS, 36-item Short-Form Health Survey—Mental Component Summary; SF36-PCS, 36-item Short-Form Health Survey—Physical Component Summary, VAS-A, visual analogue scale arm pain; VAS-N, visual analogue scale neck pain; VAS-NAS, visual analogue scale neck and arm pain

Appendix 3: Secondary Surgeries at Index and Adjacent Cervical Levels in Longer-Term Follow-Up

Table A2: Secondary Surgeries in Longer-Term Follow-Up of C-ADR– Versus Fusion–Treated One-Level Cervical Degenerative Disc Disease

C-ADR Device Author, Year Location Study Design	Follow-Up N (C-ADR/ Fusion)	Index- and Adjacent-Level Secondary Surgeries C-ADR	Index- and Adjacent-Level Secondary Surgeries Fusion	Conclusion
Bryan				
Garrido et al,[113] 2010 United States IDE RCT, 1 of 31 sites	4 yr N = 21/25	N = 1, 1 for ASD	N = 6, 3 for ASD, 1 for nonadjacent ASD, 2 for pseudarthrosis (facet neurotomy and posterior cervical fusion)	In this small study, fewer surgeries were performed in C-ADR group
Sasso et al,[95] 2011 United States IDE RCT, 31 sites	4 yr N = 242/221 (95%/75%)	Index-level 4-yr secondary surgery rate 3.7% (n = 9, 3 surgeries at > 2 yr) • Adjacent-level 4-yr surgery rate 4.1% (n = 10, 4 surgeries > 2 yr) • Other cervical surgery 0.4% (n = 1, within 2 yr) • Implant removal (n = 1)	Index-level 4-yr secondary surgery rate 4.5% (n = 10, 2 surgeries > 2 yr) • Adjacent-level 4-yr surgery rate 4.1% (n = 9, 4 surgeries > 4 yr) • Other cervical surgeries 1.4% (n = 3, all within 2 yr) • Hardware removal (n = 1)	Rates of secondary surgeries at index and adjacent cervical level were low and not significantly different between surgery groups
Sasso et al,[168] 2017 United States IDE RCT, 1 of 31 sites	10 yr N = 24/23 (86%/92%)	Cumulative 7-yr surgery rate 9% (n = 2, 1 adjacent and 1 nonadjacent) • None required reoperation at index level • 3 patients had been converted intra-operatively to fusion (too much degeneration, unable to prepare disc space, inadequate visualization)	Cumulative 7-yr surgery rate 32% (n = 8, 11 procedures) • Adjacent-level surgeries (n = 6) • Secondary surgeries included 1-level adjacent fusion (n = 4), single nonadjacent fusion (n = 1), posterior laminectomy at non-index and nonadjacent levels (n = 1), multiple fusions at adjacent levels (n = 4), posterior fusion for pseudarthrosis (n = 1)	Overall 7-yr surgical device survivorship (no secondary operations) was higher but not significantly for C-ADR than for fusion group (90.9% vs. 68%, *P* = .056)
Donk et al,[46] 2017 Netherlands Non-IDE RCT, 1 site	9 yr (range 5.6–12.2 yr) N = 50/47	Indication for index-level secondary surgery—dorsal foraminotomy (n = 1)—was recurrent symptoms from arthrotic-related compression	Index level secondary surgery (n = 1) indication was recurrent symptoms due to arthrotic related compression and underwent dorsal foraminotomy	At long-term follow-up, adjacent-level surgery was more common in fusion group

C-ADR Device Author, Year Location Study Design	Follow-Up N (C-ADR/ Fusion)	Index- and Adjacent-Level Secondary Surgeries C-ADR	 Fusion	Conclusion
		Adjacent-level surgery (n = 0)	• Adjacent-level surgery 10.6% (n = 5) • No patients with symptomatic ASD refused surgery or were treated conservatively. Median time from index surgery to reoperation for ASD was 4.1 yr (0.5–7.2 yr)	
KineFlex-C				
Coric et al,[96] 2018 United States IDE RCT, 21 sites	5 yr 68%/62%	• 5-yr index-level secondary surgeries • 8.1% (n = 11)	• 5-yr index-level secondary surgery • 8.3% (n = 11)	Index-level surgeries at 5 yr were similar between groups
		Adjacent-level NR	Adjacent-level NR	
Mobi-C				
Zhang et al,[54] 2014 China Non-IDE RCT, 11 sites	4 yr N = 55/56	4-yr index-level secondary surgery (n = 0)	4-yr adjacent-level surgery (n = 0)	First follow-up report at 4 yr on small study reported no secondary surgeries at index site for either surgery group and, at adjacent level, only in fusion group
		4-yr adjacent-level surgery (n=0)	4-yr adjacent-level surgery 7.1% (n = 4: 1 fusion, 1 Mobi-C disc, and 2 posterior cervical open-door laminoplasties)	
Hou et al,[47] 2016 China Non-IDE RCT, 11 sites	5 yr N = 56/51 91%/94%	• 5-yr adjacent-level surgery rate 2% (n = 1, for superior-level degeneration) • Surgeries occurred > 4 yr after surgery	• 5-yr adjacent-level surgery rate 14.6% (n = 7), indications NR • Surgeries occurred at 2 yr (n = 1), 3 yr (n = 2), and 4 yr (n = 4) after surgery	Second follow-up report at 5 yr on study group showed higher rate of adjacent-level secondary surgery in fusion group (*P* = .049)
Hisey et al,[88] 2015 United States IDE RCT, 23 sites	4 yr N = 164/81 78%/68%	• 4-yr index-level secondary surgery rate 3.0% (n = 5) • Surgeries included cervical laminectomy without implant removal (n = 1) and 4 surgeries with implant removal and fusion (n = 4) • Indications for index-level surgeries included nerve impingement, oversized disc for space, HO with cord impingement, bothersome feeling with movement, kyphosis due to inadequate endplate fixation	• 4-yr index-level secondary surgery rate 9.9% (n = 8) • Indications for surgery included symptomatic pseudarthrosis (n = 5), malpositioned screws (n = 1), cervical stenosis (n = 1)	• At first follow-up report, at 4 yr, secondary surgery rate for index-level surgeries was significantly lower in C-ADR group (3.6% vs. 12.3%, *P* < .05) • Only 1 surgery in each group was adjacent level for ASD

C-ADR Device Author, Year Location Study Design	Follow-Up N (C-ADR/ Fusion)	Index- and Adjacent-Level Secondary Surgeries C-ADR	Fusion	Conclusion
		4-yr adjacent-level surgery (n = 1) indication was herniated disc 17 mo after surgery	4-yr adjacent-level surgery (n = 1) indication was for ASD	
Hisey et al,[88] 2016 United States IDE RCT, 23 sites	5 yr N = 164/81 86%/79%	5-yr index-level secondary surgery rate 4.9% (n = 8, 3 for ASD)	5-yr index-level secondary surgery rate 17.3% (n = 14, 6 for ASD)	• At second follow-up report, 5-yr index-level secondary surgery rate was significantly lower in C-ADR group (4.9% vs. 17.3%, $P <$ .01) • Adjacent-level surgery rates were lower for C-ADR group (1.8% vs. 3.7%)
		5-yr adjacent-level surgery rate for ASD 1.8% (n = 3)	5-yr adjacent-level surgery rate for ASD 3.7% (n = 3)	
Jackson et al,[115] 2016 United States IDE RCT, 23 sites	5 yr N = 179/81 (included training courses) 86%/79%	Overall 5-yr surgery rate 4.5% (n = 8, 2 patients required multiple procedures): surgeries involved 4 index, 2 adjacent, 2 index and adjacent	Overall 5-yr secondary surgery rate 17.3% (n = 14, 1 patient required multiple procedures): surgeries involved 5 index, 4 adjacent, 5 index and adjacent	• At 5 yr, overall secondary surgery rate was significantly higher for fusion than for C-ADR (4.5% vs. 17.3%, P = .0012) • Index-level secondary surgery rate was significantly higher for fusion group (12.3% vs. 3.4%, P = .0097) but were still higher but not significantly higher when censoring index-level procedures for secondary ASD in fusion group (8.6% vs. 3.4%, P = .1194) • Adjacent-level surgery rate was significantly lower for C-ADR group (2.2% vs. 11.1%, P = .0043)
		Index-level secondary surgery rate 3.4% (n = 6): most common indication was radiculopathy	• Index-level secondary surgery rate 12.3% (n = 10). Most common indications were radiculopathy, neck pain, and pseudarthrosis • [Three surgeries involved plate removal at index level for adjacent-level ASD. Censoring these events, index-level surgery rate was 8.6%]	
		• Adjacent-level secondary surgery rate 2.2% (n = 4) • O Most common indications for adjacent-level secondary surgery were radiculopathy and adjacent-level disease. 2 patients had herniated discs at 19 and 52 mo after surgery	• Adjacent-level secondary surgery rate 11.1% (n = 9) • Most common indications for adjacent-level secondary surgery were adjacent level disease and neck pain. Herniated discs occurred in 6 patients at 14, 20, 34, 49.5, 52, and 59 mo after surgery	
Radcliff et al,[84] 2017	7 yr N = 164/81	• 7-yr index secondary surgery rate 3% (5/164) • Surgeries included removals (n = 4) and re-operations (n = 1)	• 7-yr index-level secondary surgery rate 12.3% (10/81) • Surgeries included removals (n = 7) and supplemental fixations (n = 3). Indication for 5 cases was pseudarthrosis • Adjusted index-level surgery rate 6.2% (5/81) (subtracting surgeries at index site involving plate removals to treat adjacent level ASD)	• Index-level secondary surgery rate was lower for C-ADR group (3% vs. 6.2%) but not significantly ($P_{KM\ log\text{-}rank}$ = .219) • Adjacent-level surgery was significantly lower for C-ADR group (3.7% vs. 13.65, $P_{KM\ log\text{-}rank}$ = .002). Disc herniation was most common indication for both surgery groups

C-ADR Device Author, Year Location Study Design	Follow-Up N (C-ADR/ Fusion)	Index- and Adjacent-Level Secondary Surgeries C-ADR	 Fusion	Conclusion
		• Adjacent-level surgery rate 3.7% (6/164) • Indications for disc herniation (n = 4)	• Adjacent-level surgery rate 13.6% (11/81) • Indications for disc herniation (n = 8)	
PCM				
Phillips et al,[89] 2015 United States IDE RCT, 24 sites	5 yr, 7-yr N = 224/192 5 yr 94%/98% 7 yr 75%/70%	5-yr overall surgery rate 8.1% (n = 17) 7-yr overall surgery rate 8.5% (n = 18) Secondary surgery (1 for ASD) ≤ 2 yr n = 11 2–5 yr n = 6 > 5 yr n = 1 Implant removals (n = 14, 6 > 2 yr for pain and 4 for device migration)	5-yr overall surgery rate 12.0% (n = 22) 7-yr overall surgery rate 13.0% (n = 24) Secondary surgery (19 for ASD) ≤ 2 yr n = 10 2–5 yr n = 12 > 5 yr n = 2 for ASD Hardware plate removal (n = 24,14 > 2 yr, 13 for ASD, 1 for nonunion)	• Overall surgery rate was not significantly different between study groups at 5-yr (P = .237) or 7-yr (P =.190) follow-up • Kaplan Meier survival curves comparing overall surgeries between study groups were not significant (P = .123), although curves began to separate after 3 yr • During 7-yr follow-up, surgeries for ASD were performed less often for C-ADR than fusion group (1 surgery vs. 13 surgeries)
Prestige-ST				
Burkus et al,[106] 2010 United States IDE RCT 31 sites (of original 32 sites)	5 yr N = 276/265 3 yr 71%/60% 5 yr 52%/49%	• Index-level 5-yr secondary surgery rate 4.0% (n = 11) • Index surgeries included revisions (n = 0), implant removals (n = 7), supplemental fixation (n = 0), and re-operation (n = 4) 5-yr adjacent-level surgery rate 2.9% (n = 8, 11 surgeries)	• Index-level 5-yr secondary surgery rate 10.9% (n = 29) • Index surgeries included revision (n = 5), implant removal (n = 13), supplemental fixation (n = 9, 5 without BGS), and re-operation (n = 2) 5-yr adjacent-level surgery rate 4.9% (n = 13, 16 surgeries)	• In this first follow-up report, index-level secondary surgery rate was significantly lower in C-ADR group (4.0% vs. 10.9%, P = .001) • Adjacent-level secondary surgery rate was lower, but not significantly, for C-ADR group (2.9% vs. 4.9%, P = .376)
Burkus et al,[106] 2014 United States IDE RCT, 31 sites	5 yr 80%/72% 7 yr 77%/69%	• Index-level 7-yr secondary surgery rate 4.8% (n = 11, 9 within 2 yr) • Index surgeries included revisions (n = 0), implant removals (n = 8), supplemental fixation (n = 0), and re-operation (n = 4) Adjacent-level 7-yr surgery rate 4.6% (n = 11)	• Index-level 7-yr secondary surgery rate 13.7% (n = 29, 19 within 2 yr) • Index-level surgeries included revisions (n = 5), implant removals (n = 8), supplemental fixation (n = 12, 7 with BGS), and re-operation (n = 4) Adjacent-level 7-yr surgery rate 11.9% (n = 24)	• In second follow-up report, 7-yr secondary surgery rate at index level (4.8% vs. 13.7%, P < .001) and at adjacent levels (4.6% vs. 11.9%, P = .008) was significantly lower for C-ADR group • Survival curves comparing overall adjacent-level surgeries up to 7 yr were significantly different between study groups, with lower surgery rates for C-ADR group ($P_{KM\ log\text{-}rank}$ = .008)

C-ADR Device Author, Year Location Study Design	Follow-Up N (C-ADR/ Fusion)	Index- and Adjacent-Level Secondary Surgeries C-ADR	Index- and Adjacent-Level Secondary Surgeries Fusion	Conclusion
Prestige-LP				
Gornet et al,[77] 2015 United States 20 site IDE matched cohorts	7 yr N = 280/265 82%/76%	• Index-level 7-yr secondary surgery rate 6.4% (n = 18) • Of 20 surgeries, 4 > 2 yr • Secondary surgeries included revisions (n = 1), removals (n = 14), supplemental fixation (n = 2), and reoperations (n = 3)	• Index level 7-yr secondary surgery rate 10.9% (n = 29) • Of 31 surgeries, 5 > 2 yr) Secondary surgeries included revisions (n = 5), removals (n = 8), supplemental fixation (n = 9, 5 without BGS), and reoperations (n = 4)	• Of index-level secondary surgeries, only supplemental fixation was significantly lower for C-ADR than for fusion group based on Cox proportional hazards model and adjusted for propensity scores ($P_{log\text{-}rank}$ = .004) • Adjacent-level 7-yr surgery rate, stand-alone or combined adjacent and index levels, was not significantly different between groups (9.6% vs. 8.3%)
		7-yr adjacent-level surgery rate 9.6% (n = 27) (stand-alone procedure or in conjunction with index level)	7-yr adjacent-level surgery rate 8.3% (n = 22) (stand-alone procedure or in conjunction with index level)	
Cincu et al,[45] 2014 Spain Non-IDE RCT, 1 site	7 yr N = 25/28	Secondary surgery site and follow-up time unclear	Secondary surgery site and follow-up time unclear	
ProDisc-C				
Delamarter et al,[82] 2010 United States IDE RCT 13 sites	4 yr N = 103/106 2 yr 98%/95% 4 yr 63%/46%	Overall 2-yr secondary surgery rate 1.9% (n = 2)	Overall 2-yr secondary surgery rate 8.5% (n = 9)	• At 4 yr, overall secondary surgery rate was significantly lower in C-ADR group (2.9% vs. 11.3%, *P* = .029) • Adjacent-level surgery rate was lower but not significantly lower in C-ADR group (0.9% vs. 5.7%) than in fusion group
		Overall 4-yr secondary surgery rate 2.9% (n = 3), 1 (0.9%) at adjacent level	Overall 4-yr secondary surgery rate 11.3% (n = 12, 6 at adjacent levels (5.7%): 3 at 1 adjacent level and 3 at both adjacent levels)	
Zigler et al,[91] 2013 United States IDE RCT, 13 sites	5 yr 73%/64%	Overall 5-yr secondary surgery rate 2.9% (n = 3, 3 procedures)	Overall 5-year secondary surgery rate 11.3% (n = 12, 16 procedures	In second follow-up report, 5-yr overall secondary surgery rate was significantly less for C-ADR than for fusion group (2.9% vs. 11.3%, *P* = .029)
Delamarter et al,[117] 2013 United States IDE RCT, 13 sites	5 yr 73%/64%	Index-level surgery rate 0.9% (n = 1), indications for ongoing pain	Index-level surgery rate 7.5% (n = 8), indications for pseudarthrosis (n = 6), foraminiferous stenosis (n = 1), and dysphagia-related plate shift (n = 1)	• In third follow-up report detailing overall re-operation rates at 5 yr, probability of not having secondary surgery (based on Anderson-Gill survival analysis) was significantly lower for C-ADR group (97.1% vs. 85.5%, *P* = .0079) • Rate of adjacent-level surgeries for ASD was significantly lower for C-ADR group (1.9% vs. 8.5%, *P* = .029)
		• Adjacent level surgeries for ASD 1.9% (n = 2) • Implant removals (n = 2)	• Adjacent-level surgery for ASD 5.7% (n = 6, 8 surgeries) • Fusion hardware plate removals (n = 8)	
		Overall 7-yr secondary surgery rate 7% (n = 7, 7 procedures)	Overall 7-yr secondary surgery rate 18% (n = 19, 30 procedures)	

C-ADR Device Author, Year Location Study Design	Follow-Up N (C-ADR/ Fusion)	Index- and Adjacent-Level Secondary Surgeries C-ADR	Fusion	Conclusion
Janssen et al,[92] 2015 United States IDE RCT, 13 sites	7 yr 92%/92%	Index-level secondary surgery rate 0.9% (n = 1)	• Index-level (only) secondary surgery rate 7.8% (n = 8) • 8 index-only procedures included 3 supplemental fixation, 1 plate removal, and 4 revisions)	• In third follow-up report, overall 7-yr secondary surgery rate was significantly lower in C-ADR group (7% vs. 18%, P = .02) • Incidence rate ratio for secondary surgery at index level was significantly greater for fusion group (3.09, 95% CI = 1.7–8.45, Pw = .0099 [Cox proportional hazard rate with Anderson-Gill model] • Incidence rate ratio for secondary surgery at adjacent level was also greater for fusion group 3.64 (P = .0103)
		• Adjacent-level surgery 5.8% (n = 6, 6 surgeries) • Procedures included 5 implant removals and conversion to fusion and 1 foraminotomy and posterior fusion)	• Adjacent-level only surgery (n = 6, 11 surgical procedures) • Procedures included 9 adjacent level fusions, 1 posterior fusion and 1 adjacent level Pro-Disc implant	
		Overall secondary surgeries (n = 0)	Overall secondary surgery rate 27% (n = 6)	
Loumeau et al, [169] 2016 United States IDE RCT 1 of 13 sites	7 yr N = 22/22 82% 86%	Index-level surgery (n = 0)	Index-level secondary surgeries (n = 2) and indications of pseudarthrosis (n = 2)	• This 7-yr follow-up at 1 site of 13 multi-site RCTs reported no secondary surgeries at index or adjacent sites in C-ADR group • There were more secondary surgeries in fusion group at both index and adjacent levels
		Adjacent surgery rate (n = 0)	Adjacent-level surgery rate 18% (n = 4) and for indications of ongoing pain	
Mixed-Device RCTs				
Coric et al,[66], 2013 1 site of 2 IDE RCTs [Bryan, Kineflex-C]	4 yr N = 33/41	Index-level secondary surgery 2.4% (n =1)	Index-level secondary surgery (n = 0)	At 4-yr follow-up, only a few secondary surgeries were performed in either study group
		Adjacent-level surgery 4.9% (n = 2) All surgeries were cervical laminoforaminotomies	Adjacent-level surgery 3.0% (n = 1) extended fusion to a second level	

Abbreviations: BGS, bone graft stimulation; HO, heterotopic ossification; IDE, investigational device exemption; ND, not defined; NR, not reported; P_B, Bayesian probability; Pw, Wald probability; rASD, radiologic adjacent-segment disease.

[a]Secondary surgery categories included revision, removal, elective removal, supplemental fixation, and reoperation. Revision was any procedure that adjusted or in any way modified original implant. Removal was any procedure that removed 1 or more components of original implant configuration without replacement using same type of device. Supplemental fixation was any procedure in which additional spine procedures were performed.

Table A3: Secondary Surgery in Longer-Term Follow-Up of C-ADR– Versus Fusion–Treated Two-Level Degenerative Disc Disease

C-ADR Device Author, Year Study Design	Duration of Follow-Up Randomized Surgery N (C-ADR/ Fusion)	Secondary Surgeries		Conclusion
		C-ADR	Fusion	
Mobi-C				
Davis et al, [93] 2005 24 sites IDE RCT	4 yr N = 225/105	• Index-level 4-yr secondary surgery rate 4.0% (n = 9, 10 procedures) • Indications for secondary surgeries included stenosis, device migration, poor endplate fixation, and persistent neck or shoulder pain	• Index-level 4-yr secondary surgery rate 15.2% (n =16, 18 procedures) • Main indication for surgery was pseudarthrosis	• At 4 yr, index-level surgery rate was significantly lower for C-ADR group (4% vs. 15.2%, *P* < .0001) • There were no adjacent-level surgeries in either study group
		Adjacent-level surgeries (n = 0)	Adjacent-level surgeries (n = 0)	
Radcliff et al, [83] 2016 24 sites IDE RCT	5 yr N = 225/105	Overall 5-yr secondary surgery rate 7.1% (n = 16)	Overall 5-yr secondary surgery rate 21% (n = 22)	• Overall 5-yr secondary surgery rate was significantly lower in C-ADR group (7% vs. 21%, *P* = .0006) • 5-yr index surgery rates were significantly lower for C-ADR group- (4.3% vs. 16.2%, *P* = .0003) • 5-yr adjacent-level surgeries were lower (3.1% vs. 11.4%) in C-ADR group but not significantly • 5-yr overall secondary operation-free survivorship was significantly higher after C-ADR ($P_{log\text{-}rank}$ = .0002)
		Index-level 5-yr secondary surgery rate 4.3% (n = 9)	Index-level 5-yr secondary surgery rate 16.2% (n = 17)	
		Adjacent-level 5-yr surgery rate 3.1% (n = 7)	Adjacent-level 5-yr surgery rate 11.4% (n = 12)	
Jackson et al, [115] 2016 24 sites IDE RCT	5 yr N = 234/105	Overall 5-yr secondary surgery rate 7.3% (n = 17, 19 procedures)	Overall 5-yr secondary surgery rate 20.9% (n = 22, 26 procedures)	• In second follow-up report, 5-yr overall secondary surgery rates were significantly lower for C-ADR group (7.3% vs. 21%, *P* = .0007) • Secondary 5-yr surgery rates were significantly lower for C-ADR group for surgeries both at index level (*P* = .0097) and at adjacent level (*P* = .0059)
		• Index-level 5-yr surgery rate 4.7% (n = 11) • Most common surgical indication was radiculopathy	• Index-level 5-yr surgery rate 18.1% (n = 19) • Modified index-level 5-yr surgery rate 12.4% (n = 13) • Indications included radiculopathy, neck pain, and pseudarthrosis • 13 of 19 surgeries were at index level and 6 were hardware removal for ASD]	
		• Adjacent-level surgery rate 3.4% (n = 8, 5 for ASD) • Indications included symptomatic ASD (n = 5), herniated disc (n = 2), and implant removals (n = 7)	• Adjacent-level surgery rate 11.4% (n = 12, 11 for ASD) • Indications included symptomatic ASD (n = 11) and herniated disc (n = 2) • Hardware plate removal (n = 12)	

C-ADR Device Author, Year Study Design	Duration of Follow-Up Randomized Surgery N (C-ADR/ Fusion)	Secondary Surgeries C-ADR	Secondary Surgeries Fusion	Conclusion
Radcliff et al, [84] 2017	7 yr N = 225/105 7 yr 74%/84%	• 7-yr index-level secondary surgery rate 4.4% (10/225) • Surgeries included removals (n=5), re-operations (n = 2), revisions (n = 2), and supplemental fixation (n = 1)	• 7-yr index-level secondary surgery rate 16.2% (17/105) • Surgeries included removals (n=8), re-operations (n = 2), revisions (n = 4), and supplemental fixation (n = 3) • Adjusted index-level secondary surgery rate 10.5% (11/ 105) (removing procedures for plate removal to treat adjacent-level disease)	• 7-yr index-level surgery rate was lower for C-ADR group ($P_{log\text{-}rank}$ = .062) • 7-yr adjacent-level surgery was significantly lower (4.4% vs. 11.4%) for C-ADR group ($P_{log\text{-}rank}$ = .009)
		• Adjacent-level surgery rate (at least 1) 4.4% (10/225) • Main indications were ASD or disc herniation (n = 10)	• Adjacent-level surgery (at least 1) 11.4% (12/105) • Main indications were ASD or disc herniation (n = 8)	
Prestige-LP				
Lanman et al, [94] 2017 30 sites IDE RCT	7 yr N = 209/188 5 yr 80%/73% 7 yr 74%/67%	Index-level 7-yr secondary surgery rate 4.2% (n = 8, 10 procedures)	• Index-level secondary surgery rate 14.7% (n = 22, 27 procedures) • Pseudarthrosis was indication for 9 patients	• 7-yr index-level secondary surgery rates were significantly lower in C-ADR group (LHR −1.29; 95% BCI −2.12 to −0.46) • Adjacent-level surgeries were not significantly lower (P_B = .942) for C-ADR (LHR −.5937; 95% BCI −1.35 to 0.156). Surgery rates, however, continued to diverge after 4-yr follow-up
		Adjacent-level surgery rate 6.5% (n = 12, 12 procedures)	Adjacent-level surgery rate 12.5% (n = 17, 22 procedures)	

Abbreviations: ASD, adjacent-segment disease; BCI, Bayesian confidence interval; C-ADR, cervical artificial disc replacement; IDE, investigational device exemption; KP, Kaplan-Meier; LHR, log hazard ratio; ND, not defined; P_B, Bayesian probability.

Appendix 4: Risk of Bias and GRADE Tables

Table A4: Risk of Bias of C-ADR Versus Fusion RCTs

Author, Year	Random Sequence Generation	Allocation Concealment	Blinding of Participants and Personnel	Blinding of Outcome Assessment	Complete Accounting of Patients and Outcome Events	Selective Reporting Bias
Heller et al,[20] 2009	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Cheng et al, 2009[44]	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Zhang et al, 2012[55]	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Donk et al, 2017[46]	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Chen et al,[43] 2013	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Rozankovic et al,[51] 2016	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Skeppholm et al,[52] 2015	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Sundseth et al,[53] 2017	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Coric et al,[36] 2011	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Hisey et al,[38] 2014	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Zhang et al,[54] 2014	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Phillips et al,[56] 2013	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Hou et al[47],[170] 2016	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Mummaneni et al,[21] 2007	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Cincu et al,[45] 2014	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Pandey et al,[50] 2017	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Murray et al,[22] 2009	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk

Author, Year	Random Sequence Generation	Allocation Concealment	Blinding of Participants and Personnel	Blinding of Outcome Assessment	Complete Accounting of Patients and Outcome Events	Selective Reporting Bias
Nabham et al,[48] 2007	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Nabham et al,[49] 2011	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Vaccaro et al,[40] 2013	Low risk	Low risk	Moderate risk[a]	Low risk	Low risk	Low risk
Gornet et al,[77] 2016	High risk[b]	Low risk	Moderate risk[a]	Low risk	Low risk	Moderate risk
Davis et al,[41] 2013	Low risk	Low risk	Moderate risk[c]	Low risk	Low risk	Low risk
Gornet et al,[42] 2017	Low risk	Low risk	Moderate risk[c]	Low risk	Low risk	Low risk
Cheng et al,[62] 2011	Low risk	Low risk	Moderate risk[c]	Low risk	Low risk	Low risk

Abbreviations: C-ADR, cervical artificial disc replacement; RCT, randomized controlled trial.

[a]Blinding of patients and investigators was not done, which may have posed a risk of bias to the main study outcome.

[b]Patients in the Gornet et al 1-level Prestige-LP study were not randomized. The study was a prospective controlled trial that employed propensity-matched fusion control patients from the prior FDA RCT of Prestige-ST; the patient selection criteria and surgeons were the same as those used in the FDA RCT of Prestige-ST.

[c]Patients and investigators were not blinded to outcome, which could have posed a risk of bias to main study outcome.

Source: Cochrane Risk of Bias Tool for Randomized Controlled Trials

Table A5: GRADE Evidence Profile for Anterior Cervical Discectomy and Artificial Cervical Disc Replacement Versus Fusion for Cervical Degenerative Disc Disease

No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	Upgrade Considerations	Quality
Overall treatment success							
7 RCTs of 1-level disease	Serious limitations (–1)[a]	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕ Moderate
2 RCTs of 2-level disease	Serious limitations (–1)[a]	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕ Moderate
Radiologic adjacent-level degenerative disease							
5 RCTs of 1-level disease	Serious limitations (–1)[a]	No serious limitations	No serious limitations	Serious limitations (–1)[b]	Undetected	NA	⊕⊕ Low
1 RCT of 2-level disease	Serious limitations (–1)[a]	No serious limitations	No serious limitations	Serious limitations (–1)[b]	Undetected	NA	⊕⊕ Low
Index-level secondary surgery							
5 RCTs of 1-level disease	Serious limitations (–1)[a]	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕ Moderate
2 RCTs of 2-level disease	Serious limitations (–1)[a]	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕ Moderate
Adjacent-level surgery for degenerative disease							
5 RCTs of 1-level disease	No serious limitations	Serious limitations (–1)[c]	No serious limitations	Serious limitations (–1)[d]	Undetected	NA	⊕⊕ Low
2 RCTs of 2-level disease	No serious limitations	Serious limitations (–1)[c]	No serious limitations	Serious limitations (–1)[d]	Undetected	NA	⊕⊕ Low
Perioperative outcomes							
9 RCTs of 1-level disease	No serious limitations	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕⊕ High
2 RCTs of 2-level disease	No serious limitations	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕⊕ High
Recovery, return to work							
9 RCTs of 1-level disease	No serious limitations	No serious limitations	No serious limitations	Serious limitations (–1)[e]	Undetected	NA	⊕⊕⊕ Moderate
2 RCTS of 2-level disease	No serious limitations	No serious limitations	No serious limitations	Serious limitations (–1)[e]	Undetected	NA	⊕⊕⊕ Moderate

No. of Studies	Risk of Bias	Inconsistency	Indirectness	Imprecision	Publication Bias	Upgrade Considerations	Quality
HRQOL							
10 RCTs of 1-level disease	No serious limitations	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕⊕ High
2 RCTs of 2-level disease	No serious limitations	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕⊕ High
Treatment satisfaction							
8 RCTs of 1-level disease	No serious limitations	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕⊕ High
2 RCTs of 2-level disease	No serious limitations	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕⊕ High
Cervical kinematics							
10 RCTs of 1-level disease	Serious limitations (–1)[a]	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕ Moderate
2 RCTs of 2-level disease	Serious limitations (–1)[a]	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕ Moderate
Long-term safety							
7 RCTs of 1-level disease	Serious limitations (–1)[a]	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕ Moderate
2 RCTs of 2-level disease	Serious limitations (–1)[a]	No serious limitations	No serious limitations	No serious limitations	Undetected	NA	⊕⊕⊕ Moderate

Abbreviations; GRADE, Grading of Recommendations Assessment, Development, and Evaluation; HRQOL, health-related quality of life; NA, none available; RCT, randomized controlled trial

[a]Serious limitations with risk of bias; trial personnel were not blinded to outcome, posing risk of bias to outcome assessment.

[b]Serious limitations with imprecision; most trials did not report radiographic findings or did not report definitions of adjacent-segment disease and evaluated condition at various follow-up points with different radiographic methods.

[c]Serious limitations with inconsistency; trials were inconsistent with differing follow-up results of adjacent-level surgeries.

[d]Serious limitations with imprecision; given longer course of cervical degenerative disease, duration of trial follow-up was inadequate and few events were observed.

[e]Serious limitations with imprecision; return-to-work outcomes varied greatly, limiting comparisons between study groups.

Table A6: AMSTAR Scores of Included Systematic Reviews

Author, Year	AMSTAR Score[a]	(1) Provided Study Design	(2) Duplicate Study Selection	(3) Broad Literature Search	(4) Considered Status of Publication	(5) Listed Excluded Studies	(6) Provided Characteristics of Studies	(7) Assessed Scientific Quality	(8) Considered Quality in Report	(9) Methods to Combine Appropriate	(10) Assessed Publication Bias	(11) Stated Conflict of Interest
Anderson et al, 2017	8	✓	✗	✓	✓	✗	✓	✓	✓	✓	✗	✓
Lehman et al, 2012	7	✓	✗	✓	✓	✗	✓	✓	✓	✓	✗	✗
Veruva et al, 2014	7	✓	✗	✓	✓	✗	✓	✓	✓	✓	✗	✗

Abbreviation: AMSTAR, A Measurement Tool to Assess Systematic Reviews.
[a]Maximum possible score is 11. Details of AMSTAR score are described in Shea et al.[29]

Appendix 5: Selected Excluded Studies

Table A7: Excluded Systematic Reviews

Systematic Review	Reason for Exclusion
Alvin MD, Abbott EE, Lubelski D, Kuhns B, Nowacki AS, Steinmetz MP, et al. Cervical arthroplasty: A critical review of the literature. Spine Journal. 2014;14(9):2231-45.	Includes all study designs and mixed follow-up periods
Alvin MD, Mroz TE. The Mobi-C cervical disc for one-level and two-level cervical disc replacement: A review of the literature. Medical Devices: Evidence and Research. 2014;7:397-403.	Restricts review to one class of C-ADR implants, all study types
Anderson PA, Hashimoto R. Total disc replacement in the cervical spine: a systematic review evaluating long-term safety. Evidencebased Spinecare Journal. 2012;3(S1):9-18.	Early search review, limited RCTs
Anderson PA, Puschak TJ, Sasso RC. Comparison of short-term SF-36 results between total joint arthroplasty and cervical spine decompression and fusion or arthroplasty. Spine (Phila Pa 1976). 2009;34(2):176-83.	Early review: 2 RCTs compare HRQOL of hip and knee to cervical arthroplasties
Anderson-Smits C, Sing D, Dmitriev A, Cheng H. A comparative analysis of secondary surgeries of six total cervical disc arthroplasty devices to cervical arthrodesis at 5-years. Pharmacoepidemiol Drug Saf. 2016;25:64.	Abstract only
Aragones M, Hevia E, Barrios C. Polyurethane on titanium unconstrained disc arthroplasty versus anterior discectomy and fusion for the treatment of cervical disc disease: a review of level I-II randomized clinical trials including clinical outcomes. Eur Spine J. 2015;24(12):2735-45.	Restricts review to one class of C-ADR implants
Bakar D, Lubelski D, Abdullah KG, Mroz TE. Artificial cervical disc arthroplasty versus anterior cervical discectomy and fusion: A systematic review. Curr Orthop Pract. 2014;25(1):9-13.	Limited search, no trial descriptions, overlapping reports
Bartels R, Donk RD, Verhagen WIM, Hosman AJF, Verbeek ALM. Reporting the results of meta-analyses: A plea for incorporating clinical relevance referring to an example. Spine Journal: Official Journal of the North American Spine Society. 2017;30:30.	Overview of other systematic reports
Bartels RHMA, Donk R, Verbeek ALM. No justification for cervical disk prostheses in clinical practice: A meta-analysis of randomized controlled trials. Neurosurgery. 2010;66(6):1153-60.	Variable reports include abstracts, limited studies, mixed follow-up periods
Botelho RV, Moraes OJ, Fernandes GA, Buscariolli YS, Bernardo WM. A systematic review of randomized trials on the effect of cervical disc arthroplasty on reducing adjacent-level degeneration. Neurosurg Focus. 2010;28(6):E5.	Early review, RCTs on only 3 C-ADR devices
Cepoiu-Martin M, Faris P, Lorenzetti D, Prefontaine E, Noseworthy T, Sutherland L. Artificial cervical disc arthroplasty: A systematic review. Spine (Phila Pa 1976). 2011;36(25):E1623-E33.	Early search period, includes all study types
Chang KE, Pham MH, Hsieh PC. Adjacent segment disease requiring reoperation in cervical total disc arthroplasty: A literature review and update. J Clin Neurosci. 2017;37:20-4.	Early review, combines rates of short- and long term follow-up
Chen C, Zhang X, Ma X. Durability of cervical disc arthroplasties and its influence factors: A systematic review and a network meta-analysis. Medicine. 2017;96(6):e5947.	Includes overlapping reports and limited long-term follow-up
Chen J, Fan SW, Wang XW, Yuan W. Motion analysis of single-level cervical total disc arthroplasty: a meta-analysis. Orthop Surg. 2012;4(2):94-100.	Early review, includes observational study and 3 RCTs
Chen J, Wang X, Bai W, Shen X, Yuan W. Prevalence of heterotopic ossification after cervical total disc arthroplasty: A meta-analysis. Eur Spine J. 2012;21(4):674-80.	Early review, mainly European studies, reports on 4 devices
Demetriades AK, Ringel F, Meyer B. Cervical disc arthroplasty: a critical review and appraisal of the latest available evidence. Adv Tech Stand Neurosurg. 2014;41:107-29.	No search strategy, includes all study types
Di Martino A, Papalia R, Albo E, Cortesi L, Denaro L, Denaro V. Cervical spine alignment in disc arthroplasty: should we change our perspective? Eur Spine J. 2015;24:810-25.	Includes observational studies and RCTs of single- and multi-level treatment
DiSilvestro KJ, Santoro AJ, Tjoumakaris FP, Levicoff EA, Freedman KB. When can I drive after orthopaedic surgery? A systematic review. Clin Orthop Relat Res. 2016;474(12):2557-70.	General review on all orthopedic surgery

Systematic Review	Reason for Exclusion
Dong L, Wang D, Chen X, Liu T, Xu Z, Tan M, et al. A comprehensive meta-analysis of the adjacent segment parameters in cervical disk arthroplasty versus anterior cervical discectomy and fusion. Clinical Spine Surgery : A Spine Publication. 2017;15:15.	Early review, reports with limited long-term follow-up
Dong L, Xu Z, Chen X, Wang D, Li D, Liu T, et al. The change of adjacent segment after cervical disc arthroplasty compared with anterior cervical discectomy and fusion: a meta-analysis of randomized controlled trials. Spine Journal: Official Journal of the North American Spine Society. 2017;15:15.	Early review, limited long-term follow-up reports
Fallah A, Akl EA, Ebrahim S, Ibrahim GM, Mansouri A, Foote CJ, et al. Anterior cervical discectomy with arthroplasty versus arthrodesis for single-level cervical spondylosis: a systematic review and meta-analysis. PLoS ONE [Electronic Resource]. 2012;7(8):e43407.	Limited search, overlapping studies, all studydesigns
Gao Y, Liu M, Li T, Huang F, Tang T, Xiang Z. A meta-analysis comparing the results of cervical disc arthroplasty with anterior cervical discectomy and fusion (ACDF) for the treatment of symptomatic cervical disc disease. Journal of Bone and Joint Surgery - Series A. 2013;95(6):555-61.	Meta-analysis updated in later report
Gao F, Mao T, Sun W, Guo W, Wang Y, Li Z, et al. An updated meta-analysis comparing artificial cervical disc arthroplasty (CDA) versus anterior cervical discectomy and fusion (ACDF) for the treatment of cervical degenerative disc disease (CDDD). Spine (Phila Pa 1976). 2015;40(23):1816-23.	Mixed follow-up periods, overlapping reports, limited long-term follow-up
Harrod CC, Hilibrand AS, Fischer DJ, Skelly AC. Adjacent segment pathology following cervical motion-sparing procedures or devices compared with fusion surgery: A systematic review. Spine (Phila Pa 1976). 2012;37(Suppl. 22):S96-S112.	Early review, limited long-term follow-up
Hu Y, Lv G, Ren S, Johansen D. Mid- to Long-Term Outcomes of Cervical Disc Arthroplasty versus Anterior Cervical Discectomy and Fusion for Treatment of Symptomatic Cervical Disc Disease: A Systematic Review and Meta-Analysis of Eight Prospective Randomized Controlled Trials. PLoS ONE [Electronic Resource]. 2016;11(2):e0149312.	Early review, limited long-term follow-up, includes mixed treatment levels
Jia Z, Mo Z, Ding F, He Q, Fan Y, Ruan D. Hybrid surgery for multilevel cervical degenerative disc diseases: A systematic review of biomechanical and clinical evidence. Eur Spine J. 2014;23(8):1619-32.	Includes biomechanical and clinical studies, no comparators, different study designs
Jiang H, Zhu Z, Qiu Y, Qian B, Qiu X, Ji M. Cervical disc arthroplasty versus fusion for single-level symptomatic cervical disc disease: a meta-analysis of randomized controlled trials. Arch Orthop Trauma Surg. 2012;132(2):141-51.	Limited search, early review search period
Jiang L, Tan M, Yang F, Yi P, Tang X, Hao Q. Comparisons of safety and clinical outcomes between multiple-level and single-level cervical disk replacement for cervical spondylosis. Clinical Spine Surgery. 2016;29(10):419-26.	Early review search period, includes different study designs
Joaquim AF, Murar J, Savage JW, Patel AA. Dysphagia after anterior cervical spine surgery: A systematic review of potential preventative measures. Spine Journal. 2014;14(9):2246-60.	Estimates preventive interventions
Joaquim AF, Riew KD. Multilevel cervical arthroplasty: current evidence. A systematic review. Neurosurg Focus. 2017;42(2):E4.	Different study designs
Kan SL, Yuan ZF, Ning GZ, Liu FF, Sun JC, Feng SQ. Cervical disc arthroplasty for symptomatic cervical disc disease: Traditional and Bayesian meta-analysis with trial sequential analysis. International Journal of Surgery. 2016;35:111-9.	Methodologic review, early search review, limited long-term follow-up
Kang J, Shi C, Gu Y, Yang C, Gao R. Factors that may affect outcome in cervical artificial disc replacement: a systematic review. Eur Spine J. 2015;24(9):2023-32.	Includes conference abstracts
Kepler CK, Brodt ED, Dettori JR, Albert TJ. Cervical artificial disc replacement versus fusion in the cervical spine: a systematic review comparing multilevel versus single-level surgery. Evidence based Spinecare Journal. 2012;3(S1):19-30.	Early search review, includes overlapping reports and conference abstracts
Kim HJ, Kelly MP, Ely CG, Riew KD, Dettori JR. The risk of adjacent-level ossification development after surgery in the cervical spine: Are there factors that affect the risk? A systematic review. Spine (Phila Pa 1976). 2012;37(Suppl. 22):S65-S74.	Early search review, compares various spine surgeries
Kong L, Cao J, Wang L, Shen Y. Prevalence of adjacent segment disease following cervical spine surgery: A PRISMA-compliant systematic review and meta-analysis. Medicine (Baltimore). 2016;95(27):e4171.	Reviews all cervical spine surgeries
Kong L, Ma Q, Meng F, Cao J, Yu K, Shen Y. The prevalence of heterotopic ossification among patients after cervical artificial disc replacement: A systematic review and meta-analysis. Medicine. 2017;96(24):e7163.	Limited inclusion of RCTs, mainly European and Asian cohorts

Systematic Review	Reason for Exclusion
Konig SS, U. Clinical outcome of anterior cervical discectomy and fusion versus total disc replacement—a meta-analysis of 2532 cases. Insights in Neurosurgery. 2016;1(2):14.	Limited database search, early search review, study characteristics unspecified
Kuang L, Chen Y, Wang B, Li L, Lu G. Cervical Disk Arthroplasty Versus Anterior Cervical Decompression and Fusion for the Treatment of 2-Level Cervical Spondylopathy: A Systematic Review and Meta-analysis. Clinical Spine Surgery. 2016;29(9):372-82.	Early review search period, different study designs, limited RCTreports
Li GL, Hu JZ, Lu HB, Qu J, Guo LY, Zai FL. Anterior cervical discectomy with arthroplasty versus anterior cervical discectomy and fusion for cervical spondylosis. J Clin Neurosci. 2015;22(3):460-7.	Early search review, limited RCTreports
Liu FY, Yang DL, Huang WZ, Huo LS, Ma L, Wang H, et al. Risk factors for dysphagia after anterior cervical spine surgery: A meta-analysis. Medicine. 2017;96(10):e6267.	Evaluates risk factors from observational studies on any spine surgery
Lu VM, Zhang L, Scherman DB, Rao PJ, Mobbs RJ, Phan K. Treating multi-level cervical disc disease with hybrid surgery compared to anterior cervical discectomy and fusion: a systematic review and meta-analysis. Eur Spine J. 2017;26(2):546-57.	Evaluates various approaches to fusion surgical
Luo J, Gong M, Huang S, Yu T, Zou X. Incidence of adjacent segment degeneration in cervical disc arthroplasty versus anterior cervical decompression and fusion meta-analysis of prospective studies. Arch Orthop Trauma Surg. 2015;135(2):155-60.	Mixed follow-up periods, no device detail
Luo J, Huang S, Gong M, Dai X, Gao M, Yu T, et al. Comparison of artificial cervical arthroplasty versus anterior cervical discectomy and fusion for one-level cervical degenerative disc disease: a meta-analysis of randomized controlled trials. European Journal of Orthopaedic Surgery and Traumatology. 2015;25:115-25.	Early search review period, limited RCT reports
Ma Z, Ma X, Yang H, Guan X, Li X. Anterior cervical discectomy and fusion versus cervical arthroplasty for the management of cervical spondylosis: a meta-analysis. Eur Spine J. 2017;26(4):998-1008.	Limited search, mixed pathology, mixed follow-up periods
Maharaj MM, Mobbs RJ, Hogan J, Zhao DF, Rao PJ, Phan K. Anterior cervical disc arthroplasty (ACDA) versus anterior cervical discectomy and fusion (ACDF): a systematic review and meta-analysis. The Journal of Spine Surgery. 2015;1(1):72-85.	Early review search period, included all study types
McAfee PC, Reah C, Gilder K, Eisermann L, Cunningham B. A meta-analysis of comparative outcomes following cervical arthroplasty or anterior cervical fusion: Results from 4 prospective multicenter randomized clinical trials and up to 1226 patients. Spine (Phila Pa 1976). 2012;37(11):943-52.	Not a SR, outcomes of selected C-ADR devices
Molinari RW, Pagarigan K, Dettori JR, Molinari RW, Dehaven KE. Return to Play in Athletes Receiving Cervical Surgery: A Systematic Review. Global spine journal. 2015;6(1):89-96.	Reviews all spine surgeries
Muheremu A, Niu X, Wu Z, Muhanmode Y, Tian W. Comparison of the short- and long-term treatment effect of cervical disk replacement and anterior cervical disk fusion: a meta-analysis. European Journal of Orthopaedic Surgery and Traumatology. 2015;25:87-100.	Early search review period, mixed follow-up periods, limited long-term RCTreports
Mummaneni PV, Amin BY, Wu JC, Brodt ED, Dettori JR, Sasso RC. Cervical artificial disc replacement versus fusion in the cervical spine: a systematic review comparing long-term follow-up results from two FDA trials. Evidencebased Spinecare Journal. 2012;3(S1):59-66.	Selected C-ADR devices with limited RCT reports
Nunley PD, Jawahar A, Cavanaugh DA, Gordon CR, Kerr IEJ, Utter PA. Symptomatic adjacent segment disease after cervical total disc replacement: Re-examining the clinical and radiological evidence with established criteria. Spine Journal. 2013;13(1):5-12.	Not a systematic review, primary analysis of 4 FDA RCTs
Rao MJ, Nie SP, Xiao BW, Zhang GH, Cao SS. Cervical disc arthroplasty versus anterior cervical discectomy and fusion for treatment of symptomatic cervical disc disease: A meta-analysis of randomized controlled trials. Arch Orthop Trauma Surg. 2015;135(1):19-28.	Early search review period, overlapping reports, mixed follow-up periods, no device details
Ren C, Song Y, Xue Y, Yang X. Mid- to long-term outcomes after cervical disc arthroplasty compared with anterior discectomy and fusion: A systematic review and meta-analysis of randomized controlled trials. Eur Spine J. 2014;23(5):1115-23.	Early review search period, limited RCTreports
Riew KD, Schenk-Kisser JM, Skelly AC. Adjacent segment disease and C-ADR: promises fulfilled? Evidencebased Spinecare Journal. 2012;3(S1):39-46.	Early search review, limited long-term RCTreports

Systematic Review	Reason for Exclusion
Saavedra-Pozo FM, Deusdara RAM, Benzel EC. Adjacent segment disease perspective and review of the literature. Ochsner Journal. 2014;14(1):78-83.	Limited search strategy and inclusion of lumbar and cervical spine
Scherman DB, Mobbs RJ, Phan K. Adjacent segment degeneration and disease following cervical arthroplasty: a systematic review and meta-analysis. The Journal of Spine Surgery. 2016;2(1):82-4.	Early search review, no study characteristics
Shangguan L, Ning GZ, Tang Y, Wang Z, Luo ZJ, Zhou Y. Discover cervical disc arthroplasty versus anterior cervical discectomy and fusion in symptomatic cervical disc diseases: A meta-analysis. PLoS ONE [Electronic Resource]. 2017;12(3):e0174822.	Restricts review to one class of C-ADR implants
Shriver MF, Lubelski D, Sharma AM, Steinmetz MP, Benzel EC, Mroz TE. Adjacent segment degeneration and disease following cervical arthroplasty: A systematic review and meta-analysis. Spine Journal. 2016;16(2):168-81.	Limited search strategy, limited long-term follow-upreports
Singh K, Phillips FM, Park DK, Pelton MA, An HS, Goldberg EJ. Factors affecting reoperations after anterior cervical discectomy and fusion within and outside of a Federal Drug Administration investigational device exemption cervical disc replacement trial. Spine Journal. 2012;12(5):372-8.	Not a systematic review
Tan W, Zhou C, Guo D, Sun J, Cao W, Yang LZ, et al. Treatment of single-level cervical spondylosis: Cervical disk arthroplasty versus anterior cervical decompression and fusion. Orthopedics. 2017;40(1):e23-e34.	Early review search period, mixed follow-up periods
Tian P, Fu X, Li ZJ, Sun XL, Ma XL. Hybrid surgery versus anterior cervical discectomy and fusion for multilevel cervical degenerative disc diseases: a meta-analysis. Sci Rep. 2015;5:134-54.	Compares fusion approaches for subset of disease
Traynelis VC, Arnold PM, Fourney DR, Bransford RJ, Fischer DJ, Skelly AC. Alternative Procedures for the Treatment of Cervical Spondylotic Myelopathy: Arthroplasty, Oblique Corpectomy, Skip Laminectomy: Evaluation of Comparative Effectiveness and Safety. Spine (Phila Pa 1976). 2013;10:10.	Includes various cervical surgical approaches ery
Traynelis VC, Leigh BC, Skelly AC. Return to work rates and activity profiles: are there differences between those receiving C-ADR and ACDF? Evidencebased Spinecare Journal. 2012;3(S1):47-52.	Early review search, limited RCT reports
Ueda H, Huang RC, Lebl DR. Iatrogenic contributions to cervical adjacent segment pathology: review article. HSS J. 2015;11(1):26-30.	Limited search strategy, limited long-term follow-up reports
Upadhyaya CD, Wu JC, Trost G, Haid RW, Traynelis VC, Tay B, et al. Analysis of the three united states food and drug administration investigational device exemption cervical arthroplasty trials: Clinical article. J Neurosurg Spine. 2012;16(3):216-28.	Not a systematic review
Verhagen AP, Van Middelkoop M, Rubinstein SM, Ostelo R, Jacobs W, Peul W, et al. Effect of various kinds of cervical spinal surgery on clinical outcomes: A systematic review and meta-analysis. Pain. 2013;154(11):2388-96.	Includes various spine surgeries
Verma K, Gandhi SD, Maltenfort M, Albert TJ, Hilibrand AS, Vaccaro AR, et al. Rate of adjacent segment disease in cervical disc arthroplasty versus single-level fusion: Meta-analysis of prospective studies. Spine (Phila Pa 1976). 2013;38(26):2253-7.	Limited search strategy, limited long-term follow-upreports
Wang Z, Liu W, Li J, Wang F, Yao Z. Safety of anterior cervical discectomy and fusion versus cervical arthroplasty in patients with cervical spondylosis: A meta-analysis of randomized controlled trials. Int J Clin Exp Med. 2016;9(10):19537-44.	Early review search, limited RCT reports
Wu AM, Xu H, Mullinix KP, Jin HM, Huang ZY, Lv QB, et al. Minimum 4-year outcomes of cervical total disc arthroplasty versus fusion: a meta-analysis based on prospective randomized controlled trials. Medicine. 2015;94(15):e665.	Early review search, limited RCT reports with long-term follow-up updated in later review
Wu TK, Liu H, Wang BY, Meng Y. Minimum four-year subsequent surgery rates of cervical disc replacement versus fusion: A meta-analysis of prospective randomized clinical trials. Orthopaedics and Traumatology: Surgery and Research. 2017;103(1):45-51.	Limited RCTreports with long-term follow-up
Wu TK, Wang BY, Meng Y, Ding C, Yang Y, Lou JG, et al. Multilevel cervical disc replacement versus multilevel anterior discectomy and fusion: A meta-analysis. Medicine. 2017;96(16):e6503.	Includes observational reports
Xia XP, Chen HL, Cheng HB. Prevalence of adjacent segment degeneration after spine surgery: a systematic review and meta-analysis. Spine (Phila Pa 1976). 2013;38(7):597-608.	Includes all spine surgeries
Xie L, Liu M, Ding F, Li P, Ma D. Cervical disc arthroplasty (CDA) versus anterior cervical discectomy and fusion (ACDF) in symptomatic cervical degenerative disc diseases (CDDDs): an updated meta-analysis of prospective randomized controlled trials (RCTs). Springerplus. 2016;5(1):1188.	Overlapping reports, mixed follow-up periods

Systematic Review	Reason for Exclusion
Xing D, Ma XL, Ma JX, Wang J, Ma T, Chen Y. A meta-analysis of cervical arthroplasty compared to anterior cervical discectomy and fusion for single-level cervical disc disease. J Clin Neurosci. 2013;20(7):970-8.	Early review search period, limited RCT reports, mixed follow-up periods
Xu B, Ma JX, Tian JH, Ge L, Ma XL. Indirect meta-analysis comparing clinical outcomes of total cervical disc replacements with fusions for cervical degenerative disc disease. Sci Rep. 2017;7(1):1740.	Search strategy not reported, includes mixed treatment levels and overlapping reports
Yang B, Li H, Zhang T, He X, Xu S. The incidence of adjacent segment degeneration after cervical disc arthroplasty (CDA): a meta analysis of randomized controlled trials. PLoS One. 2012;7(4):e35032.	Early search review period, limited RCT reports
Yao Q, Liang F, Xia Y, Jia C. A meta-analysis comparing total disc arthroplasty with anterior cervical discectomy and fusion for the treatment of cervical degenerative diseases. Arch Orthop Trauma Surg. 2016;136(3):297-304.	Early search review period, limited RCT reports, mixed follow-up period
Yin S, Yu X, Zhou S, Yin Z, Qiu Y. Is cervical disc arthroplasty superior to fusion for treatment of symptomatic cervical disc disease? A meta-analysis. Clin Orthop Relat Res. 2013;471(6):1904-19.	Early review search period, limited long-term follow-up of RCTs
Yu L, Song Y, Yang X, Lv C. Systematic review and meta-analysis of randomized controlled trials: Comparison of total disk replacement with anterior cervical decompression and fusion. Orthopedics. 2011;34(10):e651-e8.	Early review search period, limited RCTreports
Zang L, Ma M, Hu J, Qiu H, Huang B, Chu T. Comparison of hybrid surgery incorporating anterior cervical discectomy and fusion and artificial arthroplasty versus multilevel fusion for multilevel cervical spondylosis: A meta-analysis. Med Sci Monit. 2015;21:4057-67.	Compares fusion surgical approaches for subset of cervical pathology
Zechmeister I, Winkler R, Mad P. Artificial total disc replacement versus fusion for the cervical spine: A systematic review. Eur Spine J. 2011;20(2):177-84.	Early review search period, limited RCT reports, includes different study designs
Zeng J, Duan Y, Liu H, Wang B, Gong Q, Yang Y, et al. Dynamic cervical implant in treating cervical degenerative disc disease: A systematic review and meta-analysis. Int J Clin Exp Med. 2017;10(6):8700-8.	Restricts review to one class of C-ADR implants. No RCTs
Zhang J, Meng F, Ding Y, Li J, Han J, Zhang X, et al. Hybrid Surgery Versus Anterior Cervical Discectomy and Fusion in Multilevel Cervical Disc Diseases: A Meta-Analysis. Medicine. 2016;95(21):e3621.	Compares fusion surgical approaches for subset of disease
Zhang Y, Liang C, Tao Y, Zhou X, Li H, Li F, et al. Cervical total disc replacement is superior to anterior cervical decompression and fusion: a meta-analysis of prospective randomized controlled trials. PLoS ONE [Electronic Resource]. 2015;10(3):e0117826.	Early review search period, limited long-term follow-up of RCT reports of RCTs
Zhao GS, Zhang Q, Quan ZX. Mid-term efficacy and safety of cervical disc arthroplasty versus fusion in cervical spondylosis: A systematic review and meta-analysis. Biomedical Reports. 2017;6(2):159-66.	Limited RCT reports with long-term follow-up, includes overlapping reports
Zhao H, Cheng L, Hou Y, Liu Y, Liu B, Mundra JJ, et al. Multi-level cervical disc arthroplasty (CDA) versus single-level CDA for the treatment of cervical disc diseases: a meta-analysis. Eur Spine J. 2015;24(1):101-12.	No trial characteristics, overlapping reports
Zhong ZM, Li M, Han ZM, Zeng JH, Zhu SY, Wu Q, et al. Does cervical disc arthroplasty have lower incidence of dysphagia than anterior cervical discectomy and fusion? A meta-analysis. Clin Neurol Neurosurg. 2016;146:45-51.	Early review search period and overlapping RCTreports
Zhong ZM, Zhu SY, Zhuang JS, Wu Q, Chen JT. Reoperation After Cervical Disc Arthroplasty Versus Anterior Cervical Discectomy and Fusion: A Meta-analysis. Clin Orthop Relat Res. 2016;474(5):1307-16.	Overlapping RCTreports with mixed follow-up periods
Zhou HH, Qu Y, Dong RP, Kang MY, Zhao JW. Does heterotopic ossification affect the outcomes of cervical total disc replacement? A meta-analysis. Spine (Phila Pa 1976). 2015;40(6):E332-E40.	Estimates prevalence of outcome
Zhu Y, Tian Z, Zhu B, Zhang W, Li Y, Zhu Q. Bryan Cervical Disc Arthroplasty Versus Anterior Cervical Discectomy and Fusion for Treatment of Cervical Disc Diseases: A Meta-analysis of Prospective, Randomized Controlled Trials. Spine (Phila Pa 1976). 2016;41(12):E733-41.	Restricts review to one class of C-ADR implants

Systematic Review	Reason for Exclusion
Zhu Y, Zhang B, Liu H, Wu Y, Zhu Q. Cervical disc arthroplasty versus anterior cervical discectomy and fusion for incidence of symptomatic adjacent segment disease a meta-analysis of prospective randomized controlled trials. Spine (Phila Pa 1976). 2016;41(19):1493-502.	Early review search period, limited RCTs with long-term follow-up
Zou S, Gao J, Xu B, Lu X, Han Y, Meng H. Anterior cervical discectomy and fusion (ACDF) versus cervical disc arthroplasty (CDA) for two contiguous levels cervical disc degenerative disease: a meta-analysis of randomized controlled trials. Eur Spine J. 2017;26(4):985-97.	Early review search period, limited RCTs

Abbreviations: C-ADR, cervical artificial disc replacement; FDA, U.S. Food and Drug Administration; HRQOL, health-related quality of life; RCT, randomized controlled trial; SF-36, 36-item Short Form Health Survey; SR, systematic review.

Appendix 6: Results of Applicability Checklists for Studies Included in Economic Literature Review

Table A8: Applicability of Studies Assessing Cost-Effectiveness of C-ADR

Objective: To assess the cost-effectiveness of C-ADR					
Author, Year	**Is study population similar to that in the question?**	**Are interventions similar to those in the question?**	**Is health care system in which study was conducted sufficiently similar to the current Ontario context?**	**Were perspectives clearly stated and what were they?**	**Are estimates of relative treatment effect from best available source?**
Menzin et al,[128] 2010	Yes	Yes	No (United States)	Yes; societal	Partially
Qureshi et al,[129] 2013	Yes	Yes	No (United States)	Yes; private insurer payer	Partially
Warren et al,[130] 2013	Yes	Yes	No (United States)	Yes; private insurer payer	No
Lewis et al,[131] 2014	Yes	Yes	No (United States)	Yes; private insurer payer	Partially
McAnany et al,[132] 2014	Yes	Yes	No (United States)	Yes; private insurer payer	Partially
Radcliff et al,[133] 2015	Yes	Yes	No (United States)	Yes; private insurer payer	Partially
Ghori et al,[134] 2016	Yes	Yes	No (United States)	Yes; societal	Partially
Radcliff et al,[135] 2016	Yes	Yes	No (United States)	Yes; private insurer payer	Partially
Wiedenhofer et al,[136] 2017	Yes	Yes	No (Germany)	Yes; state health insurance payer	Partially
McAnany et al,[137] 2018	Yes	Yes	No (United States)	Yes; private insurer payer	Partially
Ament et al,[138] 2014	Yes	Yes	No (United States)	Yes; societal and private insurer payer	Partially
Ament et al,[139] 2016	Yes	Yes	No (United States)	Yes; societal and private insurer payer	Partially
Overley et al,[140] 2018	Yes	Yes	No (United States)	Yes; private insurer payer	Partially
Merrill et al,[141] 2018	Yes	Yes	No (United States)	Yes; private insurer payer	Partially

Abbreviation: C-ADR, cervical anterior discectomy and artificial disc replacement.

Author, Year	Are all future costs and outcomes discounted?[a] (If yes, at what rate?)	Is the value of health effects expressed in terms of quality-adjusted life-years?[a]	Are costs and outcomes from other sectors fully and appropriately measured and valued?[a]	Overall judgement[b] (directly applicable/partially applicable/ not applicable)
Menzin et al,[128] 2010	No (not reported)	No	Partially	Partially applicable
Qureshi et al,[129] 2013	Yes	Yes	NA	Partially applicable
Warren et al,[130] 2013	Yes	Yes	NA	Partially applicable
Lewis et al,[131] 2014	Unclear	Yes	NA	Partially applicable
McAnany et al,[132] 2014	Yes	Yes	NA	Partially applicable
Radcliff et al,[133] 2015	Unclear	No	NA	Partially applicable
Ghori et al,[134] 2016	Yes	No	Partially	Partially applicable
Radcliff et al,[135] 2016	Yes	Yes	NA	Partially applicable
Wiedenhofer et al,[136] 2017	Unclear	No	NA	Partially applicable
McAnany et al,[137] 2018	Yes	Yes	NA	Partially applicable
Ament et al,[138] 2014	Yes	Yes	Partially	Partially applicable
Ament et al,[139] 2016	Yes	Yes	Partially	Partially applicable
Overley et al,[140] 2018	Yes	Yes	NA	Partially applicable
Merrill et al,[141] 2018	Yes	Yes	NA	Partially applicable

Abbreviation: C-ADR, cervical artificial disc replacement; NA, not applicable.
[a]Response options were “yes,” “partially,” “no,” “unclear,” and “NA.”
[b]Response options were “directly applicable,” “partially applicable,” and “not applicable.”

Appendix 7: Artificial Disc Devices Considered in Economic Evaluation

Table A9: Devices Considered

Device	Manufacturer or Distributor	Levels	First Licence Issue Date in Canada	Licence No.
M6-C	Spinal Kinetics	1-level	2014-07-25	93628
ProDisc-C	DePuy Synthes (ProDisc product line was recently sold to Centinel Spine)	1-level	2007-10-02	74982
ProDisc-C NOVA[a]			2015-05-08	95104
ProDisc-C VIVO[a]			2017-02-16	98617
Mobi-C	Zimmer Biomet (formerly LDR)	1- and 2-level	Applied (in process)	--
Prestige-LP	Medtronic	1- and 2-level	2010-02-10	81979

[a]ProDisc-C NOVA and ProDisc-C VIVO are new models of ProDisc-C. Utilities were available for ProDisc-C and were assumed to be the same for newer models. We used price for newest model.

Table A10: Devices Excluded from Consideration

Device	Manufacturer or Distributor	Levels	First Licence Issue Date in Canada	Licence No.	Reason Excluded
Bryan	Medtronic	1	2003-05-06	62403	Device no longer sold in Canada
PCM	NuVasive Inc.	1	2007-05-29	74164	Licence expired 2017-10-25

Appendix 8: Extracted Natural History Parameters

Classes of Corrective Surgeries as Defined by the U.S. Food and Drug Administration[171]

- **Revision:** A procedure that adjusts or in any way modifies the original implant configuration (e.g., adjusting position of the original configuration, removal and replacement with the same type of study implant)
- **Removal:** A procedure that removes one or more components of the original implant configuration without replacement with the same type of trial implant. Removals include elective removals
- **Supplemental fixation:** A procedure at the involved level in which additional spinal devices not approved as part of the protocol are placed. This categorization of supplemental fixation includes supplemental therapies (i.e., external bone growth stimulators).
- **Reoperation:** A procedure that involves any surgical procedure at the involved level that does not remove, modify, or add any components and that is not considered a removal, revision, or supplemental fixation.

Table A11: Incidence of Index-Level Corrective Surgeries by Type in 7-Year Follow-Up

Device Author, Year	Levels	Type of Index-Level Corrective Surgery	Disc (n)	Fusion (n)
Prestige-LP Gornet et al,[90] 2016	1-level	Revision	1	5
		Removal (includes elective)	14	8
		Supplemental fixation	2	9
		Reoperation	3	4
		Total	**20**	**26**
ProDisc-C Janssen et al,[92] 2015	1-level	Revision	0	4
		Removal	5	1
		Supplemental fixation	1	3
		Plate removal (elective; because of adjacent surgery)	--	11
		Total	**6**	**19**
Mobi-C Radcliff et al,[84] 2017	1-level	Revision	0	0
		Removal	4	2
		Supplemental fixation	0	3
		Reoperation	1	0
		Plate removal (elective; because of adjacent surgery)	--	5
		Total	**5**	**10**
Prestige-LP Lanman et al,[94] 2017	2-level	Revision	0	1
		Removal	6	6
		Supplemental fixation	1	7
		Reoperation	3	7
		Plate removal (elective; because of adjacent surgery)	--	6
		Total	**10**	**27**
Mobi-C Radcliff et al,[84] 2017	2-level	Revision	2	4
		Removal	5	2
		Supplemental fixation	1	3
		Reoperation	2	2
		Plate removal (elective; because of adjacent surgery)	--	6
		Total	**10**	**17**

Appendix 9: Cost Details

Table A12: Other Sources Explored for Unit Prices of Artificial Discs

Source	Device	Unit Price (USD)	Standard Deviation	Comments
Ament et al,[138] 2014	Mobi-C	8,000		Prices are expressed in 2012 USD
Lewis et al,[131] 2014	Various	4,555	1,845	Prices are expressed in 2014 USD

Abbreviation: USD, U.S. dollars.

Table A13: Other Sources Explored for Unit Prices of Instruments Used in Fusion

Source	Level	Price (USD)	Standard Deviation	Comments
Menzin et al,[128] 2010	1-level	1,282 allograft + 2,393 Atlantis plate = **3,675**		Total instrument cost per surgery. Prices are expressed in 2007 USD
	2-level	**2,967** Atlantis plate		Prices are expressed in 2007 USD
Ament et al,[138] 2014	2-level	**6,372**		Total instrument cost per surgery. Prices are expressed in 2012 USD
Lewis et al,[131] 2014	1-level	1,980 cage/spacer + 2,036 fixation construct = **4,016**	360 370	Total instrument cost per surgery. Prices are expressed in 2014 USD

Abbreviation: USD, U.S. dollars.

Table A14: Surgeon, Anesthetist, and Surgical Assistant Fees for Index and Revision Surgeries, 2018

	C-ADR				Fusion			
	One-Level		Two-Level		One-Level		Two-Level	
Cost Factor	Code/Units	Fee ($)	Code/Units	Fee ($)	Code/Units	Fee ($)	Code/Units	Fee ($)
Surgeon (initial)	N526	2,040.00	N526	2,040.00	N500	918.00	N500	918.00
			E394	765.00	E365	765.00	E365	765.00
							E360	306.00
							E366	153.00
Surgeon (revision)	N526 +30%	2,652.00	N526 +30%	2,652.00	N500+30%	1,193.40	N500+30%	1,193.40
			E394	765.00	E365	765.00	E365	765.00
							E360	306.00
							E366	153.00
Anesthesia, per unit fee = $15.01	17 units	255.17	17 units	255.17	10 units	150.10	10 units	150.10
Assistant, per unit fee = $12.04	11 units	132.44	11 units	132.44	9 units	108.36	9 units	108.36
Total (initial)		**2,427.61**		**3,192.61**		**1,941.46**		**2,400.46**
Total (revision)		**3,039.61**		**3,804.61**		**2,216.86**		**2,675.86**

Abbreviation: C-ADR, cervical artificial disc replacement.
Source: Data provided by Ontario Health Insurance Plan.

Table A15: Surgeon, Anesthetist, and Surgical Assistant Fees for Removal Surgery, 2018

	Removal Followed by Fusion[a]			
	One-Level		Two-Level	
Cost Factor	Code/Units	Fee ($)	Code/Units	Fee ($)
Surgeon	N500+30%	1,193.40	N500+30%	1,193.40
	E365	765.00	E365	765.00
			E360	306.00
			E366	153.00
Anesthesia, per unit fee = $15.01	10 units	150.10	10 units	150.10
Assistant, per unit fee = $12.04	9 units	108.36	9 units	108.36
Total		**2,216.86**		**2,675.86**

[a]Secondary surgery is almost always fusion, irrespective of type of previous surgery.
Source: Data provided by Ontario Health Insurance Plan.

Table A16: Surgeon, Anesthetist, and Surgical Assistant Fees for Cervical Supplemental Fixation Surgery, 2018

	Cervical and Thoracic Posterior Spinal Arthrodesis as Sole Procedure[a]			
	One-Level		Two-Level	
Cost Factor	Code/Units	Fee ($)	Code/Units	Fee ($)
Surgeon	N515+30%	1,326.00	N515+30%	1,326.00
			E366	153.00
Anesthesia, per unit fee = $15.01	11 units	165.11	11 units	165.11
Assistant, per unit fee = $12.04	9 units	108.36	9 units	108.36
Total		**1,599.47**		**1,752.47**

[a]According to Summary of Safety and Effectiveness Data from U.S. Food and Drug Association, supplemental fixation is usually posterior arthrodesis, irrespective of type of previous surgery.
Source: Data provided by Ontario Health Insurance Plan.

Table A17: Surgeon, Anesthetist, and Surgical Assistant Fees for Reoperation Surgery, 2018

	Cervical/Thoracic Posterior Spinal Decompression[a]			
	One-Level		Two-Level	
Cost Factor	Code/Units	Fee ($)	Code/Units	Fee ($)
Surgeon	N509+30%	1,306.11	N509+30%	1,306.11
			E361	255.00
Anesthesia, per unit fee = $15.01	12 units	180.12	12 units	180.12
Assistant, per unit fee = $12.04	9 units	108.36	9 units	108.36
Total		**1,594.59**		**1,849.59**

[a]According to Summary of Safety and Effectiveness Data from U.S. Food and Drug Association, reoperation is usually posterior decompression, irrespective of type of previous surgery.
Source: Data provided by Ontario Health Insurance Plan.

Table A18: Hospital Cost for C-ADR and Fusion Surgeries in Ontario

Period	Radiculopathy and Myelopathy Cases				All Cases[a]			
	C-ADR		Fusion		C-ADR		Fusion	
	Patients (n)	Average Cost ($)	Patients (n)	Average Cost ($)	Patients (n)	Average Cost ($)	Patients (n)	Average Cost ($)
2010–2011	18	11,616	162	8,856	19	11,809	179	9,699
2011–2012	17	10,443	170	8,106	19	10,190	197	8,032
2012–2013	6	13,161	203	9,538	9	12,666	229	9,604
2013–2014	7	7,913	251	8,942	8	9,071	274	8,803
2014–2015	FOI	FOI	202	9,209	6	10,250	224	9,300
2015–2016	7	13,807	210	8,880	11	13,122	232	8,803
6-year average		**11,230**		**8,947**		**11,255**		**9,030**

Abbreviations: C-ADR, cervical artificial disc replacement; FOI, Freedom of Information.
[a]Cost includes price of artificial disc or of instruments for fusion surgery. "All Cases" was chosen for our analysis.
Source: Ontario Case Costing Initiative.[145]

Table A19: Hospital Cost for Cervical Posterior (Supplemental) Fixation and Reoperation Surgeries in Ontario

Period	Posterior (Supplemental) Fixation		Reoperation	
	Patients (n)	Average Yearly Cost[a] ($)	Patients (n)	Average Yearly Cost[a] ($)
2010–2011	18	14,610	47	14,874
2011–2012	27	12,648	67	12,073
2012–2013	36	15,512	70	11,197
2013–2014	34	15,037	73	7,233
2014–2015	19	14,464	61	7,465
2015–2016	34	13,724	87	6,936
6-year average		**14,379**		**9,577**

[a]Cost includes price of instruments required for surgery.
Source: Ontario Case Costing Initiative.[145]

Table A20: Multi-level (Lumbar + Cervical) Artificial Disc Implantation

Type of Involvement	2010	2011	2012	2013	2014	2015
Multi-level (n)	297	322	305	354	312	355
All (N)	1,477	1,528	1,629	1,753	1,724	1,895
Ratio (n/N × 100)	20%	21%	19%	20%	18%	19%

Source: Data provided by Ontario Health Insurance Plan.

Table A21: Number and Relative Distribution of Cervical Fusion Procedures in Alberta, 2004–2007

Procedure	One-Level	Two-Level	Three-Level	Total (N)
Fusion surgeries (n)	639	257	239	1,135
Relative weight, n/N × 100	**56%**	**23%**	**21%**	

Source: Alberta Health Evidence Reviews,[146] Table 17.

Table A22: Cost for C-ADR and Fusion Surgeries in Ontario

		C-ADR			Fusion		
Cost Factor	Levels	Cost, $	Weights1, %	Weights2,[a] %	Cost, $	Weights3, %	Weights4,[a] %
Hospital + device/instruments		11,255			9,030		
Device/instruments	1-level	3,836	Assumed 100	80	1,925	Assumed 100	56
	2-level	2 × 3,836	Assumed negligible	20	3,012	Assumed negligible	23
	3-level	3 × 3,836	Assumed negligible	Assumed negligible	4,975	Assumed negligible	21
Hospital (only)		With weights1: **7,392** With weights2: **6,652**			With weights3: **7,105** With weights4: **6,214**		

[a]Weights4 are taken from Alberta Health Evidence Reviews.[146] Weights2 are taken from OHIP (lumbar + cervical), Table A13. These are weights assigned to relative proportion of 1-level, 2-level, and 3-level surgeries for calculating average device cost to be deducted from total hospital cost (device included) and produce hospital-only cost (excluding device). In detail, we used the formula Hospital (only) cost = Hospital cost (including device/instruments) – Cost of device/instruments, and in calculating average cost of device/instruments used the relative shares for different number of involvement of levels based on assumptions.

Table A23: Other Sources Explored for Hospital-Only Cost

		C-ADR		Fusion		
Source	Level	Unit Price (USD)	Standard Deviation	Unit Price (USD)	Standard Deviation	Comments
Ament et al,[138] 2014	2-level	11,720		11,720		Prices are expressed in 2012 USD. Reported costs are for fusion, assumed to apply to C-ADR as well
Lewis et al,[131] 2014	1-level	10,498	643	12,248	748	Prices are expressed in 2014 USD

Abbreviations: C-ADR, cervical artificial disc replacement; USD, U.S. dollars.

Table A24: Consultation Fees for Surgeons

	Consultation		Special Surgical Consultation		Repeat Consultation		Specific Assessment	
Consultant	**Code**	**Fee ($)**	**Code**	**Fee ($)**	**Code**	**Fee ($)**	**Code**	**Fee ($)**
Neurosurgeon	A045	121.10	A935	160.00	A046	58.25	A043	58.25
Orthopedic surgeon	A065	83.10	A935	160.00	A066	51.70	A063	42.55
Average		**102.10**		**160.00**		**54.97**		**50.40**

Source: Ontario Health Insurance Plan.

Table A25: Consultations Needed for Each Intervention

			Postsurgical Follow-Up			
	Presurgical		C-ADR		Fusion	
Purpose of Appointment	**No.**	**Cost, $ (Unit Fee × n)**	**No.**	**Cost, $ (Unit Fee × n)**	**No.**	**Cost, $ (Unit Fee × n)**
Consultation	1	102.10	--	--	--	--
Specific assessment	--	--	2	100.80	3	151.20
Total		**102.10**		**100.80**		**151.20**

Abbreviations: C-ADR, cervical artificial disc replacement.
Source: Expert consultants.

Table A26: Imaging Needed for Each Intervention

				Postsurgical Follow-Up			
		Presurgical Diagnostic		C-ADR		Fusion	
Imaging Method	**Unit Cost ($)**	**n**	**Cost ($)**	**n**	**Cost ($)**	**n**	**Cost ($)**
X-ray	180.00	1	180.00	2	360.00	3	540.00
MRI	783.00	1	783.00	0	--	0	--
Total			**963.00**		**360.00**		**540.00**

Abbreviation: C-ADR, cervical artificial disc replacement; MRI, magnetic resonance imaging.
Source: Expert consultants.

Table A27: Sources Explored for Unit Prices of MRI in Ontario

Source	MRI Cost, $[a]	Source
Private provider in Ontario (website)	895	Priced per segment
Health Provider Suggestion Service (website)	420–1100	Least expensive to most expensive (average of the two = $760)
Private provider in U.S. (special price for Canadians, website)	535 USD = 694 CAD	
Average	**783**	

Abbreviation: MRI, magnetic resonance imaging; CAD, Canadian dollars; USD, U.S. dollars.
[a]Prices current as of March 2018.

Table A28: Sources Explored for Unit Prices of X-Ray in Ontario

Source	X-Ray Cost, $	Comments
Private provider in Ontario (website)	220	Priced per segment
Health Provider Suggestion Service (website)	140 (100–260)	Least expensive to most expensive
Average	**180**	

Table A29a: Sources Explored for Cost of Adverse Events

Source	Levels	Type of Adverse Events (treatment)		Cost	Comments
Lewis et al,[131] 2014	1-level	All, perioperative	C-ADR	$10,498 with no complications (SD = 643) $13,268 with complications (SD = 811) Δ = **$2,770**	2014 USD
			ACD(F)	$12,248 with no complications (SD = 748) $16,143 with complications (SD = 1,211) Δ = **$3,895**	
		All, 1-year follow-up		$1,997 with no complications (SD = 1,211) $4,734 with complications (SD = 2,874) Δ = **$2,737**	
Radcliff et al,[135] 2016	1-level	Dysphagia Infection (complex drainage of wound)		**$203** (SD = 26) **$2,336** (SD = 1340)	2014 USD
Overley et al,[140] 2018	2-level		C-ADR	$17,965 with no complications **$8,068** extra for complications	2014 USD Reported costs are for 1-level but assumed same for 2-level surgery in that publication
			ACDF	$13,025 with no complications **$3,961** extra for complications	

Abbreviation: ACD, anterior cervical discectomy (without fusion); ACDF, anterior cervical discectomy and fusion; ACD(F), anterior cervical discectomy (with or without fusion); C-ADR, cervical artificial disc replacement; SD, standard deviation; Δ, difference.

Table A29b: Adjusted Canadian Adverse Event Cost

Type of Adverse Events (treatment)	Cost, $
Dysphagia	129
Infection	1,480

Appendix 10: Primary Economic Evaluation, Scenario Analysis

Table A30: Reference Case Analysis Results for One-Level and Two-Level, Hospital-Payer Perspective[a]

Strategy	Average Total Costs, $	Incremental Cost,[b] $	Average Total Effects, QALYs	Incremental Effect,[c] QALYs	ICER, $/QALY
One-level disease					
Fusion	10,687.35		4.4448		
C-ADR	12,629.39	1,942.04	4.6416	0.1968	9,866.85
Two-level disease					
Fusion	12,746.75		4.4996		
C-ADR	16,244.54	3,497.80	4.7135	0.2139	16,348.41

Abbreviations: C-ADR, cervical artificial disc replacement; ICER, incremental cost-effectiveness ratio; QALY, quality-adjusted life-year.
[a]Results might appear incorrect because of rounding.
[b]Incremental cost = average cost (C-ADR) − average cost (fusion).
[c]Incremental effect = average effect (C-ADR) − average effect (fusion).

Appendix 11: Budget Impact, Scenario Analysis

Table A31: Predicted Yearly Number of Surgeries in Ontario Eligible for Fusion

Year	Surgeries Eligible for Fusion[a]
2018	779–3,407
2019	790–3,455
2020	800–3,499
2021	809–3,538
2022	818–3,578

[a]Incidence rates were based on published research.

Table A32: Predicted Yearly Number of Surgeries in Ontario Eligible for C-ADR

	Surgeries Eligible for C-ADR[a]	
Year	One-Level	Two-Level
2018	222–973	86–375
2019	225–987	87–380
2020	228–999	88–385
2021	231–1,010	89–389
2022	234–1,022	90–393

Abbreviation: C-ADR, cervical artificial disc replacement.
[a]Incidence rates were based on published research.

Table A33: Predicted Number of Patients Receiving C-ADR or Fusion in Ontario Annually Among Patients Eligible for Either Surgery

	Current Scenario				New Scenario							
					Standard Uptake Increase				Quick Uptake Increase			
	One-Level, n		Two-Level, n		One-Level, n		Two-Level, n		One-Level, n		Two-Level, n	
Year	C-ADR	Fusion	C-ADR	Fusion	C-ADR	Fusion	C-ADR	Fusion	C-ADR	Fusion	C-ADR	Fusion
Using lower bound of incidence[a]												
2018	18	204	7	79	22	200	9	77	22	200	9	77
2019	18	207	7	80	56	169	22	65	112	113	44	43
2020	18	210	7	81	114	114	44	44	228	0	88	0
2021	18	213	7	82	173	58	67	22	231	0	89	0
2022	19	215	7	83	234	0	90	0	234	0	90	0
Using upper bound of incidence[a]												
2018	78	895	30	345	97	876	38	337	97	876	38	337
2019	79	908	30	350	247	740	95	285	494	493	190	190
2020	80	919	31	354	500	499	192	193	999	0	385	0
2021	81	929	31	358	758	252	292	97	1,010	0	389	0
2022	82	940	31	362	1,022	0	393	0	1,022	0	393	0

Abbreviations: C-ADR, cervical artificial disc replacement.
[a]Incidence rates were based on published research.

Table A34: Net Budget Impact for C-ADR Versus Fusion for Cervical Degenerative Disc Disease

	Total Cost, $[a]					
Scenario	**2018**	**2019**	**2020**	**2021**	**2022**	**5-Yr Total**
One-level disease						
New Scenario 1: standard uptake increase						
Using lower bound	9,657	78,047	201,826	333,307	472,550	**1,095,387**
Using upper bound	45,872	345,049	882,989	1,455,799	2,066,031	**4,795,740**
New Scenario 2: quick uptake increase						
Using lower bound	9,657	193,063	441,494	458,028	472,550	**1,574,793**
Using upper bound	45,872	852,353	1,932,064	1,997,692	2,066,031	**6,894,012**
Two-level disease						
New Scenario 1: standard uptake increase						
Using lower bound	10,920	74,249	172,066	261,081	336,374	**854,689**
Using upper bound	43,679	321,744	748,720	1,135,701	1,467,077	**3,716,921**
New Scenario 2: quick uptake increase						
Using lower bound	10,920	183,147	376,685	356,810	336,374	**1,263,936**
Using upper bound	43,679	791,986	1,646,254	1,557,781	1,467,077	**5,506,777**

[a]In 2018 Canadian dollars, derived from rates in published research.

Table A35: Net Budget Impact for C-ADR Versus Fusion for Cervical Degenerative Disc Disease

	Total Cost, $[a]					
Scenario	**2018**	**2019**	**2020**	**2021**	**2022**	**5-Yr Total**
One-level disease						
New Scenario 1: standard uptake increase	7,399	78,923	197,307	320,624	443,940	**1,048,192**
New Scenario 2: quick uptake increase	7,399	194,840	431,609	439,008	443,940	**1,516,796**
Two-level disease						
New Scenario 1: standard uptake increase	5,345	64,135	160,338	267,230	374,122	**871,171**
New Scenario 2: quick uptake increase	5,345	160,338	358,089	363,433	374,122	**1,261,327**

[a]Perspective of hospital payers expressed in 2018 Canadian dollars.

Appendix 12: Letter of Information

Ontario
Health Quality Ontario

LETTER OF INFORMATION

Health Quality Ontario is conducting a review of **Cervical Artificial Disc Replacement** for patients with **degenerative disc disease**. The purpose is to understand whether this procedure should be more broadly funded in Ontario.

An important part of this review involves speaking to patients and families of those who have experience with degenerative disc disease, who may or may not have had surgical treatment. Our goal is to make sure the experiences of patients and caregivers are considered in the funding recommendations for this procedure.

WHAT DO YOU NEED FROM ME?

- ✓ 20-40 minutes of your time for a phone or in-person interview to share your story
- ✓ Permission to audio- (not video-) record the interview, if possible

WHAT YOUR PARTICIPATION INVOLVES

If you agree to share your experiences, you will be asked to have an interview with Health Quality Ontario staff. The interview will likely last 20-40 minutes. It will be held in a private location or over the telephone. With your consent, the interview will be audio-taped. The interviewer will ask you questions about you or your loved one's condition and your perspectives about degenerative disc disease and treatment options in Ontario. If you or your loved one have had surgical treatment for degenerative disc disease, including Cervical Artificial Disc Replacement, the interviewer will ask some additional questions surrounding the procedure.

Participation is voluntary. You may refuse to participate, refuse to answer any questions or withdraw before your interview. Withdrawal will in no way affect care you receive.

CONFIDENTIALITY

All information collected for the review will be kept confidential and privacy will be protected except as required by law. The results of this review will be published, however no identifying information will be released or published. Any records containing information from your interview will be stored securely.

RISKS TO PARTICIPATION:

There are no known physical risks to participating. Some participants may experience discomfort or anxiety after speaking about their lived experience. If this is the case, please contact any staff.

If you are interested in participating, please contact Health Quality Ontario staff:

Appendix 13: Interview Guide

Ontario
Health Quality Ontario

Interview for Cervical Artificial Disc Replacement HTA

Intro
Explain HQO purpose, HTA process, and purpose of interview
History of Degenerative Disc Disease - diagnosis and background (general only)

Lived- Experience
Day-to-day routine
How active is patient? What is the impact of DDD on quality of life?
(driving? Loss of independence?)
Most intrusive aspect of condition, most distressing?
Impact on loved-ones/caregivers, work, etc?

Therapies
What current therapies/treatments are used and their impact?
Is accessibility to therapies/treatments an issue (are you able to take advantage of all potential therapies?)
Expectations of current therapies?

Surgery
(What was the primary benefit you were looking for with surgery?) (if applicable)
Examples: alleviation of pain, improvement in mobility, hand function
Previous surgery (fusion, for example)? How long ago, did it meet expectations?
Post-surgery devices needed? Cost associated with this?
Result, impact, change in quality of life (if applicable)?

Cervical Artificial Disc Replacement
Information surrounding this procedure (or other surgical procedures)?
Decision-making for treatment. Was it difficult to weigh potential risks/benefits?
Expectations
Description of the procedure (if applicable)
Result, impact, change in quality of life (if applicable)

REFERENCES

(1) Prescher A. Anatomy and pathology of the aging spine. Eur J Radiol. 1998;27(3):181-95.
(2) Edwards CC 2nd, Riew KD, Anderson PA, Hilibrand AS, Vaccaro AF. Cervical myelopathy. Current diagnostic and treatment strategies. Spine J. 2003;3(1):68-81.
(3) Radhakrishnan K, Litchy WJ, O'Fallon WM, Kurland LT. Epidemiology of cervical radiculopathy: a population-based study from Rochester, Minnesota, 1976 through 1990. Brain. 1994;117:325-35.
(4) Moore AP, Blumhardt LD. A prospective survey of the causes of non-traumatic spastic paraparesis and tetraparesis in 585 patients. Spinal Cord. 1997;35(6):361-7.
(5) Nouri A, Tetreault L, Singh A, Karadimas SK, Fehlings MG. Degenerative cervical myelopathy: epidemiology, genetics, and pathogenesis. Spine (Phila Pa 1976). 2015;40(12):E675-93.
(6) Boogaarts HD, Bartels RH. Prevalence of cervical spondylotic myelopathy. Eur Spine J. 2015;24 Suppl 2:139-41.
(7) Wu JC, Ko CC, Yen YS, Huang WC, Chen YC, Liu L, et al. Epidemiology of cervical spondylotic myelopathy and its risk of causing spinal cord injury: a national cohort study. Neurosurg Focus. 2013;35(1):E10.
(8) Tetreault L, Goldstein CL, Arnold P, Harrop J, Hilibrand A, Nouri A, et al. Degenerative cervical myelopathy: a spectrum of related disorders affecting the aging spine. Neurosurgery. 2015;77 Suppl 4:S51-67.
(9) Gore DR. Roentgenographic findings in the cervical spine in asymptomatic persons: a ten-year follow-up. Spine (Phila Pa 1976). 2001;26(22):2463-6.
(10) Gore DR, Sepic SB, Gardner GM. Roentgenographic findings of the cervical spine in asymptomatic people. Spine (Phila Pa 1976). 1986;11(6):521-4.
(11) Harrison DE, Bula JM, Gore DR. Roentgenographic findings in the cervical spine in asymptomatic persons: a 10-year follow-up. Spine (Phila Pa 1976). 2002;27(11):1249-50.
(12) Yalamanchili PK, Vives MJ, Chaudhary SB. Cervical spondylotic myelopathy: factors in choosing the surgical approach. Adv Orthop. 2012;2012:783762.
(13) Shriver MF, Lewis DJ, Kshettry VR, Rosenbaum BP, Benzel EC, Mroz TE. Pseudoarthrosis rates in anterior cervical discectomy and fusion: a meta-analysis. Spine J. 2015;15(9):2016-27.
(14) Kaiser MG, Haid RW Jr, Subach BR, Barnes B, Rodts GE Jr. Anterior cervical plating enhances arthrodesis after discectomy and fusion with cortical allograft. Neurosurgery. 2002;50(2):229-36; discussion 36-8.
(15) Goffin J, Geusens E, Vantomme N, Quintens E, Waerzeggers Y, Depreitere B, et al. Long-term follow-up after interbody fusion of the cervical spine. J Spinal Disord Tech. 2004;17(2):79-85.
(16) Hilibrand AS, Carlson GD, Palumbo MA, Jones PK, Bohlman HH. Radiculopathy and myelopathy at segments adjacent to the site of a previous anterior cervical arthrodesis. J Bone Joint Surg Am. 1999;81(4):519-28.
(17) Lee JC, Lee SH, Peters C, Riew KD. Adjacent segment pathology requiring reoperation after anterior cervical arthrodesis: the influence of smoking, sex, and number of operated levels. Spine (Phila Pa 1976). 2015;40(10):E571-7.
(18) Litrico S, Lonjon N, Riouallon G, Cogniet A, Launay O, Beaurain J, et al. Adjacent segment disease after anterior cervical interbody fusion: a multicenter retrospective study of 288 patients with long-term follow-up. Orthop Traumatol Surg Res. 2014;100(6 Suppl):S305-9.

(19) Xu R, Bydon M, Macki M, De la Garza-Ramos R, Sciubba DM, Wolinsky JP, et al. Adjacent segment disease after anterior cervical discectomy and fusion: clinical outcomes after first repeat surgery versus second repeat surgery. Spine (Phila Pa 1976). 2014;39(2):120-6.
(20) Heller JG, Sasso RC, Papadopoulos SM, Anderson PA, Fessler RG, Hacker RJ, et al. Comparison of BRYAN cervical disc arthroplasty with anterior cervical decompression and fusion: clinical and radiographic results of a randomized, controlled, clinical trial. Spine (Phila Pa 1976). 2009;34(2):101-7.
(21) Mummaneni PV, Burkus JK, Haid RW, Traynelis VC, Zdeblick TA. Clinical and radiographic analysis of cervical disc arthroplasty compared with allograft fusion: a randomized controlled clinical trial. J Neurosurg Spine. 2007;6(3):198-209.
(22) Murrey D, Janssen M, Delamarter R, Goldstein J, Zigler J, Tay B, et al. Results of the prospective, randomized, controlled multicenter Food and Drug Administration investigational device exemption study of the ProDisc-C total disc replacement versus anterior discectomy and fusion for the treatment of 1-level symptomatic cervical disc disease. Spine J. 2009;9(4):275-86.
(23) Care OMoHaL-T. Quality -based pathway clinical handbook for non-emergent integrated spine care [Internet]. Toronto Ontario: Ontario Ministry of Health and Long-Term Care; 2017 2017]. Available from: http://www.health.gov.on.ca/en/pro/programs/ecfa/docs/hb_spine.pdf
(24) Canadian Spine Outcomes and Research Network. 2017 annual report [Internet]. Markdale (ON): Canadian Spine Society; 2017 [cited 2018 Jan 15]. Available from: http://spinecanada.ca/wp-content/uploads/2018/02/CSORN-Annual-Report-2017.pdf
(25) Rampersaud YR, Hall H. Referral criteria for non-emergent spinal symptoms in the neck and low back; a survey of Canadian spine surgeons. J Curr Clin Care. 2016;6(6):21-34.
(26) McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS Peer Review of Electronic Search Strategies: 2015 guideline statement. J Clin Epidemiol. 2016;75:40-6.
(27) Greenland S. Quantitative methods in the review of epidemiologic literature. Epidemiol Rev. 1987;9:1-30.
(28) Furlan AD, Malmivaara A, Chou R, Maher CG, Deyo RA, Schoene M, et al. 2015 updated method guideline for systematic reviews in the Cochrane Back and Neck Group. Spine (Phila Pa 1976). 2015;40(21):1660-73.
(29) Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol. 2007;7:10.
(30) Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol. 2009;62(10):1013-20.
(31) Guyatt GH, Oxman AD, Schunemann HJ, Tugwell P, Knottnerus A. GRADE guidelines: a new series of articles in the Journal of Clinical Epidemiology. J Clin Epidemiol. 2011;64(4):380-2.
(32) Moher D, Liberati A, Tetzlaff J, Altman DG, the PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(6):e1000097.
(33) Anderson PA, Nassr A, Currier BL, Sebastian AS, Arnold PM, Fehlings MG, et al. Evaluation of adverse events in total disc replacement: a meta-analysis of FDA summary of safety and effectiveness data. Global Spine J. 2017;7(1 Suppl):76S-83S.
(34) Lehman R, Bevevino AJ, Brewer DD, Skelly AC, Anderson PA. A systematic review of cervical artificial disc replacement wear characteristics and durability. Evid. 2012;3(S1):31-8.

(35) Veruva SY, Steinbeck MJ, Toth J, Alexander DD, Kurtz SM. Which design and biomaterial factors affect clinical wear performance of total disc replacements? A systematic review. Clin Orthop Relat Res. 2014;472(12):3759-69.
(36) Coric D, Nunley PD, Guyer RD, Musante D, Carmody CN, Gordon CR, et al. Prospective, randomized, multicenter study of cervical arthroplasty: 269 patients from the Kineflex, C artificial disc investigational device exemption study with a minimum 2-year follow-up. Clinical article. J Neurosurg Spine. 2011;15(4):348-58.
(37) Gornet MF, Burkus JK, Shaffrey ME, Argires PJ, Nian H, Harrell FE, Jr. Cervical disc arthroplasty with PRESTIGE LP disc versus anterior cervical discectomy and fusion: a prospective, multicenter investigational device exemption study. J Neurosurg Spine. 2015:1-16.
(38) Hisey MS, Bae HW, Davis R, Gaede S, Hoffman G, Kim K, et al. Multi-center, prospective, randomized, controlled investigational device exemption clinical trial comparing Mobi-C Cervical Artificial Disc to anterior discectomy and fusion in the treatment of symptomatic degenerative disc disease in the cervical spine. Int J Spine Surg. 2014;8(7).
(39) Phillips FM, Geisler F, Chaput CD, De Vine JG, Gilder K, Reah CJ, et al. A prospective, randomized clinical investigation of the PCM cervical disc: five-year results from the US IDE study. Spine J. 2013;1:92S.
(40) Vaccaro A, Beutler W, Peppelman W, Marzluff JM, Highsmith J, Mugglin A, et al. Clinical outcomes with selectively constrained SECURE-C cervical disc arthroplasty: two-year results from a prospective, randomized, controlled, multicenter investigational device exemption study. Spine (Phila Pa 1976). 2013;38(26):2227-39.
(41) Davis RJ, Kim KD, Hisey MS, Hoffman GA, Bae HW, Gaede SE, et al. Cervical total disc replacement with the Mobi-C cervical artificial disc compared with anterior discectomy and fusion for treatment of 2-level symptomatic degenerative disc disease: a prospective, randomized, controlled multicenter clinical trial. Clinical article. J Neurosurg Spine. 2013;19(5):532-45.
(42) Gornet MF, Lanman TH, Burkus JK, Hodges SD, McConnell JR, Dryer RF, et al. Cervical disc arthroplasty with the Prestige LP disc versus anterior cervical discectomy and fusion, at 2 levels: results of a prospective, multicenter randomized controlled clinical trial at 24 months. J Neurosurg Spine. 2017;26(6):653-67.
(43) Chen Y, Wang X, Lu X, Yang H, Chen D. Cervical disk arthroplasty versus ACDF for preoperative reducible kyphosis. Orthopedics. 2013;36(7):e958-65.
(44) Cheng L, Nie L, Zhang L, Hou Y. Fusion versus Bryan cervical disc in two-level cervical disc disease: a prospective, randomised study. Int Orthop. 2009;33(5):1347-51.
(45) Cincu R, Lorente F de A, Gomez J, Eiras J, Agrawal A. Long term preservation of motion with artificial cervical disc implants: a comparison between cervical disc replacement and rigid fusion with cage. Asian J Neurosurg. 2014;9(4):213-7.
(46) Donk RD, Verhagen WIM, Hosman AJF, Verbeek A, Bartels R. Symptomatic adjacent segment disease after anterior cervical discectomy for single-level degenerative disk disease. Clin Spine Surg. 2017;31(1):E50-4.
(47) Hou Y, Nie L, Pan X, Si M, Han Y, Li J, et al. Effectiveness and safety of Mobi-C for treatment of single-level cervical disc spondylosis: a randomised control trial with a minimum of five years of follow-up. Bone Joint J. 2016;98-B(6):829-33.
(48) Nabhan A, Ahlhelm F, Pitzen T, Steudel WI, Jung J, Shariat K, et al. Disc replacement using Pro-Disc C versus fusion: a prospective randomised and controlled radiographic and clinical study. Eur Spine J. 2007;16(3):423-30.
(49) Nabhan A, Ishak B, Steudel WI, Ramadhan S, Steimer O. Assessment of adjacent-segment mobility after cervical disc replacement versus fusion: RCT with 1 year's results. Eur Spine J. 2011;20(6):934-41.

(50) Pandey PK, Pawar I, Gupta J, Verma RR. Comparison of outcomes of single-level anterior cervical discectomy with fusion and single-level artificial cervical disc replacement for single-level cervical degenerative disc disease. Spine (Phila Pa 1976). 2017;42(1):E41-9.
(51) Rozankovic M, Marasanov SM, Vukic M. Cervical disk replacement with discover versus fusion in a single-level cervical disk disease: a prospective single-center randomized trial with a minimum 2-year follow-up. Clin Spine Surg. 2017;30(5):E515-22.
(52) Skeppholm M, Lindgren L, Henriques T, Vavruch L, Lofgren H, Olerud C. The Discover artificial disc replacement versus fusion in cervical radiculopathy: a randomized controlled outcome trial with 2-year follow-up. Spine J. 2015;15(6):1284-94.
(53) Sundseth J, Fredriksli OA, Kolstad F, Johnsen LG, Pripp AH, Andresen H, et al. The Norwegian Cervical Arthroplasty Trial (NORCAT): 2-year clinical outcome after single-level cervical arthroplasty versus fusion--a prospective, single-blinded, randomized, controlled multicenter study. Eur Spine J. 2017;26(4):1225-35.
(54) Zhang HX, Shao YD, Chen Y, Hou Y, Cheng L, Si M, et al. A prospective, randomised, controlled multicentre study comparing cervical disc replacement with anterior cervical decompression and fusion. Int Orthop. 2014;38(12):2533-41.
(55) Zhang X, Zhang X, Chen C, Zhang Y, Wang Z, Wang B, et al. Randomized, controlled, multicenter, clinical trial comparing BRYAN cervical disc arthroplasty with anterior cervical decompression and fusion in China. Spine (Phila Pa 1976). 2012;37(6):433-8.
(56) Phillips FM, Lee JYB, Geisler FH, Cappuccino A, Chaput CD, Devine JG, et al. A prospective, randomized, controlled clinical investigation comparing PCM cervical disc arthroplasty with anterior cervical discectomy and fusion: 2-year results from the US FDA IDE clinical trial. Spine (Phila Pa 1976). 2013;38(15):E907-18.
(57) Murrey DB, Janssen ME, Odum SM, Gottlieb JR, Spector LR, Darden BV. Two-year results of a randomized controlled clinical trial comparing ProDisc-C and anterior cervical discectomy and fusion. SAS J. 2008;2(2):76-85.
(58) Gornet MF, McConnell JR, Burkus JK, Lanman TH, Dryer RF, Hodges SD, et al. Two-level cervical disc arthroplasty with prestige LP disc versus ACDF: A prospective, randomized, controlled multicenter clinical trial with 24-month results. Spine Journal. 2015;15 (10 Supplement 1):130S.
(59) Gornet MF, Burkus JK, Shaffrey ME, Argires PJ, Nian H, Harrell FE Jr. Cervical disc arthroplasty with PRESTIGE LP disc versus anterior cervical discectomy and fusion: a prospective, multicenter investigational device exemption study. J Neurosurg Spine. 2015:1-16.
(60) Bae HW, Kim KD, Nunley PD, Jackson RJ, Hisey MS, Davis RJ, et al. Comparison of clinical outcomes of 1-and 2-level total disc replacement: four-year results from a prospective, randomized, controlled, multicenter IDE clinical trial. Spine (Phila Pa 1976). 2015;40(11):759-66.
(61) Zigler JE, Rogers RW, Ohnmeiss DD. Comparison of 1-level versus 2-level anterior cervical discectomy and fusion: clinical and radiographic follow-up at 60 months. Spine (Phila Pa 1976). 2016;41(6):463-9.
(62) Cheng L, Nie L, Li M, Huo Y, Pan X. Superiority of the Bryan disc prosthesis for cervical myelopathy: a randomized study with 3-year followup. Clin Orthop Relat Res. 2011;469(12):3408-14.
(63) Riew KD, Buchowski JM, Sasso R, Zdeblick T, Metcalf NH, Anderson PA. Cervical disc arthroplasty compared with arthrodesis for the treatment of myelopathy. J Bone Joint Surg Am. 2008;90(11):2354-64.
(64) McAfee PC, Cunningham BW, Devine J, Williams E, Yu-Yahiro J. Classification of heterotopic ossification (HO) in artificial disk replacement. J Spinal Disord Tech. 2003;16(4):384-9.

(65) Coric D, Cassis J, Carew JD, Boltes MO. Prospective study of cervical arthroplasty in 98 patients involved in 1 of 3 separate investigational device exemption studies from a single investigational site with a minimum 2-year follow-up. Clinical article. J Neurosurg Spine. 2010;13(6):715-21.
(66) Coric D, Kim PK, Clemente JD, Boltes MO, Nussbaum M, James S. Prospective randomized study of cervical arthroplasty and anterior cervical discectomy and fusion with long-term follow-up: results in 74 patients from a single site. Presented at the 2012 Joint Spine Section Meeting. Clinical article. J Neurosurg Spine. 2013;18(1):36-42.
(67) Skeppholm M, Svedmark P, Noz ME, Maguire GQ, Olivecrona H, Olerud C. Evaluation of mobility and stability in the Discover artificial disc: an in vivo motion study using high-accuracy 3D CT data. J Neurosurg Spine. 2015;23(3):383-9.
(68) McAfee PC, Cappuccino A, Cunningham BW, Devine JG, Phillips FM, Regan JJ, et al. Lower incidence of dysphagia with cervical arthroplasty compared with ACDF in a prospective randomized clinical trial. J Spinal Disord Tech. 2010;23(1):1-8.
(69) Segebarth B, Datta JC, Darden B, Janssen ME, Murrey DB, Rhyne A, et al. Incidence of dysphagia comparing cervical arthroplasty and ACDF. SAS J. 2010;4(1):3-8.
(70) Smucker JD, Bassuener SR, Sasso RC, Riew KD. Comparison of long-term differences in dysphagia: cervical arthroplasty and anterior cervical fusion. Clin Spine Surg. 2016;27:27.
(71) Skeppholm M, Olerud C. Comparison of dysphagia between cervical artificial disc replacement and fusion: data from a randomized controlled study with two years of follow-up. Spine (Phila Pa 1976). 2013;38(24):E1507-10.
(72) Nabhan A, Ahlhelm F, Shariat K, Pitzen T, Steimer O, Steudel WI, et al. The ProDisc-C prosthesis: clinical and radiological experience 1 year after surgery. Spine (Phila Pa 1976). 2007;32(18):1935-41.
(73) Nabhan A, Steudel WI, Nabhan A, Pape D, Ishak B. Segmental kinematics and adjacent level degeneration following disc replacement versus fusion: RCT with three years of follow-up. J Long Term Eff Med Implants. 2007;17(3):229-36.
(74) Qizhi S, Lei S, Peijia L, Hanping Z, Hongwei H, Junsheng C, et al. A comparison of zero-profile devices and artificial cervical disks in patients with 2 noncontiguous levels of cervical spondylosis. J Spinal Disord Tech. 2016;29(2):E61-6.
(75) Skeppholm M, Fransson R, Hammar M, Olerud C. The association between preoperative mental distress and patient-reported outcome measures in patients treated surgically for cervical radiculopathy. Spine J. 2017;17(6):790-8.
(76) Sundseth J, Jacobsen EA, Kolstad F, Sletteberg RO, Nygaard OP, Johnsen LG, et al. Heterotopic ossification and clinical outcome in nonconstrained cervical arthroplasty 2 years after surgery: the Norwegian Cervical Arthroplasty Trial (NORCAT). Eur Spine J. 2016;25(7):2271-8.
(77) Gornet MF, Burkus JK, Shaffrey ME, Nian H, Harrell FE. Cervical disc arthroplasty with Prestige LP disc versus anterior cervical discectomy and fusion: seven-year outcomes. 30th Annual Meeting of the North American Spine Society. Spine J. 2015;15(10 Suppl 1):127S-8S.
(78) Steinmetz MP, Patel R, Traynelis V, Resnick DK, Anderson PA. Cervical disc arthroplasty compared with fusion in a workers' compensation population. Neurosurgery. 2008;63(4):741-7.
(79) Kelly MP, Mitchell MD, Hacker RJ, Riew KD, Sasso RC. Single-level degenerative cervical disc disease and driving disability: results from a prospective, randomized trial. Global Spine J. 2013;3(4):237-42.
(80) Liu JJ, Cadena G, Panchal RR, Schrot RJ, Kim KD. Relief of cervicogenic headaches after single-level and multilevel anterior cervical diskectomy: a 5-year post hoc analysis. Global Spine J. 2016;6(6):563-70.

(81) Schrot RJ, Mathew JS, Li Y, Beckett L, Bae HW, Kim KD. Headache relief after anterior cervical discectomy: post hoc analysis of a randomized investigational device exemption trial. Clinical article. J Neurosurg Spine. 2014;21(2):217-22.
(82) Delamarter RB, Murrey D, Janssen ME, Goldstein JA, Zigler J, Tay BKB, et al. Results at 24 months from the prospective, randomized, multicenter Investigational Device Exemption trial of ProDisc-C versus anterior cervical discectomy and fusion with 4-year follow-up and continued access patients. SAS J. 2010;4(4):122-8.
(83) Radcliff K, Coric D, Albert T. Five-year clinical results of cervical total disc replacement compared with anterior discectomy and fusion for treatment of 2-level symptomatic degenerative disc disease: a prospective, randomized, controlled, multicenter investigational device exemption clinical trial. J Neurosurg Spine. 2016;25(2):213-24.
(84) Radcliff K, Davis RJ, Hisey MS, Nunley PD, Hoffman GA, Jackson RJ, et al. Long-term evaluation of cervical disc arthroplasty with the Mobi-C© cervical disc: a randomized, prospective, multicenter clinical trial with seven-year follow-up. Int J Spine Surg. 2017;11:31.
(85) Schroeder GD, Coric D, Kim HJ, Albert TJ, Radcliff KE. Are patient-reported outcomes predictive of patient satisfaction 5 years after anterior cervical spine surgery? Spine J. 2017;17(7):943-52.
(86) Zoega B, Karrholm J, Lind B. Outcome scores in degenerative cervical disc surgery. Eur Spine J. 2000;9(2):137-43.
(87) Hacker RJ. Cervical disc arthroplasty: a controlled randomized prospective study with intermediate follow-up results. Invited submission from the joint section meeting on disorders of the spine and peripheral nerves, March 2005. J Neurosurg. 2005;3(6):424-8.
(88) Hisey MS, Zigler JE, Jackson R, Nunley PD, Bae HW, Kim KD, et al. Prospective, randomized comparison of one-level Mobi-C cervical total disc replacement vs. anterior cervical discectomy and fusion: results at 5-year follow-up. Int J Spine Surg. 2016;10:10.
(89) Phillips FM, Geisler FH, Gilder KM, Reah C, Howell KM, McAfee PC. Long-term outcomes of the US FDA IDE prospective, randomized controlled clinical trial comparing PCM cervical disc arthroplasty with anterior cervical discectomy and fusion. Spine (Phila Pa 1976). 2015;40(10):674-83.
(90) Gornet MF, Burkus JK, Shaffrey ME, Nian H, Harrell FE Jr. Cervical disc arthroplasty with Prestige LP disc versus anterior cervical discectomy and fusion: seven-year outcomes. Int J Spine Surg. 2016;10:24.
(91) Zigler JE, Delamarter R, Murrey D, Spivak J, Janssen M. ProDisc-C and anterior cervical discectomy and fusion as surgical treatment for single-level cervical symptomatic degenerative disc disease: five-year results of a Food and Drug Administration study. Spine (Phila Pa 1976). 2013;38(3):203-9.
(92) Janssen ME, Zigler JE, Spivak JM, Delamarter RB, Darden BV 2nd, Kopjar B. ProDisc-C total disc replacement versus anterior cervical discectomy and fusion for single-level symptomatic cervical disc disease: seven-year follow-up of the prospective randomized US Food and Drug Administration investigational device exemption study. J Bone Joint Surg Am. 2015;97(21):1738-47.
(93) Davis RJ, Nunley PD, Kim KD, Hisey MS, Jackson RJ, Bae HW, et al. Two-level total disc replacement with Mobi-C cervical artificial disc versus anterior discectomy and fusion: a prospective, randomized, controlled multicenter clinical trial with 4-year follow-up results. J Neurosurg Spine. 2015;22(1):15-25.
(94) Lanman TH, Burkus JK, Dryer RG, Gornet MF, McConnell J, Hodges SD. Long-term clinical and radiographic outcomes of the Prestige LP artificial cervical disc replacement at 2 levels: results from a prospective randomized controlled clinical trial. J Neurosurg Spine. 2017;27(1):7-19.

(95) Sasso RC, Anderson PA, Riew KD, Heller JG. Results of cervical arthroplasty compared with anterior discectomy and fusion: four-year clinical outcomes in a prospective, randomized controlled trial. Orthopedics. 2011;34(11):889.
(96) Coric D, Guyer RD, Nunley PD, Musante D, Carmody C, Gordon C, et al. Prospective, randomized multicenter study of cervical arthroplasty versus anterior cervical discectomy and fusion: 5-year results with a metal-on-metal artificial disc. J Neurosurg Spine. 2018:1-10.
(97) Burkus JK, Traynelis VC, Haid RW, Mummaneni PV. Clinical and radiographic analysis of an artificial cervical disc: 7-year follow-up from the Prestige prospective randomized controlled clinical trial. J Neurosurg Spine. 2014;21(4):516-28.
(98) Janssen M, Kopjar B, Zigler JE, Spivak JM, Murrey DB. Long-term efficacy and safety of anterior cervical discectomy and fusion in single-level cervical disk disease: 7 years follow-up of food and drug administration investigational exemption ProDisc-C study. Global Spine J. 2014;4(1 Suppl):1376678.
(99) Park DK, Lin EL, Phillips FM. Index and adjacent level kinematics after cervical disc replacement and anterior fusion: in vivo quantitative radiographic analysis. Spine (Phila Pa 1976). 2011;36(9):721-30.
(100) Auerbach JD, Anakwenze OA, Milby AH, Lonner BS, Balderston RA. Segmental contribution toward total cervical range of motion: a comparison of cervical disc arthroplasty and fusion. Spine (Phila Pa 1976). 2011;36(25):E1593-9.
(101) Powell JW, Sasso RC, Metcalf NH, Anderson PA, Hipp JA. Quality of spinal motion with cervical disk arthroplasty: computer-aided radiographic analysis. J Spinal Disord Tech. 2010;23(2):89-95.
(102) Sasso RC, Best NM. Cervical kinematics after fusion and bryan disc arthroplasty. J Spinal Disord Tech. 2008;21(1):19-22.
(103) Sasso RC, Metcalf NH, Hipp JA, Wharton ND, Anderson PA. Sagittal alignment after Bryan cervical arthroplasty. Spine (Phila Pa 1976). 2011;36(13):991-6.
(104) Sasso RC, Best NM, Metcalf NH, Anderson PA. Motion analysis of Bryan cervical disc arthroplasty versus anterior discectomy and fusion: results from a prospective, randomized, multicenter, clinical trial. J Spinal Disord Tech. 2008;21(6):393-9.
(105) Hisey MS, Bae HW, Davis RJ, Gaede S, Hoffman G, Kim KD, et al. Prospective, randomized comparison of cervical total disk replacement versus anterior cervical fusion. J Spinal Disord Tech. 2015;28(4):E237-43.
(106) Burkus JK, Haid RW Jr, Traynelis VC, Mummaneni PV. Long-term clinical and radiographic outcomes of cervical disc replacement with the prestige disc: results from a prospective randomized controlled clinical trial. Presented at the 2009 joint spine section meeting. J Neurosurg Spine. 2010;13(3):308-18.
(107) Park JJ, Quirno M, Cunningham MR, Schwarzkopf R, Bendo JA, Spivak JM, et al. Analysis of segmental cervical spine vertebral motion after Prodisc-C cervical disc replacement. Spine (Phila Pa 1976). 2010;35(8):E285-9.
(108) Kelly MP, Mok JM, Frisch RF, Tay BK. Adjacent segment motion after anterior cervical discectomy and fusion versus Prodisc-C cervical total disk arthroplasty: analysis from a randomized, controlled trial. Spine (Phila Pa 1976). 2011;36(15):1171-9.
(109) Anakwenze OA, Auerbach JD, Milby AH, Lonner BS, Balderston RA. Sagittal cervical alignment after cervical disc arthroplasty and anterior cervical discectomy and fusion: results of a prospective, randomized, controlled trial. Spine (Phila Pa 1976). 2009;34(19):2001-7.
(110) Peng CWB, Quirnoa M, Bendo JA, Spivak JM, Goldstein JA. Effect of intervertebral disc height on postoperative motion and clinical outcomes after Prodisc-C cervical disc replacement. Spine J. 2009;9(7):551-5.

(111) Jawahar A, Cavanaugh DA, Kerr IEJ, Birdsong EM, Nunley PD. Total disc arthroplasty does not affect the incidence of adjacent segment degeneration in cervical spine: results of 93 patients in three prospective randomized clinical trials. Spine J. 2010;10(12):1043-8.
(112) Nunley PD, Jawahar A, Kerr EJ, Gordon CJ, Cavanaugh DA, Birdsong EM, et al. Factors affecting the incidence of symptomatic adjacent-level disease in cervical spine after total disc arthroplasty: 2- to 4-year follow-up of 3 prospective randomized trials. Spine (Phila Pa 1976). 2012;37(6):445-51.
(113) Garrido BJ, Taha TA, Sasso RC. Clinical outcomes of Bryan cervical disc arthroplasty: a prospective, randomized, controlled, single site trial with 48-month follow-up. J Spinal Disord Tech. 2010;23(6):367-71.
(114) Kettler A, Rohlmann F, Neidlinger-Wilke C, Werner K, Claes L, Wilke HJ. Validity and interobserver agreement of a new radiographic grading system for intervertebral disc degeneration: part II. Cervical spine. Eur Spine J. 2006;15(6):732-41.
(115) Jackson RJ, Davis RJ, Hoffman GA, Bae HW, Hisey MS, Kim KD, et al. Subsequent surgery rates after cervical total disc replacement using a Mobi-C cervical disc prosthesis versus anterior cervical discectomy and fusion: a prospective randomized clinical trial with 5-year follow-up. J Neurosurg Spine. 2016;24(5):734-45.
(116) Blumenthal SL, Ohnmeiss DD, Guyer RD, Zigler JE. Reoperations in cervical total disc replacement compared with anterior cervical fusion: results compiled from multiple prospective Food and Drug Administration investigational device exemption trials conducted at a single site. Spine (Phila Pa 1976). 2013;38(14):1177-82.
(117) Delamarter RB, Zigler J. Five-year reoperation rates, cervical total disc replacement versus fusion: results of a prospective randomized clinical trial. Spine (Phila Pa 1976). 2013;38(9):711-7.
(118) Hacker RJ. Very late complications of cervical arthroplasty: results of two controlled randomized prospective studies. J Neurosurg. 2012;117(2):A420-1.
(119) Kurtz SM, Toth JM, Siskey R, Ciccarelli L, Macdonald D, Isaza J, et al. The latest lessons learned from retrieval analyses of ultra-high molecular weight polyethylene, metal-on-metal, and alternative bearing total disc replacements. Semin Spine Surg. 2012;24(1):57-70.
(120) Van Der Straeten C, Grammatopoulos G, Gill HS, Calistri A, Campbell P, De Smet KA. The 2012 Otto Aufranc Award: the interpretation of metal ion levels in unilateral and bilateral hip resurfacing. Clin Orthop Relat Res. 2013;471(2):377-85.
(121) Lehman R, Bevevino AJ, Brewer DD, Skelly AC, Anderson PA. A systematic review of cervical artificial disc replacement wear characteristics and durability. Evid Based Spine Care J. 2012;3(S1):31-8.
(122) Slim K, Nini E, Forestier D, Kwiatkowski F, Panis Y, Chipponi J. Methodological index for non-randomized studies (minors): development and validation of a new instrument. ANZ J Surg. 2003;73(9):712-6.
(123) Lebl DR, Cammisa FP Jr, Girardi FP, Wright T, Abjornson C. The mechanical performance of cervical total disc replacements in vivo: prospective retrieval analysis of prodisc-C devices. Spine (Phila Pa 1976). 2012;37(26):2151-60.
(124) National Institute for Health and Care Excellence. Process and methods guides. Appendix I quality appraisal checklist–economic evaluations [Internet]. London (UK): The Institute; 2012 [cited 2017 Dec 12]. Available from: https://www.nice.org.uk/process/pmg4/chapter/appendix-i-quality-appraisal-checklist-economic-evaluations
(125) Qureshi S, Goz V, McAnany S, Cho SK, Hecht AC, Delamarter RB, et al. Health state utility of patients with single-level cervical degenerative disc disease: comparison of

anterior cervical discectomy and fusion with cervical disc arthroplasty. J Neurosurg Spine. 2014;20(5):475-9.
(126) Ament JD, Yang Z, Chen Y, Green RS, Kim KD. A novel quality-of-life utility index in patients with multilevel cervical degenerative disc disease: comparison of anterior cervical discectomy and fusion with total disc replacement. Spine (Phila Pa 1976). 2015;40(14):1072-8.
(127) McAnany SJ, Merrill RK, Brochin RL, Overley SC, Kim JS, Qureshi SA. Comparing the 5-year health state utility value of cervical disc replacement and anterior cervical discectomy and fusion. Global Spine J. 2017;8(1):6-10.
(128) Menzin J, Zhang B, Neumann PJ, Lines LM, Polly DW, Barnett-Myers S, et al. A health-economic assessment of cervical disc arthroplasty compared with allograft fusion. Tech Orthop. 2010;25(2):133-7.
(129) Qureshi SA, McAnany S, Goz V, Koehler SM, Hecht AC. Cost-effectiveness analysis: comparing single-level cervical disc replacement and single-level anterior cervical discectomy and fusion. Clinical article. J Neurosurg Spine. 2013;19(5):546-54.
(130) Warren D, Andres T, Hoelscher C, Ricart-Hoffiz P, Bendo J, Goldstein J. Cost-utility analysis modeling at 2-year follow-up for cervical disc arthroplasty versus anterior cervical discectomy and fusion: a single-center contribution to the randomized controlled trial. Int J Spine Surg. 2013;7:e58-66.
(131) Lewis DJ, Attiah MA, Malhotra NR, Burnett MG, Stein SC. Anterior surgical management of single-level cervical disc disease: a cost-effectiveness analysis. Spine (Phila Pa 1976). 2014;39(25):2084-92.
(132) McAnany SJ, Overley S, Baird EO, Cho SK, Hecht AC, Zigler JE, et al. The 5-year cost-effectiveness of anterior cervical discectomy and fusion and cervical disc replacement: a Markov analysis. Spine (Phila Pa 1976). 2014;39(23):1924-33.
(133) Radcliff K, Zigler J, Zigler J. Costs of cervical disc replacement versus anterior cervical discectomy and fusion for treatment of single-level cervical disc disease: an analysis of the Blue Health Intelligence database for acute and long-term costs and complications. Spine (Phila Pa 1976). 2015;40(8):521-9.
(134) Ghori A, Konopka JF, Makanji H, Cha TD, Bono CM. Long term societal costs of anterior discectomy and fusion (ACDF) versus cervical disc arthroplasty (CDA) for treatment of cervical radiculopathy. Int J Spine Surg. 2016;10(1).
(135) Radcliff K, Lerner J, Yang C, Bernard T, Zigler JE. Seven-year cost-effectiveness of ProDisc-C total disc replacement: results from investigational device exemption and post-approval studies. J Neurosurg Spine. 2016;24(5):760-8.
(136) Wiedenhofer B, Nacke J, Stephan M, Richter W, Carstens C, Eichler M. Is total disk replacement a cost-effective treatment for cervical degenerative disk disease? Clin Spine Surg. 2017;30(5):E530-4.
(137) McAnany SJ, Merrill RK, Overley SC, Kim JS, Brochin RL, Qureshi SA. Investigating the 7-year cost-effectiveness of single-level cervical disc replacement compared to anterior cervical discectomy and fusion. Global Spine J. 2018;8(1):32-9.
(138) Ament JD, Yang Z, Nunley P, Stone MB, Kim KD. Cost-effectiveness of cervical total disc replacement vs. fusion for the treatment of 2-level symptomatic degenerative disc disease. JAMA Surg. 2014;149(12):1231-9.
(139) Ament JD, Yang Z, Nunley P, Stone MB, Lee D, Kim KD. Cost utility analysis of the cervical artificial disc vs. fusion for the treatment of 2-level symptomatic degenerative disc disease: 5-year follow-up. Neurosurgery. 2016;79(1):135-45.
(140) Overley SC, McAnany SJ, Brochin RL, Kim JS, Merrill RK, Qureshi SA. The 5-year cost-effectiveness of two-level anterior cervical discectomy and fusion or cervical disc replacement: a Markov analysis. Spine J. 2018;18(1):63-71.

(141) Merrill RK, McAnany SJ, Albert TJ, Qureshi SA. Is two-level cervical disc replacement more cost-effective than anterior cervical discectomy and fusion at 7 years? Spine (Phila Pa 1976). 2018;43(9):610-6.
(142) Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health. 2013;16(2):231-50.
(143) Canadian Agency for Drugs and Technologies in Health (CADTH). Guidelines for the economic evaluation of health technologies: Canada. 4th ed. Ottawa (ON): The Agency; 2017.
(144) Ministry of Health and Long-Term Care. Schedule of benefits: physician services under the Health Insurance Act. Toronto (ON): Queen's Printer for Ontario; 2015. p. Z2-Z11.
(145) Ministry of Health and Long-Term Care. Health Data Branch Web Portal: Ontario Case Costing [Internet]. Toronto (ON): Queen's Printer for Ontario. c2018 [cited 2018 Mar 10]. Available from: https://hsim.health.gov.on.ca/HDBPortal
(146) HTA Unit. Artificial disc arthroplasty (ACDA). Calgary (AB): University of Calgary; 2009. 114 p.
(147) Quirno M, Goldstein JA, Bendo JA, Kim Y, Spivak JM. The incidence of potential candidates for total disc replacement among lumbar and cervical fusion patient populations. Asian Spine J. 2011;5(4):213-9.
(148) Medical Services Advisory Committee (MSAC). Use of artificial disc replacement in patients with cervical degenerative disc disease. Canberra, Australia: MSAC; 2011.
(149) Canadian Institute for Health Information. Discharge Abstract Database Metadata [Internet]. Ottawa (ON): The Institute. c1996-2018 [cited 2018 Mar 10]. Available from: https://www.cihi.ca/en/discharge-abstract-database-metadata
(150) Statistics Canada. Population estimates on July 1st, by age and sex: Table 17-10-0005-01 [Internet]. Ottawa (ON): Statistics Canada. 2018 [cited 2018 Mar 10]. Available from: https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000501
(151) Ministry of Finance. Ontario Population Projections Update, 2017–2041 [Internet]. Toronto (ON): Queen's Printer for Ontario. c2010 [cited 2018 Mar 10]. Available from: https://www.fin.gov.on.ca/en/economy/demographics/projections/table1.html
(152) Boselie TFM, Willems PC, van Mameren H, de Bie R, van Santbrink H. Arthroplasty versus fusion for single-level cervical degenerative disc disease. Cochrane Database Syst Rev. 2016(9).
(153) Barham L. Public and patient involvement at the UK National Institute for Health and Clinical Excellence. Patient. 2011;4(1):1-10.
(154) Messina J, Grainger DL. A pilot study to identify areas for further improvements in patient and public involvement in health technology assessments for medicines. Patient. 2012;5(3):199-211.
(155) Ontario Health Technology Advisory Committee Public Engagement Subcommittee. Public engagement for health technology assessment at Health Quality Ontario—final report from the Ontario Health Technology Advisory Committee Public Engagement Subcommittee [Internet]. Toronto (ON): Queen's Printer for Ontario; 2015 Apr [cited 2018 Apr 30]. Available from: http://www.hqontario.ca/Portals/0/documents/evidence/special-reports/report-subcommittee-20150407-en.pdf
(156) Kvale S. Interviews: an introduction to qualitative research interviewing. Thousand Oaks (CA): Sage Publications; 1996.
(157) Kuzel AJ. Sampling in qualitative inquiry. In: Miller WL, Crabtree BF, editors. Doing qualitative research. Thousand Oaks (CA): Sage Publications; 1999. p. 33-45.

(158) Morse J. Emerging from the data: cognitive processes of analysis in qualitative research. In: Morse J, editor. Critical issues in qualitative research methods. Thousand Oaks (CA): Sage Publications; 1994. p. 23-41.
(159) Patton MQ. Qualitative research and evaluation methods. 3rd ed. Thousand Oaks (CA): Sage Publications; 2002.
(160) Strauss AL, Corbin JM. Basics of qualitative research: techniques and procedures of developing a grounded theory. 2nd ed. Thousand Oaks (CA): Sage Publications; 1998.
(161) Health Technology Assessment International (HTAi). FAQ for HTA agencies and policy makers [Internet]. Edmonton (AB): HTAi; 2015 [cited 2018 Apr 30]. Available from: http://www.htai.org/interest-groups/patient-and-citizen-involvement/resources/for-hta-agencies-and-policy-makers/faq-for-hta-agencies-and-policy-makers.html#c3229
(162) Strauss AL, Corbin JM. Grounded theory research: procedures, canons, and evaluative criteria. Qual Sociol. 1990;13(1):3-21.
(163) Strauss AL, Corbin JM. Grounded theory methodology: an overview. In: Denzin NK, Lincoln YS, editors. Handbook of qualitative research. Thousand Oaks (CA): Sage Publications; 1994. p. 273-85.
(164) Anderson PA, Sasso RC, Riew KD. Comparison of adverse events between the Bryan artificial cervical disc and anterior cervical arthrodesis. Spine (Phila Pa 1976). 2008;33(12):1305-12.
(165) Sasso RC, Smucker JD, Hacker RJ, Heller JG. Artificial disc versus fusion: a prospective, randomized study with 2-year follow-up on 99 patients. Spine (Phila Pa 1976). 2007;32(26):2933-40.
(166) Sasso RC, Smucker JD, Hacker RJ, Heller JG. Clinical outcomes of BRYAN cervical disc arthroplasty: a prospective, randomized, controlled, multicenter trial with 24-month follow-up. J Spinal Disord Tech. 2007;20(7):481-91.
(167) Garrido BJ, Wilhite J, Nakano M, Crawford C, Baldus C, Riew KD, et al. Adjacent-level cervical ossification after Bryan cervical disc arthroplasty compared with anterior cervical discectomy and fusion. J Bone Joint Surg Am. 2011;93(13):1185-9.
(168) Sasso WR, Smucker JD, Sasso MP, Sasso RC. Long-term clinical outcomes of cervical disc arthroplasty: a prospective, randomized, controlled trial. Spine (Phila Pa 1976). 2017;42(4):209-16.
(169) Loumeau TP, Darden BV, Kesman TJ, Odum SM, Van Doren BA, Laxer EB, et al. A RCT comparing 7-year clinical outcomes of one level symptomatic cervical disc disease (SCDD) following ProDisc-C total disc arthroplasty (TDA) versus anterior cervical discectomy and fusion (ACDF). Eur Spine J. 2016;25(7):2263-70.
(170) Porchet F, Metcalf NH. Clinical outcomes with the Prestige II cervical disc: preliminary results from a prospective randomized clinical trial. Neurosurg Focus. 2004;17(3):E6.
(171) U.S. Food and Drug Administration (FDA). Clinical data presentations for orthopedic device applications: guidance for industry and FDA staff. Silver Spring (MD): FDA; 2004.

About Health Quality Ontario

Health Quality Ontario is the provincial lead on the quality of health care. We help nurses, doctors and other health care professionals working hard on the frontlines be more effective in what they do – by providing objective advice and data, and by supporting them and government in improving health care for the people of Ontario.

We focus on making health care more effective, efficient and affordable through a legislative mandate of:

- Reporting to the public, organizations, government and health care providers on how the health system is performing,
- Finding the best evidence of what works, and
- Translating this evidence into clinical standards; recommendations to health care professionals and funders; and tools that health care providers can easily put into practice to make improvements.

Health Quality Ontario is governed by a 12-member Board of Directors with a broad range of expertise – doctors, nurses, patients and from other segments of health care – and appointed by the Minister of Health and Long-Term Care.

In everything it does, Health Quality Ontario brings together those with first-hand experience to hear their experiences and views of how to make them better. We partner with patients, residents, families and caregivers to be full participants in designing our programs and services, to ensure they are aligned to their needs and priorities. We work collaboratively with organizations across the province to encourage the spread of innovative and proven programs to support high quality care, while also saving money and eliminating redundancy. And, we work with clinicians on the frontlines to use their collective wisdom and experience to bring about positive change in areas important to Ontario – such as addressing the challenges of hallway health care and mental health.

> For example, 29 Ontario hospitals participated in a pilot program last year that reduced infections due to surgery by 18% – which in turn reduces the number of patients returning to hospital after surgery and alleviating some of the challenges faced in hallway health care. This program enabled surgeons to see their surgical data and how they perform in relation to each other and to 700 other hospitals worldwide. We then helped them identify and action improvements to care. Forty-six hospitals across Ontario are now part of this program, covering 80% of hospital surgeries.
>
> Health Quality Ontario also develops quality standards for health conditions that demonstrate unnecessary gaps and variations in care across the province, such as in major depression or schizophrenia. Quality standards are based on the best evidence and provide recommendations to government, organizations and clinicians. They also include a guide for patients to help them ask informed questions about their care.
>
> In addition, Health Quality Ontario's health technology assessments use evidence to assess the effectiveness and value for money of new technologies and procedures, and incorporate the views and preferences of patients, to make recommendations to government on whether they should be funded.

> Each year, we also help hospitals, long-term care homes, home care and primary care organizations across the system create and report on the progress of their annual Quality Improvement Plans, which is their public commitment on their priorities to improve health care quality.

Health Quality Ontario is committed to supporting the development of a quality health care system based on six fundamental dimensions: efficient, timely, safe, effective, patient-centred and equitable.

Our goal is to challenge the status quo and to focus on long-lasting pragmatic solutions that improve the health of Ontarians, enhance their experience of care, reduce health care costs, and support the well-being of health care providers. A quality health system results in Ontarians leading healthier and more productive lives, and a vibrant society in which everyone benefits.

About the Ontario Health Technology Advisory Committee

About the Ontario Health Technology Assessment Series

How to Obtain Reports From the Ontario Health Technology Assessment Series

Disclaimer

Health Quality Ontario
130 Bloor Street West, 10th Floor
Toronto, Ontario
M5S 1N5
Tel: 416-323-6868
Toll Free: 1-866-623-6868
Fax: 416-323-9261
Email: EvidenceInfo@hqontario.ca
www.hqontario.ca

ISSN 1915-7398 (online)
ISBN 978-1-4868-3125-8 (PDF)