

Artificial disc replacement (ADR)

Assessing signals for update

January 22, 2016

Health Technology Assessment Program (HTA)

Washington State Health Care Authority

PO Box 42712

Olympia, WA 98504-2712

(360) 725-5126

www.hca.wa.gov/about-hca/health-technology-assessment

shtap@hca.wa.gov

Artificial Disc Replacements (ADR): Assessing Signals for Update

Provided by:



Spectrum Research, Inc.

Prepared by:

Joseph R. Dettori, PhD, MPH

Krystle Pagarigan, BS

January 22, 2016

Table of Contents

1. Introduction	2
HTCC Coverage Determination.....	2
HTCC Reimbursement Determination	2
Committee Conclusions.....	3
2. Purpose of Report	4
3. Methods	5
3.1 Literature Searches.....	5
3.2 Study selection.....	5
4. Results	6
4.1 Search	6
4.2 Identifying signals for re-review	7
5. Conclusions	13
REFERENCES.....	14
APPENDIX A. SEARCH STRATEGIES.....	15
APPENDIX B. SUMMARY OF INCLUDED SYSTEMATIC REVIEWS.....	16
APPENDIX C. SYSTEMATIC REVIEWS EXCLUDED AT FULL TEST REVIEW	17

1. Introduction

A Health Technology Assessment titled: **Artificial Disc Replacement**, was published on September 19, 2008 by the Health Care Authority. Findings and Coverage Decision was released on October 17, 2008 and adopted on March 20, 2009. The Committee's Coverage Decision is summarized below.

HTCC Coverage Determination

Cervical and Lumbar Artificial Disc Replacement is a covered benefit only under criteria identified in the reimbursement determination

HTCC Reimbursement Determination

Limitations of Coverage:

Lumbar Artificial Disc Replacement (L-ADR)

- 1) Patients must first complete a structured, intensive, multi-disciplinary program for management of pain, if covered by the agency;
- 2) Patients must be 60 years or under;
- 3) Patients must meet FDA approved indications for use and not have any contra-indications. FDA approval is device specific but includes:
 - Failure of at least six months of conservative treatment
 - Skeletally mature patient
 - Replacement of a single disc for degenerative disc disease at one level confirmed by patient history and imaging

Artificial Disc Replacement FDA general contra-indications:

Active systemic infection or infection localized to site of implantation
Allergy or sensitivity to implant materials
Certain bone and spine diseases (e.g. osteoporosis, spondylosis)

Cervical Artificial Disc Replacement (C-ADR):

- 1) Patients must meet FDA approved indications for use and not have any contra-indications. FDA approval is device specific but includes:
 - Skeletally mature patient
 - Reconstruction of a disc following single level discectomy for intractable symptomatic cervical disc disease (radiculopathy or myelopathy) confirmed by patient findings and imaging.

Artificial Disc Replacement FDA general contra-indications:

- Active systemic infection or infection localized to site of implantation
- Allergy or sensitivity to implant materials
- Certain bone and spine diseases (e.g. severe spondylosis or marked cervical instability)

Non-Covered Indications

Non-FDA approved uses

Committee Conclusions

Having made findings as to the most significant and relevant evidence regarding health outcomes, key factors and identified evidence related to those factors, primarily based on the evidence based technology assessment report, the committee concludes:

1. Evidence availability and technology features

The committee concludes that the best available evidence on artificial disc replacement has been collected and summarized.

- 1.1. There is moderate evidence from 5 randomized controlled trials and about 40 uncontrolled studies about several important health outcomes for artificial disc replacement. The randomized trials have shared limitations: some methodological flaws, fusion as only comparator, non-inferiority design, lack of long term data, and measure/definition of success.
- 1.2. The controlled studies compare surgical options only. Fusion surgery as a treatment for spine pain is still not established a clearly superior option, so the lack of inclusion of optimized medical management severely limits the results.
- 1.3. As compared to fusion, a currently approved alternative, the overall evidence is moderate and demonstrates at least equivalence of ADR in short term safety and efficacy.
- 1.4. Longer follow up data, especially around safety events and reoperation rates is needed (often this evidence comes from non RCT data such as registries). Also, the post approval FDA studies requiring up to seven year follow up should be monitored.

2. Is it safe?

The committee concludes that the comprehensive evidence reviewed shows that the technology has been proven at least equally safe as a currently offered alternative, fusion. Key factors to the committee's conclusion include:

- 2.1. Moderate evidence demonstrated that L-ADR has a similar safety profile as lumbar anterior or circumferential fusion two years following surgery. Longer term safety on L-ADR is not known.
- 2.2. Moderate evidence demonstrated that C-ADR tends to be safer than fusion as measured by the risk of device failure and surgical complications up to two years following surgery. Longer term safety on C-ADR is not known.

3. Is it effective?

The committee concludes that the comprehensive evidence reviewed shows that the technology has been proven equally or more effective as a currently offered alternative, fusion. Key factors to the committee's conclusion include:

- 3.1. While there is no evidence comparing ADR with non-operative care, there are five moderate quality, controlled studies comparing ADR with a currently performed alternative, fusion. Based on the limited comparator and other evidence limitations, the evidence of efficacy should not be generalized beyond carefully selected patients that match trial and FDA indications.
- 3.2. Moderate evidence demonstrated that the efficacy/effectiveness of L-ADR is comparable with fusion up to two years following surgery based on a composite measure for FDA approval of overall clinical success, pain improvement, an ODI and SF-36 improvement..

- 3.3. Moderate evidence demonstrated that the efficacy/effectiveness of C-ADR is equal to fusion for pain and function and potentially superior to fusion for neurological and overall success up to two years following surgery.
- 3.4. There is insufficient evidence to draw conclusions regarding the safety and efficacy of ADR in special populations or populations outside those studied for FDA approval. Thus, coverage should be limited to studied indications.

4. Is it cost-effective?

The Committee concludes that the comprehensive evidence review does not show that the technology is more cost effective. Although cost-effectiveness was not a major decision factor, the committee concluded cost-effectiveness is unproven because of insufficient evidence.

4.1. The cost analyses were limited by short time horizons, comparators chosen, and differences with US health system, and provided mixed answers. For L-ADR, one assessment showed an increase in cost based on the device cost and another showed similar or possibly reduced cost based primarily on shorter hospital stays for L-ADR. For C-ADR, one cost analysis showed similar surgical costs, but higher total cost with C-ADR due to device cost.

5. Medicare Decision and Expert Treatment Guidelines

The committee deliberations included a discussion of National Medicare Decisions and expert treatment guidelines, and an understanding that the committee must find substantial evidence to support a decision that is contrary. RCW 70.14.110. The independent evidence report identified a national Medicare coverage decision on lumbar fusion and no expert treatment guidelines. The committee's conditional coverage is consistent with the national Medicare decision to not cover L-ADR for patients older than 60 years of age.

2. Purpose of Report

The purpose of this literature update is to determine whether or not there is sufficient evidence published after the original report to conduct a re-review of this technology based on the presence of preset signal criteria. The key questions included the following:

Key question 1

What is the evidence of efficacy and effectiveness of ADR compared with comparative therapies (including non-operative therapy; spinal fusion; other surgery)?

Key Question 2

What is the evidence related to the ADR safety profile? (including device failure, reoperation)

Key Question 3

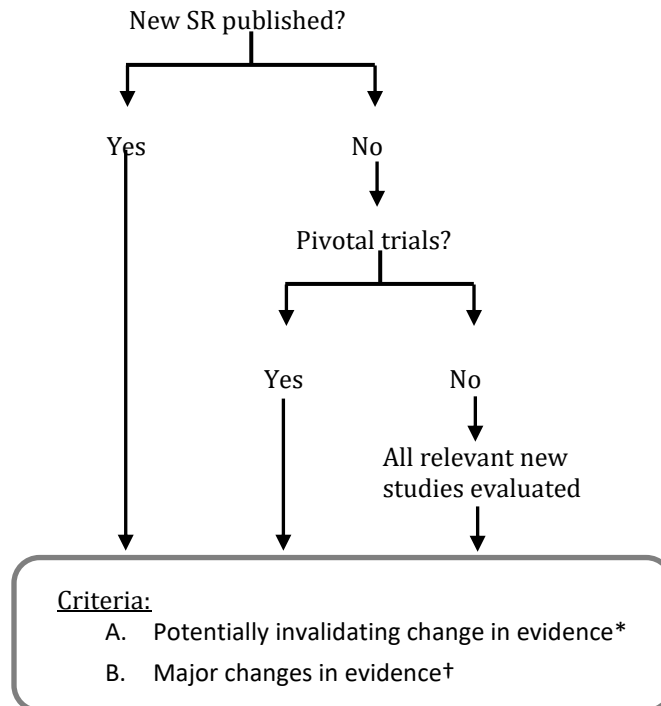
What is the evidence of differential efficacy or safety issues amongst special populations (including but not limited to the elderly and workers compensation populations)?

Key Question 4

What are the cost implications and cost effectiveness for ADR?

3. Methods

To determine the need for systematic review update, the following algorithm was followed:
Figure 1. Algorithm of the modified Ottawa Method of Identifying Signals for SR Updates



- *A-1. Opposing findings: Pivotal trial or SR including at least one new trial that characterized the treatment in terms opposite to those used earlier
- A-2. Substantial harm: Pivotal trial or SR whose results called into question the use of the treatment based on evidence of harm or that did not proscribe use entirely but did potentially affect clinical decision making
- A-3. Superior new treatment: Pivotal trial or SR whose results identified another treatment as significantly superior to the one evaluated in the original review, based on efficacy or harm.
- †B-1. Important changes in effectiveness short of “opposing findings”
- B-2. Clinically important expansion of treatment
- B-3. Clinically important caveat
- B-4. Opposing findings from discordant meta-analysis or nonpivotal trial

3.1 Literature Searches

We conducted a limited electronic literature of Medline for systematic reviews with meta-analysis during the period January 1, 2008 through January 8, 2016 using search terms used for the original report. Appendix A includes the search methodology for this topic. In addition, we searched the FDA website to determine if there was approval of new indications for ADR. Finally, we searched for individual cost-effectiveness studies for KQ 4.

3.2 Study selection

We sought systematic reviews of randomized controlled trials (RCTs) of efficacy and safety with meta-analysis that included articles that met inclusion and exclusion criteria similar to the original report. In addition we sought systematic reviews reflecting updates or new advances for the technology. Secondary to the large number of citations returned, we focused on screening only systematic reviews and meta-analyses of RCTS published between 2012 and 2015. Although quality of systematic reviews

was not formally evaluated for this report, we chose two systematic reviews, one for the lumbar and one for the cervical spine that that were the most comprehensive and of high quality based on the following: report of search strategies (two or more data bases and description of dates searched), number of included relevant RCTs, pre-stated inclusion and exclusion criteria, information on methodologies used for synthesis of data, inclusion of patient reported or safety outcomes and evaluation of the strength of the body of literature using GRADE or another analogous system. A summary of the two SRs is found in Appendix B.

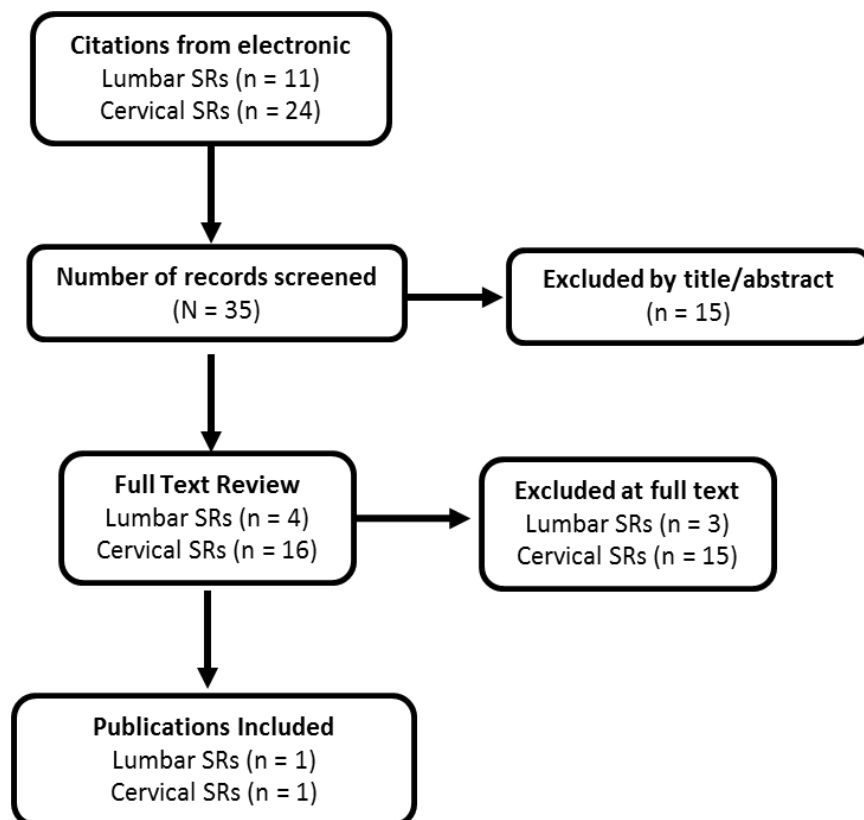
4. Results

4.1 Search

We identified 11 lumbar and 24 cervical systematic reviews from the electronic search that addressed in part or in full key questions 1 and 2, Figure 2. We reviewed the full text of four lumbar and 16 cervical studies. We chose one systematic review for each anatomical region (lumbar and cervical) that we felt most closely met the inclusion criteria (see excluded studies and the reasons for exclusion in Appendix C). There were no systematic reviews on differential efficacy or safety (key questions 3). We found three cervical cost-effectiveness studies (Key Question 4) where there were none in the previous report.

The FDA approved one device (Mobi-C) for two- level cervical disc reconstruction since our initial report.

Figure 2. Electronic search results for systematic reviews



4.2 Identifying signals for re-review

Table 1 shows the original key questions, the conclusions of the original report, the new sources of evidence, the new findings, and the recommendations of Spectrum Research, Inc. (SRI) regarding the need for update.

Table 1. ADR Summary Table for Key Question 1.

Key Question 1. What is the evidence of efficacy and effectiveness of ADR compared with comparative therapies (including nonoperative therapy, spinal fusion, other surgery)?			
Conclusions from CER Executive Summary	New Sources of Evidence	New Findings	Conclusion from SRI
<p><u>L-ADR vs. nonoperative care</u> No evidence available</p>	<p>Systematic Review Jacobs et al¹</p>	<ul style="list-style-type: none"> A systematic review identified one study that compared disc replacement against rehabilitation and found a statistically significant advantage in ODI in favor of surgery, which, however, did not reach the predefined threshold for clinical relevance. 	<p>This section of the report is NOT valid. A new comparison group is added and the report needs updating</p>
<p><u>L-ADR vs. lumbar fusion</u></p> <ul style="list-style-type: none"> There is moderate evidence that the efficacy of L-ADR as measured by the composite measure of overall clinical success, Oswestry Disability Index (ODI) improvement, pain improvement, neurological success, SF-36 improvement, and patient satisfaction is comparable with anterior lumbar interbody fusion or circumferential fusion up to two years following surgery. This evidence is based on two moderate quality randomized controlled trials conducted as FDA Investigational Device Exemption non-inferiority trials. Overall clinical success (a composite measure considering most or all of the following: ODI improvement, device failure, complications, neurological change, SF-36 change and radiographic success) was achieved in 56% of patients receiving L-ADR and 48% receiving lumbar fusion. Though the results suggest that 24 month outcomes for L-ADR are similar to lumbar fusion, it should be noted that a non-inferiority trial requires that the reference 	<p>Systematic Review Jacobs et al¹</p>	<ul style="list-style-type: none"> A systematic review (Jacobs) included 39 publications, describing six unique RCT's. The follow-up of the studies was 24 months, with only one extended to five years. Five studies had a low risk of bias, although there is a risk of bias in the included studies due to sponsoring and absence of any kind of blinding. The six studies found that the mean improvement in VAS back pain was 5.2 mm (of 100 mm) higher (two studies, 676 patients; 95% confidence interval (CI) 0.18 to 10.26) with a low quality of evidence, while from the same studies leg pain showed no difference. The improvement of Oswestry score at 24 months in the disc replacement group was 4.27 points more than in the fusion group (five studies; 1207 patients; 95% CI 1.85 to 6.68) with a low quality of evidence. Both upper bounds of the confidence intervals for VAS back pain and Oswestry score were below the predefined clinically relevant difference. Choice of control group (circumferential or anterior fusion) did not appear to result in different outcomes. 	<p>This section of the report is still valid and does not need updating.</p>

Key Question 1. What is the evidence of efficacy and effectiveness of ADR compared with comparative therapies (including nonoperative therapy, spinal fusion, other surgery)?			
Conclusions from CER Executive Summary	New Sources of Evidence	New Findings	Conclusion from SRI
<p>treatment have an established efficacy or that it is in widespread use. For the lumbar spine, the efficacy of the comparator treatment, lumbar fusion, for degenerative disc disease remains uncertain, especially when it is compared with nonoperative care. Given what is known about lumbar fusion as a comparator and having evidence that only compares L-ADR with lumbar fusion limits the ability to fully answer the efficacy/effectiveness question.</p> <ul style="list-style-type: none"> There are no (medium-) or long-term follow-up data assessing efficacy/effectiveness from the two index RCTs at this time 			
<p><u>C-ADR vs. nonoperative care</u> No evidence available</p>	No new evidence	No new evidence	This section of the report is still valid and does not need updating.
<p><u>C-ADR vs. cervical fusion</u></p> <ul style="list-style-type: none"> There is moderate evidence for the cervical spine that C-ADR is superior to ACDF with respect to overall clinical success (77% versus 68%) and neurological success (92% versus 86%), and is comparable with ACDF with respect to Neck Disability Index (NDI), and pain up to two years following surgery. The evidence is based on two moderate quality randomized controlled FDA Investigational Device Exemption non-inferiority trials. An interim analysis of approximately 65% of a third RCT was reported in an FDA Panel Executive Summary. If the results following completion of the trial are similar to the interim results of that same trial, the confidence in the evidence that C-ADR is superior to ACDF will increase. There is evidence that segmental motion is maintained or improved up to three years in the L-ADR patients and up to four years in C-ADR patients compared with 	<p>Systematic Review Zhang et al²</p>	<p>19 RCTs (n = 4516) Short-term follow-up (2-3 years)</p> <ul style="list-style-type: none"> The C-ADR group had statistically lower NDI scores (SMD, -0.34; 95% CI: -0.68 to 0.00, P = 0.05) than the ACDF group. However, there existed a substantial heterogeneity. In sensitivity analysis, the result also showed that C-ADR group had better NDI scores (SMD, -0.13; 95% CI: -0.25 to -0.02, P = 0.02) compared with ACDF group. The C-ADR group had a statistically higher NDI success rate than the ACDF group (OR, 0.72; 95% CI: 0.54 to 0.95, P = 0.02). A higher neurological success rate was seen in the C-ADR group than in the ACDF group (OR, 0.62; 95% CI: 0.45 to 0.85, P = 0.003). C-ADR group had significantly lower neck pain scores in three studies using numerical rating scales (SMD, -0.14; 95% CI: -0.27 to -0.01) and lower neck (SMD -1.28; 95% 	<p>This section of the report is NOT valid. There are new data for medium-term follow-up of 4-5 years.</p>

Key Question 1. What is the evidence of efficacy and effectiveness of ADR compared with comparative therapies (including nonoperative therapy, spinal fusion, other surgery)?			
Conclusions from CER Executive Summary	New Sources of Evidence	New Findings	Conclusion from SRI
<p>preoperative motion. It is unclear the true extent to which preserving segmental motion by using ADR instead of fusion influences rates of adjacent segment disease (ASD). Whether ASD is a continuation of a disease process necessitating fusion or a result of fusion continues to be disputed. Furthermore, there continues to be debate on whether the presence of ASD is clinically important given that patients with marked radiographic ASD often have no symptoms.</p>		<p>CIO: -2.16 to 0.40) and arm pain scores (SMD -0.19; 95% CI: -0.35 to -0.03) vs. ACDF in three studies using VAS.</p> <ul style="list-style-type: none"> The C-ADR group presented a significantly higher overall composite success rate (OR, 0.59; 95% CI: 0.48 to 0.74, P < 0.00001) <p>Medium-term (4-5 years)</p> <ul style="list-style-type: none"> NDI scores in the C-ADR group were lower than those of the ACDF group in two studies (SMD, -0.31; 95% CI: -0.47 to -0.15, P = 0.0002). Neurological success from two studies occurred more frequently in the C-ADR group than in the ACDF group (OR, 0.55; 95% CI: 0.30 to 1.01, P = 0.05). Neck (SMD, -0.28; 95% CI: -0.44 to -0.12, P = 0.0008) and arm pain scores (SMD, -0.19; 95% CI: -0.35 to -0.03, P = 0.02) were lower in two studies using NRS scores. 	

Table 2. ADR Summary Table for Key Question 2.

Key Question 2: What is the evidence related to the ADR safety profile (including complications, adverse events, device failure, reoperation)?			
Conclusions from CER Executive Summary	New Sources of Evidence	New Findings	Conclusion from SRI
<p><u>L-ADR vs. nonoperative care</u> No evidence available</p>	<p>Systematic Review Jacobs et al¹</p>	<ul style="list-style-type: none"> A systematic review (Jacobs) identified one study that compared disc replacement against rehabilitation. Among those receiving L-ADR, six patients (8%) had complications resulting in impairment at two year follow-up, and the reoperation rate was 6.5% (n=5). 	<p>This section of the report is NOT valid. A new comparison group is added and the report needs updating</p>
<p><u>L-ADR vs. lumbar fusion</u></p> <ul style="list-style-type: none"> There is moderate evidence that L-ADR results in a similar proportion of device-related complications (7 to 18%) compared with lumbar fusion (4 to 20%) There is moderate evidence that L-ADR results in a similar proportion of major complications (0 to 1%) compared with lumbar fusion (0 to 1%) There are no (medium-) or long-term follow-up data assessing safety from the two index RCTs at this time 	<p>Systematic Review Jacobs et al¹</p>	<ul style="list-style-type: none"> There were 63 of 810 (7.8%) re-operations in the total disc replacement group and 35 of 384 (9.1%) in the fusion group. There is very low quality evidence from five studies that the difference in re-operations up to 24 months was not statistically significant. Only one secondary publication of a low risk of bias study reported neurological complications and found no difference between the two groups. There is very low quality evidence from one low risk of bias study that the difference in adjacent segment degeneration at 24 months was not statistically different. This one study only marginally reported adjacent segment degeneration mentioning six of 72 cases of fusion and only one of 80 cases of total disc replacement with adjacent segment problems. There is very low quality of evidence from one low risk of bias study that the occurrence of facet joint degeneration is not statistically significantly different. 	<p>This section of the report is still valid and does not need updating.</p>
<p><u>C-ADR vs. nonoperative care</u> No evidence available</p>	<p>No new evidence</p>	<p>No new evidence</p>	<p>This section of the report is still valid and does not need updating.</p>
<p><u>C-ADR vs. cervical fusion</u></p> <ul style="list-style-type: none"> Complication rates varied among the studies but generally device related or device/surgical 	<p>Systematic Review</p>	<p>Short-term follow-up (2-3 years)</p>	<p>This section of the report is NOT valid. There are new data for</p>

Key Question 2: What is the evidence related to the ADR safety profile (including complications, adverse events, device failure, reoperation)?			
Conclusions from CER Executive Summary	New Sources of Evidence	New Findings	Conclusion from SRI
<p>procedure related complications or adverse events occurred less frequently among the C-ADR patients (5%) than anterior fusion patients (10%).</p> <ul style="list-style-type: none"> There are no (medium-) or (medium-) or long-term follow-up data assessing safety from the five index RCTs at this time 	<p>Zhang et al²</p>	<ul style="list-style-type: none"> Adverse events occurred more frequently in the ACDF group than in the C-ADR group (OR, 0.58; 95% CI: 0.43 to 0.80, P = 0.0007) in eight studies. Secondary surgical procedures were defined as any hardware removal, revisions, supplemental fixations, and reoperations. They were typically used to resolve persistent neck or shoulder pain, dysphagia, prosthesis flexibility or adjacent level degeneration. Secondary surgical procedures were recorded at the index level and the adjacent level. C-ADR group had significantly fewer secondary surgical procedures at the index (OR, 0.32; 95% CI: 0.19 to 0.53, P < 0.00001) and the adjacent level (OR, 0.28; 95% CI: 0.11 to 0.72, P = 0.008). <p>Medium-term (4-5 years)</p> <ul style="list-style-type: none"> Only one study with 74 patients had valid adverse-event data for midterm follow-up, no data given for this study. The rate of secondary surgical procedures at the adjacent level (OR, 0.76; 95% CI: 0.47 to 1.22, P = 0.25) was not significantly different between the groups in five studies. There were significantly fewer secondary surgical procedures related to the index level in the C-ADR group in five studies (OR, 0.45; 95% CI: 0.29 to 0.68, P = 0.0002). 	<p>medium-term follow-up of 4-5 years.</p>

Table 3. ADR Summary Table for Key Questions 3 and 4.

Key Question 3: What is the evidence of differential efficacy or safety issues amongst special populations (including but not limited to the elderly and workers compensation populations)?			
Conclusions from CER Executive Summary	New Sources of Evidence	New Findings	Conclusion from SRI
There is insufficient evidence to draw conclusions regarding the safety and efficacy of LADR in the few special populations studied (elderly, smokers, athletes). No studies or sub-analyses were found on the use of C-ADR in special or subpopulations.	No new evidence	No new evidence	This section of the report is still valid and does not need updating.
Key Question 4: What are the cost implications and cost effectiveness for ADR?			
Conclusions from CER Executive Summary	New Sources of Evidence	New Findings	Conclusion from SRI
There are inadequate data from partial economic studies reflecting short time horizons for L-ADR and no economic studies for C-ADR to truly assess the potential cost effectiveness of ADR technology. One report and one previously done HTA suggest that the type of fusion may influence complication rates and therefore costs.	1 lumbar ³ ; 2 cervical (1 single- ⁴ and 1 two-level replacement ⁵)	<p>Lumbar: L-ADR was cost-effective compared with multi-disciplinary rehab after 2 years when using EQ-5D for assessing QALYs gained and a willingness to pay. L-ADR was not cost-effective when SF-6D was used. Longer follow-up is needed to accurately assess cost-effectiveness of L-TDR.</p> <p>Cervical: Single level – One study suggests that a non-significant added benefit versus ACDF comes at a reasonable cost, whether actual hospital costs or Medicare reimbursement values are used. Two-level – One study concludes that the incremental cost-effectiveness ratio of C-ADR compared with traditional ACDF is lower than the commonly accepted threshold of \$50,000 per QALY.</p>	This section of the report is NOT valid. Studies of cost-effectiveness are now available for 1-level L-ADR versus conservative care, 1-level C-ADR versus 1-level cervical fusion, and 2-level C-ADR versus 2-level cervical fusion. Therefore, this section of the report needs updating.

5. Conclusions

L-ADR

- There are several systematic reviews that include new RCTs since the publication of the original ADR report. From a review of these systematic reviews, there is one new RCT that evaluates L-ADR versus conservative (non-operative) care. This is the first study making this comparison and warrants an update of the section comparing efficacy and safety of ADR versus a treatment other than ACDF, (criteria B-3, Figure 1).
- New studies comparing the efficacy and safety of L-ADR versus ACDF are consistent with the original ADR HTA. This section does not need an update.
- One study on cost effectiveness of L-ADR intervention has been published since the original HTA comparing L-ADR versus conservative (non-operative) care. Therefore, this section of the report needs updating.

C-ADR

- There are no new data for C-ADR versus new comparisons other than cervical fusion.
- One C-ADR, the Mobi-C, has been approved by the FDA for 2-level fusion. This is a new indication since the original report. There is at least 1 RCT (the FDA trial) that reports 2 year results on 2-level C-ADR. This warrants an update of the section of the report on efficacy and safety of C-ADR, (criteria B-2, Figure 1).
- The results of integrating new RCTs (total number: 19 RCTs, 4,516 patients) are similar to the original report with respect to pain and function for the short-term (24 months). However, there are new efficacy and safety data for medium-term (4-5 years) that were not present in the original report. Therefore, this section needs updating for both efficacy and safety.
- There were no new studies on differential efficacy or safety. This section of the report does not need updating.
- Two studies on cost effectiveness of C-ADR intervention have been published since the original HTA; 1-level C-ADR versus 1-level cervical fusion, and 2-level C-ADR versus 2-level cervical fusion. Therefore, this section of the report needs updating.

REFERENCES

1. Jacobs W, Van der Gaag NA, Tuschel A, et al. Total disc replacement for chronic back pain in the presence of disc degeneration. *Cochrane Database Syst Rev* 2012; **9**: CD008326.
2. Zhang Y, Liang C, Tao Y, et al. Cervical total disc replacement is superior to anterior cervical decompression and fusion: a meta-analysis of prospective randomized controlled trials. *PLoS One* 2015; **10**(3): e0117826.
3. Johnsen LG, Hellum C, Storheim K, et al. Cost-effectiveness of total disc replacement versus multidisciplinary rehabilitation in patients with chronic low back pain: a Norwegian multicenter RCT. *Spine (Phila Pa 1976)* 2014; **39**(1): 23-32.
4. 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.
5. 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.

APPENDIX A. SEARCH STRATEGIES

Below is the search strategy for PubMed.

(artificial[TI] OR TOTAL[TI] OR ARTHROPLASTY[TI] OR PROSTHETIC*[TI] OR PROSTHES*[TI]) AND (DISC[TI] OR DISK[TI]) AND (LOW BACK[TIAB] OR LUMBAR[TIAB]) AND META-ANALYS*

(artificial[TI] OR TOTAL[TI] OR ARTHROPLASTY[TI] OR PROSTHETIC*[TI] OR PROSTHES*[TI]) AND (DISC[TI] OR DISK[TI]) AND (NECK[TIAB] OR CERVICAL[TIAB]) AND META-ANALYS*

(artificial[TI] OR TOTAL[TI] OR ARTHROPLASTY[TI] OR PROSTHETIC*[TI] OR PROSTHES*[TI]) AND (DISC[TI] OR DISK[TI]) AND (LOW BACK[TIAB] OR LUMBAR[TIAB]) AND COST*[TI]

(artificial[TI] OR TOTAL[TI] OR ARTHROPLASTY[TI] OR PROSTHETIC*[TI] OR PROSTHES*[TI]) AND (DISC[TI] OR DISK[TI]) AND (NECK[TIAB] OR CERVICAL[TIAB]) AND COST*[TI]

APPENDIX B. SUMMARY OF INCLUDED SYSTEMATIC REVIEWS.

Assessment (year) Search dates	Purpose	Condition	Treatments vs. controls	Primary Outcomes	Evidence-base Used	Primary Conclusions
Jacobs (2012) Database inception to 12/2011	To assess the effect of total disc replacement for chronic low-back pain in the presence of lumbar disc degeneration	Chronic low-back pain	Lumbar total disc replacement vs. lumbar fusion	Pain, overall improvement, patient satisfaction, back-specific function status, quality of life	5 RCTs (1,301 patients)	Total disc replacement has slightly better outcomes in terms of back pain and function than those who had fusion surgery, but these differences were not clinically significant.
Zhang (2015) Database inception to 12/2014	To determine if cervical total disc replacement is superior to cervical fusion.	Symptomatic cervical disc disease	Cervical total disc replacement vs. anterior cervical decompression and fusion	Pain, function, quality of life, adverse events, overall success	19 RCTs (4,516 patients)	At short- and mid-term follow-up, cervical total disc replacement is superior to anterior cervical decompression and fusion with regards to efficacy and safety. However, longer-term multicenter studies are needed to better evaluate the long-term efficacy and safety.

APPENDIX C. SYSTEMATIC REVIEWS EXCLUDED AT FULL TEST REVIEW**Excluded systematic reviews, lumbar spine.**

Citation	Reason for exclusion
Nie H, Chen G, Wang X, Zeng J. Comparison of Total Disc Replacement with lumbar fusion: a meta-analysis of randomized controlled trials. <i>J Coll Physicians Surg Pak.</i> 2015;25(1):60-67.	Not comprehensive; lacks inclusion/exclusion criteria; no GRADE
Wei J, Song Y, Sun L, Lv C. Comparison of artificial total disc replacement versus fusion for lumbar degenerative disc disease: a meta-analysis of randomized controlled trials. <i>Int Orthop.</i> 2013;37(7):1315-1325.	Lacks inclusion/exclusion criteria; no GRADE
Rao MJ, Cao SS. Artificial total disc replacement versus fusion for lumbar degenerative disc disease: a meta-analysis of randomized controlled trials. <i>Arch Orthop Trauma Surg.</i> 2014;134(2):149-158.	Combined studies with short and medium f/u; no GRADE

Excluded systematic reviews, cervical spine.

Citation	Reason for exclusion
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. <i>Spine (Phila Pa 1976).</i> 2015.	Not comprehensive; metaanalyses of one manufacturer's disc results
Jee YM, Bak JS, Weinlander E, Anderson PA. Comparing Nonrandomized Observational Studies With Randomized Controlled Trials in Cervical Disc Arthroplasty: A Meta-analysis. <i>Spine (Phila Pa 1976).</i> 2015.	Not comparison of interest; RCT vs observational studies
Wu AM, Xu H, Mullinix KP, et al. Minimum 4-year outcomes of cervical total disc arthroplasty versus fusion: a meta-analysis based on prospective randomized controlled trials. <i>Medicine (Baltimore).</i> 2015;94(15):e665.	No comprehensive; only looked at 4+ year f/u
Rao MJ, Nie SP, Xiao BW, Zhang GH, Gan XR, 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. <i>Arch Orthop Trauma Surg.</i> 2015;135(1):19-28.	Combined studies with short and medium f/u; no GRADE
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. <i>Arch Orthop Trauma Surg.</i> 2015.	No GRADE
Luo J, Huang S, Gong M, 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. <i>Eur J Orthop Surg Traumatol.</i> 2015;25 Suppl 1:S115-125	No GRADE
Zhao H, Cheng L, Hou Y, et al. Multi-level cervical disc arthroplasty (CDA) versus single-level CDA for the treatment of cervical disc diseases: a meta-analysis. <i>Eur Spine J.</i> 2015;24(1):101-112.	Not comparison of interest; single vs. multilevel
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. <i>Eur Spine J.</i> 2014;23(5):1115-1123.	Not comprehensive; 4 year only

Citation	Reason for exclusion
Verma K, Gandhi SD, Maltenfort M, et al. Rate of adjacent segment disease in cervical disc arthroplasty versus single-level fusion: meta-analysis of prospective studies. <i>Spine (Phila Pa 1976)</i> . 2013;38(26):2253-2257.	Not comprehensive; adjacent segment disease as primary outcome
Gao F, Mao T, Sun W, 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). <i>Spine (Phila Pa 1976)</i> . 2015;40(23):1816-1823	Not comprehensive; limited studies; no GRADE
Boselie TF, Willems PC, van Mameren H, de Bie R, Benzel EC, van Santbrink H. Arthroplasty versus fusion in single-level cervical degenerative disc disease. <i>Cochrane Database Syst Rev</i> . 2012;9:CD009173.	Not comprehensive; limited studies
Nunley PD, Jawahar A, Cavanaugh DA, Gordon CR, Kerr EJ, 3rd, Utter PA. Symptomatic adjacent segment disease after cervical total disc replacement: re-examining the clinical and radiological evidence with established criteria. <i>Spine J</i> . 2013;13(1):5-12.	Not comprehensive; adjacent segment disease as primary outcome
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. <i>Arch Orthop Trauma Surg</i> . 2015;135(2):155-160.	Not comprehensive; adjacent segment disease as primary outcome
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. <i>PLoS One</i> . 2012;7(4):e35032	Not comprehensive; adjacent segment disease as primary outcome
Chen J, Wang X, Bai W, Shen X, Yuan W. Prevalence of heterotopic ossification after cervical total disc arthroplasty: a meta-analysis. <i>Eur Spine J</i> . 2012;21(4):674-680.	Not comprehensive; heterotopic ossification as primary outcome