Health Technology Assessment Program

Selected Technologies 2018 - Revised

List of contents

[This page left intentionally blank.]

STATE OF WASHINGTON
HEALTH CARE AUTHORITY
626 8th Avenue, SE • P.O. Box 45502 • Olympia, Washington 98504-5502

June 14, 2018

To whom it may concern:

SUBJECT: Health Technology Assessment Topic Selection, 2018 - REVISED

As the Director of the Health Care Authority and per the Health Technology Assessment law (70.14 RCW), I select technologies for review by the program in consultation with other agencies and the Health Technology Clinical Committee. Technologies are selected when there are concerns about safety, efficacy or value (cost-effectiveness), when state expenditures are or could be high, and there is adequate evidence to conduct a review. Technologies are selected for re-review when new evidence is available that could change a previous determination. In addition, anyone may petition for a technology review.

For the 2018 selection cycle, I have reviewed the proposed topics as well as the comments received from the interested individuals and groups who responded in the first comment period (March 5-19, 2018). Based on the information provided by the Health Technology Assessment program, and the recommendations from staff at the Health Care Authority, Department of Labor and Industries, and Department of Corrections, I have selected the following technologies for review:

	Technology	Safety	Efficacy	Cost
1	**Wearable cardiac defibrillators (WCD)**	**Medium**	**Med/ High**	**High**
	Policy context/reason for selection: Wearable defibrillators are externally worn devices that can monitor heart function and provide electrical shock (defibrillation) if a life-threatening cardiac arrhythmia is detected. Wearable defibrillators may offer a temporary alternative treatment to more invasive treatments or hospitalization. The topic is proposed based on concerns related to the safety, efficacy and value for wearable defibrillators.			
2	**Peripheral nerve ablation (PNA) for the treatment of limb pain**	**High**	**High**	**Med/High**
	Policy context/reason for selection: Ablation, or the severing of nerves transmitting pain signals from joints or other origins, is a potential treatment for discomfort caused by osteoarthritis and other conditions. This procedure can be used for upper and lower limb pain including, pain in the shoulder or knee. Nerve ablation for osteoarthritis and other limb and joint pain appears to be an emerging medical intervention with recent publications evaluating the treatment. Initial topic research resulted in the addition of upper limb ablation to the scope of this review. The topic is proposed based on concerns related to the safety, efficacy and value of the intervention for treatment.			

	Technology	Safety	Efficacy	Cost
3	**Renal denervation (RDN)**	**High**	**High**	**Med/ High**
	Policy context/reason for selection: Renal denervation (RDN) is a treatment for chronic high blood pressure (hypertension) that does not respond adequately to drug or other treatment. This procedure has been evaluated in good quality studies including comparisons to sham procedures. Questions remain regarding the long-term efficacy of the procedure, as well as selection of optimal treatment populations. The topic is proposed based on high levels of concern related to the safety and efficacy of the intervention.			
4	**Sacroiliac joint fusion**	**High**	**High**	**High**
	Policy context/reason for selection: Sacroiliac joint fusion is a surgical treatment sometimes used to address pain that may be originating from the joint between bones in the spine and hip (sacrum and ilium). There are both open and less invasive or minimally invasive procedures developed to address sacroiliac joint dysfunction. The topic is proposed based on high concerns related to the safety, efficacy, and value of this treatment.			

[1] *Link to Primary Criteria Ranking.*

Additionally, I have selected the following topics for re-review based on newly available published evidence:

- *Proton beam therapy*
- *Optune/ Novocure*
- *Lymphoma – Positron Emission Tomography (PET)*

Upon publication of the selected list of technologies, a 30-day comment period will begin whereby any interested person or group may provide information relevant to review of these topics. HTA will begin work to review these technologies following this comment period.

Should you have any questions or concerns, please contact Josh Morse, Health Technology Assessment Program Director by telephone at 360-725-0839 or via email at Josh.morse@hca.wa.gov.

Sincerely,

[redacted]

Susan E. Birch, MBA, BSN, RN
Director

cc: Josh Morse, HTA Program Director, CQCT, HCA

Technologies selected

	Technology	Primary criteria ranking: Safety	Efficacy	Cost
1	**Wearable cardiac defibrillators (WCD)**	**Med**	**Med/ High**	**High**
	Policy Context/Reason for Selection: Wearable defibrillators are externally worn devices that can monitor heart function and provide electrical shock (defibrillation) if a life-threatening cardiac arrhythmia is detected. Wearable defibrillators may offer a temporary alternative treatment to more invasive treatments or hospitalization. The topic is proposed based on concerns related to the safety, efficacy and value for wearable defibrillators.			
2	**Peripheral nerve ablation (PNA) for the treatment of limb pain**	**High**	**High**	**Med/ High**
	Policy context/reason for selection: Ablation, or the severing of nerves transmitting pain signals from joints or other origins, is a potential treatment for discomfort caused by osteoarthritis and other conditions. This procedure can be used for upper and lower limb pain including, pain in the shoulder or knee. Nerve ablation for osteoarthritis and other limb and joint pain appears to be an emerging medical intervention with recent publications evaluating the treatment. Initial topic scoping resulted in the addition of upper limb ablation to the topic. The topic is proposed based on concerns related to the safety, efficacy and value of the intervention for treatment.			
3	**Renal denervation (RDN)**	**High**	**High**	**Med/ High**
	Policy Context/Reason for Selection: Renal denervation (RDN) is a treatment for chronic high blood pressure (hypertension) that does not respond adequately to drug or other treatment. This procedure has been evaluated in good quality studies including comparisons to sham procedures. Questions remain regarding the long-term efficacy of the procedure, as well as selection of optimal treatment populations. The topic is proposed based on high levels of concern related to the safety and efficacy of the intervention.			
4	**Sacroiliac joint fusion**	**High**	**High**	**High**
	Policy Context/Reason for Selection: Sacroiliac joint fusion is a surgical treatment sometimes used to address pain that may be originating from the joint between bones in the spine and hip (sacrum and ilium). There are both open and less invasive or minimally invasive procedures developed to address sacroiliac joint dysfunction. The topic is proposed based on high concerns related to the safety, efficacy, and value of this treatment.			

Topics considered, not proposed

	Technology
1	Non-pharmacologic treatment of urinary incontinence
2	Alzheimer's and in vivo tau imaging re-review
3	Non-invasive testing for fibrosis in patients with chronic hepatitis C
4	Stereotactic radiation therapy for liver tumors
5	Skin testing for allergic rhinitis
6	Transcatheter aortic valve replacement for treatment of aortic valve stenosis

Technologies considered for re-review:

Technologies are considered for re-review at least every eighteen months based on availability of new evidence that may change the decision. *(Detailed criteria are included below*). All technologies with determinations beyond eighteen months since the final determination previously reviewed by the Health Technology Clinical Committee (HTCC) are listed below, along with information on whether they have been selected for re-review.

	Technology	Originally reviewed	Recommended for re-review
1	**Proton beam**	**March 2014**	**Yes**
	New evidence and indications are identified that support re-reviewing the evidence for proton beam.		
2	**Optune/ Novocure**	**January 2016**	**Yes**
	New evidence identified that supports re-reviewing the evidence for tumor treatment fields.		
3	**Lymphoma – Positron Emission Tomography (PET)**	**September 2011**	**Yes**
	Evidence supports re-reviewing PET scans for lymphoma.		
4	**Femoroacetabular impingement (FAI) Syndrome**	**September 2011**	**No**
	Signal search conducted in 2018 (attached). New information does not support re-review at this time.		
5	**Facet neurotomy**	**March 2014**	**No**
	Signal search conducted in 2018 (attached). New information does not support re-review at this time.		
6	**Osteochondral allograft transplantation surgery (OATS)**	**November 2011**	**No**
	Signal search conducted in 2018 (attached). New information does not support re-review at this time.		

For the current period, the program has not received or identified new evidence to support review of the following:

	HTA decisions	Latest review/ scan
1	Arthroscopic knee surgery	October 2008
2	Computed Tomographic Angiography (CTA)	May 2009
3	Calcium scoring	May 2010
4	Knee joint replacement or knee arthroplasty	December 2010
5	Positron Emission Tomography (PET) scans for lymphoma	November 2011
6	Microprocessor-controlled lower limb prosthetics	March 2012
7	Osteochondral allograft / autograft transplantation	March 2012
8	Sleep apnea diagnosis and treatment	May 2012
9	Bone Morphogenetic Protein (BMP)	May 2012
10	Upright / positional MRI	June 2012
11	Hip resurfacing	August 2012
12	Robotic assisted surgery	September, 2012
13	Upper endoscopy for GERD and GERD-like symptoms	September 2012
14	Virtual colonoscopy or Computed Tomographic Colonography (CTC)	December 2012
15	Vitamin D screening and testing	March 2013
16	Hyperbaric oxygen for wound healing	May 2013
17	Cervical spinal fusion for degenerative disc disease	May 2013
18	Ablation procedures for supraventricular tachycardia	September 2013
19	Cochlear implants	September 2013
20	Discography	November 2013
21	Implantable infusion pumps	November 2013
22	Electrical Neural Stimulation (ENS)	November 2013
23	Hyaluronic acid / viscosupplementation	November 2013
24	Routine ultrasound for pregnancy	November 2013
25	Intensity modulated radiation therapy	November 2013
26	Carotid artery stenting	November 2013
27	Cardiac nuclear imaging	November 2013
28	Spinal cord stimulators	January 2014
29	Non-pharmacological treatments for treatment-resistant depression	March 2014
30	Facet neurotomy	March 2014
31	Proton beam therapy	May 2014
32	Screening and monitoring tests for osteopenia/osteoporosis	November 2014

	HTA decisions	Latest review/ scan
33	Functional neuroimaging for primary degenerative dementia or mild cognitive impairment	November 2014
34	Appropriate imaging for breast cancer screening in special populations	January 2015
35	Testosterone testing	March 2015
36	Imaging for rhinosinusitis	May 2015
37	Bariatric surgery	May 2015
38	Tympanostomy tubes in children	November 2015
39	Lumbar fusion for degenerative disc disease	November 2015

Primary criteria ranking:

HTA created a process and tools based on the legislative requirements and criteria that are widely used in technology assessment priority setting. Identification of criteria and use of priority tools makes the process explicit and increases transparency and consistency across decision-makers. The tools are intended to be used by agency liaisons when making recommendations and by the clinical committee when making comments or selections of technologies. The primary criteria are directly linked to the legislative mandates for the program to focus technology reviews where there are concerns about safety, efficacy, or cost effectiveness, especially relative to existing alternatives. See RCW 70.14.100. These criteria are also common to other technology assessment programs. The primary criteria ranking tool is available on the website.

Re-review topic criteria:

Re-review criteria are directly linked to the legislative mandate that technologies shall be selected for re-review only where evidence has since become available that could change a previous determination. Technologies are considered for re-reviews at least once every 18 months. Re-reviews consider only evidence made available since the previous determination. See RCW 70.14.100. The re-review criterion is directed at identifying those situations where a technology requires a re-review to consider new evidence that was not available when the initial review was completed and the likelihood that the new evidence could result in a change to a previous determination.

Safety and Efficacy of Femoroacetabular Impingement Syndrome Procedures: Assessing Signals for Update

Provided by:

Aggregate Analytics, Inc.

Prepared by:

Andrea C. Skelly, PhD, MPH
Aaron J.R. Ferguson, BA

February 12, 2018

Table of Contents

Previous Coverage Decision

A Health Technology Assessment titled: ***Hip Surgery Procedures for Treatment of Femoroacetabular Impingement Syndrome***, was originally released on August 26, 2011 by the Washington State Health Technology Clinical Committee. Additionally, an update signal assessment was published in December 29, 2014. The Committee's Coverage Decision for the original report is summarized below, followed by the main conclusions of the 2014 Signal Update review.

Health Technology Background

The Hip Surgery for Femoroacetabular Impingement Syndrome (FAI) topic was selected and published in December 2010 to undergo an evidence review process. The evidence based technology assessment report indicated that FAI syndrome is a recently recognized diagnosis in primarily younger individuals where relatively minor abnormalities in the joint (orientation or morphology) are thought to cause friction/impingement and pain. It is theorized that FAI starts the breakdown of cartilage, leading to osteoarthritis. There are two types of FAI: cam impingement (nonspherical femoral head or abnormality at the head-neck junction) and pincer impingement (deep or retroverted acetabulum resulting in over coverage of the femoral head). Proponents believe that surgical correction of the impinging deformities will alleviate the symptoms and retard the progression of OA degeneration.

Hip surgery is an invasive procedure to correct FAI using either an open surgery or arthroscopic approach. The surgeon resects abnormal outgrowths of bone, removes damaged cartilage, and reshapes the femoral neck to ensure that there is sufficient clearance between the rim of the joint socket and the neck of the femur. Labral debridement and labral repair are surgical treatment options for treating damaged labral tissue when addressing FAI. After corrective surgery, avoidance of weight bearing for several weeks to months and rehabilitation is required. Surgery to correct FAI includes arthroscopy, open dislocation of the hip, and arthroscopy combined with a mini-open approach.

Health Technology Clinical Committee's Findings and Coverage Decision

Topic: Hip Surgery for Femoroacetabular Impingement Syndrome (FAI)
Meeting Date: September 16th, 2011
Final Adoption: November 18th, 2011

HTCC Coverage Determination
Hip Surgery for Femoroacetabular Impingement Syndrome (FAI) is **not** a covered benefit.

Committee Findings
Having considered the evidence based technology assessment report and the written and oral comments, the committee identified the following key factors and health outcomes, and evidence related to those health outcomes and key factors:

(1) Evidence availability and technology features
The evidence based technology assessment report indicates:

- The evidence based technology assessment report stated that there are two types of FAI: cam impingement (non-spherical femoral head or abnormality at the head-neck junction) and pincer

impingement (deep or retroverted acetabulum resulting in over-coverage of the femoral head). Proponents believe that surgical correction of the impinging deformities will alleviate the symptoms and retard the progression of OA degeneration.

- The evidence based technology assessment report indicated that surgery to correct FAI includes arthroscopy, open dislocation of the hip and arthroscopy combined with a mini-open approach. The purpose of the surgery is to remove abnormal outgrowths of bone and damaged cartilage, and to reshape the femoral neck to ensure that there is sufficient clearance between the rim of the acetabulum and the neck of the femur.
- The committee also reviewed information provided by the state agencies, and public members; and heard comments from the evidence reviewer, clinical expert, HTA program, agency medical directors and the public.

(2) Is the technology safe?

The committee discussed multiple key factors and health outcomes that were important for consideration in their overall decision on whether the technology is safe. Summary of committee considerations follows.

- The evidence based technology assessment reported that six comparative studies, 31 case-series and three case-reports were found that reported complications following surgical treatment for FAI in non- or recreational athletes. Altogether, 20 studies reported on arthroscopy, ten on open dislocation and seven on the mini-open procedure.
- The evidence based technology assessment report indicated reoperation for reasons other than a conversion to a total hip arthroplasty occurred 3.8% in patients undergoing arthroscopy, 4.4% in those receiving open dislocation and 8.7% in patients following a mini-open procedure. There was only one reported head-neck fracture (<0.1%) and no reports of AVN, osteonecrosis or trochanteric nonunion. Heterotopic ossification occurred in 2% to 3% of those receiving arthroscopy or mini-open, and 6% in those receiving open dislocation.
- The evidence based technology assessment report indicated neurological complications (nerve palsy, paresthesia, and neuropraxia) were rare in those receiving arthroscopy or open dislocation; however, they occurred in 22% of 258 hips undergoing a mini-open procedure. Most were transient in nature. Three case-reports described an occurrence of extravasation of fluid into the abdomen/chest during arthroscopic treatment of FAI. In one case, the fluid extravasation resulted in

(3) Is the technology effective?

The committee discussed multiple key factors and health outcomes that were important for consideration in their overall decision on whether the technology is effective. Summary of committee considerations follows.

- *Hip surgery (open or arthroscopic) compared with no surgery for FAI:* The evidence based technology assessment report indicated that no randomized controlled trials (RCTs) comparing surgery with conservative care for FAI or comparing different surgical treatments for FAI was found.

- *Hip surgery (open or arthroscopic) compared with no surgery for FAI:* The evidence based technology assessment report identified one study that retrospectively compared conservatively treated patients versus those receiving FAI surgery versus patients having a total hip arthroplasty in the short-term (<5 year follow-up). In addition, the report identified four comparative studies which investigated the effectiveness of various surgical treatments for FAI: labral debridement versus labral refixation (two studies) and osteoplasty versus no osteoplasty (two studies). The first study poorly describes the selection of patients so that it was not possible to tell how the treatment and control groups were obtained. The last four studies use historical controls. There was no evidence identified that one specific treatment resulted in better outcomes than another (surgery versus no surgery, labral debridement versus refixation, osteoplasty versus no osteoplasty).
- *Hip surgery (open or arthroscopic) compared with no surgery for FAI:* The evidence based technology assessment report identified 27 case series that reported on clinical outcomes following treatment for FAI in non- or recreational athletes. All studies report improvement in pain, patient-reported and clinician-reported hip outcomes scores, patient satisfaction and return to normal activities following FAI surgery.
- *Hip surgery (open or arthroscopic) compared with no surgery for FAI:* The evidence based technology assessment report stated that approximately 8% of patients diagnosed with FAI who undergo surgery in published series go on to have a total hip arthroplasty within 3 years. There are no long-term (≥10 years) data available to assess long-term effectiveness of FAI surgery. There are no data yet published to test the hypothesis that FAI surgery prevents or delays hip osteoarthritis or the need for total hip arthroplasty.
- *Hip surgery for FAI compared with no surgery:* The evidence based technology assessment reported six comparative studies, 31 case-series and three case-reports were found that reported complications following surgical treatment for FAI in non- or recreational athletes. Altogether, 20 studies reported on arthroscopy, ten on open dislocation and seven on the mini-open procedure.

(4) Special Populations?

- The evidence based technology assessment report indicated no studies were found comparing the differential effectiveness of surgery versus nonsurgical care in FAI patients. However, five studies were identified that looked at outcomes following surgical treatment for FAI in two subpopulations, those with varying degrees of osteoarthritis as assessed by the Tönnis grade and patients with varying degrees of chondral damage assessed during surgery.
- The evidence based technology assessment report indicated that outcomes following FAI surgery were consistently worse in patients with greater preoperative osteoarthritis compared with those with less osteoarthritis. In one study, the relative risk of a conversion to total hip arthroplasty (THA) in those with preoperative Tönnis grade 2–3 was 58 (95% CI: 8, 424) compared with Tönnis grade 0-1. There was no reported difference in outcomes in patients with varying degrees of chondral

(5) Is the technology cost-effective?

The committee discussed multiple key factors that were important for consideration in their overall decision on whether the technology has value and is cost-effective. Summary of committee considerations follows.

- The evidence based technology assessment report indicated no cost effectiveness, cost utility or costing studies were found on FAI surgery.

(6) Medicare Decision and Expert Treatment Guidelines

Committee reviewed and discussed the expert guidelines as identified and reported in the technology assessment report.

- The Centers for Medicare and Medicaid Services have no national or local coverage determinations or policies regarding the surgical treatment of FAI syndrome.
- Guidelines – a search of the core sources and relevant specialty groups identified three guidelines.
    - National Institute for Health and Clinical Excellence (NICE), 2007: The National Institute for Health and Clinical Excellence (NICE), (which provides guidance on health technologies and clinical practice for the National Health Service in England and Wales) concluded in 2007 that current evidence on the efficacy and safety of both arthroscopic surgery for the treatment of FAI syndrome "does not appear adequate for these procedures to be used without special arrangements for consent and for audit or research"; further publications of safety and efficacy outcomes will be needed. NICE stated that only surgeons with specialist expertise in arthroscopic hip surgery should perform this procedure for FAI and that the natural history of FAI syndrome and the selection of patients for this procedure are uncertain; further research on these issues will be useful.
    - National Institute for Health and Clinical Excellence (NICE), 2011: In July 2011, NICE published an updated report on arthroscopy for FAI syndrome in the form of a rapid review of the medical literature and specialist opinion. The review is based on approximately 1126 patients from three non-randomized controlled trials, five case-series, and one case-report. Several short-comings in the available literature were addressed such as overall poor study quality, limited prospective data collection in case-series, variability of outcome assessment scales used and lack of validation of these scales, heterogeneity in treatments making comparison between studies difficult, and descriptions of hip impingement pathology/lesions not well defined in all studies. The specialists' concluded that "there is no proof yet that this procedure is efficacious, but the technique may have a place in preventing the development of osteoarthritis of the hip in some patients". They also stated that use of this procedure will become more widespread, but should remain with the confines of the specialist dealing with hip disorders in young adults.
    - National Institute for Health and Clinical Excellence (NICE), 2011: NICE published an updated guidance report on open surgery for FAI in July 2011 stating that "current evidence on the efficacy of open femoroacetabular surgery for hip impingement syndrome is adequate in terms of symptom relief in the short and medium term. With regard to safety, there are well recognized complications. Therefore this

procedure may be used provided that normal arrangements are in place for clinical governance, consent and audit with local review of outcomes.

Committee Decision

Based on the deliberations of key health outcomes, the committee decided that it had the most complete information: a comprehensive and current evidence report, public comments, and agency and state utilization information. The committee concluded that the current evidence on Femoroacetabular Impingement Syndrome (FAI) demonstrates that there is insufficient evidence to cover. The committee considered all the evidence and gave greatest weight to the evidence it determined, based on objective factors, to be the most valid and reliable. Based on these findings, the committee voted to not cover Femoroacetabular Impingement Syndrome (FAI).

Conclusions of the 2014 Signals for Update Assessment - FAI

1. There are several systematic reviews that include new literature on FAI since the publication of the HTA. From a review of these systematic reviews, there are no opposing findings or important changes in results for key questions 1-5. Furthermore, there continues to be no randomized controlled trials of efficacy of surgical treatment of FAI compared with non-operative treatment, or whether osteochondroplasty improves outcomes compared with no osteochondroplasty.
2. There are no studies to evaluate the efficacy of surgical intervention in reducing hip osteoarthritis in patients with a diagnosis of FAI.
3. There are a number of recent studies, mostly non-randomized studies, which compare labral repair with labral debridement in FAI patients. These studies suggest that labral repair may result in better outcome. However, the evidence base for this is low.
4. There are four ongoing randomized controlled trials in patients with FAI. Three will help to answer the question of surgical versus non-surgical treatment, and one will help to answer the question of the efficacy of osteochondroplasty. These studies are due to be completed in 2014 (n=1), 2016 (n=1) and 2017 (n=2).
5. Three studies on cost effectiveness of surgical intervention have been published since the original HTA. Two conclude that hip arthroscopy could be cost effective in non-arthritic patients depending on the accuracy of assumptions. One concludes that the mini-open approach may be more cost effective than open dislocation or arthroscopy. These new reports don't meet the criteria that would trigger an updated report.

1. Purpose of Report

A prior update report was completed in December 2014. The purpose of this update is to determine whether or not there is sufficient evidence published subsequent to the last signal assessment to conduct a re-review of this technology. The key questions from the original report are listed below. **For this signal update, updated searches were only performed for Key Questions 3-6.**

Key question 1
Is there a consistent or agreed upon case definition for FAI? What is the evidence of reliability and validity of these case definitions?

Key question 2
What are the expected treatment outcomes of hip surgery for FAI? Are there validated instruments related to hip surgery outcomes? Have clinically meaningful improvement in outcomes been defined for FAI?

Key question 3
What is the evidence of efficacy and effectiveness of hip surgery (open or arthroscopic) compared with no surgery for FAI?

Key question 4
What is the evidence of the safety of hip surgery for FAI compared with no surgery?

Key question 5
What is the evidence that hip surgery for FAI compared with no surgery has differential efficacy or safety issues in sub populations?

Key question 6
What evidence of cost implications and cost-effectiveness of hip surgery compared with no surgery exists for FAI?

2. Methods

2.1 Literature Searches

We conducted an electronic literature search for the period August 1, 2014 through January 11, 2018 using identical search terms used for the original report for key questions 3 through 6. This search included three main databases: PubMed/Medline, Cochrane Library, and EMBASE. Additionally, we reviewed ClinicalTrials.gov for relevant ongoing studies. Appendix A reports the search methodology for this topic.

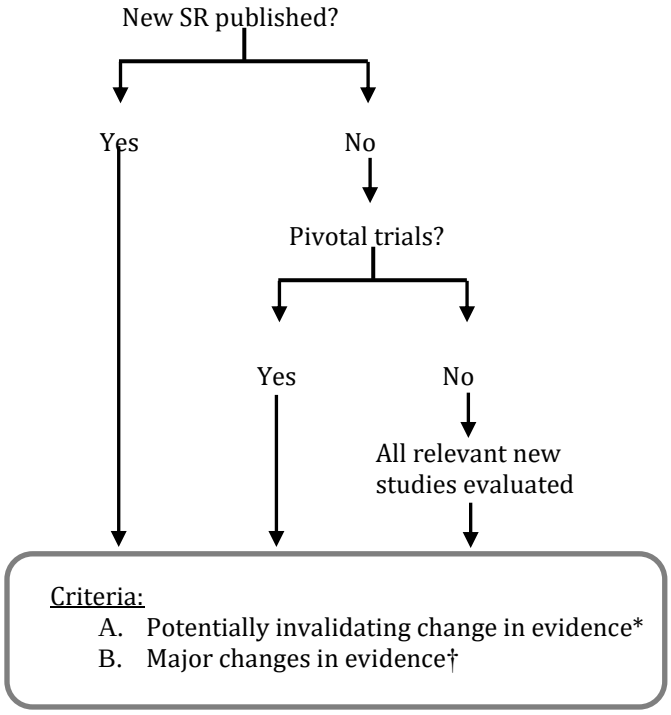
2.2 Study selection

We used the same inclusion and exclusion criteria as the original HTA and the 2014 Signal Update Review for Key Questions 3-6.

2.3 Compilation of Findings and Conclusions

For this assessment we constructed a summary table that included the key questions 3-6, the original conclusions, new sources of evidence, new findings, and new conclusions based on available signals. To assess whether the conclusions might need updating, we used an algorithm based on a modification of the Ottawa method, Figure 1.

Figure 1. Algorithm of the modified Ottawa Method of Identifying Signals for SR Update



*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. Results

3.1 Search

From 121 citations returned from the updated search, 107 were excluded at title/abstract review. Of the 14 reviewed at full text, 6 systematic reviews that addressed in part or in full key questions 3 through 6, were retained (Figure 2). We identified no new cost-effectiveness studies for inclusion. A full list of excluded studies and the reasons for exclusions can be found in Appendix C.

3.2 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 Aggregate Analytics, Inc. (AAI) regarding the need for update. Table 2 shows updated information on currently ongoing trials assessing arthroscopic surgery versus non-surgical interventions. Appendix B details data abstraction and summaries for included systematic reviews and recent comparative studies. Appendix C includes a list of Systematic Reviews excluded at full-text review.

Figure 2. Flow chart showing results of literature search

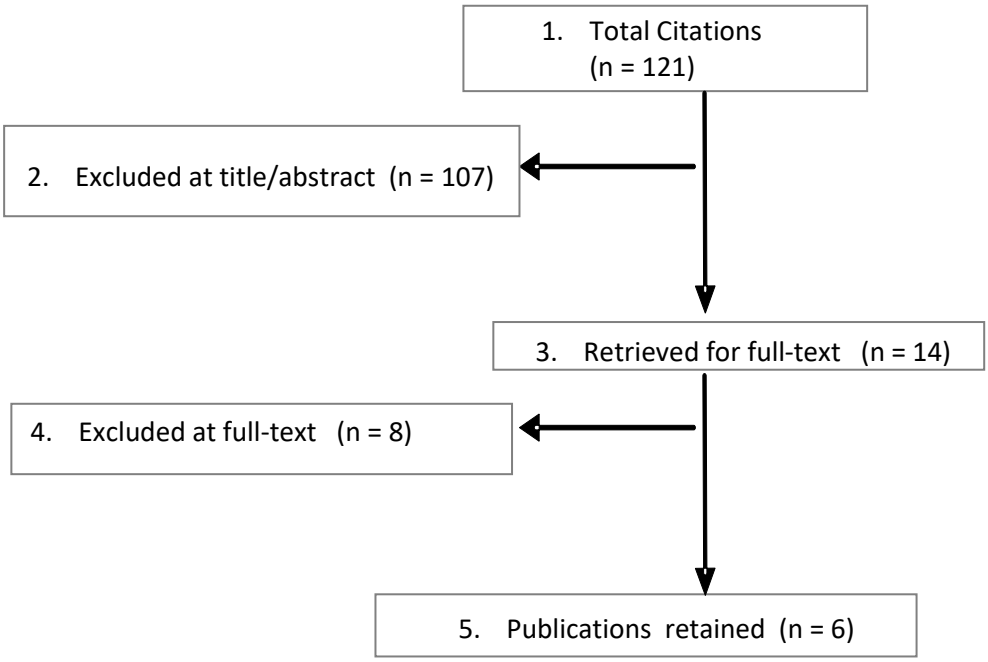


Table 1. Summary Table of Key Questions 1-6

Conclusions from CER Executive Summary	Conclusions from 2014 Signal Update	New Sources of Evidence	New Findings	Conclusion from AAI
Key Question 1. Is there a consistent or agreed upon case definition for FAI? What is the evidence of reliability and validity of these case definitions?				
<u>Case definition</u> • The most consistent case definition of FAI (cam or mixed) as defined by inclusion/exclusion criteria in prospective studies of treatment effectiveness includes hip/groin pain, positive clinical impingement test, and an □-angle >50-55º • There is no evidence that the diagnosis of FAI can be obtained from clinical exam in one small study. One clinical test, the impingement sign, had a positive and negative predictive value of 86% and 79% in one study where the prevalence of FAI was 50%; however, in another study, the reliability of the impingement sign was only moderate. • Even though the α-angle showed moderate to high interobserver reliability in several studies, it had poor diagnostic value in identifying FAI. Other imaging tests assessing abnormalities of the femur and acetabulum had variable degrees of reliability, but no others were tested for diagnostic validity.	This section of the report is still valid and does not need updating	Not sought for 2018 update	N/A	N/A
Key Question 2: What are the expected treatment outcomes of hip surgery for FAI?				
<u>Patient- and clinician reported outcomes</u> • Seven hip outcomes measures were used commonly in FAI patients. Three have undergone psychometric analysis in FAI (HOS-D, M-WOMAC) or young hip-pain (HOS, NAHS) patient populations. • Only one, the Non-arthritic Hip Score (NAHS), of the three	This section of the report is still valid. However, there are at least two new outcomes that have been developed since the original report that may become more frequent in	Not sought for 2018 update	N/A	N/A

Conclusions from CER Executive Summary	Conclusions from 2014 Signal Update	New Sources of Evidence	New Findings	Conclusion from AAI
instruments was adequately tested for validity, and it was performed in a young hip-pain patient population. • Reliability was inadequately tested for all three instruments. • The MCID was defined to be 9 points for the ADL subscale and 6 points for the sports subscale of the HOS-D in FAI patients. The MCID has not been defined for any other outcome measures in FAI or young hip-pain patients.	future studies of FAI.			
Key Question 3: What is the evidence of efficacy and effectiveness of hip surgery compared with no surgery for FAI?				
Efficacy • There are no data available to assess the short- or long-term efficacy of FAI surgery compared with no surgery	This section of the report is still valid and does not need updating.	**Systematic Review:** Wall 2014[5] Fairley 2016[1]	Efficacy • One Cochrane systematic review (Wall 2014) found no randomized or quasi-randomized trials that compared surgical intervention with no surgery; review did not compare surgical interventions with other surgical interventions. • Another systematic review (Fairley 2016) primarily reviewing cohort studies, found no studies comparing surgical and non-surgical treatment, and no overarching conclusions regarding the relative efficacy of one surgical approach over another were made. No quantitative analyses were provided.	This section of the report remains valid and does not need updating.
Effectiveness (short term) • There is no evidence that one specific treatment resulted in better outcomes than another (surgery versus no surgery, labral debridement versus refixation, osteoplasty versus no osteoplasty). • Several case series report improvement in pain, patient reported and clinician reported hip outcome scores, patient satisfaction and return to normal activities following FAI surgery. However, whether this improvement is a result of the surgery, or the postoperative rehabilitation, or the change in	Though there is a suggestion that labral fixation may have slightly better MHHS scores than debridement, other outcomes have mixed results. This section of the report is still valid and does not need updating.	**Systematic Reviews:** Forster-Horvath 2016[2] Kierkegaard 2017[4]	Effectiveness (short-term) • There is no evidence comparing outcomes between surgery and no surgery from two systematic reviews. • One systematic review (Forster-Horvath 2016) indirectly compared surgical interventions (labral debridement/segmental resection and labral reconstruction) using evidence primarily from case series. No quantitative synthesis was reported. Authors concluded that clinical outcomes were comparable for labral debridement/segmental resection and labral reconstruction. One available comparative study reported a significantly greater mean change (improved function) in the Non-Arthritic	Comparisons between surgical interventions were indirect. This section of the report remains valid and does not need updating.

Conclusions from CER Executive Summary	Conclusions from 2014 Signal Update	New Sources of Evidence	New Findings	Conclusion from AAI
activity subsequent to the surgery or placebo is not known. • Approximately 8% of patients diagnosed with FAI who undergo surgery in published series go on to have a total hip arthroplasty within 3 years.			Hip Score (P = 0.046) and Hip Outcome Score-Activities of Daily Living (P = 0.045) favoring labral reconstruction over labral debridement/ segmental resection. • One systematic review (Kierkegaard 2017) found hip pain reduction and Activities of Daily Living Function improvements between 3 and 6 months post-arthroscopic surgery, and sport function improvements between 6 and 12 months post-surgery. The overall low level of evidence (primarily case series) and lack of comparative studies indicate that further evidence is needed to determine comparative effectiveness.	
Key Question 3: What is the evidence of efficacy and effectiveness of hip surgery compared with no surgery for FAI? (continued)				
Effectiveness (long term) • There are no data available to assess long-term effectiveness of FAI surgery compared with no surgery. • There are no data yet published to test the hypothesis that FAI surgery prevents or delays hip osteoarthritis or the need for total hip arthroplasty.	This section of the report is still valid and does not need updating.	**Systematic Review:** Kierkegaard 2017[4]	• There are no data available to assess long-term effectiveness of FAI surgery compared with no surgery. • One systematic review of primarily case series (Kierkegaard 2017) found that in hip pain reduction, and improvements in ADL function and sport function were evident at least up to 3 years after hip arthroscopy in patients with FAI, however, authors that report lower average scores after hip arthroscopy than patient's healthy counterparts indicated residual mild hip pain and/or impaired hip function during ADL and sport. The overall low level of evidence (primarily case-series) and lack of comparative studies indicate that further evidence is needed to determine relative effectiveness.	This section of the report remains valid and does not need updating.
Key Question 4: What is the evidence of the safety of hip surgery for FAI compared with no surgery?				
Safety • The risk of reoperation (other than conversion to THA) occurred in 4% (arthroscopy and open dislocation) and 9% of the patients (mini-open). • There was only one reported head-neck fracture (0.1%) and no reports of AVN, osteonecrosis or trochanteric nonunion.	This section of the report is still valid and does not need updating.	**Systematic Reviews:** Zhang 2016[6] Forster-Horvath 2016[2]	Safety • One systematic review (Zhang 2016) found a higher risk of reoperation (relative risk: 0.40, 95% CI: 0.17–0.95, P= 0.04) for open surgical dislocation than for hip arthroscopy across four cohort studies (n=292 hips). No statistical difference in complications between arthroscopy and open surgical dislocation was found. • One systematic review (Forster-Horvath 2016) mostly reviewing case series found	This section of the report remains valid and does not need updating.

Conclusions from CER Executive Summary	Conclusions from 2014 Signal Update	New Sources of Evidence	New Findings	Conclusion from AAI
• Heterotopic ossification occurred in 2% to 3% of those receiving arthroscopy or mini-open, and 6% in those receiving open dislocation. • Neurological complications (nerve palsy, paresthesia, and neuropraxia) were rare in those receiving arthroscopy or open dislocation; however, they occurred in 22% of 258 hips undergoing a mini-open procedure. Most were transient in nature.			an overall range of conversion to hip arthroplasty of 0% to 30% across debridement and refixation groups. Patients who underwent labral debridement/ segmental resection were not found to transition to Total Hip Arthroplasty more frequently than those who underwent labral reconstruction.	
Key Question 5: What is the evidence that hip surgery for FAI compared with no surgery has differential efficacy or safety issues in sub populations?				
<u>Differential efficacy, effectiveness or safety</u> • We found no studies comparing the differential efficacy, effectiveness or safety of surgery versus nonsurgical care in FAI patients. • Outcomes following FAI surgery were consistently worse in patients with greater preoperative osteoarthritis compared with those with less osteoarthritis. • There was no reported difference in outcomes in patients with varying degrees of chondral damage assessed during surgery. • No data from other subpopulations were found.	This section of the report is still valid and does not need updating.	**Systematic Review:** Griffin 2017[3]	• We found no studies comparing the differential efficacy, effectiveness or safety of surgery versus nonsurgical care in FAI patients. • Study designs were not conducive to evaluation of differential efficacy, effectiveness or safety of surgery intervention versus another surgical interventions; only information on subpopulations was available. • One systematic review (Griffin 2017) reviewing cohort studies and case series concluded that hip arthroscopy was safe and efficacious procedure across studies among patients older than 40 who did not have significant underlying degenerative changes. Authors do not report on comparative effectiveness of arthroscopy versus other treatment options. • Overall reoperation rate (excluding conversion to arthroplasty) was 2.3% (arthroscopy) among adults older than 40. This review found an overall complication rate of 5.1% (8/157) of patients across five studies (cohort studies and case series). Complications included: 1 deep venous thrombosis, 1 case of heterotopic ossification (HO), 1 superficial wound infection resolved with oral antibiotics, 1 deep wound infection, 3 cases of psoas tendinitis, and 2 cases of transient sensory neurapraxia (perineum and foot).	This section of the report remains valid and does not need updating.

Conclusions from CER Executive Summary	Conclusions from 2014 Signal Update	New Sources of Evidence	New Findings	Conclusion from AAI
Key Question 6: What evidence of cost implications and cost-effectiveness of hip surgery compared with no surgery exists for FAI?				
Cost-effectiveness There were no cost-effectiveness, cost utility or costing studies found on FAI surgery.	There are new data that would update this section of the report. However, the findings from these studies don't meet the criteria that would trigger an updated report.	No new sources of evidence.	• We found no additional cost-effectiveness, cost utility or costing studies that would change the conclusions of the previous signal update	This section of the report remains valid and does not need updating.

3.3 Current ongoing clinical trials

We identified one additional ongoing clinical trial registered since the 2014 signal update report, Table 2. Along with the details of the additional trial, the status of the other four trials has been updated. No analyses of the four trials were identified.

The newly identified trial will compare arthroscopic surgery with sham surgery (diagnostic arthroscopy). One previously identified trial will test whether osteochondroplasty will provide improved clinical results versus arthroscopic lavage, while the other three will test whether arthroscopic surgery in FAI patients will result in better clinical outcomes compared with non-operative care that includes physical therapy.

Table 2. Characteristics of current ongoing studies registered in clinical trials.gov assessing the efficacy of FAI treatment.

Study Author, NCT ID, Completion date	Purpose	Inclusion/exclusion	Intervention	Outcomes
New Ongoing Trial Identified Since 2014 Report				
Risberg NCT: 02692807 Last Update: May 16, 2017 Completion: currently recruiting, completion unknown	The primary aim of this study is to determine the efficacy of hip arthroscopic surgery compared to a sham surgery (diagnostic arthroscopy only) for patients with symptomatic and radiological findings related to impingement (FAI) and/or labral tears using a randomized controlled design (HIPARTI Study: Primary aim and the main paper:	Inclusion Criteria: 1. Adult men or women ages 18 to 65 years 2. hip pain during daily and/or sporting activities; 3. intra-articular hip pain with radiological signs of FAI and/or labral tears eligible for hip arthroscopy (to be determined in a pragmatic fashion by the surgeon based on clinical examination and imaging 4. the patient is able to give written informed consent and to participate fully in the interventions and follow-up procedures Exclusion Criteria: 1. pain that is not confirmed by physical examination of the hip 2. evidence of preexisting osteoarthritis, defined as Tonnis grade >1, or less than 3mm superior joint space width on AP pelvic radiograph	Intervention: Arthroscopic surgery Control: Sham Surgery (Diagnostic Arthroscopy)	Primary: • International Hip Outcome Tool (IHOT-33) Secondary: • Expectations of Surgery Questionnaire • Hip Dysfunction and Osteoarthritis (HOOS) • Arthritis Self-Efficacy Scale (ASES) • Tampa Scale of Kinesiophobia Fear of Movement Questionnaire • Hip Sports Activity Scale (HSAS)

Study Author, NCT ID, Completion date	Purpose	Inclusion/exclusion	Intervention	Outcomes
	primary end point: iHOT 1 year follow-up)).	3. center edge angle on radiograph <25°; (v) previous known hip pathology such as Perthes' disease, slipped upper femoral epiphysis or avascular necrosis 4. previous hip injury such as acetabular fracture, hip dislocation or femoral neck fracture 5. previous hip surgery 6. medical conditions complicating surgery (ASA 3); (ix) inflammatory joint disease (RA, Bechterew etc) 7. physical inability to undertake testing procedures; expected lack of compliance such as cognitive impairment, drug abuse or similar; inability to understand the written and spoken language of the treatment centre; contra-indications to placebo surgery, which will include large loose body, chondral flap >1cm2 detached at 3 sides, complete labral radial flap tear and labral bucket-handle tear with complete avulsion >1.5cm long		• Work place Activity Limitation Survey (WALS) • Patient Specific Functional Scale • Measures of hip physical impairment • Hip Muscle Strength • Single Leg Squat Performance • Total Hip Replacement
Status of Trials Identified in Previous Report				
Ayeni NCT: 01623843 Last Update: June 7, 2017 Completion: unknown	To determine whether surgical correction of hip impingement morphology via arthroscopic osteochondroplasty (shaving of bone) will provide improved clinical results (decreased pain and improved function) in adult patients with FAI compared to arthroscopic lavage (washing out of painful inflammation debris) and treating obvious damage of the hip joint.	Inclusion Criteria: 1. Adult men or women ages 18 to 50 years 2. Hip pain for greater than 6 weeks with no relief from non-operative means (physiotherapy, non-steroidal anti-inflammatory medication, rest) 3. CAM or Mixed Type FAI as diagnosed on x-rays and magnetic resonance imaging (MRI) or magnetic resonance arthrogram (MRA) 4. Temporary relief from an intra-articular hip injection 5. Informed consent from participant 6. Ability to speak, understand and read in the language of the clinical site Exclusion Criteria: 1. Previous inclusion in a study involving FAI 2. Evidence of hip dysplasia (centre edge angle less than 20) 3. Presence of advanced hip osteoarthritis (Tonnis Grade 2 or 3) 4. Previous trauma to the affected hip 5. Previous surgery on the affected hip 6. Isolated Pincer lesion 7. Immunosuppressive medication use 8. Chronic pain syndromes 9. Significant medical co-morbidities (requiring daily assistance for ADLs 10. History of paediatric hip disease (Legg-Calve-Perthes; SCFE)	Intervention: Osteochondroplasty Control: Arthroscopic Lavage	Primary: Pain (VAS) Secondary: • HRQoL (SF-12) • Function (HOS, iHOT-12) • Health utility (EQ-5D) • Sexual and urinary function (IIEF, FSFI, ICIQ- MLUTS, ICIQ-FLUTS) • Complications/AE

Study Author, NCT ID, Completion date	Purpose	Inclusion/exclusion	Intervention	Outcomes
		11. Ongoing litigation or compensation claims secondary to hip problems 12. Any other reasons given to exclude the patient		
Glyn-Jones NCT: 01893034 Last Update: December 3, 2013 Completion: Unknown	To compare the effectiveness of arthroscopic surgery versus physical therapy and activity modification for the treatment of FAI.	Inclusion Criteria: 1. Adult men or women ages 18 to 60 years 2. Symptomatic patients 3. Clinical and radiological evidence of FAI 4. Competent to consent Exclusion Criteria: 1. Prior hip surgery 2. Established osteoarthritis (Kellgren-Lawrence ≥ 2) 3. Hip dysplasia (Centre-Edge angle < 20 degrees on radiograph 4. Completion of physical therapy program targeting FAI within the past year 5. Comorbidities that mean surgical intervention is not possible/safe 6. Contraindication to MRI 7. Pregnancy	Intervention: Arthroscopic surgery Control: Conservative management	Primary: • Hip Outcome Score Secondary: • Patient reported outcome measures • Morphological and physiological MRI ○ Morphological parameters ○ Measures of osteoarthrits
Naudie NCT: 01621360 Last Update: February 8, 2013 Completion: Unpublished	To determine if patients with FAI who undergo arthroscopic hip surgery experience similar outcomes at 2 years post-operative with respect to physical function, pain, and health related quality of life, compared to similar patients who receive conservative management, including medication and physical therapy.	Inclusion Criteria: 1. Adult men or women ages 18 to 60 years 2. Patients with FAI of the hip 3. Grade 1, 2 or 3 radiographic severity of osteoarthritis as defined by the Tonnis classification scale Exclusion Criteria: 1. Identified isolated labral tear 2. Inflammatory or post-infectious arthritis 3. Previous arthroscopic treatment for hip osteoarthritis 4. Previous major hip trauma 5. Tonnis grade 4 osteoarthritis in two compartments in persons over 60 years of age 6. Patients with a major neurologic deficit, serious medical illness (life expectancy less than 2 years or high intraoperative risk) or those who are unable to provide informed consent or who are deemed unlikely to comply with follow-up	Intervention: Arthroscopic surgery Control: Conservative management	Primary: • Hip Outcome Score Secondary: • Non-arthritic hip score (NAHS) • Modified Harris Hip Score • SF-12 • Range of motion
Mansell NCT: 01993615 Last Update: December 20, 2016 Completion: Recruitment Completed	To compare the outcomes for patients that receive two different treatments used for FAI. The programs are 1) a 6-week supervised physical therapy program and 2)	Inclusion Criteria: 1. Adult men or women ages 18 to 65 years 2. Tricare beneficiaries 3. Diagnosis of FAI and/or labral pathology confirmed by a combination of the following: • Pain at anterior hip or groin • Pain with hip flexion • Positive FADIR test • Patient reported relief of pain after intra-articular injection 4. Surgical candidate for hip arthroscopy defined by both:	Intervention: Arthroscopic surgery Control: Physical therapy	Primary: • Hip Outcome Score • International Hip Outcome Score (iHOT33) Secondary: • Global Rating of Change (GROC) • Self-Motivation Inventory

Study Author, NCT ID, Completion date	Purpose	Inclusion/exclusion	Intervention	Outcomes
	arthroscopic surgery.	• No less than 2 mm of joint space based on imaging (CT scan, radiographs, and MR arthrogram) • Positive crossover sign and/or alpha angle >50^{o} based on imaging (CT scan, radiographs, and MR arthrogram) 5. Failed 6 weeks of conservative management <u>Exclusion Criteria:</u> 1. Pregnancy 2. Has other concurrent systemic disease that may affect the condition (cancer, rheumatoid arthritis, or other systemic arthralgia/arthritis) 3. Has had surgery on the same hip that will be analyzed in the study 4. Diagnosis of hip osteoarthritis is more likely 5. Clearing the lumbar spine reproduces the patient's hip symptoms 6. Plans to move/relocate out of the local area within 6 months 7. Pending litigation for their hip condition 8. Unable to give formal consent to participate in the study		• Pain Catastrophizing Scale (PCS)

4. Conclusions

4.1. There are no new systematic reviews that include new studies that compare surgical interventions with non-surgical interventions for the treatment of FAI since the publication of the HTA or previous signal update report. (Criteria A-1, A-3, B-1-4)

4.2. There are no new comparative studies to evaluate the efficacy of surgical intervention in reducing hip osteoarthritis in patients with a diagnosis of FAI. (Criteria A-1, A-3, B-1-4)

4.3. There are a number of recent non-randomized studies that indirectly compare labral repair with labral debridement in FAI patients. Although some suggest that labral repair may result in better outcome, the evidence base for this is low and does not meet the criteria that would trigger an updated report. (Criteria A-1, A-3, B-1-4)

4.4. One systematic review reviewing five cohort studies found a higher risk of reoperation for open surgical dislocation than for hip arthroscopy across four studies and 292 hips. No statistical difference in complications between arthroscopy and open surgical dislocation was found. Another systematic review, primarily reviewing case series, found an overall range of conversion to hip arthroplasty of 0% to 30% for both debridement and refixation groups. New safety evidence does not meet the criteria that would trigger an updated report. (Criterion A-2)

4.5. Although one new systematic review described outcomes for arthroscopic treatment of FAI in people over 40, the evidence does not meet criteria that would trigger an updated report. (Criteria A-1, A-3, B-1-4)

4.6. We identified no new cost-effectiveness, cost utility or costing studies that would change the conclusions of the previous signal update. (Criteria B-1-3)

4.7. One new ongoing trial was identified that is in the process of recruiting patients. No published data or completion timelines are evident from the four ongoing trials identified in the previous signal update.

REFERENCES

1. Fairley J, Wang Y, Teichtahl AJ, et al. Management options for femoroacetabular impingement: a systematic review of symptom and structural outcomes. Osteoarthritis and Cartilage 2016;24:1682-96.
2. Forster-Horvath C, von Rotz N, Giordano BD, Domb BG. Acetabular Labral Debridement/Segmental Resection Versus Reconstruction in the Comprehensive Treatment of Symptomatic Femoroacetabular Impingement: A Systematic Review. Arthroscopy : the journal of arthroscopic & related surgery : official publication of the Arthroscopy Association of North America and the International Arthroscopy Association 2016;32:2401-15.
3. Griffin DW, Kinnard MJ, Formby PM, McCabe MP, Anderson TD. Outcomes of Hip Arthroscopy in the Older Adult: A Systematic Review of the Literature. The American journal of sports medicine 2017;45:1928-36.
4. Kierkegaard S, Langeskov-Christensen M, Lund B, et al. Pain, activities of daily living and sport function at different time points after hip arthroscopy in patients with femoroacetabular impingement: a systematic review with meta-analysis. British journal of sports medicine 2017;51:572-9.
5. Wall PD, Brown JS, Parsons N, Buchbinder R, Costa ML, Griffin D. Surgery for treating hip impingement (femoroacetabular impingement). The Cochrane database of systematic reviews 2014:Cd010796.
6. Zhang D, Chen L, Wang G. Hip arthroscopy versus open surgical dislocation for femoroacetabular impingement: A systematic review and meta-analysis. Medicine 2016;95:e5122.

APPENDIX A. SEARCH STRATEGIES

Below is the search strategy for PubMed (August 1, 2014 – January, 11 2018). Parallel strategies were used to search other electronic databases listed below together with the search dates. Keyword searches were conducted in the other listed resources. Updated searches for Key Questions 1 and 2 were not conducted.

Key Question 1

	Search Terms
1.	FEMOROACETABULAR IMPINGEMENT* OR FEMORACETABULAR IMPINGEMENT* OR "Femoracetabular Impingement"[Mesh] OR ((HIP OR ACETABUL* OR FEMUR OR FEMORAL) AND IMPINGMENT*) OR “femoral osteochondroplasty” OR “femoral osteoplasty”
2.	SENSITIVITY[TIAB] OR SPECIFICITY[TIAB] OR PREDICT*[TIAB] OR "Reproducibility of Results"[Mesh] OR RELIAB*[TI] OR VALID* OR INTERTEST* OR INTEROBSERV* OR INTRATEST* OR INTRAOBSERV* OR INTERRAT* OR INTRARAT* OR “Validation Studies" [Publication Type] OR "Reproducibility of Results"[Mesh]
3.	PROSPECTIV*
4.	#1 AND #2 AND SYSTEMATIC REVIEW (LIMIT ENGLISH)
5.	#1 AND #3 AND SYSTEMATIC REVIEW (LIMIT ENGLISH) AND English 9

Key Question 2

	Search Terms
6.	FEMOROACETABULAR IMPINGEMENT* OR FEMORACETABULAR IMPINGEMENT* OR "Femoracetabular Impingement"[Mesh] OR ((HIP OR ACETABUL* OR FEMUR OR FEMORAL) AND IMPINGMENT*) OR “femoral osteochondroplasty” OR “femoral osteoplasty”
7.	"Merle d'Aubigné" OR “HARRIS HIP SCORE” OR “Western Ontario and McMaster Universities Osteoarthritis Index” OR WOMAC OR “NON ARTHRITIC HIP SCORE” OR “NONARTHRITIC HIP SCORE” OR “HIP OUTCOME SCORE” OR “OUTCOME SCORE”
8.	"Reproducibility of Results"[Mesh] OR RELIAB*[TI] OR VALID* OR INTERTEST* OR INTEROBSERV* OR INTRATEST* OR INTRAOBSERV* OR INTERRAT* OR INTRARAT*) OR “Validation Studies" [Publication Type]) OR "Reproducibility of Results"[Mesh]
9.	#6 AND #7 AND #8 AND SYSTEMATIC REVIEW (LIMIT ENGLISH)

Key Question 3, 5

	Search Terms	**Number of Articles**
10.	FEMOROACETABULAR IMPINGEMENT* OR FEMORACETABULAR IMPINGEMENT* OR "Femoracetabular Impingement"[Mesh] OR ((HIP OR ACETABUL* OR FEMUR OR FEMORAL) AND IMPINGMENT*) OR “femoral osteochondroplasty” OR “femoral osteoplasty”	2,586
11.	"Research Design/classification"[Mesh] OR "Research Design/epidemiology"[Mesh] OR "Research Design/methods"[Mesh] OR "Comparative Study" [Publication Type] OR "Clinical Trial" [Publication Type] OR RANDOM*[TIAB] OR "Treatment Outcome”]	3,474,807
12.	#10 AND #11 AND SYSTEMATIC REVIEW (LIMIT ENGLISH)	97

Key Question 4

	Search Terms	Number of Articles
13.	FEMOROACETABULAR IMPINGEMENT* OR FEMORACETABULAR IMPINGEMENT* OR "Femoracetabular Impingement"[Mesh] OR ((HIP OR ACETABUL* OR FEMUR OR FEMORAL) AND IMPINGMENT*) OR “femoral osteochondroplasty” OR “femoral osteoplasty”	2,586
14.	"Reoperation"[Mesh] OR "Femur Head Necrosis"[Mesh] OR "Arthroplasty, Replacement, Hip"[Mesh] OR REOPERATION REATTACHMENT OR AVN OR AVASCULAR NECROSIS OR TOTAL HIP OR TOTAL JOINT OR ARTHROPLASTY OR INFECTION* OR DEATH OR COMPLICATION* OR ADVERSE EVENT OR "Intraoperative Complications"[Mesh] OR SCIATIC* OR NERVE OR NEURO* OR FRACTURE* OR INTRAABDOM* OR CARDIAC ARREST OR THROMBO* OR EMBOL* OR INSTABILITY	850,549
15.	#13 AND #14 AND SYSTEMATIC REVIEW (LIMIT ENGLISH)	82

Key Question 6

	Search Terms	Number of Articles
16.	FEMOROACETABULAR IMPINGEMENT* OR FEMORACETABULAR IMPINGEMENT* OR "Femoracetabular Impingement"[Mesh] OR ((HIP OR ACETABUL* OR FEMUR OR FEMORAL) AND IMPINGMENT*) OR “femoral osteochondroplasty” OR “femoral osteoplasty”	2,016
17.	COST[TIAB] OR "Cost-Benefit Analysis"[Mesh] OR DECISION ANALYSIS [TIAB]	98,867
18.	#16 AND #17 (LIMIT ENGLISH)	12

Parallel strategies were used to search the Cochrane Library and others listed below. Keyword searches were conducted in the other listed resources.

Electronic Database Searches

The following databases have been searched for relevant information:

Agency for Healthcare Research and Quality (AHRQ) (August 1, 2014 through January, 11 2018)

Cochrane Database of Systematic Reviews (August 1, 2014 through January, 11 2018, Issue 1)

Database of Abstracts of Reviews of Effects (DARE - Cochrane Library) (August 1, 2014 through January, 11 2018 Issue 1)

Informational Network of Agencies for Health Technology Assessment (INAHTA) (Database Inception through January, 11 2018)

EMBASE (August 1, 2014 through January, 11 2018)

PubMed (August 1, 2014 through January, 11 2018)

APPENDIX B. SUMMARY OF INCLUDED STUDIES

Appendix Table B1. Summary of Included Systematic Reviews

Assessment Search dates	Purpose	Condition	Treatment vs. comparators	Primary Outcomes	Evidence- base Used	Primary Conclusions
Fairley et al. 2016 Systematic Review *Jan 2000 to July 2015*	The optimal therapy for femoroacetabular impingement (FAI) is unclear. The aim of this systematic review was to examine the evidence for surgical and non-surgical treatment of FAI on symptom and structural outcomes.	Femoroacetabular impingement (FAI)	Surgical and non-surgical treatment, Open Surgery vs. arthroscopy, Different arthroscopic techniques with each other, Different open surgical techniques with each other	Symptoms assessed by validated tools, hip bone shape (radiographic measures, joint degeneration, or progression to joint replacement	18 studies (16 cohort studies, 2 RCTs)	Although evidence supports improvement in symptoms after surgery in FAI, no studies have compared surgical and non-surgical treatment. Therefore no conclusion regarding the relative efficacy of one approach over the other can be made. Surgery improves alpha angle but whether this alters the risk of development or progression of hip OA is unknown. This review highlights the lack of evidence for use of surgery in FAI. Given that hip geometry may be modified by non-surgical factors, clarifying the role of non-surgical approaches vs surgery for the management of FAI is warranted.
Forster-Horvath et al. 2016[2] Systematic Review *Database inception through April 2016*	To perform a systematic review comparing outcomes of labral debridement/segmental resection with labral reconstruction as part of a comprehensive treatment strategy for femoroacetabular impingement.	Femoroacetabular impingement (FAI)	Acetabular Labral Debridement/Segmental Resection vs. Reconstruction		20 studies (12 case series or case-control studies, 1 RCT, 7 cohort studies)	**Results:** After an exhaustive search of the available literature, 20 publications were included. Twelve studies explored outcomes after labral debridement/resection in a total of 400 hips, whereas 7 studies reported on outcomes after labral reconstruction in a total of 275 hips. One additional matched-pair control study compared labral resection (22 hips) with reconstruction (11 hips). The surgical intervention was a revision in 0% to 100% for group 1 versus 5% to 55% for group 2. A direct anterior approach was not performed in group 2, and cam-type impingement appeared to make up a larger percentage of group 1. The Tönnis grade ranged from 0 to 1 for group 1 versus 0.3 to 1.1 for group 2. Joint replacements were

Assessment Search dates	Purpose	Condition	Treatment vs. comparators	Primary Outcomes	Evidence- base Used	Primary Conclusions
						performed in 0% to 30% and 0% to 25%, respectively. The modified Harris Hip Score was the most widely used patient-reported outcome measure and suggested that labral reconstruction was not inferior to labral debridement/segmental resection. Clinical outcomes after labral debridement/segmental resection versus labral reconstruction were found to be comparable. In the setting of unsalvageable labral pathology, labral reconstruction was used more frequently as a revision option whereas debridement may be more commonly used in the index setting. **Reoperation:** Of the patients, 0% to 25% underwent conversion to THA. Outcomes after revision labral treatment in the setting of FAI have consistently been shown to be inferior to those of primary surgical procedures in the literature. There were more patients in group 2 who underwent labral reconstruction as a revision procedure. Therefore, these patients may have exhibited more extensive chondral wear, capsular scarring, or injury, and compensatory myotendinous adaptations or neurogenic pain modulation may have developed through the chronicity of their hip disease. A sophisticated labral procedure may have been inadequate to resolve these layered challenges. **Conversion:** Overall, for both groups, the range of conversion to hip arthroplasty was 0% to 30%. Because one study did not

Assessment Search dates	Purpose	Condition	Treatment vs. comparators	Primary Outcomes	Evidence- base Used	Primary Conclusions
						stratify the type of labral procedure (debridement/ segmental resection vs refixation), it is difficult to make precise conclusions on the THA conversion rate. Nonetheless, patients who underwent labral debridement/ segmental resection were not found to transition to THA more frequently than those who underwent labral reconstruction.
Griffin et al. 2017[3] Systematic Review *Database inception to June 2016*	To review the outcomes of hip arthroscopy in older adults and identify factors associated with treatment failures.	Femoroacetabular impingement (FAI)	Noncomparative	Patient-reported Outcomes (validated), Quality of Life, Range of Motion, Reoperation, Complications	8 studies (3 cohort studies and 5 case series)	**Complications:** Overall complication rate of 5.1% (8/157 patients) across five studies. 1 deep venous thrombosis, 1 case of heterotopic ossification (HO), 1 superficial wound infection resolved with oral antibiotics, 1 deep wound infection, 3 cases of psoas tendinitis, and 2 cases of transient sensory neurapraxia (perineum and foot). **Reoperation:** Seven of 8 studies reported reoperation rates. Excluding conversion to arthroplasty, the rate of reoperation was 2.3% (8/351 patients). The majority of reoperations were repeat hip arthroscopy for continued pain and/or labral tear identified on postoperative MRI. There were 3 additional reoperations: 1 for excision of HO, 1 irrigation and debridement for deep wound infection, and 1 lysis of adhesions. When including arthroplasty, the total reoperation rate increased to 20.8%.
Kierkegaard et al. 2017[4] Systematic Review	To investigate pain, activities of daily living (ADL) function, sport function, quality of life	Femoroacetabular impingement (FAI)	Noncomparative	Preoperative and postoperative hip pain and/or hip function during ADL	26 studies (primarily 22 case series, 3 cohort studies, 1 RCT – comparative studies included	Clinically relevant pain and ADL function improvements were first reported between 3 and 6 months, and sport function improvements between 6 months and 1 year after surgery. It is not clear when

Assessment Search dates	Purpose	Condition	Treatment vs. comparators	Primary Outcomes	Evidence- base Used	Primary Conclusions
Database inception to Sept 2015	and satisfaction at different time points after hip arthroscopy in patients with femoroacetabular impingement (FAI).			and sport and/or quality of life and/or postoperative satisfaction absolute scores	comparisons of revision surgery versus surgery)	quality of life improvements were first achieved. On average, residual mild pain and ADL and sport function scores lower than their healthy counterparts were reported by patients following surgery. Postoperative patient satisfaction ranged from 68% to 100%. **Function and Pain:** In patients with FAI, hip pain reduction and ADL function improvements may be achieved between 3 and 6 months after surgery, while sport function improvements occurs between 6 months and 1 year after hip arthroscopy. Hip pain, ADL and sport function improvements are evident at least up to 3 years after hip arthroscopy in patients with FAI. Average scores from patients indicate residual mild hip pain and/or hip function during ADL and sport lower than their healthy counterparts after hip arthroscopy. In patients with FAI, hip pain reduction and ADL function improvements may be achieved between 3 and 6 months after surgery, while sport function improvements occurs between 6 months and 1 year after hip arthroscopy. Hip pain, ADL and sport function improvements are evident at least up to 3 years after hip arthroscopy in patients with FAI. Average scores from patients indicate residual mild hip pain and/or hip function during ADL and sport lower than their healthy counterparts after hip arthroscopy.
Wall et al. 2014[5] Cochrane Review	To determine the benefits and safety of surgery for femoroacetabul	Femoroacetabular impingement (FAI)	Operative treatment for FAI versus placebo, no treatment or	Proportion of participants with 30% or more reduction in	0 randomized or quasi-randomized included	No studies that met the inclusion criteria, with 11 studies that were excluded following detailed review. There were four ongoing studies identified that may

Assessment Search dates	Purpose	Condition	Treatment vs. comparators	Primary Outcomes	Evidence- base Used	Primary Conclusions
Database inception to Nov 2013	ar impingement.		non-operative treatment	pain, preferred pain measures, hip function measures, Quality of Life, Participant global assessment of treatment success, the adverse events		meet the inclusion criteria when they are completed; the results from these ongoing studies may begin to become available within the next five years.
Zhang et al. 2016[6] Systematic Review *Database inception to Aug 2016*	This meta-analysis aims to evaluate the efficacy and safety of hip arthroscopy versus open surgical dislocation for treating femoroacetabular impingement (FAI) through published clinical trials	Femoroacetabular impingement (FAI)	hip arthroscopy versus open surgical dislocation	Alpha angle improvement, Nonarthritic Hip Scores (NAHS), modified Harrison Hip Score (mHHS), Hip Outcome Score-Activities of Daily Living (HOS-ADL), Hip Outcome Score-Sport Specific Subscale (HOS-SSS), reoperation rates, complications	5 cohort studies	Hip arthroscopy resulted in higher NAHS and lower reoperation rates, but had less improvement in alpha angle in patients with cam osteoplasty, than open surgical dislocation. **Reoperation Rate**: Data reporting on reoperation rate are described in 4 studies that included a total of 292 hips. This meta-analysis demonstrated that more additional operations were required after open surgical dislocation than after hip arthroscopy (relative risk [RR]: 0.40, 95% CI: 0.17–0.95, P= 0.04, I2=0%; Fig. 4A). **Complications:** Data reporting on complications are described in 2 studies that included a total of 61 hips. This meta-analysis demonstrated no statistical difference in complications between hip arthroscopy and open surgical dislocation (RR: 0.76, 95% CI: 0.12–4.63, P= 0.76, I2=0%; Fig. 4B).

Appendix Table B2. Summary of Comparative Studies Published after 2014 in Included Systematic Reviews

Author, N, *Study Type*	Treatment vs. comparators	Author Conclusions
Domb et al. 2014 (n=23) *prospective matched-pair cohort*	Segmental labral resection vs. reconstruction	Arthroscopic labral reconstruction is an effective and safe procedure that provides good short-term clinical outcomes in hips with insufficient and nonfunctional labra in the setting of FAI. There was no statistically significant difference between groups regarding the preoperative NAHS (P = .697), any of the other preoperative PROs, or demographic and radiographic data. The mean change in the NAHS was 24.8 ± 16.0 in the RECON group and 12.5 ± 16.0 in the RESEC group. The mean change in the HOS–activities of daily living (HOS-ADL) was 21.7 ± 16.5 in the RECON group and 9.5 ± 15.5 in the RESEC group. Comparison of the amount of change between groups showed greater improvement in the NAHS and HOS-ADL for the RECON group (P = .046 and .045, respectively). There was no statistically significant difference in the mean changes in the rest of the PROs, although there were trends in all in favor of the RECON group. All PROs in both groups showed a statistically significant improvement at follow-up compared with preoperative levels. Chondral Lesion: NR OA (Tonnis Grade): 0 in 15, 1 in 7 Complications: SSI in 2, adhesive capsulitis in 1 Reoperations: 3 rearthroscopies (9-36 mo)
Larson et al. 2014 (n=90) **Matched-pair cohort**	Revision hip arthroscopy vs. primary hip arthroscopy	Arthroscopic hip revision surgery for residual FAI yielded significantly improved outcome measures, but these were inferior to those after primary arthroscopic FAI corrective surgery. Improved femoral head-neck offset after cam decompression, identification and treatment of subspine/AIIS impingement, labral preservation/reconstruction, and capsular preservation/plication may be paramount to achieve satisfactory outcomes.
Skendzel et al. 2014 (n=323) n=323 *Prospective Cohort*	labral repair vs labral debridement	Significant improvement in outcome scores with surgical intervention, with better results seen in some parameters with labral repair compared with debridement.
Frank et al. 2014 (n=64) *Cohort study*	T-capsulotomy with partial capsular repair vs. complete capsular repair	While significant improvements were seen at 6 months, 1 year, and 2.5 years of follow-up regardless of the closure technique, patients who underwent CR of the hip capsule demonstrated superior sport-specific outcomes compared with those undergoing PR. There was a 13% revision rate in the PR group, but no patients in the CR group required revision surgery. While longer term outcome studies are needed to determine if these results are maintained over time, these data suggest improved outcomes after CR compared with PR at 2.5 years after hip arthroscopic surgery for FAI.

Author, N, *Study Type*	Treatment vs. comparators	Author Conclusions
Redmond et al. 2015 (n=174) *Cohort study*	arthroscopic acetabuloplasty and labral refixation without labral detachment vs. with labral detachment	Treatment of pincer- and combined-type impingement with arthroscopic acetabuloplasty and labral refixation without detachment, when possible, resulted in similar patient outcomes compared with acetabuloplasty with labral detachment. We may conclude that in cases where the chondrolabral junction remains intact, acetabuloplasty and labral refixation without detachment is a viable option.
Botser et al. 2014 (n=23) *Prospective Cohort*	surgical hip dislocation vs hip arthroscopy	Improvement in both groups with no significant between group differences.

APPENDIX C. SYSTEMATIC REVIEWS EXCLUDED AT FULL TEXT REVIEW

Citation	Reason for exclusion
Casartelli NC, Leunig M, Maffiuletti NA, Bizzini M. Return to sport after hip surgery for femoroacetabular impingement: a systematic review. Br J Sports Med. 2015 Apr 3:bjsports-2014.	Search strategy only extends thru Oct 2014
de Sa D, Horner NS, MacDonald A, Simunovic N, Slobogean G, Philippon MJ, Belzile EL, Karlsson J, Ayeni OR. Evaluating healthcare resource utilization and outcomes for surgical hip dislocation and hip arthroscopy for femoroacetabular impingement. Knee surgery, sports traumatology, arthroscopy: official journal of the ESSKA. 2016 Dec;24(12):3943.	Not a formal economic analysis
Gillespie JA, Patil SR, Meek RD. Clinical outcome scores for arthroscopic femoral osteochondroplasty in femoroacetabular impingement: a quantitative systematic review. Scottish medical journal. 2015 Feb;60(1):13-22.	Lack of new evidence not captured in previous HTA, Signal Update or newer included reviews.
Gupta A, Redmond JM, Hammarstedt JE, Schwindel L, Domb BG. Safety measures in hip arthroscopy and their efficacy in minimizing complications: a systematic review of the evidence. Arthroscopy. 2014 Oct 1;30(10):1342-8.	Search strategy only extends through 2013
Horner NS, Vikas K, MacDonald AE, Naendrup JH, Simunovic N, Ayeni OR. Femoral neck fractures as a complication of hip arthroscopy: a systematic review. Journal of hip preservation surgery. 2017 Jan 9;4(1):9-17.	Patient population composed of less than 80% FAI
Levy DM, Kuhns BD, Chahal J, Philippon MJ, Kelly BT, Nho SJ. Hip arthroscopy outcomes with respect to patient acceptable symptomatic state and minimal clinically important difference. Arthroscopy. 2016 Sep 1;32(9):1877-86.	Patient population composed of less than 80% FAI
Nakano N, Lisenda L, Jones TL, Loveday DT, Khanduja V. Complications following arthroscopic surgery of the hip: a systematic review of 36 761 cases. Bone Joint J. 2017 Dec 1;99(12):1577-83.	Patient population composed of less than 80% FAI
Weber AE, Harris JD, Nho SJ. Complications in Hip Arthroscopy: A Systematic Review and Strategies for Prevention. Sports medicine and arthroscopy review 2015;23:187-93.	Patient population composed of less than 80% FAI

Facet Neurotomy:
Assessing Signals for Update

Provided by:

Aggregate Analytics, Inc.

Prepared by:

Andrea C. Skelly, PhD, MPH
Erika D. Brodt, BS
Cassandra Winter, BS
Aaron Ferguson, BA

March 14th, 2018

Table of Contents

1. Previous Coverage Decision

A Health Technology Assessment titled: *Facet Neurotomy*, was published on February 21st, 2014 by the Health Care Authority. Findings and Coverage Decision was adopted on May 16th, 2014. The Committee's Coverage Decision is summarized below.

HTCC Coverage Determination

Facet Neurotomy is a **covered benefit with conditions** consistent with the criteria identified in the reimbursement determination.

HTCC Reimbursement Determination

Lumbar Facet Neurotomy is a **covered benefit with the following conditions:**

- Patient(s) must be over 17 years of age, and:
- Has at least six months of continuous low back pain referable to the facet join
- The pain is non-radicular pain
- Condition is unresponsive to other therapies including conservative care
- There are no other clear structural cause of back pain
- There is no other pain syndrome affecting the spine
- For identification, diagnosis, and treatment:
    - Patient must be selected by at least 80% improvement in pain after each of two differential medial branch blocks, one short-acting; on long-acting
    - One or two joints per each intervention, with documented, clinically significant improvement in pain and/or function for six months before further neurotomy at any level

Cervical Facet Neurotomy for cervical pain is a **covered benefit with the following conditions:**

- Limited to C3–4, through C6–7
- Patient(s) over 17 years of age, and:
- Has at least six months of continuous neck pain referable to the facet joint
- The pain is non-radicular
- Condition is unresponsive to other therapies including conservative care
- There are no other clear structural cause of neck pain
- No other pain syndrome affecting the spine
- For identification, diagnosis and treatment:
    - Patients must be selected by 100% improvement in pain after each of two differential medial branch blocks, one short-acting; one long-acting
    - One joint per each intervention, with documented, clinically significant improvement in pain and/or function for size months before further neurotomy at any level

Facet Neurotomy for the thoracic spine **is not covered.**

Facet Neurotomy for headache **is not covered.**

Committee Decision

Based on the deliberations of key health outcomes, the committee decided that it had the most complete information: a comprehensive and current evidence report, public comments, and agency and state utilization information. The committee concluded that the current evidence on Facet Neurotomy demonstrates that there is sufficient evidence to cover with conditions. The committee considered all the evidence and gave greatest weight to the evidence it determined, based on objective factors, to be the most valid and reliable. Based on these findings, the committee voted to cover with conditions Facet Neurotomy.

The committee reviewed selected payer coverage policies from Aetna, Cigna and Health Net. The committee also reviewed practice guidelines from The American Pain Society, National Institute for Health and Clinical Excellence/ National Collaborating Centre for Primary Care, American College of Occupational and Environmental Medicine; American Society of Interventional Pain Physicians; Colorado Division of Workers' Compensation, American College of Occupational and Environmental Medicine, Institute of Health Economics, Work Loss Data Institute, Institute for Clinical Systems Improvement and American Society of Regional Anesthesia and Pain Medicine.

The committee Chair directed HTA staff to prepare a Findings and Decision document on Facet Neurotomy reflective of the majority vote for final approval at the next public meeting.

Medicare Decision and Expert Treatment Guidelines

CMS does not have a national coverage determination (NCD) for Facet Neurotomy, but has a decision on nerve ablation. The committee considered this decision and determined there was no data shown supporting the decision, and HTCC's determination did not conflict with this NCD.

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 (see Figure 1). The key questions in the included original report are listed below.

Key question 1

1. What is the evidence that the use of diagnostic blocks (i.e., medial branch blocks or intraarticular injections with local anesthetic) to select patients for facet neurotomy improves clinical outcomes following facet neurotomy? Consider each of the following:
    a. Diagnostic block versus alternative diagnostic test (e.g., physical examination, radiological examination)
    b. Type of diagnostic block (i.e., medial branch block versus intra-articular injection) for patient selection
    c. Use of a single diagnostic block versus two or more controlled diagnostic blocks (i.e., use of a short- versus a long-acting local anesthetic, or use of a local anesthetic versus saline)
    d. Degree and duration of pain reduction from diagnostic block (e.g., pain relief of ≥ 30% versus ≥ 50%, or ≥ 50% versus ≥80%)
    e. Unilateral versus bilateral diagnostic block
    f. Diagnostic block of single versus multiple levels

Key Question 2

2. With different regions of the spine (lumbar, thoracic, and cervical) considered separately, what is the evidence of short- and long-term comparative efficacy and effectiveness of facet neurotomy (FN) compared with alternatives (e.g., sham neurotomy, therapeutic intra-articular injections, etc.)?
    a. What is the evidence of the short- and long-term comparative efficacy and effectiveness of different types of facet neurotomy (e.g., radiofrequency, pulsed (cooled), chemical, cryoablation, laser)
    b. What is the evidence of the short- and long-term comparative efficacy of repeat neurotomy procedures at the same level and the same side as the initial successful procedure?
    c. Is there evidence of differential effectiveness when conducting unilateral versus bilateral facet neurotomy?
    d. Is there evidence of differential effectiveness when conducting facet neurotomy on single versus multiple spinal levels?

Key Question 3

3. With different regions of the spine (lumbar, thoracic, and cervical) considered separately, what is the comparative evidence regarding adverse events and complications during the periprocedural period and longer term for facet neurotomy?

Key Question 4

4. With different regions of the spine (lumbar, thoracic, and cervical) considered separately, is there evidence of differential efficacy or safety compared with other treatment options in subpopulations? Include consideration of age, gender, race, ethnicity, disability, and workers compensation.

Key Question 5

5. What is the evidence of cost effectiveness of facet neurotomy compared with other treatment options?

3. Methods

3.1 Literature Searches

We conducted an electronic literature search for the period July 1, 2013 to the present using identical search terms used for the original report for key questions 1 through 5. This search included 3 main databases: PubMed, Cochrane Library, and EMBASE. Additional electronic databases were searched; see Appendix A for search methodology and additional details. In addition, we searched the FDA website for updated information on such products.

3.2 Study selection

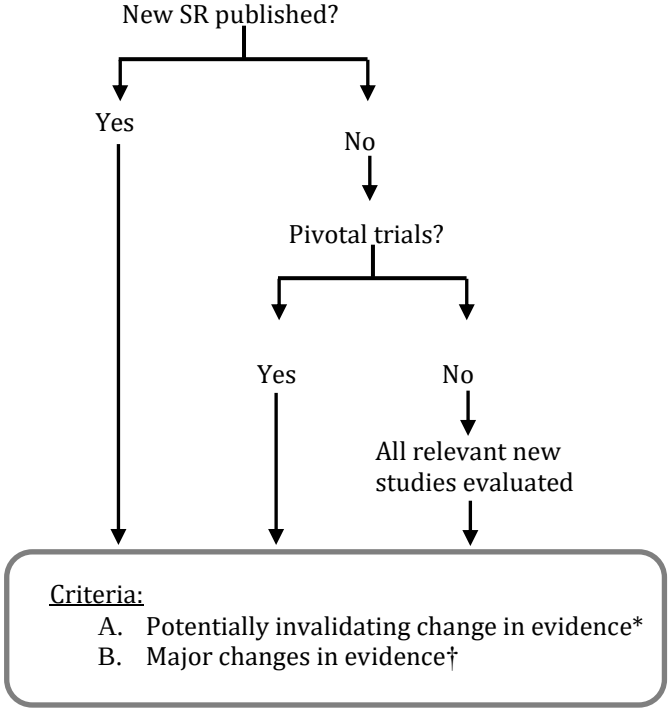
We sought systematic reviews (SR) of randomized controlled trials (RCTs) of efficacy and safety with meta-analysis that included articles that met inclusion and exclusion criteria similar to the original report. In addition we sought systematic reviews reflecting updates or new advances for the technology. Although quality of systematic reviews was not formally evaluated for this report, we chose systematic reviews of head to head trials for efficacy that were the most comprehensive and of higher quality based on the following: report of search strategies (two or more databases 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. Only systematic reviews of RCTs were included for efficacy. Systematic reviews focused on longer-term safety outcomes may include nonrandomized studies. A summary of the included SRs and RCTs is found in Appendix B.

3.3 Compilation of Findings and Conclusions

For this assessment we constructed a summary table that included the key questions, the original conclusions, new sources of evidence, new findings, and conclusions based on available signals. To assess whether the conclusions might need updating, we used an algorithm based on a modification of the Ottawa method, Figure 1.

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

4. Results

4.1 Search

The literature search identified 269 citations. After title and abstract review, 249 articles were excluded and 20 articles that addressed in part or in full the key questions were reviewed at full text. A total of 10 articles were retained for the signal update, Figure 2. A full list of excluded studies and the reasons for exclusions can be found in Appendix C.

We identified two systematic reviews that addressed in part or in full the key questions. Systematic reviews were excluded if they did not include study types of interest and/or if they were not the most comprehensive and of the highest quality, Appendix B. Two systematic reviews related to efficacy were retained. No systematic reviews for safety and no full health technology assessments were identified. No systematic review described results for differential safety (key question 3). We found no cost-effectiveness studies (Key Question 5); there were none in the previous report. Eight new RCTs were identified. No follow-up publications of RCTs included in the previous report were also identified. Clinicaltrials.gov was searched for currently ongoing comparative clinical trials, Appendix D.

The FDA has approved one new lesion probe device for facet neurotomy since the publication of the initial report (Table 1).

Table 1. FDA-Approved Neurotomy Devices approved since the publication of the original report

Manufacturer	Device Name	510(k) Number	Indications for Use	Year of Approval	Recalls?
Stryker Instruments, Kalamazoo, MI, USA	MultiGen 2 RF Generator System	K170242	Intended for coagulation of soft tissues in orthopedic, spinal, and neurosurgical applications. Examples include but are not limited to: Facet Denervation, Trigeminus Neuralgia, Peripheral Neuralgia, and Rhizotomy.	2017	None

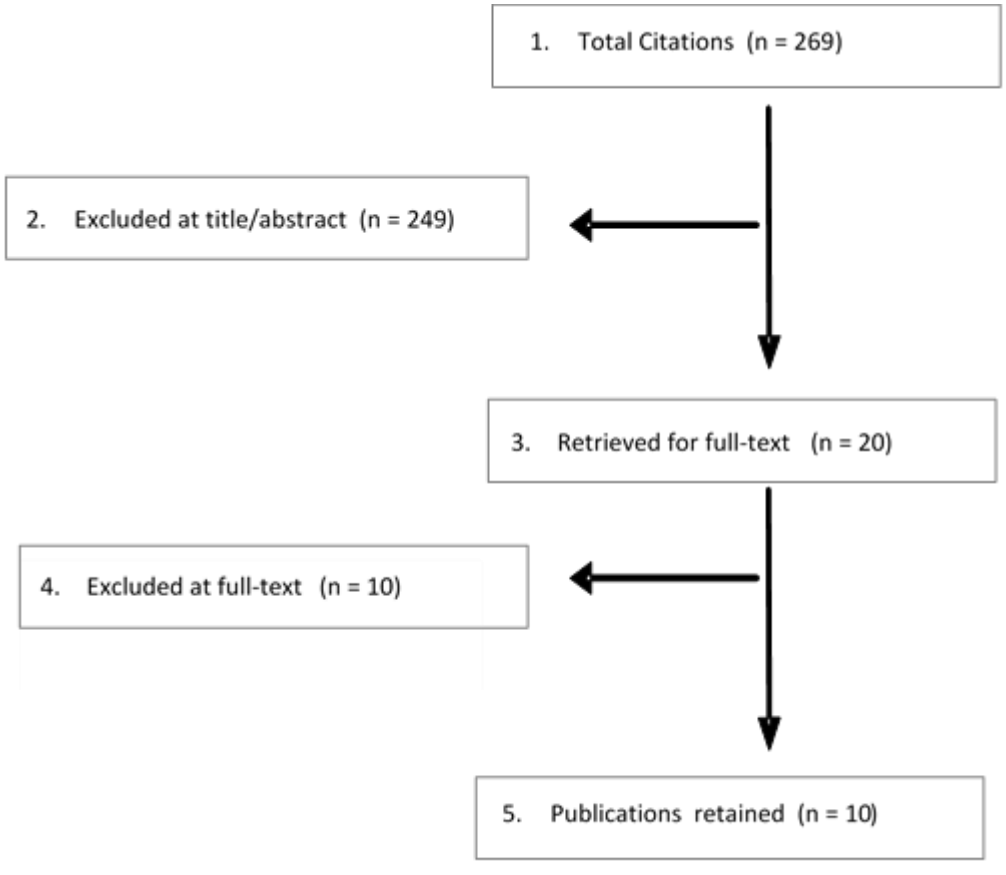


Figure 2. Flow chart showing results of literature search

4.2 Identifying signals for re-review

Tables 2-7 show the original key questions, the conclusions of the original report, the new sources of evidence, the new findings, and the recommendations of Aggregate Analytics, Inc. (AAI) regarding the need for update (Figure 1).

Table 2. Summary Table for Key Question 1.

Key Question 1. What is the evidence that the use of diagnostic blocks (i.e., medial branch blocks or intra-articular injections with local anesthetic) to select patients for facet neurotomy improves clinical outcomes following facet neurotomy? Consider each of the following:			
Conclusions from CER Executive Summary	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
Key Question 1a. Diagnostic block versus alternative diagnostic test (e.g., physical examination, radiological examination)			
Diagnostic block versus physical examination: ***Lumbar spine (LOW evidence)*** • 1 RCT: Neurotomy selection based on clinical exam (n = 51) or one medial branch block ≥50% pain (n = 19) relief and positive GPE • 1 and 3 months: No difference between diagnostic groups in the percentage of patients who achieved "success" (≥50% pain relief and a positive global perceived effect). ***Cervical or Thoracic Spine:*** No evidence	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)
Diagnostic block versus radiological examination: • No evidence in the cervical, lumbar or thoracic spine.	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)
Key Question 1b. Type of diagnostic block (i.e., medial branch block versus intra-articular injection) for patient selection			
Diagnostic medial branch block versus pericapsular block: ***Lumbar spine: (LOW evidence)*** • 1 RCT: Cryodenervation selection based on positive response (≥50% pain relief) to either a diagnostic medial branch block (n = 13) or pericapsular block (n = 13) • No difference between groups in the mean improvement in back pain or function ***Cervical or Thoracic Spine:*** No evidence	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)

Key Question 1. What is the evidence that the use of diagnostic blocks (i.e., medial branch blocks or intra-articular injections with local anesthetic) to select patients for facet neurotomy improves clinical outcomes following facet neurotomy? Consider each of the following:			
Conclusions from CER Executive Summary	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
Other diagnostic block comparators: • ***Cervical, Lumbar or Thoracic Spine:*** No evidence	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)
Key Question 1c. Use of a single diagnostic block versus two or more controlled diagnostic blocks (i.e., use of a short- versus a long-acting local anesthetic, or use of a local anesthetic versus saline)			
Lumbar spine: (LOW evidence) • 1 RCT: RF Neurotomy selection based on positive response (≥50% pain relief) to single diagnostic medial branch block (n=19) or two comparative diagnostic medial branch blocks (n=14). • Short term (1, 3 months): No difference between groups on "success" (≥50% pain relief and a positive global perceived effect)	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)
Key Question 1d. Degree of pain reduction from diagnostic block (i.e., pain relief of ≥30% versus ≥50%, or ≥50% versus ≥80%)			
Lumbar spine (Insufficient evidence) • 4 cohort studies: diagnostic groups based on the pain relief thresholds required to proceed with neurotomy of 50-79% and ≥80% • Taken together, the suggested that pain relief and function may be better following RF neurotomy in those patients who achieved a minimum of 80% pain relief following diagnostic media; branch block though this was not consistently shown across all studies. • Pain at 3 months, 6 months: one study showed no difference between groups, another reported more "success" (≥50% pain relief and a positive global perceived effect) in the higher diagnostic pain relief threshold (≥80%) group. • Function (≥50% improvement in activity level) at 6 months: One retrospective study reported significantly better function in the higher diagnostic pain relief threshold (≥80%) group.	**Lumbar Spine** ***Systematic Review:*** Lee 2017[5] (7 trials) ***RCTs:*** Do 2017,[2] Moussa 2016,[7] Zhou 2016[10] **Cervical or Thoracic Spine**: No systematic reviews or RCTs	***SR*:** Lee et al.'s analysis of equivocal diagnostic block response (≥50% pain relief) and best response (≥80% pain relief, "significant relief" or "near complete relief") indicates that best responders demonstrated better pain relief versus controls at all time points. Meta regression suggests modification by diagnostic block responder type, suggesting that equivocal responders show no difference versus controls or better pain relief with control treatment. A formal test of interaction is not provided. ***RCTs:***	New SR and RCT data suggest that response to diagnostic block may impact pain outcome; additional new trials allow for pooling. These data support the previous HTA's conclusions. A re-review may not be warranted. (Criteria B-1).

Key Question 1. What is the evidence that the use of diagnostic blocks (i.e., medial branch blocks or intra-articular injections with local anesthetic) to select patients for facet neurotomy improves clinical outcomes following facet neurotomy? Consider each of the following:			
Conclusions from CER Executive Summary	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
Cervical or Thoracic Spine: No evidence		Preliminary pooled effect estimates combining data from 3 new RCTs (Do, Moussa, Zhou) with data from trials included in the previous HTA (See Appendix E) provide RCT support for the conclusion of the previous report: • Regardless of the comparator (sham or steroid), trials requiring ≥80% relief (to include "complete or near complete" or "significant" relief) with diagnostic block generally showed better pain improvement compared with those requiring ≥50% relief (to include "good" relief"). • A formal test of interaction is not done.	
Key Question 1e. Unilateral versus bilateral diagnostic block			
No studies were identified which met our inclusion criteria.	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)
Key Question 1f. Single versus multiple level diagnostic block			
No studies were identified which met our inclusion criteria.	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)

GPE: global perceived effect; RCT: randomized controlled trial; RF: radiofrequency

Table 3. Summary Table for Key Questions 2.

Key Question 2. What is the evidence of short- and long-term comparative efficacy and effectiveness of facet neurotomy compared with alternatives (e.g., sham neurotomy, therapeutic intraarticular injections, etc.)?			
Conclusions from CER Executive Summary (Strength of Evidence)	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
KQ 2. Radiofrequency Neurotomy (RFN) versus Sham Neurotomy			
Efficacy: Lumbar spine (LOW Evidence) • Six RCTs; Neurotomy selection criteria varied. Three studies performed diagnostic medial branch block(s) and required ≥50% (2 trials) or ≥80% (1 trial pain relief following the block(s) the three remaining studies employed one or two intraarticular block(s); one specified the percentage of pain relief required. Taken together, the results suggest that outcomes *may* be better following RF neurotomy compared with sham neurotomy, though in many instances there were no differences between treatment groups. Measures of pain and function varied across trials. • Pain, Short-term (1-6 months): ○ Success: One RCT (N = 81) reported no difference for VAS back pain between groups at three months when defined as ≥50% pain relief but marginally significant improvement when defined as (≥50% improvement in GPE of back pain ○ Mean change from baseline, VAS back pain: Four RCTs found no difference between groups in VAS back pain, 1 found no difference in McGill Pain scores at 3-6 months; however, two RCTs favored neurotomy, describing improvement in VAS back pain. ○ Leg and generalized pain; difference in mean change from baseline on leg pain, favored neurotomy in two trials, one of which reported no difference in "success" (≥50% improvement in VAS Scores) ○ The one small trial (N=40) which used 2 MBBs and ≥80% pain relief as a criteria for neurotomy selection consistently reported improve pain with RFN across measures. • Pain, Long-term (12 months) (1 RCT N = 40): significantly improved VAS back pain following RF neurotomy	**Lumbar Spine** ***Systematic Review:*** Lee 2017[5] (7 trials) ***RCTs*** Moussa 2016[7] (N = 120) (in Lee 2017 SR) van Tilburg 2016[9] (N = 60)	***SRs*:** Lee reported results for pain only and pooled across studies of RF neurotomy vs. any comparator (sham or steroid injection). Authors did not report pooled estimates separately for the comparison of RFN versus sham alone. Across comparators for 6 trials (7 publications, N =454 patients), RF neurotomy was not associated with improvement in VAS pain at 1-3 months. At 6 months, RFN was associated with a small improvement in pain (5 trials pooled MD 1.5 95% CI 0.15, 2.8) compared with sham or steroid injection but the difference is not likely to be clinically significant. There was substantial heterogeneity at both time periods. At 12 months, one new study (Moussa) favored RF neurotomy over sham (MD 5.1, 95% CI 4.8, 5.4). Analysis of RF neurotomy groups only suggests that point estimates for pain improvement generally meet an MCID (≥ 3 point improvement in 0-10 VAS), however the lowest confidence interval bound did not exceed the MCID at 3 or 6 months.	Findings from new trials and one systematic review are consistent with the previous report with respect to mean difference in pain improvement, and function for RF neurotomy vs. sham. Additional data on pain success at 6 and 12 months from one new trial significantly favored RF neurotomy versus sham that would update the report. (Criteria B1)

Key Question 2. What is the evidence of short- and long-term comparative efficacy and effectiveness of facet neurotomy compared with alternatives (e.g., sham neurotomy, therapeutic intraarticular injections, etc.)?

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
• Function, short-term (1-6 months): Across 3 trials, ODI scores were improved favoring RFN, however no differences in other functional outcomes were seen in two other trials. • Function, long-term (12 months): Improved ODI scores favoring RFN were reported in 1 trial. • Success on composite scores: No differences between RFN and sham were identified. ***No evidence for any of the following:*** • Efficacy or effectiveness of other types of neurotomy versus sham neurotomy in the lumbar spine. • Effectiveness of neurotomy versus sham neurotomy in the lumbar spine		***RCTs*:** Preliminary pooled effect estimates combining data from two new RCTs (Moussa 2016 and van Tilburg 2016) with data from trials included in the previous HTA (See Appendix E) suggest results were generally consistent with those of the previous report for mean back and leg pain and function. For pain success at 6 and 12 month pooled estimates including one new trial provide additional evidence favoring RFN at 6 and 12 months: • Back pain (improvement in VAS scores): no difference between RF neurotomy and sham at 3 months (1 new trial, van Tilburg) but at 6 months, the pooled estimate tended to favor RF neurotomy but did not reach statistical significance and heterogeneity was substantial (1 new trial, Moussa). One new trial with 12 month data is consistent with the old trial showing statistically greater improvement with RF neurotomy versus sham. • Leg pain (improvement in VAS scores): no difference between groups at 3 months and a tendency to favor RF neurotomy vs. sham at 6 months (1 new trial, Moussa). Longer-term data is available at 24 and 36 months from one new trial also showing a tendency	

Key Question 2. What is the evidence of short- and long-term comparative efficacy and effectiveness of facet neurotomy compared with alternatives (e.g., sham neurotomy, therapeutic intraarticular injections, etc.)?			
Conclusions from CER Executive Summary (Strength of Evidence)	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
		to favor RF neurotomy versus sham but the differences did not reach statistical significance. • Pain "success" (various definitions): the addition of one new trial (Moussa) provides additional evidence. While there was no difference between groups at 3 months (consistent with the previous report), RF neurotomy was substantially favored at both 6 months (RR 2.9, 95% CI 1.6, 5.1) and 12 months (5.0, 95% CI 2.1 to 12.1) compared with sham. • Function (improvement in ODI scores): pooled estimates at 6 and 12 months with the addition of one new trial (Moussa) tended to favor RFN but did not reach statistical significance. Moussa was significant at both time points but due to substantial heterogeneity, pooled estimates are not reliable. Moussa required "complete or near complete" reduction of pain following diagnostic block; van Tilberg required only a decrease ≥ 2 on a 0 to 10 point NRS scale.	

Key Question 2. What is the evidence of short- and long-term comparative efficacy and effectiveness of facet neurotomy compared with alternatives (e.g., sham neurotomy, therapeutic intraarticular injections, etc.)?			
Conclusions from CER Executive Summary (Strength of Evidence)	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
		Intermediate and long term: Moussa reported sustained pain relief at 6, 12, 24 and 36 months	
Efficacy: Cervical spine (Insufficient Evidence) • 1 RCT; Neurotomy selection criteria, 100% pain relieve with anesthetics; 3 MBBs used • More FN patients achieved "Freedom from accustomed pain" compared with sham at 6 months ***No evidence for the following:*** • Effectiveness of neurotomy versus sham neurotomy in the cervical spine • Efficacy or effectiveness of other types of neurotomy compared with sham neurotomy in the cervical spine	**Cervical Spine** No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)
KQ 2. RF Neurotomy versus Spinal Injections/Epidural Block			

Efficacy: Lumbar spine (LOW Evidence)
Taken together, the results suggest that outcomes are similar following RF neurotomy and spinal injections
• Two RCTs; Neurotomy selection, one RCT ≥50% pain relief following a diagnostic medial branch block, other RCT used intra-articular injection, pain relief threshold not described.
• Pain relief
  ○ Success (≥50% pain relief from baseline, 1 RCT): more RFN patients achieved success at 6 and 12 months vs. spinal injections.
  ○ VAS score improvement (2 RCTs): No difference between groups at 6 or 12 months.
• Function (1 RCT): No differences between treatment groups on ODI or Roland-Morris scores at 6 months.
	Lumbar Spine
	
	Systematic Reviews: Lee 2017[5] (7 trials); Piso 2016[8] (4 trials)
	
	RCTs:
	Zhou 2016[10] (N = 80) (in Lee 2017 SR)
	Do 2017[2] (N = 60)
	Hashemi 2014[3] (in Piso 2016 SR) (N = 80)
		SRs:
		Two SRs were identified which included one new trial each.
		
		As stated above, Lee et al. did not provide pooled estimates separately by comparator. One included new trial (Zhou 2016; N=80) reported pain improvement with RFN at 3 months (MD 2.3, 95%CI 1.8, 2.8) and 6 months (4.2, 95% CI 3.7, 4.8) versus injections.
		
		Piso et al. reported significant improvement in VAS pain scores across three trials over all timepoints measured: ≤1 month (pooled MD -1.8, 95% CI -3.1 to -0.6, 2 trials), ≥6 months to <12 months (pooled MD -2.1, 95% CI -3.5 to -0.8, 3
			There are new data that would update the report. New evidence suggests that RF neurotomy may be associated with improved pain relief versus steroid injections. A re-review may be warranted. (Criteria B-1).

Key Question 2. What is the evidence of short- and long-term comparative efficacy and effectiveness of facet neurotomy compared with alternatives (e.g., sham neurotomy, therapeutic intraarticular injections, etc.)?			
Conclusions from CER Executive Summary (Strength of Evidence)	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
		trials), and ≥12 months (pooled MD -2.7, 95% CI -3.4 to -1.9, 2 trials); two of the trials were included in our previous report and one trial was excluded from our previous report. The included new trial (Hashemi 2014; N=80) did not provide detailed data and therefore was not included in the pooled analyses above. This trial reported improvement in both pain (MD in NRS change scores -5) and function (MD in ODI change scores -56.3%) favoring pulsed RFN at 6 months; results were also significant at 3 months but not at 1.5 months. ***RCTs*** None reported on long-term pain. Pain relief: Across the three new trials, results were mixed. Short-term, Do reported significant improvement in back pain favoring intra-articular steroid injection over RFN. Hashemi reports improvement in back pain at 3 months and Zhou reports improvement in leg pain at 1 month. At 6 months, Do reports no difference between RFN and steroid injection; Hashemi and Zhou report sustained improvement in pain compared with steroid injection. Zhou required ≥ 80% pain relief from diagnostic block, Hashemi didn't specify and Do used ≥ 50% pain relief as a threshold.	

Key Question 2. What is the evidence of short- and long-term comparative efficacy and effectiveness of facet neurotomy compared with alternatives (e.g., sham neurotomy, therapeutic intraarticular injections, etc.)?			
Conclusions from CER Executive Summary (Strength of Evidence)	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
		Function: Hashemi reports no improvement in ODI at 1.5 months, but statistically significant improvement at 3 and 6 months. The others did not report on function.	
Efficacy: Cervical spine (LOW Evidence) Taken together, results suggest no difference between RFM and occipital nerve injection. • One RCT, no diagnostic blocks used; RFN compared with occipital nerve injection in patients with cervicogenic headache. • At 2 months, no difference in headache relief (VAS score improvement) or a composite measure 20% reduction in pain (as measured on the VAS scale) or a global perceived effect (GPE) score of +2 or +3 ("much better" or "complete relief"). ***No evidence for any of the following*** • Efficacy or effectiveness of other types of neurotomy compared with spinal injections in the cervical spine. • Neurotomy compared with spinal injections in the thoracic spine	**Cervical spine** No new SRs; ***RCTs*:** Lim 2017[6] (N = 40)	Lim 2017 reports no difference in pain relief between intraarticular RFN and steroid injection in patients with cervical *facet joint pain* at either 3 or 6 months; ≥50% pain relief following diagnostic block was required.	There is limited new evidence that would update the report; however the findings from this small trial are not sufficient to trigger an updated report. (Criterion B1)
KQ 2. RF Neurotomy Plus exercise versus Exercise			
No studies in previous report	**Lumbar spine** No new SRs; ***RCT*:** Juch 2017[4] (N=251)	Radiofrequency denervation combined with a standardized exercise program resulted in either no improvement or no clinically important improvement in chronic low back pain compared with a standardized exercise program alone. There were no differences between treatment groups in mean NRS pain scores at any time up to 12 months and no statistical differences between groups in the proportion of patients achieving	There are new data that would update the report. New evidence suggests that RF neurotomy combined with exercise is not associated with improved pain or function compared with exercise alone. A re-review may be warranted (Criteria B-1).

Key Question 2. What is the evidence of short- and long-term comparative efficacy and effectiveness of facet neurotomy compared with alternatives (e.g., sham neurotomy, therapeutic intraarticular injections, etc.)?			
Conclusions from CER Executive Summary (Strength of Evidence)	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
		>30% pain reduction. There was no difference between groups for function measured via ODI or for Global Perceived Effect.	

CI: confidence interval; GPE: Global Perceived Effect; HTA: Health Technology Assessment; MBB: medial branch block; MCID: minimal clinically important difference; MD: mean difference; NRS: numerical rating scale; ODI: Oswestry Disability Index; RCT: randomized controlled trial; RF: radiofrequency; RFN: radiofrequency neurotomy; VAS: visual analog scale.

Table 4. Summary Table for Key Questions 2a - d.

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
Key Question 2a. What is the evidence of the short- and long-term comparative efficacy and effectiveness of different types of facet neurotomy (e.g., radiofrequency, pulsed (cooled), chemical, cryoablation, laser			
KQ 2a. Conventional versus Pulsed RF Neurotomy:			
Efficacy: Lumbar spine (LOW Evidence) Taken together, results suggest that outcomes are similar with conventional and pulsed RFN • Two RCTs; Neurotomy selection based on ≥50% pain relief following diagnostic MBB. • Pain, short-term (3, 6 months, 2 RCTs): No difference between groups for improvement on VAS scores. Long term, (12 months) 1 RCT favored conventional RFN • Function, short-term (3, 6 months, 2 RCTs) and long term (12 months, 1RCT): No difference between groups for improvement on ODI. ***No evidence for any of the following:*** • Effectiveness of conventional versus pulsed RF neurotomy in the lumbar spine • Efficacy or effectiveness of conventional versus pulsed RF neurotomy in the cervical or thoracic spine	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)
KQ 2a. RF Neurotomy versus Alcohol Ablation:			
Efficacy: Lumbar spine (LOW Evidence) Long-term, outcomes *may* favor alcohol ablation, though there was no difference between treatment groups in the short-term results. • One RCT (N = 40); Neurotomy selection based on 2 diagnostic blocks, degree of pain relief NR. • Composite "success" outcome (VAS score <7 and a revised ODI score <22%) no differences between ablation types at 9 months; alcohol ablation favored between 12 and 24 months. ***No evidence for any of the following:*** • Effectiveness of RF neurotomy vs. alcohol ablation in the lumbar spine • Efficacy or effectiveness of RF neurotomy vs. alcohol ablation in the cervical or thoracic spine	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
KQ 2a: OTHER COMPARISONS			
• No studies meeting the inclusion criteria were identified	***RCTs:*** Aranious 2016[1]; thermal radio-frequency ablation (TRF) alone vs. pulsed dose radio-frequency (PDRF) immediately followed by TRF (N = 55)	Aranious et al.: Although patients receiving PDFR followed TRF demonstrated statistically significant pain scores the morning post-procedure Day 1, there were no differences between groups the evening of Day 1 or on Day 2. An improvement of ≥ 80% following diagnostic block was required for inclusion. Function was not reported.	There is limited new evidence that would update the report; however the findings from this small trial comparing combined use of TRF (continuous) and PDRF with TRF alone is not sufficient to trigger an updated report. (Criterion A1, B1).
KQ 2b. What is the evidence of the short- and long-term comparative efficacy of repeat neurotomy procedures at the same level and the same side as the initial procedure?			
Repeat neurotomy: Lumbar spine (Insufficient evidence) • Six case series; Taken together, results suggest that patients undergoing a second or third procedure may have similar results to those achieved during the first procedure. ***Repeat neurotomy: Cervical spine (Insufficient evidence)*** • Two case series; Taken together, results suggest that patients undergoing a second or third procedure may have similar results to those achieved during the first procedure. ***Repeat neurotomy: Thoracic spine (Insufficient evidence)*** • No studies met inclusion criteria.	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)
KQ2c: Is there evidence of differential effectiveness when conducting unilateral versus bilateral facet neurotomy?			
Unilateral vs. bilateral RF neurotomy effectiveness: Lumbar spine (LOW Evidence) • One retrospective cohort: No difference between treatment groups for the percentage of procedures that resulted in back	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
pain "success" (≥50% pain relief or complete elimination of pain) at a mean of 5.6 months			
KQ2d: Is there evidence of differential effectiveness when conducting facet neurotomy on single versus multiple spinal levels?			
• No studies meeting the inclusion criteria were identified	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)

ODI: Oswestry Disability Index; PDRF: pulsed dose radiofrequency; RCT: randomized controlled trial; RF: radiofrequency; TRF: thermal radiofrequency; VAS: visual analog scale.

Table 5. Summary Table for Key Question 3

Key Question 3: What is the comparative evidence regarding adverse events and complications during the periprocedural period and longer term for facet neurotomy?			
Conclusions from CER Executive Summary (Strength of Evidence)	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
KQ 3: RF Neurotomy versus Sham Neurotomy			
Safety: Lumbar spine (LOW Evidence) • 1 RCT (N=81): no differences between treatment groups in treatment-related pain, change of sensibility, or loss of motor function during the periprocedural period. • 4 RCTs (N=191 total) stated only that no adverse events or complications occurred in either treatment group during the periprocedural period. • No nonrandomized comparative studies or case series met inclusion criteria. ***Safety: Cervical spine (LOW Evidence)*** • 1 RCT (N=24): significantly higher frequency of procedure-related numbness following RF neurotomy vs. sham neurotomy (38% vs. 0%); no differences between groups for all other safety outcomes reported. • No nonrandomized comparative studies or case series met inclusion criteria. ***Safety: Thoracic spine*** • No evidence	**Lumbar spine** ***RCTs:*** van Tilburg 2016[9] (N = 60) **Cervical and Thoracic spine:** no new evidence	van Tilburg 2016 stated that no serious adverse events were encountered during the trial. Four patients withdrew for the following reasons: increased pain after diagnostic test (n=1) and painful procedure despite local anesthetic (n=3); however, the group to which patients were randomized was not reported.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4) Findings from the new trial and are consistent with the previous report with respect to frequency of adverse events.
KQ 3: RF Neurotomy versus Spinal Injections			
Safety: Lumbar spine (LOW Evidence) • 1 RCT (N=100), vs. medial branch block: no difference between treatment groups in any of the following adverse events over 6 months: infection, new motor deficit, new sensory deficit, superficial burns, and increase in lower back pain; a second RCT reported vaguely on adverse events but did not define which specific outcomes they examined. • No harms data in one retrospective cohort; no case series met inclusion criteria	**Lumbar spine** ***RCTs:*** Do 2017,[2] pulsed RF neurotomy (N=60); Zhou 2016,[10] RF neurotomy	**Lumbar spine** Do 2017 reported no adverse events in the pulsed RF group vs. one event (hyperglycemia) in the steroid injection group; Zhou 2016 reported no adverse events in either group.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4) Findings from the new trials and are consistent with the previous report with respect to frequency

Key Question 3: What is the comparative evidence regarding adverse events and complications during the periprocedural period and longer term for facet neurotomy?			
Conclusions from CER Executive Summary (Strength of Evidence)	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
Safety: Cervical and Thoracic spine • No evidence	(N = 80) (in Lee 2017 SR); **Cervical spine:** ***RCTs:*** Lim 2017,[6] pulsed RF neurotomy (N = 40) **Thoracic spine:** no new evidence	**Cervical spine** Lim 2017 reported no adverse events in the pulsed RF group vs. two events in the steroid injection group (1 case each of facial flushing and hyperglycemia).	of adverse events following neurotomy in the lumbar spine. For the cervical spine, there is limited new evidence that would update the report; however the findings from one small trial are not sufficient to trigger an updated report. (Criterion A2, B)

RCT: randomized controlled trial; RF: radiofrequency.

Table 6. Summary Table for Key Question 4

Key Question 4: Is there evidence of differential efficacy or safety compared with other treatment options in subpopulations? Include consideration of age, gender, race, ethnicity, disability, and workers compensation?			
Conclusions from CER Executive Summary (Strength of Evidence)	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
KQ 4: Heterogeneity of treatment effect			
Lumbar spine (LOW evidence) • 1 RCT (N=81); RF neurotomy vs. sham neurotomy; patient selection by either diagnostic medial branch block or clinical exam alone. • None of the following subgroups had differential treatment effect in terms of the composite outcome "success" or GPE pain relief "success": sex, age (18-40 versus >40), duration of pain (≤5 versus > 5 years), employment status (unemployed versus employed), and previous low back surgery. ***Cervical and Thoracic spine*** • No evidence	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)
KQ 4: Comparative efficacy of RF Neurotomy: patients selected on the basis of ≥50% pain relief following medial branch block *In Key Question 1, no direct evidence was identified that type of diagnostic block (i.e., medial branch block versus intra-articular block) affected patient outcomes following facet neurotomy. As a result, no restrictions were placed on type of diagnostic block used for patient selection for studies included in Key Question 2. However, during the public comment period, a peer reviewer (Paul Dreyfuss, MD) indicated that the methods by which patients are selected for facet neurotomy affects the efficacy of the procedure. Specifically, he suggested that patients should be selected on the basis of ≥50% pain relief following one or more diagnostic medial branch block(s). In order to address this concern, we provided the results from on a subgroup studies included in Key Question 2 that selected patients on the basis of ≥50% pain relief following medial branch block.*			
RF Neurotomy vs. Sham Neurotomy: efficacy following medial branch block			
Lumbar spine (LOW evidence) Taken together, the results suggested that outcomes favored RF neurotomy over sham neurotomy. • 3 RCTs (N=111 total); patient selection based on ≥50% or ≥80% pain relief following diagnostic medial branch block. • Pain, Short-term (2-6 months): ○ VAS back pain, mean change from baseline: Two RCTs (N=71 total) favored RF neurotomy, describing significant	**Lumbar Spine** ***Systematic Review:*** Lee 2017[5] **Cervical and Thoracic Spine:**	Lee reported results for pain only and pooled across studies of RF neurotomy vs. any comparator (sham or steroid injection) (6 trials [7 publications], N=454 patients). Authors did not report pooled estimates separately for the comparison of RF neurotomy versus sham alone, or	This section of the report is still valid and does not need updating. (Criteria A1, B-1-4)

Key Question 4: Is there evidence of differential efficacy or safety compared with other treatment options in subpopulations? Include consideration of age, gender, race, ethnicity, disability, and workers compensation?			
Conclusions from CER Executive Summary (Strength of Evidence)	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
improvement in back pain VAS scores over 2-6 months; the third RCT (N=40) found no difference between groups. ○ VAS leg and generalized pain, mean change from baseline (1 RCT, N=40); significantly improved leg and generalized pain VAS scores following RF neurotomy at 6 months. ○ The one small trial (N=40) which used two medial branch blocks and ≥80% pain relief as a criteria for neurotomy selection consistently reported improve pain with RFN across measures. • Pain, Long-term (12 months) (1 RCT, N = 40): significantly improved VAS back pain scores following RF neurotomy • Function, Short-term (2-6 months): Two RCTs (N=71 total) reported significant improvement in ODI scores favoring RF neurotomy. A third trial (N=31) found no difference between groups for improvement in Waddell scores at 2 months. • Function, Long-term (12 months) (1 RCT, N=40): significant improvement in ODI scores favoring RF neurotomy ***Cervical spine (INSUFFICIENT evidence)*** • 1 RCT (N=24); patient selection based on three medial branch blocks and 100% pain relief following diagnostic blocks (i.e. anesthetic) and 0% pain relief when saline was injected. • Back pain, Short-term (6 months): significantly more patients in the RF neurotomy group had achieved freedom from "accustomed pain" compared with those in the sham group. ***Thoracic spine*** • No evidence	no new evidence	for the type of diagnostic block used (medial branch, intraarticular). Authors' analysis of equivocal diagnostic block response (≥50% pain relief) and best response (≥80% pain relief, "significant relief" or "near complete relief") indicates that best responders demonstrated better pain relief versus controls at all time points. Meta regression suggests modification by diagnostic block responder type, suggesting that equivocal responders show no difference versus controls or better pain relief with control treatment. A formal test of interaction is not provided. As stated above, results were not reported by type of diagnostic block.	
RF Neurotomy vs. Spinal injection: efficacy following medial branch block			
Lumbar spine (LOW evidence) • 1 RCT (N=56); patient selection based on ≥50% pain relief following diagnostic medial branch block.	**Lumbar Spine** ***RCT***: Zhou 2016[10] (N =	Zhou selected patients based on ≥80% pain relief following diagnostic medial branch block *or* intraarticular injection;	This section of the report is still valid and does not need updating. (Criteria A1, B-1-4)

Key Question 4: Is there evidence of differential efficacy or safety compared with other treatment options in subpopulations? Include consideration of age, gender, race, ethnicity, disability, and workers compensation?			
Conclusions from CER Executive Summary (Strength of Evidence)	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
• Pain and Function, Short-term (6 months): no difference between treatment groups for improvement in VAS back pain scores and ODI or Roland Morris scores. ***Cervical and Thoracic spine*** • No evidence	80) (in Lee 2017 SR) **Cervical and Thoracic Spine:** no new evidence	however, results were not reported by type of diagnostic block. Authors report results for pain only. Greater improvement in VAS pain scores was seen with RF neurotomy at 3 months (MD 2.3, 95% CI 1.8, 2.8) and 6 months (MD 4.2, 95% CI 3.7, 4.8) versus injections. A formal test of interaction is not provided. As stated above, results were not reported by type of diagnostic block.	

CI: confidence interval; GPE: Global Perceived Effect; MD: mean difference; ODI: Oswestry Disability Index; RCT: randomized controlled trial; RF: radiofrequency; RF: radiofrequency; VAS: visual analog scale.

Table 7. Summary Table for Key Question 5

Key Question 5: What is the evidence of cost effectiveness of facet neurotomy compared with other treatment options?			
Conclusions from CER Executive Summary (Strength of Evidence)	**New Sources of Evidence**	**New Findings**	**Conclusion from AAI**
• No studies meeting the inclusion criteria were identified	No systematic reviews or RCTs	No studies meeting inclusion criteria were identified.	This section of the report is still valid and does not need updating. (Criteria A1, B-1-4)

RCT: randomized controlled trial

5. Conclusions

Tables 2-7 show the original key questions, the conclusions of the original report, the new sources of evidence, the new findings, and the conclusions of Aggregate Analytics, Inc. (AAI) with respect to the criteria that identify a trigger for an update (Figure 1).

5.1 Key Question 1 (Diagnostic):

- 1a-c, e-f: Comparisons of diagnostic block versus alternative diagnostic test; type of diagnostic block; use of a single versus two or more controlled diagnostic blocks; unilateral versus bilateral diagnostic block; and single versus multiple level diagnostic block:
    - No new systematic reviews or RCTs published since the previous HTA that evaluated whether the use of diagnostic blocks (considering the above comparisons) to select patients for facet neurotomy improves clinical outcomes following facet neurotomy were identified (Criteria A-1, A-3, B-1-4). These sections do not need updating.
- 1d: Comparison of response to diagnostic block:
    - New SR and RCT evidence suggest that response to diagnostic block (e.g., ≥50% vs. ≥80% relief) may impact pain outcome; additional new trials allowed for a preliminary pooled analysis. These data support the previous HTA's conclusions that pain relief may be better in patients achieving a greater degree (e.g., ≥80%) of relief with diagnostic block. A re-review may not be warranted. (Criteria B-1).

5.2 Key Question 2 (Efficacy): For the comparison of RF neurotomy versus sham in the lumbar spine, findings from new trials and one systematic review are consistent with the previous report with respect to mean pain improvement and function, however, additional evidence from pooled estimates that include one new trial significantly favored RF neurotomy versus sham on pain success at 6 and 12 months and would update the report (Criteria B-1). There is new evidence (from 2 SRs, 3 RCTs) suggesting that RF neurotomy may be associated with improved pain relief versus steroid injections in the lumbar spine. Additionally, a new comparator was identified for the lumbar spine: new evidence from one RCT suggests that RF neurotomy combined with a standardized exercise program is not associated with improved pain or function compared with exercise alone. There are new data that would update this section of the report. A re-review may be warranted. (Criterion B-1).

No new evidence was identified for the comparison of RF neurotomy versus sham in the cervical spine. There is limited new evidence that would update the report for the comparison of RF neurotomy versus steroid injection; however the findings from this small trial alone are not sufficient to trigger an updated report (Criterion B-1).

5.3 Key Question 2a-d (Efficacy):

- 2a: Comparison of different types of facet neurotomy
    - Conventional versus pulsed RF neurotomy and RF neurotomy versus alcohol ablation: no new systematic reviews or RCTs published since the previous HTA were identified which met inclusion criteria. (Criteria A-1, A-3, B-1-4)

    - One new RCT compared thermal RF neurotomy alone versus pulsed dose RF neurotomy immediately followed by thermal RF and showed no difference in pain between groups the evening of Day 1 or on Day 2 (function was not reported). However, findings from one small trial alone are not sufficient to trigger an updated report (Criterion B-1). This section does not need updating.
- 2b-d: Comparisons of repeat neurotomy procedures (same level and side as initial successful procedure); unilateral versus bilateral facet neurotomy; and facet neurotomy on single versus multiple spinal levels.
    - No new systematic reviews or RCTs published since the previous HTA that evaluated the above comparisons were identified which met inclusion criteria. (Criteria A-1, A-3, B-1-4). These sections do not need updating.

5.4 Key Question 3 (Safety): New evidence from three RCTs of the lumbar spine (1 comparing RF neurotomy with sham neurotomy and 2 comparing conventional or pulsed RF with steroid injections) does not change the conclusions from the previous report (criteria A-1-3); there are not any major changes in the evidence base (criteria B-1-4). For the cervical spine, there is limited new evidence from one RCT (pulsed RF neurotomy vs. steroid injection); however the findings from one trial are not sufficient to trigger an updated report (criteria B-2, 3). This section does not need updating.

5.5 Key Question 4 (Differential efficacy or safety): No new systematic reviews or RCTs published since the previous HTA were identified which met inclusion criteria and evaluated heterogeneity of treatment effect for facet neurotomy compared with other treatment options in subpopulations (e.g., age, sex, race, ethnicity, disability, and workers compensation) (Criteria A-1, A-3, B-1-4). This section does not need updating.

5.6 Key Question 5 (Cost-effectiveness): No new systematic reviews (that included new studies) or RCTs published since the previous HTA were identified which met inclusion criteria that evaluated the cost effectiveness of facet neurotomy compared with other treatment options (Criteria A-1, A-3, B-1-4). This section does not need updating.

REFERENCES

1. Arsanious D, Gage E, Koning J, et al. Pulsed Dose Radiofrequency Before Ablation of Medial Branch of the Lumbar Dorsal Ramus for Zygapophyseal Joint Pain Reduces Post-procedural Pain. Pain physician 2016;19:477-84.
2. Do KH, Ahn SH, Cho YW, Chang MC. Comparison of intra-articular lumbar facet joint pulsed radiofrequency and intra-articular lumbar facet joint corticosteroid injection for management of lumbar facet joint pain: A randomized controlled trial. Medicine 2017;96:e6524.
3. Hashemi M, Hashemian M, Mohajerani SA, Sharifi G. Effect of pulsed radiofrequency in treatment of facet-joint origin back pain in patients with degenerative spondylolisthesis. European spine journal : official publication of the European Spine Society, the European Spinal Deformity Society, and the European Section of the Cervical Spine Research Society 2014;23:1927-32.
4. Juch JNS, Maas ET, Ostelo R, et al. Effect of Radiofrequency Denervation on Pain Intensity Among Patients With Chronic Low Back Pain: The Mint Randomized Clinical Trials. Jama 2017;318:68-81.
5. Lee CH, Chung CK, Kim CH. The efficacy of conventional radiofrequency denervation in patients with chronic low back pain originating from the facet joints: a meta-analysis of randomized controlled trials. The spine journal : official journal of the North American Spine Society 2017;17:1770-80.
6. Lim JW, Cho YW, Lee DG, Chang MC. Comparison of Intraarticular Pulsed Radiofrequency and Intraarticular Corticosteroid Injection for Management of Cervical Facet Joint Pain. Pain physician 2017;20:E961-e7.
7. Moussa WM, Khedr W. Percutaneous radiofrequency facet capsule denervation as an alternative target in lumbar facet syndrome. Clinical neurology and neurosurgery 2016;150:96-104.
8. Piso B, Reinsperger I, Rosian K. Radiofrequency denervation for sacroiliac and facet joint pain. Vienna2016.
9. van Tilburg CW, Stronks DL, Groeneweg JG, Huygen FJ. Randomised sham-controlled double-blind multicentre clinical trial to ascertain the effect of percutaneous radiofrequency treatment for lumbar facet joint pain. The bone & joint journal 2016;98-b:1526-33.
10. Zhou Q, Zhou F, Wang L, Liu K. An investigation on the effect of improved X-rays-guided radiofrequency thermocoagulation denervation on lumbar facet joint syndrome. Clinical neurology and neurosurgery 2016;148:115-20.

APPENDIX A. SEARCH STRATEGIES

Search strategy for PubMed—Search dates: 07/01/13 to present

	Search terms	Number of articles
#1	Facet OR Zygapophyseal OR "Zygapophyseal Joint"[Mesh] OR "medial branch"	4,264
#2	Neurotomy OR "Rhizotomy"[Mesh] OR Rhizotomy OR "Articular rhizolysis" OR rhizolysis OR "Radiofrequency neurotomy" OR "Radiofrequency denervation" OR (radiofrequency AND "denervation"[MeSH Terms]) OR Denervation OR "Radiofrequency neurolysis" OR "Radiofrequency facet denervation" OR "Pulsed Radiofrequency Treatment"[Mesh] OR "Cooled radiofrequency ablation" OR "cooled ablation" OR ablat* OR chemodenervation OR "Chemical facet neurolysis" OR "cryosurgery"[MeSH Terms] OR Cryoablation OR radiofrequency	45,289
#3	#1 AND #2	222
#4	(In Vitro[TI] OR Cadaver*[TIAB] OR Case Reports[Publication Type] OR rat[TI] OR rats[TI] OR mouse[TI] OR mice[TI] OR dog[TI] OR dogs[TI] OR sheep[TI] OR rabbit[TI] OR "experimental model"[TI])	
#5	#3 NOT #4	189
#6	Additional references identified from hand searching	0

Search strategy for Cochrane—Search dates: 2013 to 03/02/18

	Search terms	Number of articles
#1	Facet OR Zygapophyseal OR "Zygapophyseal Joint"(Mesh) OR "medial branch"	565
#2	Neurotomy OR "Rhizotomy"(Mesh) OR Rhizotomy OR "Articular rhizolysis" OR rhizolysis OR "Radiofrequency neurotomy" OR "Radiofrequency denervation" OR (radiofrequency AND "denervation"(MeSH Terms)) OR Denervation OR "Radiofrequency neurolysis" OR "Radiofrequency facet denervation" OR "Pulsed Radiofrequency Treatment"(Mesh) OR "Cooled radiofrequency ablation" OR "cooled ablation" OR ablat* OR chemodenervation OR "Chemical facet neurolysis" OR "cryosurgery"(MeSH) OR Cryoablation OR radiofrequency	3870
#3	#1 AND #2	50
#4	(In Vitro(ti) OR Cadaver*(ab,ti) OR Case Reports(Publication Type) OR rat(ti) OR rats(ti) OR mouse(ti) OR mice(ti) OR dog(ti) OR dogs(ti) OR sheep(ti) OR rabbit(ti) OR "experimental model"(ti))	
#5	#3 NOT #4	46* (19 unique citations)

*Other reviews, technology assessments, and economic evaluations were not included in title abstract triage—all citations were abstracts and/or were not study types of interest

EMBASE search strategy—Search dates: 2013 to 03/02/2018

	Search terms	Number of articles
#1	'facet joint' OR 'zygapophyseal joint' OR 'medial branch'	1,759
#2	'neurotomy' OR 'rhizotomy' OR 'radiofrequency' OR 'denervation' OR ablation	71,137
#3	#1 AND #2	270
#4	Article/lit OR review/lit	
#5	#3 AND #4	151 (60 unique citations)

Additional electronic databases were searched using key words and included ClinicalTrials.gov, AHRQ, National Guideline Clearinghouse and INAHTA for eligible studies, including health technology assessments (HTAs), systematic reviews, primary studies and FDA reports. Original search was performed through October 4, 2013. The updated search goes from July 1, 2013 to the present.

The first twenty related PubMed articles of all newly included studies were evaluated for inclusion. Bibliographies of included systematic reviews were reviewed for relevant articles

APPENDIX B. SUMMARY OF INCLUDED STUDIES

Appendix Table B1. Summary of systematic reviews included for efficacy

Assessment *Search dates*	Purpose	Condition	Treatment vs. comparators	Primary Outcomes	Evidence-base Used	Primary Conclusions
Lee 2017 *Database inception to October 12, 2016*	To elucidate the precise effects of RF in patients with low back pain originating from the facet joints relative to those obtained using control treatments, with particular attention to consistency in the denervation protocol.	Facet joint disease of the lumbar spine	RF denervation vs sham or epidural nerve block	Pain VAS	7 RCTs (2 new RCTs: Moussa 2016, Zhou 2016)	**RFN vs control, pain:** At a short term follow-up (1-3 months), a pooled analysis across comparators for 6 trials reported no difference in pain VAS scores between RFN vs sham. At a 1 to 3 month follow-up across comparators for 6 trials, RFN was not associated with pain VAS improvement. At an intermediate follow-up at 6 months, RFN was associated with a small improvement in pain VAS (5 trials pooled; MD 1.5 95% CI 0.15, 2.8) compared to sham or steroid injection but the difference was not likely to be clinically significant. At both a short and intermediate term follow-up, there was substantial heterogeneity. At a long term follow-up of 12 months, one new study (Moussa 2016) found a statistically significant difference favoring RFN over sham (MD 5.1, 95% CI 4.8, 5.4). An analysis of the RN group suggested that point estimates for pain improvement generally meet an MCID (≥ 3 points improvement in 0-10 VAS), however the lowest confidence interval bound did not exceed MCID at either 3 or 6 months. **RFN equivocal diagnostic block response or best diagnostic block response vs control, success on pain VAS:** The authors' analysis of equivocal diagnostic block response (≥50% pain relief) and best response (≥80% pain relief, "significant relief", or "near complete relief") indicates that best responders demonstrated better pain relief compared to controls at all time points. A meta regression analysis suggests modification by diagnostic block responder type, suggesting that equivocal

Assessment *Search dates*	Purpose	Condition	Treatment vs. comparators	Primary Outcomes	Evidence-base Used	Primary Conclusions
						responders show no difference versus controls or better pain relief with control treatment. A formal test of interaction is not provided. **RFN vs spinal injections:** Authors did not provide pooled estimates separately by comparator. One included new trial (Zhou 2016) reported pain improvement with RFN over injections at 3 months (MD 2.3, 95% CI 1.8, 2.8) and 6 months (MD 4.2. 95% CI 3.7, 4.8).
Piso 2016	To compare RF denervation to placebo or other treatments in patients with chronic facet joint pain and a positive response to the diagnostic block.	Chronic facet joint pain	RFN vs steroid injections	Pain, functional status, global improvement, HR-QoL; ability to work, satisfaction with treatment, safety complications	RFN vs steroid injections: 4 RCTs	**RFN vs spinal injections/epidural blocks, pain:** Authors report improvement in VAS pain scores across three trials over all timepoints measured: ≤1 month (pooled MD -1.8, 95% CI -3.1 to -0.6, 2 trials), ≥6 months to <12 months (pooled MD -2.1, 95% CI -3.5 to -0.8, 3 trials), and ≥12 months (pooled MD -2.7, 95% CI -3.4 to -1.9, 2 trials); two of the trials were included in our previous report and one trial was excluded from our previous report. The included new trial (Hashemi 2014; N=80) did not provide detailed data and therefore was not included in the pooled analyses above

CI: confidence interval; HR-QoL: health-related quality of life; MCID: minimal clinically important difference; MD: mean difference; RF: radiofrequency; RFN: radiofrequency neurotomy; VAS: visual analog score

Appendix Table B2. Study characteristics and results of new RCTs

Author (Year)	Demographics	Results	Conclusions	Comments
RFN vs. sham neurotomy, lumbar				
Moussa 2016	N=80 *RFN vs sham neurotomy* Age, mean: 56.5 vs 55.9 years Female: 72.5% Mean duration of procedure: 35 minutes *Total* F/U: 3, 6, 12, 24, and 36 months RFN at facet joints procedure description (n=40): After sensory and motor tests, radiofrequency was delivered at 85°C for 90 seconds at both the medial and lateral sides of the facet joint *Device used:* RFG-1A Sham neurotomy procedure description (n=40): Same procedure but without delivering current to the electrode. *Device used:* RFG-1A	Pain: RFN vs sham neurotomy* VAS, mean improvement (SD): • Baseline: NR vs NR • 3 months: 6.0 (1.0) vs 5.4 (1.1), p=0.01 • 6 months: 6.0 (1.1) vs 2.1 (0.4), p<0.001 • 12 months: 5.8 (1.0) vs 0.7 (0.3), p<0.001 • 24 months: 2.3 (0.4) vs 0.5 (0.1), p<0.001 • 36 months: 2.2 (0.8) vs 0.4 (0.2), p<0.001 Pain reduction >50%, n/N (%): • 3 months: 30/40 (75%) vs 23/40 (57.5%), p=0.008 • 6 months: 24/40 (60%) vs 8/40 (20%), p<0.001 • 12 months: 18/40 (45%) vs 3/40 (7.5%), p<0.001 • 24 months: 7/40 (17.5%) vs 1/40 (2.5%), p=0.01 • 36 months: 5/40 (12.5%) vs 1/40 (2.5%), p=0.052 Function: RFN vs sham neurotomy ODI, mean change: • 3 months: 44.3 vs 39.8 • 6 months: 40.3 vs 10.3 • 12 months: 31.6 vs 5.9 • 24 months: 12.3 vs 3.2 • 36 months: 8.2 vs 2.9	**Pain:** Based on calculations performed by AAI using the reported data, RFN had a statistically significant better effect on pain VAS at all time points. The RFN group had a statistically significant higher percent of patients reaching >50% reduction in pain than the sham neurotomy group at all time points except 36 months. **Function:** The authors did not provide enough information to draw conclusions on the impact of RFN compared to sham neurotomy on functional outcomes.	Authors report no conflict of interest Authors report that no funding was received for the research
Van Tilburg 2016	N=60 *RFN group vs sham neurotomy group* Age, median (IQR): 65 (12) vs 58 (12) years	Pain: RFN vs sham neurotomy VAS, mean (SD): • Baseline: 7.2 (1.4) vs 7.4 (0.8) • 1 month: 5.3 (1.8) vs 5.5 (1.9), p NS	**Pain:** The authors reported no differences in pain VAS scores at a 1 month follow-up	Authors state that no benefits in any form have been received or will be received from a commercial part related

Author (Year)	Demographics	Results	Conclusions	Comments
	Total Female: 57% BMI, mean (SD): 29.6 (5.3) Caucasian: 100% F/U: 1 and 3 months <u>RFN procedure description (n=30)</u>: 1 mL of 2% lidocaine was infiltrated into skin. After sensory and motor tests, RF heat lesion delivered at 80°C for 60 seconds per level. *Device used:* NT2000, Neurotherm <u>Sham neurotomy procedure description (n=30)</u> Sham group underwent the same procedure but without RF lesions *Device used:* NT2000, Neurotherm			directly or indirectly to the subject of this article. Funding NR
RFN vs spinal injections/epidural block, lumbar				
Do 2017	N=60 *PRF vs ICI* Age, mean (SD): 67 (9.6) vs 63 (10.9) years Female: 60% *Total* F/U: 2 weeks, 1, 3 and 6 mos. <u>PRF procedure description (n=30)</u>: Treatment was administered at 5Hz with a 5-millisesond pulsed width for 360 seconds, at 55V. Electrode tip temperature did not exceed 42°C. *Device used:* Cosman G4 radiofrequency generator	<u>Pain: PRF vs ICI</u> NRS, mean change (SD): • Baseline: 4.9 (0.8) vs. 5.0 (0.8) • 2 weeks: 2.3 (1.4) vs 1.4 (0.8), p<0.001 • 1 month: 2.5 (1.4) vs 1.8 (1.2), p=0.011 • 3 months: 2.5 (1.3) vs 2.9 (1.4), p=NS • 6 months: 2.7 (1.5) vs 3.2 (NR), p=NS <u>Percent of patients with pain relief of ≥50%, PRF vs ICI</u> • 6 months: 50.0% (15/30) vs 46.7% (14/30), p=NS	**Pain:** Authors report statistically significant improvement in pain VAS score for the PRF group over the ICI group at 2 weeks and 1 month. The difference was not significant at 3 and 6 months. There was no difference in the percent of patients with pain relief ≥50% at 6 months.	The authors declare no conflict of interest Funding NR

Author (Year)	Demographics	Results	Conclusions	Comments
	ICI procedure description (n=30): 10mg (0.25mL) of dexamethasone mixed with 0.25mL of 0.125% bupivacaine injected.			
Hashemi 2014	N=80 *PRF group vs steroid injection* Age: 64.3 (13.3) vs 63.9 (11.5) years Female: 42% vs 44% BMI: 23.4 (5.3) vs 22.6 (4.8) Duration of Low Back Pain: 3.4 (2.3) vs 3.8 (2.4) years History of Smoking: 34% vs 38% *Total* F/U: 1.5, 3 and 6 months PRF procedure description (n=40): Local anesthesia was administered. After sensor and motor tests were performed, radiofrequency was delivered in 2 x 20 ms/s duration 120 seconds with 45 v with silent time 480 ms. Skin temperature did not exceed 42°C. *Device used:* NeuroTherm radiofrequency generator Steroid injection procedure description (n=40): Injection of 1 mL (40 mg) of triamcinolone and 0.5 mL bupivacaine (0.5%)	Pain: PRF vs Steroid Injection NRS, mean (SD)*: • Baseline: 7.4 (1.1) vs. 8.1 (1.0) • 1.5 months: 2.5 (0.8) vs 3.2 (0.8), p NS • 3 months: 2.9 (0.9) vs 5.9 (0.8), p<0.05 • 6 months: 2.4 (1.9) vs 7.4, (1.2) p<0.05 Function: PRF vs Steroid Injection ODI%, mean (SD)†: • Baseline: 75.6 (14.3) vs. 74.0 (NR) • 1.5 months: 2.5 (NR) vs 3.2 (NR), p NS • 3 months: 2.9 (NR) vs 5.9 (NR), p=0.022 • 6 months: 19.3 (9.5) vs 7.4 (NR), p<0.03	**Pain:** Authors report no difference in pain VAS scores between groups at 1.5 months. At 3 and 6 months follow-up, authors report the PRF group had statistically significant better pain VAS scores than the steroid injection group. **Function:** Authors report no difference in ODI scores between groups at 1.5 mos. At 3 and 6 months follow-up, authors report the PRF group had statistically significant better ODI scores than the steroid injection group.	Conflict of interest NR Funding NR
Zhou 2016	N=80 *RF-T group vs spinal injection* Age, mean (SD): 56.5 (8.7) vs 54.6 (7.5) years	Pain: RFN vs spinal injection VAS, mean (SD‡): • Baseline: 6.7 (0.9) vs 6.8, p = NS • 1 week: 1.4 (0.3) vs 1.9 (0.2), p = NS	**Pain:** Authors report no difference in treatments at 1 week but found that the RFN had statistically	Authors declare no conflicts of interest Funding NR

Author (Year)	Demographics	Results	Conclusions	Comments
	Female: 42.5% vs 47.5% *Total* F/U: 1 week, 1 month and 6 months <u>RFN procedure description (n=40)</u>: 3 mL of 2% lidocaine was injected followed by sensory and motor tests. RF-T was delivered at 80°C was performed for 90 s. *Device used:* Smith-Nephew Electrothermal 20s Spine System Radiofrequency Device <u>Spinal injection procedure description (n=40)</u>: 5 mL solution containing 1 mL of betamethasone and 1 mL of 2% lidocaine (diluted with normal saline) into facet joint cavity and medial branch of the spinal nerve. Infiltration block was also performed around the facet joint.	• 1 month: 1.4 (1.2) vs 3.6 (0.9), p < 0.05 • 6 months: 1.7 (1.6) vs 5.8 (1.1), p < 0.01 <u>Efficacy: RFN vs spinal injection</u> Proportion of patients with 'excellent§' rating: • 6 months: 62.5% vs 12.5%, p<0.01	significant lower pain VAS scores compared to the spinal injection group at 1 and 6 months. **Efficacy:** The authors report that a statistically significant higher proportion of patients in the RFN group had an efficacy rating of excellent.	
RFN vs PRF neurotomy + RFN, lumbar				
Arsanious 2016	N=55 Age, mean (SD): 51.3 (10.5) years Female: 77% BMI, mean (SD): 36.0 (10.0) F/U: 1 day AM, 1 day PM, 2 days AM, 2 days PM <u>Procedure description, RFN</u>: After sensory and motor tests, RF heat lesions were delivered at 80°C for 90 seconds at each level treated.	<u>Pain: RFN vs PRF neurotomy + RFN</u> VAS, mean (SD): • Day 1 AM: 4.43 (2.9) vs 2.38 (2.4), p = 0.01 • Day 1 PM: 4.80 (3.2) vs 3.08 (2.8), p = 0.06 • Day 2 AM: 3.86 (2.8) vs 2.31 (2.7), p = 0.06 • Day 2 PM: 3.90 (2.7) vs 2.60 (2.4), p = 0.09	**Pain:** Authors reported a statistically significant difference favoring PRF neurotomy+RFN in pain VAS scores at post-procedure Day 1 AM, but no differences were observed in Day 1 PM or on Day 2.	Authors declare no conflicts of interest Funding: No external funding was provided

Author (Year)	Demographics	Results	Conclusions	Comments
	Device used: NeuroTherm NT2000iX Procedure description, PRF+RFN: After sensory and motor tests, RF heat lesions were delivered at 80°C for 90 seconds at each level treated. Immediately after, pulsed RF waves at 42°C at 2 Hz for 240 pulses were delivered. *Device used:* NeuroTherm NT2000iX			
RFN+exercise vs exercise, lumbar				
Juch 2017	N=251 *RFN+Exercise vs Exercise Alone* Age, mean (SD): 52.9(11.4) vs 52.6 (10.8) years Female: 55.5% vs 51.7% Pain Duration, median: 146 vs 100.3 months *Total* F/U: 1 week, 1 and 6 months Exercise alone procedure description (n=126): All patients received standardized 3 month (8-12 hours) exercise program based on Dutch physical therapy guidelines, focusing on quality of movement and behavior. RFN+exercise procedure description (n=125): Within 1 week of the first exercise session, patients underwent RFN. Sensory and motor tests were	Pain: RFN+Exercise vs. Exercise NRS, mean (95%CI): • Baseline, mean (SD): 7.14 (1.38) vs. 7.19 (1.29) • 3 weeks: 5.17 (4.73 to 5.61) vs. 5.92 (5.58 to 6.26); MD −0.41 (−1.02 to 0.19), p=0.18 • 1.5 months: 5.19 (4.76 to 5.61) vs 5.90 (5.53 to 6.26); MD −0.38 (−0.96 to 0.20), p=0.20 • 3 months: 5.01 (4.59 to 5.43) vs 5.44 (5.03 to 5.85); MD −0.18 (−0.76 to 0.40), p=0.55 • 6 months: 4.61 (4.18 to 5.04) vs 4.84 (4.38 to 5.30); MD −0.04 (−0.63 to 0.56) p=0.91 • 9 months: 4.66 (4.20 to 5.00) vs 4.73 (4.24 to 5.22); MD 0.19 (−0.41 to 0.80), p=53 • 12 months: 4.49 (4.00 to 4.97) vs 4.44 (3.94 to 4.94); MD 0.47(−0.14 to 1.07), p=0.13	**Pain:** The authors report no difference between groups in pain NRS scores or in the percentage of patients with a reduction of pain greater than 30% at any time point. **Function:** The authors report no difference between groups in function scores at any time point.	Conflict of interest: One author received grant funding from the Netherlands Organization for Scientific Research and Scientific Association Physiotherapy. One author received funding to his institution from professional organizations, travel expenses by the professional organizations when speaking at conferences, and honoraria for reviewing grant proposals from Swedish and Canadian governmental grant agencies. Funding: The study was funded by grant 171202013 from the

Author (Year)	Demographics	Results	Conclusions	Comments
	performed followed by a 1-2 mL injection of 2% lidocaine. RF was performed at 90°C for 90 seconds	<u>Pain Intensity Reduction >30%: RFN+Exercise vs. Exercise</u> NRS, %(n/N): • 3 weeks: 39% (40/102) vs 27% (27/100); RR 1.33 (0.80 to 1.97) p=0.25 • 1.5 months: 40% (45/112) vs 31.5% (36/114); RR 1.13(0.70 to 1.63) p=0.59 • 3 months: 45.6% (52/114) vs 36% (40/111); RR 1.16 (0.76 to 1.60), p=0.46 • 6 months: 55.5% (60/108) vs 50.4% (53/105) RR 1.02 (0.71 to 1.33) p=0.88 • 9 months: 51% (52/102) vs 49% (50/102) RR 1.09(0.75 to 1.42) p=0.60 • 12 months: 47% (47/100) vs 53.5% (53/99); RR 0.78 (0.50 to 1.09) p=.16 <u>Function: RFN+Exercise vs Exercise</u> ODI, mean (95%CI): • Baseline, mean (SD): 35.07 (14.66) 34.39 (12.24) • 3 months: 26.03(23.01 to 29.06) vs 28.67(26.06 to 31.84); MD −2.45 (−5.93 to 1.03), p=0.17 • 6 months: 25.38(22.45 to 28.30) vs 27.15(24.07 to 30.23); MD −0.60(−4.13 to 2.92), p=0.74 • 9 months: 25.74(22.74 to 28.73) vs 24.52(21.49 to 27.54); MD 2.26 (−1.29 to 5.82), p=0.21 • 12 months: 24.59(21.39 to 27.79) vs 25.04 (21.77 to 28.31); MD 1.48(−2.09 to 5.06), p=0.42		Netherlands Organization for Health Research and Development, by the Society for Anesthesiology, and the Dutch Health insurance companies.

Author (Year)	Demographics	Results	Conclusions	Comments
RFN vs spinal injections/epidural block, cervical				
Lim 2017	N=40 *Pulsed RF group vs ICI* Age, mean (SD): 52.8(12.1) vs 52.7(14.8) years Female: 65% vs 50% Pain Duration: 15.1(14.1) vs 11.1(10.8) months *Total* F/U: 1 week, 1 and 6 months <u>PRF procedure description (n=20)</u>: Treatment was administered at 5Hz with a 5-millisesond pulsed width for 360 seconds, at 55V. Electrode tip temperature did not exceed 42°C. *Device used:* Cosman G4 radiofrequency generator <u>ICI procedure description (n=20)</u>: 10mg (0.25mL) of dexamethasone mixed with 0.25mL of 0.125% bupivacaine injected.	<u>Pain: PRF vs ICI</u> NRS, mean (SD): • Baseline: 5.6 (1.3) vs. 5.8 (1.4), p=NS • 1.5 months*: 2.4 (1.6) vs 1.7 (0.9), p=NS • 3 months*: 3.0 (1.7) vs 2.4 (1.5), p=NS • 6 months: 3.2 (1.7) vs 2.7 (1.5), p=NS <u>Percent of patients with pain relief of ≥50%, PRF vs ICI</u> • 6 months: 50.0% (10/20) vs 60% (12/20), p=NS	**Pain:** The authors report no statistically significant differences between groups in pain VAS at any time point or in the percent of patients with pain relief ≥50%.	Authors report no conflict of interest Funding: 2016 Yeungnam University Research Grant (Level 2)

CI: confidence interval; F/U: follow-up; IA: intra-articular; ICI: intra-articular corticosteroid injection; IQR: interquartile range; MD: mean difference; NR: not reported; NRS: numerical rating scale; ODI: Oswestry Disability Index; PRF: pulsed radiofrequency; RF: radiofrequency; RFN: radiofrequency neurotomy; RR: risk ratio; SD: standard deviation; VAS: visual analog scale.

*P values calculated by AAI

†NRS means and SDs for both groups at 1.5 and 3 months, and for ICI at baseline and 6 months were estimated from graphs

‡All SD's estimated from graph

§Excellent rating defined as patient's pain disappearing, lumbar range of motion partly restored, and the patient returning to normal work and life

Appendix Table B3. Safety information from new RCTs

Author (Year)	Safety outcomes
RFN vs sham neurotomy, lumbar	
Moussa 2016	NR
Van Tilburg 2016	Withdrawals*, reason (n of patients): • increased pain after diagnostic test (1) • painful procedure despite local anesthetic (3) No serious adverse events were encountered during the trial
RFN vs spinal injections/epidural block, lumbar	
Do 2017	No adverse events in PRF group, 1 event of hyperglycemia in ICI
Hashemi 2014	NR
Zhou 2016	No adverse events reported
RFN vs P RF neurotomy + RFN, lumbar	
Arsanious 2016	NR
RFN+exercise vs exercise, lumbar	
Juch 2017	None reported
RFN vs spinal injections/epidural block, cervical	
Lim 2017	No adverse events in PRF group, 2 adverse events in ICI group (1 report of facial flushing, 1 report of hyperglycemia)

ICI: intra-articular corticosteroid injection; NR: not reported; PRF: pulsed radiofrequency; RFN: radiofrequency neurotomy
*Group that withdrawals were in was not reported

APPENDIX C. ARTICLES EXCLUDED AT FULL TEST REVIEW

Appendix Table C1. Excluded systematic reviews

Citation	Reason for exclusion
Al-Najjim M, Shah R, Rahuma M, Gabbar OA. Lumbar facet joint injection in treating low back pain: Radiofrequency denervation versus SHAM procedure. Systematic review. Journal of orthopaedics 2018;15:1-8.	No new RCTs included
Boswell MV, Manchikanti L, Kaye AD, et al. A Best-Evidence Systematic Appraisal of the Diagnostic Accuracy and Utility of Facet (Zygapophysial) Joint Injections in Chronic Spinal Pain. Pain physician 2015;18:E497-533.	No new RCTs included
Engel A, Rappard G, King W, Kennedy DJ. The Effectiveness and Risks of Fluoroscopically-Guided Cervical Medial Branch Thermal Radiofrequency Neurotomy: A Systematic Review with Comprehensive Analysis of the Published Data. Pain medicine (Malden, Mass) 2016;17:658-69.	No new RCTs included
Facchini G, Spinnato P, Guglielmi G, Albisinni U, Bazzocchi A. A comprehensive review of pulsed radiofrequency in the treatment of pain associated with different spinal conditions. The British journal of radiology 2017;90:20150406.	No new RCTs included
Leggett LE, Soril LJ, Lorenzetti DL, et al. Radiofrequency ablation for chronic low back pain: a systematic review of randomized controlled trials. Pain research & management 2014;19:e146-53.	No new RCTs included
Maas ET, Ostelo RW, Niemisto L, et al. Radiofrequency denervation for chronic low back pain. The Cochrane database of systematic reviews 2015:Cd008572.	No new RCTs included
Manchikanti L, Hirsch JA, Kaye AD, Boswell MV. Cervical zygapophysial (facet) joint pain: effectiveness of interventional management strategies. Postgraduate medicine 2016;128:54-68.	No new RCTs included
Manchikanti L, Hirsch JA, Falco FJ, Boswell MV. Management of lumbar zygapophysial (facet) joint pain. World journal of orthopedics 2016;7:315-37.	No new RCTs included
Poetscher AW, Gentil AF, Lenza M, Ferretti M. Radiofrequency denervation for facet joint low back pain: a systematic review. Spine 2014;39:E842-9.	No new RCTs included

RCTs: randomized controlled trials.

Appendix Table C2. Excluded observational studies

Citation	Reason for exclusion
Cohen SP, Moon JY, Brummett CM, White RL, Larkin TM. Medial Branch Blocks or Intra-Articular Injections as a Prognostic Tool Before Lumbar Facet Radiofrequency Denervation: A Multicenter, Case-Control Study. Regional anesthesia and pain medicine 2015;40:376-83.	Case-control design – previous report had RCT data to answer KQ1

KQ1: Key Question 1; RCT: randomized controlled trial.

APPENDIX D. ONGOING COMPARATIVE CLINICAL STUDIES ASSESSING RADIOFREQUENCY FACET NEUROTOMY

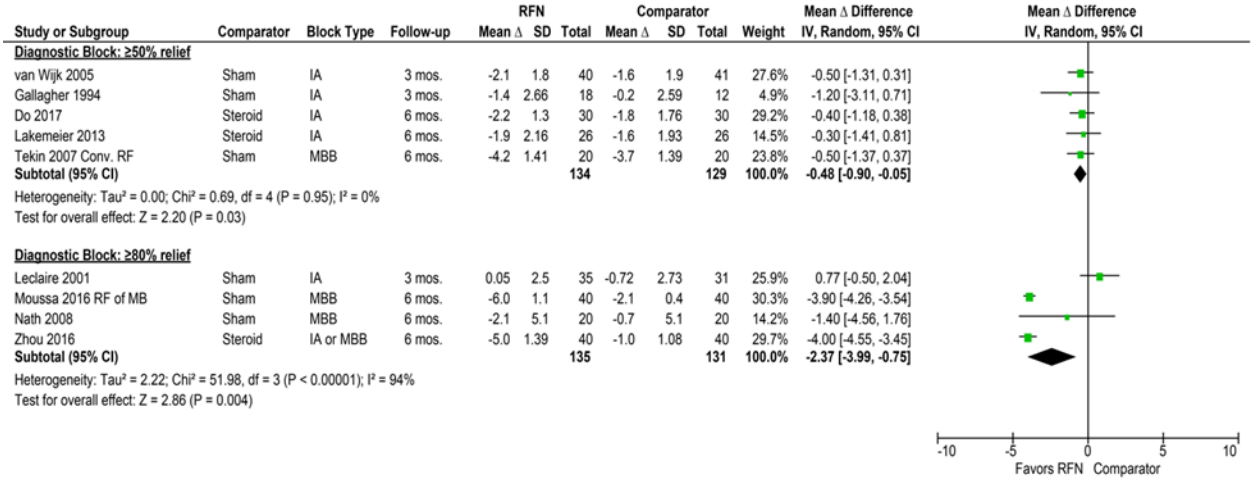
Appendix Table D1. Ongoing clinical trials evaluating facet neurotomy indexed in CLINICALTRIALS.GOV*

NCT number	Title	Status	Conditions	Study type (N)	Interventions	Comparator	Sponsor	State date	Estimated completion date
NCT01300715	An alternative technique for lumbar medial branch radiofrequency: Comparison with the empirical technique	Unknown	Low back pain, lumbar facet joint pain, arthropathy	RCT (N = 100)	Modified lumbar MBRF	Tunnel vision lumbar MBRF	Seoul National University Bundang Hospital	November 2010	May 2011
NCT01743326	RFD versus cervical medial branch blocks in chronic degenerative neck pain	Unknown	Facet joint arthritis	RCT (N = 84)	Radiofrequency denervation	Local anesthesia	Maastricht University Medical Center	November 2012	June 2015
NCT03066960	Radiofrequency neurotomy for chronic facet joint related neck pain	Not yet recruiting	Neck pain	RCT (N = 44)	Radiofrequency neurotomy	Sham radiofrequency neurotomy	Oslo University Hospital	August 2017	April 2019
NCT03039296	EuroPainClinics® Study IV	Recruiting	Low back pain, facet joint pain	Cohort (N = 150)	Unilateral endoscopic rhizotomy	Bilateral endoscopic rhizotomy	Europainclinics z.u.	February 2017	December 2021
NCT02478437	A trial of cooled radiofrequency ablation of medial branch nerves for the treatment of lumbar facet syndrome	Recruiting	Low back pain	RCT (N = 40)	Cooled radiofrequency ablation	Conventional radiofrequency ablation	Northwestern University	June 2015	September 2017
NCT02148003	Effect of the temperature used in thermal radiofrequency ablation	Recruiting	Low back pain	RCT (N = 237)	Radiofrequency ablation at 90°C	Radiofrequency ablation at 80°C	The Cleveland Clinic	May 2014	February 2018

*accessed March 8, 2018.

APPENDIX E. PRELIMINARY META-ANALYSES CONDUCTED BY AAI

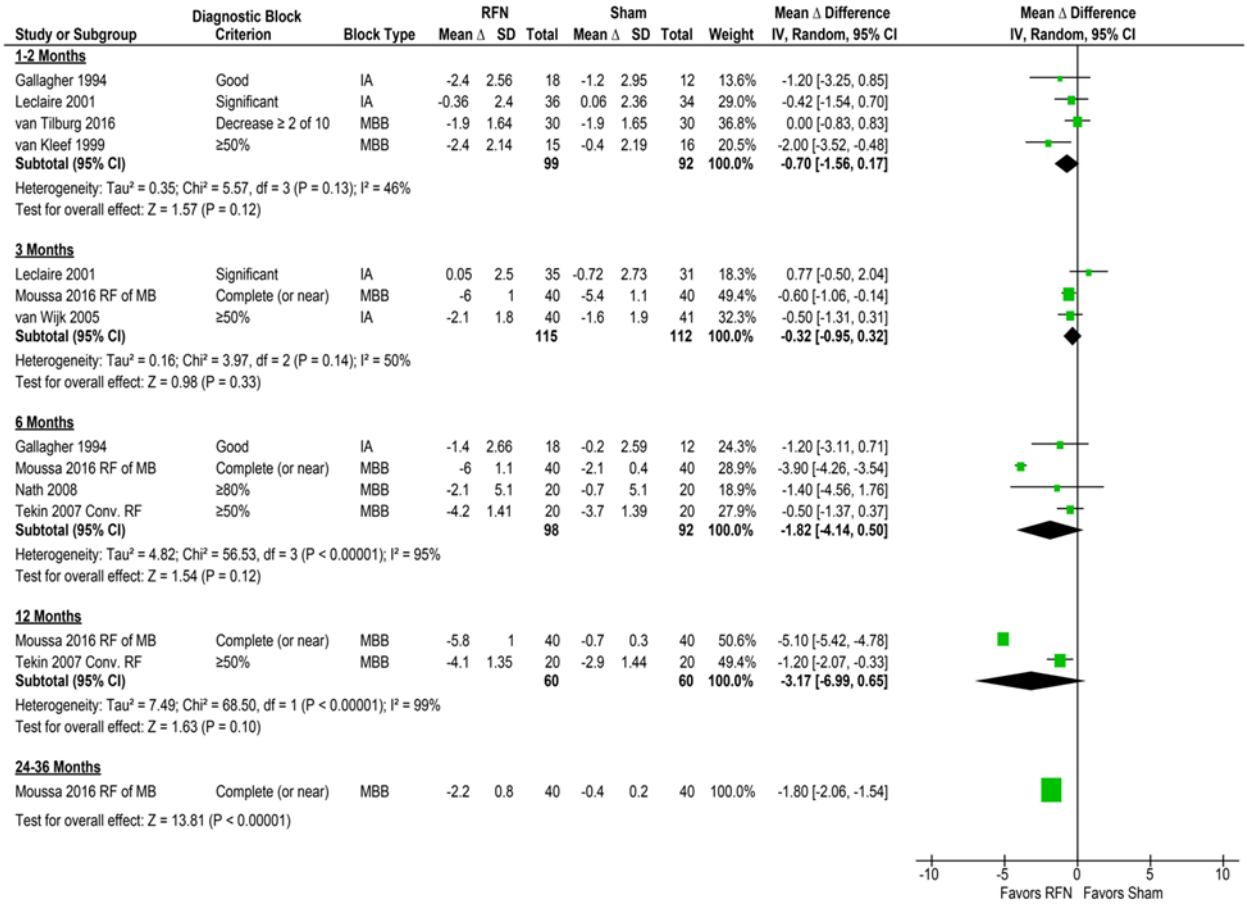
Appendix Figure E1. Improvement in VAS pain following RF neurotomy stratified by degree of pain relief achieved from diagnostic block*



CI: confidence interval; IA: intraarticular; MBB: medial branch block; RFN: radiofrequency neurotomy; SD: standard deviation; VAS: visual analog scale.

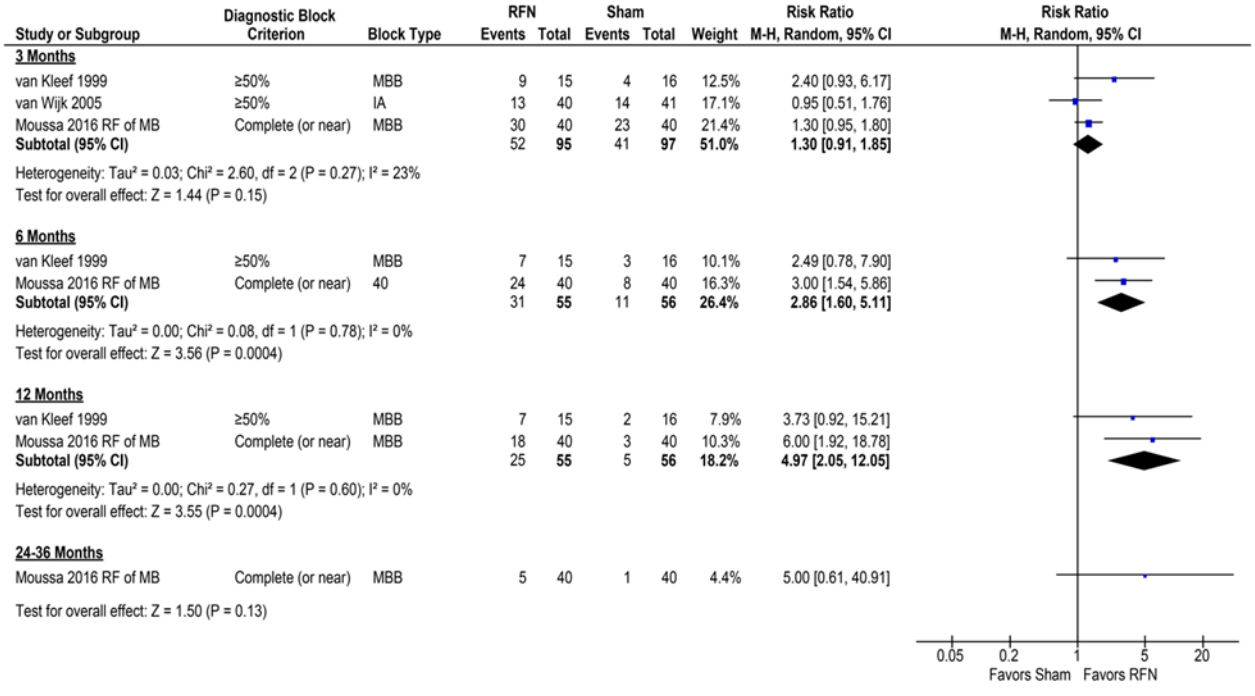
*Three trials did not describe the required response to diagnostic block using a % cut-off. Gallagher 1994 stated that a "good" response was required for inclusion; we decided it was most appropriate to group this trial with the "≥50% relief" trials. Leclaire 2001 require a "significant" response and Moussa 2016 required "a complete or near complete" response; we grouped these with the "≥80% relief" trials.

Appendix Figure E2. RF Neurotomy versus Sham: Back pain improvement (change in VAS scores)*



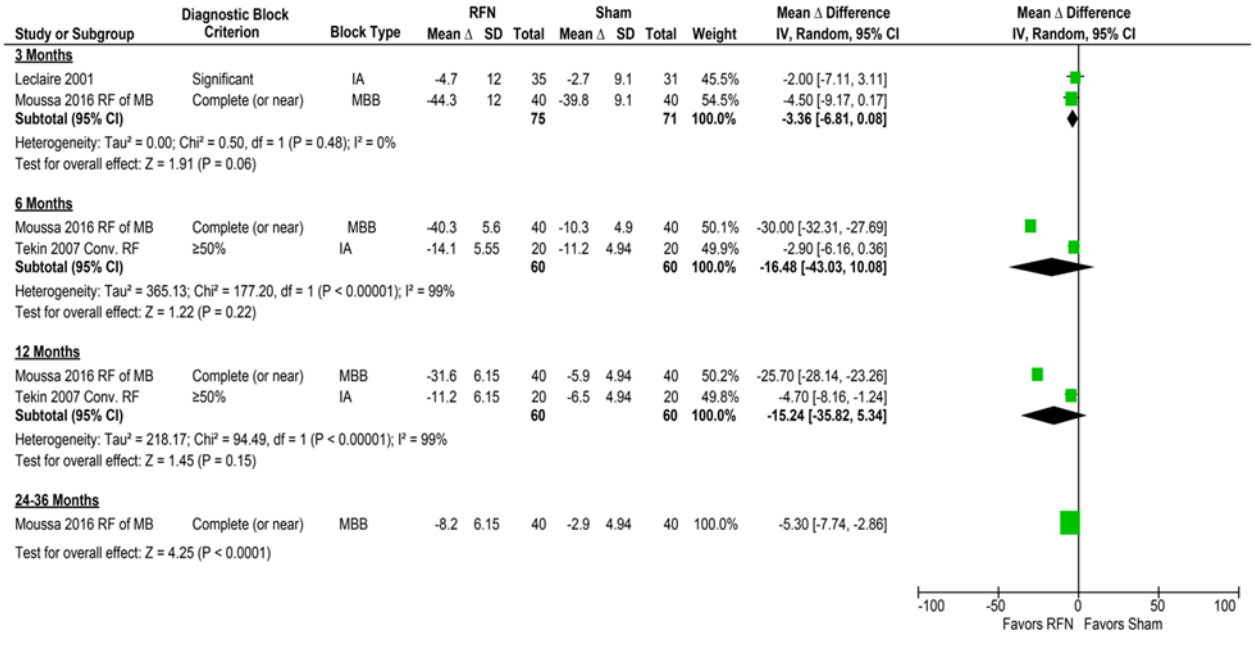
CI: confidence interval; IA: intraarticular; MBB: medial branch block; RFN: radiofrequency neurotomy; SD: standard deviation; VAS: visual analog scale.
*Van Tilburg 2016 and Moussa 2016 are new trials.

Appendix Figure E3. RF Neurotomy versus Sham: Pain relief "success"*



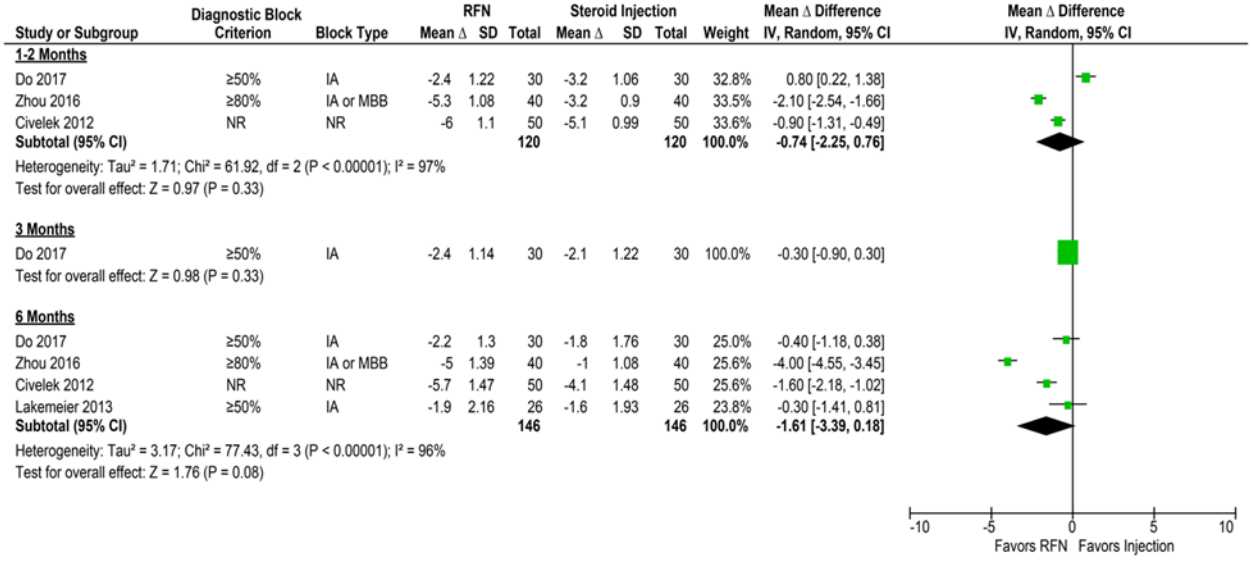
CI: confidence interval; IA: intraarticular; MBB: medial branch block; RFN: radiofrequency neurotomy; SD: standard deviation.
*Definitions of pain success included: 1) Visual analog scale (VAS) pain reduction of ≥50% (van Wijk 2005, Moussa 2016 [new trial]), and 2) Both 2-point reduction on VAS and ≥50% pain reduction on global perceived effect (van Kleef 1999).

Appendix Figure E4. RF Neurotomy versus Sham: Function improvement (change in ODI scores)*



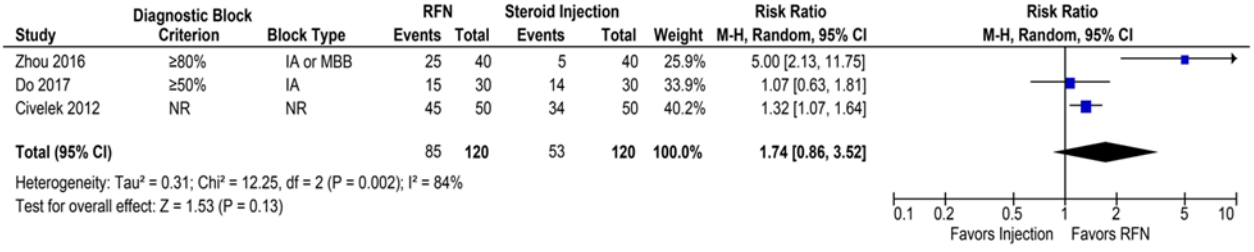
CI: confidence interval; IA: intraarticular; MBB: medial branch block; ODI: Oswestry Disability Index; RFN: radiofrequency neurotomy; SD: standard deviation.
*Moussa 2016 is a new trial.

Appendix Figure E5. RF Neurotomy versus Steroid Injection: Back pain improvement (change in VAS scores)*



CI: confidence interval; IA: intraarticular; MBB: medial branch block; NR: not reported; RFN: radiofrequency neurotomy; SD: standard deviation; VAS: visual analog scale.
*Do 2017 and Zhou 2016 are new trials.

Appendix Figure E6. RF Neurotomy versus Steroid Injection: Pain relief "success"*



CI: confidence interval; IA: intraarticular; MBB: medial branch block; NR: not reported; RFN: radiofrequency neurotomy; SD: standard deviation.

*Definitions of pain success included: 1) Visual analog scale (VAS) pain reduction of ≥50% (Civelek 2012, Do 2017 [new trial]), and 2) Complete relief of pain, lumbar range of motion restored, and patient returned to normal work life (Zhou 2016 [new trial])

Osteochondral Allograft/Autograft Transplantation (OAT): Assessing Signals for Update

Provided by:

Aggregate Analytics, Inc.

Prepared by:

Andrea C. Skelly, PhD, MPH
Erika D. Brodt, BS
Cassie Winter, BS

January 31, 2018

Table of Contents

1. Previous Coverage Decision

A Health Technology Assessment titled: *Osteochondral Allograft/Autograft Transplantation (OAT)*, was published on November 18, 2011 by the Health Care Authority. Findings and Coverage Decision was adopted on March 16, 2012. The Committee's Coverage Decision is summarized below.

Health Technology Background

Osteochondral Allograft/Autograft Transplantation (OAT) was selected in December 2010 to undergo an evidence review process. The evidence based technology report indicates that OAT referred to the use of cylindrical, dowel-shaped or geometric-shaped plugs of osteochondral material that are press fit into a defect and do not require the use of screws, pins, plates, or other fixation devices. Mosaicplasty, which involves multiple cylindrical plugs, was also included in the report. Osteochondral autograft (or allograft) transplantation or mosaicplasty involve transplantation of cartilage and subchondral bone into the defect to facilitate the growth or new tissue. These procedures can be done open or arthroscopically and are sometimes combined with other joint operations such as arthroscopic debridement or anterior cruciate ligament (ACL) repair.

Osteochondral autograft transplantation involves harvesting bone and intact articular cartilage from a non-weight bearing portion of a joint from the patient to fill a defect in the weight-bearing portion of the joint. This is a technically demanding procedure and is limited to treating defects < 4 cm^2 because of donor tissue limitations. Osteochondral allograft transplantation involves the transplantation of a piece of cartilage and subchondral bone from a source outside of the patient to fill the osteochondral defect. Osteochondral allografts are regulated by the FDA as Human Cell or Tissue Products (HCT/P), as defined in section 361 of the Public Health and Service Act.

HTCC Coverage Determination

Osteochondral Allograft/Autograft Transplantation (OAT) is a **covered benefit with conditions.**

Osteochondral Allograft/Autograft Transplantation (OAT) for joints other than the knee is a **not covered benefit with conditions.**

HTCC Reimbursement Determination

Limitations of Coverage

Osteochondral Allograft/Autograft Transplantation for the knee is a covered benefit when the following conditions are met:

- Age <50, older at the discretion of the agency;
- Excluding malignancy, degenerative and inflammatory arthritis in the joint; and
- Single focal full-thickness articular cartilage defect

Non-Covered Indicators

Osteochondral Allograft/Autograft Transplantation for joints other than the knee are not covered.

Committee Decision

Based on the deliberations of key health outcomes, the committee decided that it had the most complete information: a comprehensive and current evidence report, public comments, and agency and state utilization information. The committee concluded that the current evidence on Osteochondral

Allograft/Autograft Transplantation (OAT) for the knee demonstrates that there is sufficient evidence to cover with conditions. The committee concluded that the current evidence on Osteochondral Allograft/Autograft Transplantation (OAT) for joints other than the knee demonstrates that there is insufficient evidence to cover. The committee considered all the evidence and gave greatest weight to the evidence it determined, based on objective factors, to be the most valid and reliable. Based on these findings, the committee voted to not cover Osteochondral Allograft/Autograft Transplantation (OAT) for joints other than the knee. Based on these findings, the committee voted to cover with conditions Osteochondral Allograft/Autograft Transplantation (OAT) for the knee.

Medicare Decision and Expert Treatment Guidelines
The committee reviewed the clinical guidelines and Medicare decision. The Centers for Medicare and Medicaid Services have no published national coverage determinations (NCD) for Osteochondral Allograft/Allograft Transplantation (OAT).

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 (see Figure 1). The key questions in the included original report are listed below. **For the signal update, updated searches were performed only for Key Questions 3, 4, 5 and 6.**

Key question 1

1. What is the case definition of a patient suitable for OATS/mosaicplasty surgery, and are there measures of reliability and validity for case identification?
    a. What are the maximum, minimum, and optimum size (volume) of the damage that is suitable for repair using OATS/mosaicplasty?
    b. What are the maximum and optimum number of lesions that can be repaired in a single OATS/mosaicplasty procedure?
    c. Are there other considerations that make OATS/mosaicplasty suitable or unsuitable (age, mobility, comorbidities, BMI)?
    d. Is there a distinction between OATS and mosaicplasty, and a related case definition difference between the two?

Key Question 2

2. What are the expected treatment outcomes of OATS/mosaicplasty, and are there validated instruments and scores to measure clinically meaningful improvement?

Key Question 3

3. What is the evidence of efficacy and effectiveness of OATS/mosaicplasty (open or arthroscopic)? Including consideration of short term and long term:
    a. Delay or avoidance of progression to osteoarthritis
    b. Impact on function, pain, range of motion, quality of life, activities of daily living and return to work
    c. Longevity of treatment effect
    d. Need for continuing and/or subsequent intervention
    e. Need for extended or continuing physical therapy
    f. Recovery time considering harvest site recovery issues
    g. Differential results from multiple versus single grafts, patterning for multiple
    h. grafts (linear arrangement versus circular arrangement)

i. Differential results between allograft and autograft procedures
j. Differential results between open and arthroscopic procedures
k. Differential results in centers of excellence

Key Question 4

4. What is the evidence of the safety of OATS surgery? Including consideration of:
   a. Adverse events type and frequency (peri-operative, cartilage plug detachment, cartilage rejection, graft fit, harvest site issues, development of fibrocartilage, mortality, other major morbidity such as DVT, deep infection, and excessive intraarticular bleeding)
   b. Revision/re-operation rates (if not addressed in efficacy)

Key Question 5

5. What is the evidence that OATS surgery has differential efficacy or safety issues in sub-populations? Including consideration of:
   a. Gender
   b. Age
   c. Psychological or psychosocial co-morbidities
   d. Baseline functional status: e.g. type of injury or lesion, extent of cartilage damage, specific damage site size, number of damage sites
   e. Other patient characteristics or evidence based patient selection criteria,
   f. especially comorbidities of diabetes and high BMI
   g. Provider type, setting or other provider characteristics
   h. Payer/ beneficiary

Key Question 6

6. What is the evidence of cost implications and cost-effectiveness for OATS/mosaicplasty? Including consideration of:
   a. Costs (direct and indirect) and cost effectiveness
   b. Short term and long term

3. Methods

3.1 Literature Searches

We conducted an electronic literature search for the period March 1, 2011 through January 10th, 2018 using identical search terms used for the original report for key questions 3 through 6. This search included 3 main databases: PubMed, Cochrane Library, and EMBASE. Additional electronic databases were searched; see Appendix A for search methodology and additional details. Osteochondral allografts are regulated by the FDA as Human Cell or Tissue Products. In addition, we searched the FDA website for updated information on such products.

3.2 Study selection

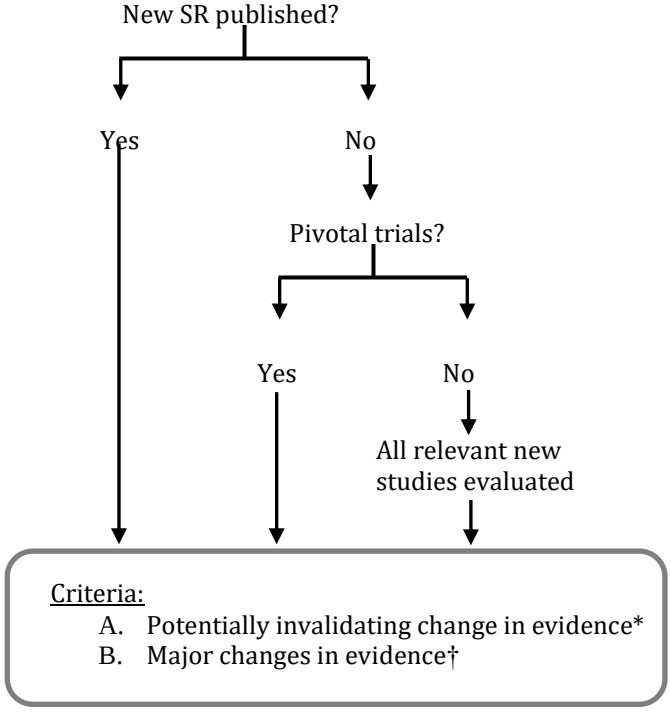
We sought systematic reviews (SR) of randomized controlled trials (RCTs) of efficacy and safety with meta-analysis that included articles that met inclusion and exclusion criteria similar to the original report. In addition we sought systematic reviews reflecting updates or new advances for the technology. Secondary to the large number of citations returned, we focused on screening systematic reviews and meta-analyses of RCTS published between 2011 and 2018. Although quality of systematic reviews was not formally evaluated for this report, we chose systematic reviews of head to head trials

for efficacy that were the most comprehensive and of higher quality based on the following: report of search strategies (two or more databases 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. Only systematic reviews of RCTs were included for efficacy. Systematic reviews focused on longer-term safety outcomes may include nonrandomized studies. A summary of the included SRs and RCTs is found in Appendix B.

3.3 Compilation of Findings and Conclusions

For this assessment we constructed a summary table that included the key questions, the original conclusions, new sources of evidence, new findings, and conclusions based on available signals. To assess whether the conclusions might need updating, we used an algorithm based on a modification of the Ottawa method, Figure 1.

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

4. Results

4.1 Search

The literature search identified 1,755 titles. After title and abstract review, 1,724 articles were excluded and 31 articles remained that addressed in part or in full key questions 3, 4, 5, and/or 6. A total of 16 articles were retained for the signal update, Figure 2. A full list of excluded studies and the reasons for exclusions can be found in Appendix C.

We identified 20 systematic reviews that addressed in part or in full key questions 3, 4, and/or 5. Systematic reviews were excluded if they did not include study types of interest and/or if they were not the most comprehensive and of the highest quality, Appendix B. Two systematic reviews related to efficacy and four systematic reviews focused on safety were retained, of which one systematic review was included for both efficacy and safety. No full health technology assessments were identified; however a 2017 Canadian Agency for Drugs and Technologies in Health Rapid Review is summarized in Appendix B for informational purposes only. One systematic review described results for differential safety (key question 5). We found two cost-effectiveness studies (Key Question 6); there were none in the previous report. Six new RCTs were identified; none were considered pivotal. Two follow-up publications of RCTs included in the previous report were also identified and included.

The FDA still regulates osteochondral allografts as Human Cell or Tissue Products (HCT/P) as defined in section 361 of the Public Health and Service Act. No updates on FDA approval have been published since our initial report.

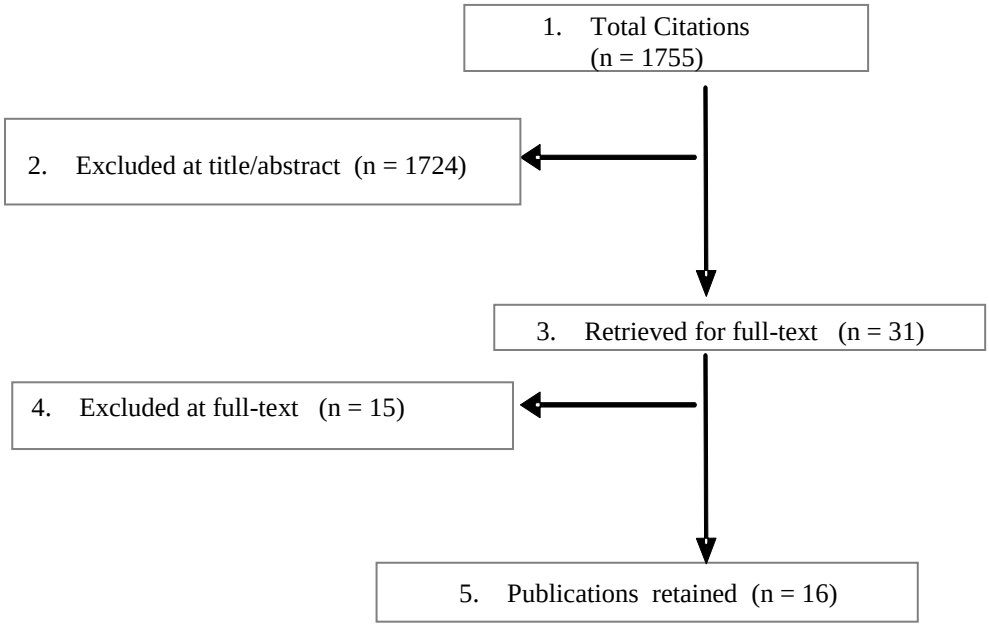


Figure 2. Flow chart showing results of literature search

4.2 Identifying signals for re-review

Tables 1- 4 show the original key questions, the conclusions of the original report, the new sources of evidence, the new findings, and the recommendations of Aggregate Analytics, Inc. (AAI) regarding the need for update (Figure 1). **For the signal update, updated searches were performed only for Key Questions 3, 4, 5 and 6.**

Table 1. Osteochondral Allograft/Autograft Transplantation Summary Table for Key Question 1. [NO UPDATED SEARCH FOR SIGNAL UPDATE]

Conclusions from CER Executive Summary	New Sources of Evidence	New Findings	Conclusion from AAI
Key Question 1. What is the case definition of a patient suitable for OATS/mosaicplasty surgery, and are there measures of reliability and validity for case identification?			
Consistent or agreed-upon case definitions: There is variability with respect to the terms used to describe the various procedures and how they are defined. No specific agreed-upon case definitions were found. Treatment algorithms (only available for the knee) cite case series. Lesion size and classification appear to be key criteria for assessing treatment options (after ligament and meniscus stability, lesion location and other factors have been determined). **Autograft (OAT or mosaicplasty)**: Based on inclusion/exclusion criteria for randomized studies for knee lesions, the most consistent characteristics defining cases for inclusion were: symptomatic (5/5 studies), isolated (4/5 studies) full-thickness lesions or Outerbridge or ICRS grades 3 or 4 lesions (4/5 studies). Exclusion criteria in three of the five studies included knee joint instability or ligamentous deficiency. The mean ages of participants in all studies was <45 years old. **Osteochondral allograft (dowel, cylinder, plug)**: No prospective comparative studies were found and limited information is available from three case series. Cases were defined as symptomatic in all three studies. Studies designed to evaluate clinical decision-making based on patient or lesion characteristics were not found **Talus**: Only one comparative study was available. Pain and presence of a full thickness lesion as inclusion criteria are consistent with criteria described above for the knee. No studies pertaining to other anatomical regions meeting the inclusion criteria were found.	NOT SOUGHT	N/A	N/A
<u>Evidence of validity and reliability (lesion classification systems):</u> • No validity studies of the Outerbridge or ICRS lesion grading systems in the population of interest were found. • Overestimation of lesion size by arthroscopy compared with open evaluation was reported in one clinical study. Inexperienced clinicians had less accurate measures. • Two clinical studies evaluated the reliability of the ICRS grading system using arthroscopy. One study reported 80.9% agreement between arthroscopic and open assessment of grade. Only one study (the smallest) reported chance-adjusted agreement between raters and suggests that there is only fair to slight agreement between raters. • Inter-rater reliability of the Outerbridge classification was evaluated in one study. The overall agreement beyond chance for the video tapes where surgeons were to discriminate between grades 2 and 3 was moderate (κ range 0.41-0.57). The authors did not apparently	NOT SOUGHT	N/A	N/A

Conclusions from CER Executive Summary	New Sources of Evidence	New Findings	Conclusion from AAI
evaluate grade 4 lesions to any large extent and thus, application to a case definition which may focus on grades 3 and 4 lesions is not clear. • No studies for anatomical regions other than the knee were found.			

Table 2. Osteochondral Allograft/Autograft Transplantation Summary Table for Key Question 2. [NO UPDATED SEARCH FOR SIGNAL UPDATE]

Conclusions from CER Executive Summary	New Sources of Evidence	New Findings	Conclusion from AAI
Key Question 2: What are the expected treatment outcomes of OATS/mosaicplasty, and are there validated instruments and scores to measure clinically meaningful improvement?			
• Review of the properties of outcomes measures used in included comparative studies is limited to those measures that were examined in samples drawn from the target population (patients with articular cartilage damage). Of these measures, five have been validated in this population: • International Cartilage Repair Society (ICRS) cartilage repair assessment • Lysholm Knee Scoring Scale (LKSS) • Modified Cincinnati Knee Rating System (MCRS) • International Knee Documentation Committee subjective knee form (IKDC SKF) • Knee Injury Osteoarthritis Outcome Score (KOOS) Four patient-reported and one clinician-based outcomes measures commonly used in studies of patients with cartilage defects in the knee have undergone psychometric analysis in these patients: • None of the five instruments were adequately tested for validity. Content validity was inadequate for all instruments, primarily because patients with chondral lesions were not involved in item selection in that particular study. Criterion validity was not tested in these studies for any instruments, likely because of the lack of a gold standard criterion. Tests of construct validity were hampered by definitional problems and small sample sizes. • Reliability was inadequately tested for the three outcome measures that were tested for internal consistency. None of the studies performed factor analysis to assess potential dimensions. While good internal consistency was shown for the KOOS and the ICRS, internal consistency for these instruments was inadequate as too few patients/raters were tested. Similarly, high values	NOT SOUGHT	N/A	N/A

Conclusions from CER Executive Summary	New Sources of Evidence	New Findings	Conclusion from AAI
for reproducibility were found for the IKDC, the LKSS, and the MCKRS in samples that were too small to meet quality criteria. • Studies that assessed responsiveness showed strong effect sizes for change from preoperative to post-operative scores on the IKDC, MCKS, LKSS, and KOOS. However, quality criteria also require that these effect sizes be supported by comparison of the minimally important clinical difference with the smallest detectable difference, analysis of receiver operating curves, or other supporting analysis. Only one study, which analyzed the IKDC and MCKS, met this criterion. • The minimal clinically important difference (MCID) for pre-op to post-op improvement was determined in one study to be from 6.3 points (6 months follow-up) to 16.7 points (12 month follow-up) on the IKDC and 14.0 points (6 months) and 26.0 points (12 months) on the MCKRS. The MCID was not reported for any other measures in patients with cartilage damage.			

Table 3. Osteochondral Allograft/Autograft Transplantation Summary Table for Key Questions 3 and 4. [UPDATED SEARCH RESULTS]

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
Key Question 3: What is the evidence of efficacy and effectiveness of OATS/mosaicplasty (open or arthroscopic)?			
Autograft versus microfracture, drilling or debridement alone			
Knee			
Efficacy: Knee (Low Evidence) • Two poor quality RCTs (N=104 total), one in young athletes, the other in children. • **Function:** OAT was associated with statistically better patient-reported and clinician-reported outcomes. • **Longevity of treatment effect**: Differences between treatments remained significant up to the last follow-up (maximum 48 months). Functional scores in young athletes improved for OAT recipients up to 36 months. In children following initial improvement at 12 months, ICRS scores decreased slightly, but remained stable up to 48 months. • **Return to activity**: A greater proportion of patients treated by OAT versus MF had returned to pre-injury activity levels at pre-specified time points.	**Systematic reviews** Graticelli 2016[1] Cochrane review (3 RCTs total, includes new RCTs Ulstein and Lim) Pareek 2016[2] (6 RCTs total, includes 3 new RCTs Gudas 2013, Ultstein and Lim and Gudas 2012 f/u) **RCTs** **Follow-up publications:** Gudas 2012[3] (*follow-up to Gudas 2005)* **New RCTS:** Solheim 2017[4] (Mosaicplasty) Ulstein 2014[5] Lim 2012[6] Gudas 2013[7] (OAT vs. MF and vs. debridement only)	**Function** *Systematic reviews*: Pareek SR: Subjective patient outcomes (International Knee Documentation Committee score [IKDC], Lysholm knee scoring scale) at 3 years favored OAT (3 trials by Gudas, one is new, SMD 0.40, 95%CI 1.04, 0.70, p = 0.008); No SOE provided. *New RCTs:* Solheim 2017 (N=40): Clinically and statistically significant difference in Lysholm score favoring OAT at 1 year and all subsequent time points. **Longevity** Gracitelli SR: Pooled mean difference from two small new trials (Lim, Ulstein, total N 72) for the Lysholm score at ≥5 years: showed no difference between OAT mosiacplasty and microfracture (pooled MD – 1.01, 95%CI -4.54, 2.33, p = 0.53) whereas 1 older trial included in prior report favored OAT on the IKDC score (MD13.97, 95%CI 13.25, 14.69. SOE was reported as very low (insufficient) Pareek SR: No difference in subjective scores (IKDC, Lysholm) at 5-10 years (3 trials pooled SMD 0.92, 95%CI -1.07, 2.9), but substantial heterogeneity is noted, only Gudas 10 year follow –up was significant. No SOE provided. Gudas 2012, 10 year follow-up to Gudas 2005 in young athletes: Authors report function continued to	Pooled data including new trials suggest no difference between OAT autograft and MF at ≥5 years for function or Tegner score. Data are from small non-pivotal trials and the evidence base is likely low or insufficient. This section of the report is still valid and does not need updating.(Criteria A1, B-1-4)

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
		be significantly better(ICRS and Tegner scores) with OAT vs. MF; *New RCTs:* Inconsistent findings at 5 years in Lysholm Score: Significant difference favoring OAT reported in Solheim (N= 40, MD 10, 955CI 0.57 to 19.4); Lim 2012 difference was not significant but point estimate tended to favor MF (N, 47, MD -2.8, 95%CI -6.64, 0.94). At 10 years there were no differences between treatments in Solheim 2017 or Ulstein 2014 (N=25) , but point estimates were in opposite directions; sample sizes are small. **Return to Activity** *Systematic reviews*: Gracitelli SR: Mean Tegner Activity Score ≥5 years was not significant for either new trial (Lim, Ulstein). Continuation of sport in older trial (Gudas 2005) and 3 years was more common with OAT vs. MF (RR 3.24 , 95%CI 1.77, 5.92) but not statistically different at 10 years (RR 2.07, 95%CI 0.81, 5.30) Authors SOE: very low (insufficient) Pareek SR: Tegner Activity score (3 – 10 years); OAT associated with better scores (3 trials, MD 0.47, 95%CI 0.14 to 0.80); Trials summarized were Gudas 2012, Gudas 2013 and Lim 2012; individually, only Gudas 2013 reached statistical significance; no SOE provided *RCT follow-up*: Gudas 2012 continuation of sport at same level at 10 years (N = 41): Mean duration of previous sport activity was statistically longer in the OAT vs. MF group.	

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
		New RCTs: In Gudas 2013, OAT was favored over debridement alone; authors do not report whether there was a statistical difference between groups with regard proportions who returned to activity.	
Effectiveness: Knee (No Evidence) • No nonrandomized comparative studies were found.	NOT SOUGHT	N/A	N/A
ANKLE			
Efficacy: Ankle (No Evidence) • No randomized controlled trials were found so efficacy could not be evaluated.	**New RCT** Sun 2016[8] (N=153)	**Function and pain at 2.3 years** Authors report no difference in changes scores of AOFAS, TAS, or Mazur ankle scoring system values between OAT and MF or in VAS pain; both OAT and MF resulted in improved AOFAS, TAS, Mazur ankle scoring and VAS pain compared with drilling.	There are new data that would update the report; however the findings from one non-pivotal trial are not sufficient to trigger an updated report. (Criterion A1)
Effectiveness Ankle (Very Low Evidence) **Function:** One small poor quality cohort (N= 32) reported differences in functional outcomes (assessed by AOFAS or SANE Scores) between OAT and chondroplasty or OAT and microfracture; however, 24-hour post-operative pain was greater among patients treated by OAT.	Not Sought	N/A	N/A
Autograft versus autologous chondrocyte implantation (ACI)			
Efficacy (Low Evidence) • Two poor quality RCTs in general (older) populations were found. One enrolled >40% of participants who had prior surgeries (N =140 total). In the other RCT, ≥50% of persons did not receive treatment (n treated = 23/44 randomized), as authors reported "spontaneous improvement" in the six months following initial debridement. • **Function:** Patient-reported outcomes were better for OAT/mosaicplasty but statistical	**Systematic Reviews** None **RCTs** **Follow-up publications** Bentley 2012[9] (*follow-up to Bentley 2003*) **New RCTS:** Lim 2012[6]	**Knee** **Function** *New RCT:* Lim 2012 (N = 40 knees): The authors reported no differences in Lysholm, Tegner, or HSS scores at a follow-up up to a mean of 5.7 years. 10 year follow-up of previously included trial, Bentley 2012: ACI continued to demonstrate a statistically significant better results than OAT in the modified Cincinnati score, however there appeared to be differential of data for this measure at 10 years that may	This section of the report is still valid and does not need updating. (Criteria A1, B-1, B-4)

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
significance was not uniformly achieved in the two small RCTS. In the largest RCT (n = 100) a significantly smaller proportion of participants receiving mosaicplasty had excellent or good outcomes (author's modification of the Cincinnati Rating Scale) and one of the smaller RCTs reported no significant differences in the Meyer score. Both these studies included substantial proportions of participants who had prior surgeries. Differences in outcomes measures used makes comparison across studies difficult. • **Longevity of treatment effect**: In one study (N =40), functional scores for both OAT and ACI increased over time for the Lysholm, Tegner and Myers scores; only for the Lysholm Knee Scoring Scale were significant differences between treatment sustained over time favoring OAT.		bias findings; 15 of 42 patients in the OAT group were evaluated for functional outcomes at the 10 year follow-up, versus 48 of 58 patients in the ACI group.	
Effectiveness (No Evidence) • No nonrandomized comparative studies were found.	Effectiveness, not sought	N/A	N/A
Autograft versus other treatments			
Efficacy: Ankle (No Evidence) • No randomized controlled trials were found so efficacy cannot be evaluated.	**New RCT** **Autograft vs. allograft** Ahmad 2016[10] (N=40)	There were no differences between autograft and allograft with regard to function or pain at a mean of 3.2 years. Similarly there were no differences in graft union or need for operative revision procedures.	There are new data that would update the report, however the findings from one small non-pivotal trial are not sufficient to trigger an updated report. (Criteria A1, B2)
Effectiveness Knee (Very Low Evidence) • Four small, poor quality nonrandomized studies compared OAT alone or in combination with other procedures. Confounding by indication was present in all and heterogeneity across studies precludes effective comparison across them.	Effectiveness, not sought	N/A	N/A

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
• **Function:** For most functional outcomes, there were no differences between treatment groups. • In one small (N =18) study, post-operative mean Modified Lysholm score was significantly less for OAT versus matrix assisted chondrocyte transplantation (MACT). • Range of motion appeared to be substantially greater among patients treated by OAT with realignment versus realignment alone in another study (n =49)			
Allograft : Osteochondral allograft using primarily press-fit dowel/cylinder or plug (not requiring hardware)			
Efficacy: No Evidence • No randomized controlled trials were found. **Effectiveness: Knee and Ankle (Very Low Evidence)** • **Comparative studies:** No statistically significant differences between treatment groups were reported for most outcomes measures across two small studies (N = 70 total). Tegner scores were improved for OA recipients compared with loose body removal and arthroscopic reduction and internal fixation in one study, and SF-12 Mental Component Scores were significantly improved in patients who received OA and MAT (meniscal allograft transplantation) compared with OA and ACI in the other. • Case series of >19 patients which primarily used press-fit plugs (dowel/cylinder/geometric) without use of fixation • **Function and QoL:** Various patient-reported, clinician based outcomes and quality of life measures were used across studies and generally indicated improved	**No new RCTs** Effectiveness not sought	No new efficacy evidence	N/A

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
function and quality of life following the allograft procedure compared with pre-operative values. • One study reported a 91% survival rate of grafts at 5 years and 76% at both 10 and 15 years (N =65).			
Key Question 4: What is the evidence of the safety of OATS surgery? Including consideration of: a. **Adverse events type and frequency (peri-operative, cartilage plug detachment, cartilage rejection, graft fit, harvest site issues, development of fibrocartilage, mortality, other major morbidity such as DVT, deep infection, and excessive intraarticular bleeding)** b. **Revision/re-operation rates (if not addressed in efficacy)**			
Autograft			
Safety: Knee and Ankle (Low Evidence) • Data from three RCTs (all knee), 3 nonrandomized comparative studies (2 knee, 1 ankle), and 15 case series of osteochondral autograft transfer (9 knee, 4 ankle, and 2 both knee and ankle) were used • Surgical complications (infection, deep vein thrombosis, and hemarthrosis) are infrequent (<7%). • In 3 RCTs, revisions of OAT procedures in the knee were performed significantly less often than revisions following microfracture (1% vs. 33%; 2 trials, mean 3-4 year follow-up). There was no clear difference for OAT compared with ACI in one trial at 2 years (0% vs. 5%, respectively). Re-operations following OATs were 17% across seven case series of the knee and 34% across three case series of the ankle (variety of procedures, unclear timeframes). • Rates of donor site morbidity were 10% in two RCTs in the knee,10% across three case series in the knee, 7% across two case series in the ankle, and 9% in one case series at both sites.	**Systematic reviews** **Knee** Pareek 2016[2] (6 RCTs total, includes 3 new RCTs Gudas 2013, Ulstein 2014 and Lim 2012 and Gudas 2012 f/u, Gudas 2009; Autograft vs. MF) Andrade 2016 [11](11 studies, includes 1 RCT, 1 prospective cohort, 4 retrospective cohorts, and 5 case-series; Autograft only) **Ankle** Andrade 2016[11] (10 studies, no new RCTs, 3 retrospective cohorts, and 7 case-series; Autograft only) **RCTs, knee** **Follow-up publications:** Gudas 2012[3], *follow-up to Gudas 2005* (Autograft vs. MF) **New RCTs:** Ulstein 2014 [5](Autograft vs. MF)	**Knee** **Surgical complications** *Systematic reviews and RCTs*: NR **Failure (as defined by authors)** *Systematic reviews*: Pareek SR, Autograft vs. MF: MF had 2.4 times the risk of failure when compared with Autograft in 4 trials (RR 2.4, 95% CI 1.05, 5.52), p=0.036; N=180; Gudas 2009, Gudas 2012, Lim 2012, Ulstein 2014) over mean follow-up of 5.6 years (range, 3-10 years). *RCTs*: Autograft vs. MF: Gudas (N=57): Significantly lower risk of failure with Autograft (14% vs. 38%, p<0.05) over a mean follow-up of 10.4 years (range, 9-11 years) Autograft vs. ACI: Bentley (N=100), large defects: Significantly greater risk of failed cartilage repair (surgical intervention) with Autograft (55% vs. 17%, p<0.001) at a minimum follow-up of 10 years (range, 10-12 years) **Reoperation (as defined by authors)** *Systematic reviews*: NR	Knee: This portion of the report is still valid. New evidence at longer term continues to suggest that OAT autograft is associated with less failure and fewer reoperations compared with microfracture. Long-term follow-up for OAT vs. ACI from one large trial still suggests OAT may have greater failure vs. ACI. (Criteria A-2, B-4) Ankle This portion is still valid as primary evidence is still from case-series.

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
• No deaths directly attributable to OAT were found in the studies reviewed.	Bentley 2012[9], *follow-up to Bentley 2003* (Autograft vs. ACI) Lim 2012[6] (Autograft vs. MF and vs. ACI) **RCTs, Ankle** **New RCTs:** Ahmad 2016[10] (Autograft vs. Allograft)	*RCTs*: Autograft vs. MF: Ulstein (N=25): 36% vs. 54% at median follow-up of 9.8 years (range, 5-11 years), p=NS Lim (52 knees): 5% vs. 10% at a mean follow-up of 5 years (range, 3-10 years), p=NS Autograft vs. ACI: Lim (40 knees): 5% vs. 11% at a mean follow-up of 5 years (range, 3-10 years), p=NS **Donor-site morbidity** *Systematic reviews*: Andrade SR (N=1472 knee patients): The pooled estimate for knee-to-knee transplantation was 5.9% (range 0%-92% across 11 studies) over follow-up periods ranging from 1 to 9.6 years. The most common donor-site complaints were patellofemoral disturbances (23%) (3 studies), crepitation (31%) (2 studies) and post-operative effusion (9%) (2 studies) **Ankle** **Revision** *Systematic reviews*: NR*RCTs*: Ahmad, Autograft vs. Allograft (N=40): 10% vs. 13% at a mean follow-up of 3.1 years (range, 1-6.4 years), p=NR **Graft nonunion** *Systematic reviews*: NR *RCTs*: Ahmad, Autograft vs. Allograft (N=40): 10% vs. 19% at a mean follow-up of 3.1 years (range, 1-6.4 years), p=NR **Donor-site morbidity** *Systematic reviews*: Andrade SR (N=254 ankle patients): The pooled estimate for knee-to-ankle transplantation was 19.6% (range 0%-55% across 10 studies) over follow-	

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
		up periods ranging from 0.5 to 6.3 years. The most common donor-site complaints were pain or instability during daily living or sports activities (44%) (3 studies) and persistent pain (13%) (2 studies)	
Allograft versus various treatments			
Safety: Knee (Low Evidence) • Rates of all re-operations following OATs using allograft were 12.5% across seven studies (2 cohorts, 5 case-series). • Rate of graft failure was 21% in two case series that used radiographs. • One case of infection (4%) was reported in one case series. • Allograft transplantation carries an extremely small potential risk of disease transmission. No study of disease transmission related to osteochondral allograft was found in our search.	**Systematic reviews** **Knee** Familiari 2017[12] (19 studies total, 1 prospective cohort, 1 retrospective cohort and 17 case series; Allograft only) Assenmacher 2016[13] (5 studies total, 1 prospective cohort and 4 case series; Allograft only) **No new RCTs**	**Knee** **Reoperation (as defined by authors)** *Systematic reviews*: Familiari SR: Mean reoperation rate across 17 studies was 30.2% (range 0%-63%) over a mean follow-up of 8.7 years. Assenmacher SR: Mean reoperation rate across all studies was 36% over a mean follow-up of 12.3 years. The most common reoperations included unicompartmental or total knee arthroplasty (37%), debridement due to symptoms (24%) and graft-related surgery (removal, fixation, and revision) (14%). **Failure (as define by authors)** *Systematic reviews*: Familiari SR: Mean failure rate across 17 studies was 18.2% (range 0%-31%) over a mean follow-up of 8.7 years. Assenmacher SR: Mean failure rate across all studies was 25% over a mean follow-up of 12.3 years A total of 72% of the failures were conversion to total (68%) or unicompartmental (4%) knee arthroplasty, and 28% involved graft removal, graft fixation, and graft revision. **Survivorship** *Systematic reviews*: Familiari SR: Kaplan-Meier analysis of mean survivorship across the included 12 studies was	New evidence does not change the conclusions from the previous report (criteria A-1 or A3), nor provide major changes in the evidence (criteria B-1 – B4) for either autograft or allograft. This section does not need updating

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
		86.7% at 5 years, 78.7% at 10 years, 72.8% at 15 years and 67.5% at 20 years. Assenmacher SR: Kaplan-Meier analysis of mean survivorship was reported by 3 studies and was 94% at 5 years, 84% at 10 years, 71% at 15 years, and 45% at 20 years. **Post-operative infection** *Systematic reviews*: Assenmacher SR: One case of deep infection (1 study) and one case of superficial cellulitis (1 study). **Disease transmission** *Systematic reviews*: NR	

Table 4. Osteochondral Allograft/Autograft Transplantation Summary Table for Key Questions 5 and 6

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
Key Question 5: What is the evidence that OATS surgery has differential efficacy or safety issues in sub populations?			
Autograft			
Efficacy: Knee (Low Evidence) • Direct comparisons within RCTs are limited and may suggest that age, defect size, and defect location may influence outcomes • Indirect comparison of factors is challenging given differences in the populations studied, study quality the comparators used.	**Systematic reviews:** Pareek 2016[2] (includes 2 new trials –Gudas 2013, Lim 2012 and 10 year follow-up, Gudas 2012)	**Systematic reviews** Pareek SR: There was no effect modifications for Tegner Activity score (3 – 10 years) by defect size (< 3cm^2, > 3cm^2), p (interaction) = 0.134 **RCTs** No formal tests for interaction were reported for subanalyses related to patient characteristics or lesion characteristics.	There are new data that would update this section of the report. However, the findings from these studies don't meet the criteria that would trigger an updated report. (Criteria B1-4)
Effectiveness: Knee and Ankle (Very Low Evidence) • No direct comparisons for any factor were made in nonrandomized comparative studies • Indirect comparisons based on case series of autograft OATS/mosaicplasty suggest that younger patients may experience better function and be better able to return to sports. Better functional outcomes may occur with one plug versus multiple plugs based on two small studies. Lesion location may influence outcome. • Allograft: Limited information from two case series is conflicting with regarding the influence of gender.	Effectiveness Not Sought	N/A	N/A
Autograft and Allograft			
Safety: Knee and Ankle(Very Low Evidence) • No comparative studies of autograft or allograft transplantation assessed differential safety • Results of case series of autograft and allograft transplantation suggested that older patients may have more risk of graft failure and that grafts of larger lesions were more likely to fail. • No full economic studies directly addressing the cost-effectiveness of either autograft or allograft	**Systematic Reviews** Pareek 2016[2] **No New RCTs**	**Systematic reviews** Pareek: There was no effect modifications for failure by lesion type (osteochondritis dissecans (OCD) and articular cartilage defect (ACD), p(interaction) =0.101	There are new data that would update this section of the report. However, the findings from these studies don't meet the criteria that would trigger an updated report. (Criteria B1-4)

Conclusions from CER Executive Summary (Strength of Evidence)	New Sources of Evidence	New Findings	Conclusion from AAI
osteochondral transplantation as described in this report were found.			
Key Question 6: What is the evidence of cost implications and cost-effectiveness for OATS/mosaicplasty?			
Knee and Ankle (No Evidence) No full economic studies directly addressing the cost-effectiveness of either autograft or allograft osteochondral transplantation as described in this report were found.	CADTH 2017 Rapid review[14]: No economic studies for shoulder, ankle **<u>New cost effectiveness analysis</u>** **Knee** Miller 2015[15] Schronk 2017[16]	2 studies for isolated distal femoral lesions based on systematic reviews of level 1 or 2 studies; age range 15-55 years old. **Miller 2015** (N = 134 patients) OAT vs. microfracture for mean lesion size of 2.7 cm^2 (1.0 to 6.0 cm^2): Results for cost per point improvement pre-op to post-op in functional measures based on outcomes measure used. Only the International Cartilage Repair Society (ICRS) functional measure showed statistically significant difference (difference $98.29/per point improvement; OAT $308.50 vs. microfracture $ 406.79). Authors report that cost to return patients back to their previous level of sport at 1, 3, and 10 years, demonstrated OAT to be more cost-effective than microfracture for all years. Authors' conclusion: Microfracture was found to be more cost effective by the Lysholm and HSS scores, whereas OAT was more cost-effective by the Tegner and ICRS scores. Given similar clinical outcomes, microfracture and OAT are both viable, cost-effective first-line treatment options for these injuries. **Schronk 2017** (N = 730 knees) OAT, microfracture, ACI-1 (First-Generation Autologous Chondrocyte Implantation). Mean lesion sizes ranged from 1.9 cm^2 to 5.1 cm^2, mean follow-up ranged from 36.7 to 38.3 months. The costs per point functional outcome change were OAT $313.84, MF $200.59, AC-1 $536.59. Author' conclusions: All treatments led to an increase in functional outcome scores postoperatively MF was found to be the most cost effective treatment option and ACI-1 the least cost-effective.	There are new data that would update this section of the report. However, the findings from these studies don't meet the criteria that would trigger an updated report.

5. Conclusions

Tables 1, 2, 3 and 4 show the original key questions, the conclusions of the original report, the new sources of evidence, the new findings, and the conclusions of Aggregate Analytics, Inc. (AAI) with respect to the criteria that identify a trigger for an update (Figure 1). This report focuses on Key questions 3-6.

5.1 Key Question 1: NOT PART OF SIGNAL UPDATE

5.2 Key Question 2: NOT PART OF SIGNAL UPDATE

5.3 Key Question 3 (Efficacy):

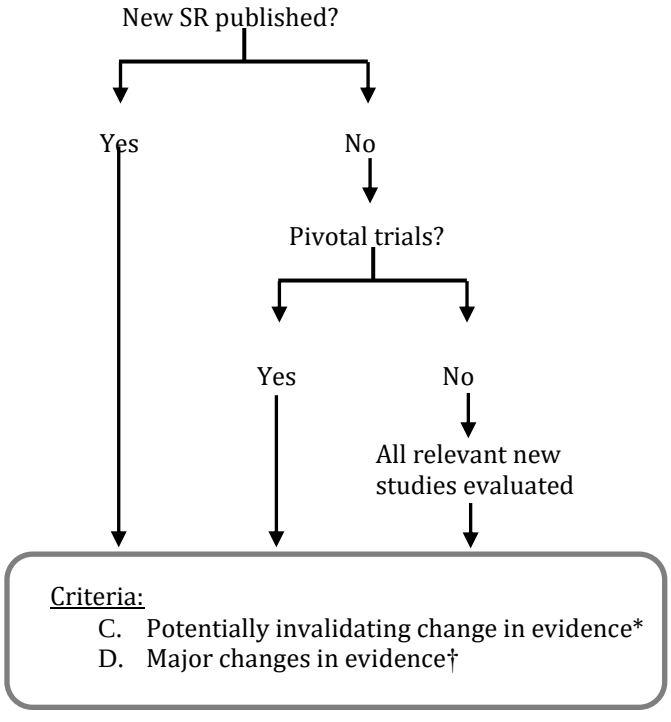
- OAT/mosaicplasty vs. microfracture, drilling or debridement alone
  - **Knee:** Two systematic reviews incorporating new RCTs and one additional RCT (not incorporated in to systematic reviews) comparing OAT and microfracture and describing longer-term outcomes were identified. Pooled data including new trials suggest no difference between OAT autograft and microfracture at ≥5 years for function or Tegner score. Data are from small non-pivotal trials; the evidence base is likely low or insufficient. This section of the report is still valid and does not need updating. (Criteria A1, B-1-4)
  - **Ankle:** One new RCT comparing OAT with microfracture and with drilling was identified, however the trial is not considered pivotal and doesn't meet the criteria that would trigger a report update. (Criteria A-1, A-3, B2).
- OAT/mosaicplasty vs. ACI (Knee)
  - One new, small RCT and 10 year follow-up from a previously included trial comparing OAT with ACI were identified. Results are consistent with the previous report; there are no major changes in evidence (criteria B 1-4). This section does not need updating.
- Autograft vs. Allograft (Ankle)
  - One small new trial evaluating OAT autograft with allograft in the ankle/talus was identified but is not considered pivotal. The findings don't meet the criteria that would trigger an updated report (criterion A-1).

5.4 Key Question 4 (Safety): New evidence does not change the conclusions from the previous report (criteria A-1 or A3); there are not any major changes in the evidence base (criteria B-1 – B4) for either autograft or allograft. This section does not need updating

5.5 Key Question 5 (Differential efficacy or safety): There is limited information from one systematic review suggesting that lesion size or type do not modify treatment with regard to the outcomes of activity or implant failure. However, the findings don't meet the criteria that would trigger an updated report (Criteria B1-4).

5.6 Key Question 6: Two cost-effectiveness studies comparing OAT with microfracture have been published since the previous report. However, the findings don't meet the criteria that would trigger an updated report.

Figure 2. 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

REFERENCES

1. Gracitelli GC, Moraes VY, Franciozi CE, Luzo MV, Belloti JC. Surgical interventions (microfracture, drilling, mosaicplasty, and allograft transplantation) for treating isolated cartilage defects of the knee in adults. The Cochrane database of systematic reviews 2016; 9: Cd010675.
2. Pareek A, Reardon PJ, Macalena JA, et al. Osteochondral Autograft Transfer Versus Microfracture in the Knee: A Meta-analysis of Prospective Comparative Studies at Midterm. Arthroscopy 2016; 32(10): 2118-30.
3. Gudas R, Gudaite A, Pocius A, et al. Ten-year follow-up of a prospective, randomized clinical study of mosaic osteochondral autologous transplantation versus microfracture for the treatment of osteochondral defects in the knee joint of athletes. Am J Sports Med 2012; 40(11): 2499-508.
4. Solheim E, Hegna J, Strand T, Harlem T, Inderhaug E. Randomized Study of Long-term (15-17 Years) Outcome After Microfracture Versus Mosaicplasty in Knee Articular Cartilage Defects. Am J Sports Med 2017: 363546517745281.
5. Ulstein S, Aroen A, Rotterud JH, Loken S, Engebretsen L, Heir S. Microfracture technique versus osteochondral autologous transplantation mosaicplasty in patients with articular chondral lesions of the knee: a prospective randomized trial with long-term follow-up. Knee Surg Sports Traumatol Arthrosc 2014; 22(6): 1207-15.
6. Lim HC, Bae JH, Song SH, Park YE, Kim SJ. Current treatments of isolated articular cartilage lesions of the knee achieve similar outcomes. Clin Orthop Relat Res 2012; 470(8): 2261-
7. Gudas R, Gudait A, Mickevi?ius T, et al. Comparison of osteochondral autologous transplantation, microfracture, or debridement techniques in articular cartilage lesions associated with anterior cruciate ligament injury: a prospective study with a 3-year follow-up. Arthroscopy, 2013.
8. Sun JL, Li LL, Yang JX, Chen Y, Ma WL. A comparison of surgical approaches for osteochondral lesions of the talus associated with ankle fractures. International journal of clinical and experimental medicine, 2016.
9. Bentley G, Biant LC, Vijayan S, Macmull S, Skinner JA, Carrington RW. Minimum ten-year results of a prospective randomised study of autologous chondrocyte implantation versus mosaicplasty for symptomatic articular cartilage lesions of the knee. The Journal of bone and joint surgery British volume 2012; 94(4): 504-9.
10. Ahmad J, Jones K. Comparison of Osteochondral Autografts and Allografts for Treatment of Recurrent or Large Talar Osteochondral Lesions. Foot & ankle international 2016; 37(1): 40-50.
11. Andrade R, Vasta S, Pereira R, et al. Knee donor-site morbidity after mosaicplasty - a systematic review. J Exp Orthop 2016; 3(1): 31.
12. Familiari F, Cinque ME, Chahla J, et al. Clinical Outcomes and Failure Rates of Osteochondral Allograft Transplantation in the Knee: A Systematic Review. Am J Sports Med 2017: 363546517732531.
13. Assenmacher AT, Pareek A, Reardon PJ, Macalena JA, Stuart MJ, Krych AJ. Long-term Outcomes After Osteochondral Allograft: A Systematic Review at Long-term Follow-up of 12.3 Years. Arthroscopy 2016; 32(10): 2160-8.
14. CADTH Rapid Response Reports. The Use of Osteochondral Allograft for the Ankle, Knee, and Shoulder: Clinical Effectiveness and Cost-Effectiveness. Ottawa (ON): Canadian Agency for Drugs and Technologies in Health Copyright (c) 2017 Canadian Agency for Drugs and Technologies in Health.; 2017.
15. Miller DJ, Smith MV, Matava MJ, Wright RW, Brophy RH. Microfracture and osteochondral autograft transplantation are cost-effective treatments for articular cartilage lesions of the distal femur. Am J Sports Med 2015; 43(9): 2175-81.

16. Schrock JB, Kraeutler MJ, Houck DA, McQueen MB, McCarty EC. A Cost-Effectiveness Analysis of Surgical Treatment Modalities for Chondral Lesions of the Knee: Microfracture, Osteochondral Autograft Transplantation, and Autologous Chondrocyte Implantation. Orthop J Sports Med 2017; 5(5): 2325967117704634.

APPENDIX A. SEARCH STRATEGIES

Search strategy for PubMed: Search dates: March 1, 2011 through January 10, 2018

	Search terms	**Number of articles**
#1	("osteochondral autograft transfer" OR "mosaicplasty" OR "mosaicplasties")	197
#2	(chondral OR osteochondral) OR ("Cartilage, Articular"[MeSH] OR "Osteochondritis Dissecans"[MeSH] OR "osteochondritis dissecans")	9202
#3	#1 OR #2	9218
#4	(transplant OR transplants OR transplantation* OR implant OR implants OR implantation* OR graft OR grafts OR grafting OR autograft* OR autologous OR autotransplant* OR ("Transplantation, Autologous"[MeSH]) OR allograft* OR allogeneic OR homograft* OR allotransplant* OR ("Transplantation, Homologous"[MeSH]))	374,476
#5	#3 AND #4	2519
#6	rabbit* OR "mouse" OR "mice" OR "rat" OR "rats" OR "dog" OR "dogs" OR "Models, Animal"[MeSH] OR (Animals[MeSH] NOT "Humans"[MeSH])	1,100,143
#7	("Case Reports"[Publication Type] OR "case report")	375,523
#8	#6 OR #7	1,467,032
#9	#1 OR #5 NOT #8	1668

Search strategy for Cochrane: Search dates: March 1, 2011 through January 10, 2018

	Search terms	**Number of articles**
#1	("osteochondral autograft transfer" OR "mosaicplasty" OR "mosaicplasties")	15
#2	(chondral OR osteochondral) OR ("Cartilage, Articular"(MeSH) OR "Osteochondritis Dissecans"(MeSH) OR "osteochondritis dissecans")	182
#3	#1 OR #2	187
#4	(transplant OR transplants OR transplantation* OR implant OR implants OR implantation* OR graft OR grafts OR grafting OR autograft* OR autologous OR autotransplant* OR ("Transplantation, Autologous"(MeSH)) OR allograft* OR allogeneic OR homograft* OR allotransplant* OR ("Transplantation, Homologous"(MeSH)))	30988
#5	#3 AND #4	89
#6	rabbit* OR "mouse" OR "mice" OR "rat" OR "rats" OR "dog" OR "dogs" OR "Models, Animal"(MeSH) OR (Animals(MeSH) NOT "Humans"(MeSH))	4248
#7	("Case Reports"(Publication Type) OR "case report")	4556
#8	#6 OR #7	8609
#9	#1 OR #5 NOT #8	76*

*4 technology assessments and 1 economic evaluation were excluded. All were either structured or provisional abstracts and/or were not study types of interest

Additional electronic databases were searched using key words and included EMBASE, ClinicalTrials.gov, AHRQ, National Guideline Clearinghouse and INAHTA for eligible studies, including health technology assessments (HTAs), systematic reviews, primary studies and FDA reports.
Additional searches yielded 11 articles not previously captured but none met inclusion criteria.

APPENDIX B. SUMMARY OF INCLUDED STUDIES

Appendix Table B1. Summary of systematic reviews included for efficacy

Assessment *Search dates*	Purpose	Condition	Treatment vs. comparators	Primary Outcomes	Evidence- base Used	Primary Conclusions
Gracitelli 2016 Cochrane review *Database inception to February 5th, 2016*	To assess the relative effects (benefits and harms) of different surgical interventions (microfracture, drilling, mosaicplasty, and allograft transplantation) for treating isolated cartilage defects of the knee in adults.	Knee cartilage defects	OAT vs MF	Knee function assessed by validated tools, QoL measures, failure of treatment and adverse effects	3 RCTs (2 new RCTs: Lim 2012, Ulstein 2014) (n=133)	**OAT vs MF efficacy:** In a pooled analysis of Lysholm scores at a follow-up of 5 years or longer (SoE very low/insufficient), the authors report no difference in outcomes (2 new trials, pooled difference -1.10, 95% CI -4.54 to 2.33). One older trial (included in prior report) favored OAT on the IKDC score (MD13.97, 95%CI 13.25, 14.69) (SoE very low/insufficient) **OAT vs MF return to activity:** Mean Tegner score at a follow-up of 5 years or longer was not statistically significant for either new trial (Lim 2012, Ulstein 2014). Continuation of sport in Gudas 2005 was more common in OAT (RR 3.24, 95% CI 0.81, 5.40) but statistical significance was not reached at 10 years (RR 2.07, 95%CI 0.81, 5.30) (SoE very low/insufficient) **OAT vs MF safety:** Across 3 trials at a follow-up of five years or longer, authors report failure of treatment and adverse events occurred at a statistically significant lower rate in OAT (SoE very low/insufficient) (3 trials [2 new trials], pooled RR 0.47, 95% CI 0.24, 0.9)
Pareek 2016 *January 1st 1995 to May 1st 2015*	To compare OAT and MF surgical techniques to determine postoperative activity level, subjective patient outcomes, failure rates, and assess if any lesion characteristics favored one	Knee articular cartilage damage	OAT vs MF	Activity related scores, subjective clinical scores, and failure rate	6 RCTs (3 new RCTs: Lim 2012, Ulstein 2014, Gudas 2013; 1 new follow-up publication: Gudas 2012) (n=249)	**OAT vs MF efficacy:** In a pooled analysis of subjective scores at a follow-up of 3 years (SoE not reported), OAT demonstrated statistically significant improved scores (3 trials [1 new trial], pooled SMD 0.40, 95% CI 0.10 to 0.70). The difference was not statistically significant at a follow-up of 5 to 10 years (3 trials [3 new trials], pooled SMD 0.92, 95% CI -1.07 to 2.90) but substantial heterogeneity was noted, only the 10 year follow-up reported in Gudas 2013 was significant.

Assessment *Search dates*	Purpose	Condition	Treatment vs. comparators	Primary Outcomes	Evidence- base Used	Primary Conclusions
	technique over the other.					**OAT vs MF return to activity:** Across 3 trials with a follow-up of 3 to 10 years (SoE NR), authors report statistically significant better Tegner scores in OAT (3 trials [3 new trials], pooled SMD 0.469, 95% CI 0.140 to 0.798). In Gudas 2012, mean duration of previous sport activity was statistically longer in OAT compared to MF. In a subgroup analysis of lesion size, the authors found that OAT performed statistically significantly better in lesions > 3 cm^2 (2 trials [2 new trials], pooled SMD 0298, 95% CI -0.076 to 0.673) but not in lesions < 3 cm^2 (1 trial [1 new trial], SMD 0.768, 95% CI 0.281 to 1.256). No modification by defect size was found (p (interaction) = 0.134).

ACI: autologous chondrocyte implantation; CI: confidence interval; MF: microfracture; OA: osteoarthritis; OAT: osteochondral autologous transplantation; QoL: quality of life; RCT: randomized controlled trial; RR: risk ratio; SMD: standardized mean difference; SoE: Strength of Evidence

Appendix Table B2. Summary of systematic reviews included for safety

Assessment *Search dates*	Purpose	Condition	Population	Primary Outcomes	Evidence- base Used	Primary Safety Conclusions
Andrade 2016 *Database inception to October 2016*	To provide an overview of donor-site morbidity associated with harvesting osteochondral plugs from the knee joint in mosaicplasty procedure*	Full-thickness cartilage lesions of weight-bearing joints in the knee or ankle*	N=21 articles (N=1726 patients) Knee**:** 11 articles (N=1472 patients, mean age 33.2 years, follow-up 12 to 115 months) Ankle**:** 10 articles (N=254 patients, mean age 34.8 years, follow-up 12 to 76 months)	Presence of donor-site morbidity after mosaicplasty	**Level I:** n=1 (1 knee) **Level II:** n=1 (1 knee) **Level III:** n=7 (4 knee, 3 ankle) **Level IV:** n=12 (5 knee, 7 ankle)	The donor-site morbidity for knee-to-ankle (19.6%, range across studies 0%-55%) was greater than knee-to-knee (5.9%, range across studies 0%-92%) mosaicplasty procedures, without any significant correlation between rate of donor-site morbidity and size of the defect, number and size of the plugs. Most common donor-site morbidity complaints for the knee were patellofemoral disturbances (23 %) and crepitation (31%); post-op effusion (9%). For the ankle, complaints were pain or instability during daily living or sports activities (44 %), patellofemoral disturbances (13 %), knee stiffness (13 %) and persistent pain (13 %)
Assemacher 2016 *January 1, 1995 to June 1, 2015*	To evaluate long-term clinical outcome scores, reoperation, and failure rates of osteochondral allograft and to examine if certain factors predispose patients to worse outcomes	Full-thickness cartilage defects of articular cartilage and subchondral bone in the knee	N=5 studies (N=291 patients, 55% male, age 34.8 years, mean 12.3 years follow-up, 10 to 17.1 years)	Clinical outcomes, reoperation rates, failure rates	**Level II**: n=1 study **Level IV:** n=4 studies	Across all studies at final follow-up, mean failure rate was 25% and mean reoperation rate was 36%. Post-operative infection was reported by 2 studies: one reported 1 case of deep infection and one reported a case of superficial cellulitis. Survivorship was reported by 3 studies and was 94% at 5 years, 84% at 10 years, 71% at 15 years, and 45% at 20 years. Results are similar to failure and reoperation rates for alternative cartilage restoration techniques. Reoperation for patellofemoral grafts was significantly higher (83%), but most of the procedures (mean 1.8, range 0 to 6) were for debridement and hardware removal. Femoral condyle grafts have slightly improved survivorship. Patellofemoral grafts are less successful than tibial and femoral grafts, as seen in this review
Familiari 2017 *1980 to March 2017*	To review clinical outcomes and failure rates after osteochondral allograft	Chondral defects of the knee	N=19 studies (N=1036 patients, mean 31.5 years (10-82), mean	Clinical outcomes and failure rates	**Prospective cohort:** n=1 study **Retrospective cohort:** n=1 study	OCA transplantation of the knee yielded good survival rates at 5 to 10 year follow-up. Mean 5-year survival rate across the studies included in this review was 86.7%, while the mean 10-year survival rate was 78.7%. The survival rates were 72.8% at 15

Assessment *Search dates*	Purpose	Condition	Population	Primary Outcomes	Evidence- base Used	Primary Safety Conclusions
	transplantation in the knee at a mean 2 years' follow-up		follow-up 8.7 years (2-32 years))		**Case series:** n=17 studies	years and, subsequently, 67.5% at 20 years. OCA transplantation was associated with considerable reoperation (30.2%, range 0%-63%) (17 studies) and failure (18.2%, range 0%-31%) (17 studies) rates at final follow-up.
Pareek 2016 *January 1st 1995 to May 1st 2015*	To compare OAT and MF surgical techniques to determine postoperative activity level, subjective patient outcomes, failure rates, and assess if any lesion characteristics favored one technique over the other.	Knee articular cartilage damage	N=6 studies (N=249)	Activity related scores, subjective clinical scores, and failure rate	**Randomized controlled trials**: n=6 trials	Pooled analysis of failure of treatment across 4 trials found OAT had a statistically significant lower rate of failure† (4 trials [3 new trials], pooled SMD 2.417, 95% CI 1.059 to 5.519). The difference remained statistically significant when a sub-analysis was performed on trials reporting on both articular cartilage defect and osteochondritis dissecan lesions (3 trials [3 new trials], pooled SMD 1.959, 95% CI 1.033 to 3.713), compared to osteochondritis dissecan lesions alone (1 trial, SMD 21.478, 95% CI 0.476 to 39.703).

CI: confidence interval; MF: microfracture; OAT: osteochondral autologous transplantation; OCA: osteochondral allograft transplantation; SMD: standardized mean difference

*All donor sites were in the knee but cartilage lesions occurred in either the knee or ankle

†Authors note in discussion that an important limitation of the meta-analysis is the variability in the definition of "failure" between studies

Appendix Table B3. Summary of CADTH Rapid Review

Assessment *Search dates*	Purpose	Condition	Treatment vs. comparators	Primary Outcomes	Evidence- base Used	Primary Conclusions
The Use of Osteochondral Allograft for the Ankle, Knee, and Shoulder: Clinical Effectiveness and Cost-Effectiveness CADTH Rapid Response Report (2017) *January 1, 2012 to January 10, 2017*	The report aimed to provide evidence on the clinical benefits, harms, and cost-effectiveness of the use of fresh, prolonged fresh, or frozen osteochondral allografts for the lesions of the ankle, knee, and shoulder.	Painful lesion of the ankle, knee, or shoulder involving cartilage or cartilage with bone that has failed non-operative and primary treatment	**Knee** Osteochondral allograft transplantation* vs before the operation. Three SRs reported some patients were treated with concomitant procedures† **Ankle** Osteochondral allograft transplantation‡ vs before the operation **Shoulder** Osteochondral allograft transplantation vs before the operation	Clinical effectiveness, functional outcomes, and cost-effectiveness	**Knee** 4 SRs of cohort or case-series studies **Ankle** 2 SRs of case-series and/or other nonrandomized studies **Shoulder** 1 SR of case-series studies No economic evaluations were identified	**Knee** Function*:* All SRs reported improved functional outcomes compared to before surgery. Return to activity*:* One SR reported patients returned to full activity 5.9 months on average after surgery. Another SR reported most patients returned to sports and preinjury-level performance by 30 months and 9.6 months, respectively. SRs (number not reported) reported improved Tegner scores compared to before surgery. Pain: All SRs reported improved pain outcomes compared to before surgery Patient satisfaction: One SR reported that 86% of patients were extremely or mostly satisfied with the operation. Failure and reoperation: One SR reported that 36% of patients had reoperations. Two SRs reported that 18% to 25% of all operations were considered failures, requiring conversion to knee arthroplasty or graft revision or removal. **Ankle** Function: SRs reported improved functional scores after surgery. Pain: One SR reported improved VAS scores compared to before surgery. Patient satisfaction: One SR reported that 71% of patients reported good to excellent satisfaction with the operation. Failure and reoperation: One SR reported that 25% of patients required at least one reoperation of any kind and that 13% of all operations were considered failures §. **Shoulder** Function: The one SR reported higher shoulder stability after surgery and the no recurrence of

Assessment *Search dates*	Purpose	Condition	Treatment vs. comparators	Primary Outcomes	Evidence- base Used	Primary Conclusions
						shoulder instability. Range of motion was restored or increased compared to before the operation. Complication rates: SR reported 74% of patients with shoulder instability had complications after the operation**.

CADTH: Canadian Agency for Drugs and Technologies in Health; SR: systematic review

*One SR included only fresh allografts, another included fresh, prolonged-fresh, and fresh-frozen allografts, and the remaining two SRs did not specify restrictions on the type of allograft
†Concomitant procedures included tibial tubercle transfer and extensor mechanism realignment, osteotomy, meniscal transplantation, ligamentous reconstruction, and retinacular release
‡One SR included only fresh allografts; the other SR did not specify any restriction on allograft type
§Defined as postoperative graft nonunion, resportion, or persistence of symptoms leading to subsequent arthrodesis or arthroplasty
**Complications included spontaneous avascular necrosis and collapse, persistent pain, clicking, catching, stiffness, and flattening

Appendix Table B4. Study characteristics of new RCTs and new follow-up publications

Author (Year)	Demographics	Results	Conclusions	Comments
Knee: Autograft vs. MF or Debridement alone				
Solheim 2017	N= 40 Age at surgery, mean (IQR): 32 (18-48) years % Male: 70% F/U: 1, 5, 10, 15 years Lesion size: 3.5 cm^2 Lesion description: full-thickness articular chondral defects on the condyles or trochlea Area: NR OAT (Autograft , mosaicplasty) After arthroscopic evaluation and debridement to subchondral bone, grafts were harvested from the periphery of the patellofemoral joint and transplanted into corresponding bur holes in the defect MF After arthroscopic evaluation and debridement to subchondral bone, angled awls were used to make holes in the subchondral bone plate were made 3 to 4 mm apart.	Function: OAT vs. MF Lysholm Score, mean (SD), p: • Baseline: 56 (15) vs 50 (16), p = 0.2 • 1 year: 85 (12) vs. 72 (22), p = 0.015 • 5 years: 83 (9) vs 67 (18), p < 0.001 • 10 years: 81 (16) vs 65 (22), p = 0.020 • ≥15 years: 77 (17) vs 61 (22), p = 0.011 (difference of >9 points considered clinically significant) Minimum 15 years success, n/N (%), p: • Lysholm <64 (poor outcome): 4/20 (20%) vs 13/20 (65%), p = 0.004 • Lysholm ≥80 (good outcome): 12/20 (60%) vs 4/20 (20%), p = 0.010 Later surgical procedure • Knee replacement, n (%): 3 (15%) vs. 1 (5%), p =0.292	**Function**: at all time points through minimum of 15 years, mosaicplasty was associated with a statistically and clinically relevant improvement in function with more mosiacplasty patients reporting good outcome at 15 years. Additional Surgery: No significant differences between groups; any other safety or adverse outcomes were reported.	The authors declare no conflict of interest Funding NR
Ulstein 2014	N= 25 Age, mean (SD): 32.3 (7.7) years % Male: 56% F/U, median (IQR): 9.8 (4.9 to 11.4) years Lesion size, mean (range): 2.8 (2.0 to 6.0) cm^2 Lesion description: chondral or osteochondral lesion of ICRS grade III-IV Area, n/N (%) trochlea vs n/N (%) medial vs n/N (%) lateral: 2/25 (8%) vs 20/25 (80%) vs 3/25 (12%)	Function: OAT vs MF Lysholm Score mean change (95% CI), (MD, 95% CI), p: • 9.8 (4.9-11.4) years: 13.4 (0.9 to 25.8) vs 21.6 (3.7 to 39.4), (MD 8.2, 95% CI -11.7 to 28.1), p NS KOOS pain mean change (95% CI), (MD, 95% CI), p: • 9.8 (4.9-11.4) years: 11.8 (-2.8 to 26.4) vs 20.6 (2.8 to 38.3), (MD 8.8, 95% CI -12.7 to 30.3), p NS KOOS symptoms mean change (95% CI), (MD, 95% CI), p:	**Function**: There were no significant differences in Lysholm score or KOOS, at median follow-up of 9.8 years.	Included in Gracitelli 2016 Cochrane Review and Pareek 2016 SR Restricted shuffling approach for randomization may not be true randomization (see Schulz 2002) Authors declare no conflict of interest

Author (Year)	Demographics	Results	Conclusions	Comments
Knee: Autograft vs. MF or Debridement alone				
	OAT (Autograft mosaicplasty) Procedure was performed though medial parapatellar arthrotomy or a mini-invasive arthrotomy. Osteochondral grafts from periphery of the femoral condyles were transplanted using "press-fit" method into lesion site MF Procedure was done arthroscopically. Debridement of all damaged/unstable cartilage was done. An arthroscopic awl was used to make multiple holes 3 to 4 mm apart	• 9.8 (4.9-11.4) years: 8.5 (-3.5 to 20.6) vs 17.4 (2.6 to 32.2), (MD 8.9, 95% CI -8.9 to 26.7), p NS KOOS activities in daily living mean change (95% CI), (MD, 95% CI), p: • 9.8 (4.9-11.4) years: 7.5 (-4.3 to 19.3) vs 13.0 (-3.8 to 29.8), (MD 5.5, 95% CI -13.4 to 24.4), p NS KOOS function in sport and recreation mean change (95% CI), (MD, 95% CI), p: • 9.8 (4.9-11.4) years: 41.3 (23.7 to 58.9) vs 32.4 (13.3 to 51.6), (MD -8.9, 95% CI -33.4 to 15.7) p NS KOOS quality of life mean change (95% CI), (MD, 95% CI), p: • 9.8 (4.9-11.4) years: 25.0 (10.6 to 39.3) vs 34.6 (15.1 to 54.0), (MD 9.6, 95% CI -12.7 to 31.9), p NS Reoperation: OAT vs MF Reoperation, n/N (%), p: 5/14 (36%) vs 6/11 (54%), p NS		Funding: Grant from Akershus University Hospital and the Foundation of Sophies Minde
Gudas 2013	N= 136 (102 randomized, 34 matched controls) Age, mean: 32.7 years w/o control, 32.0 w/control Male: 63% w/o control, 65% w/control F/U: Lesion size, mean (SD): 2.9 (4.2) Lesion description: articular cartilage damage grades III-IV in the femoral condyle Area: medial OAT (Autograft) (n=34) Performed under arthroscopic control simultaneously with ACL reconstruction. Eight mm plugs from	Function: OAT vs MF IKDC subjective score, mean (SD): • Preoperative: 45.5 vs 46.5 • 3 years*: 86.8 (2.6) vs 86.0 (3.5), p = 0.024 Return to activity: OAT vs MF Tegner score, mean (SD): • Preoperative: 2.5 vs 2.7 • 3 years: 7.1 vs 6.9 Function: OAT vs debridement IKDC subjective score, mean (SD): • Preoperative: 45.5 vs 47.1 • 3 years*: 86.8 (2.6) vs 84.5 (2.6), p = 0.018	**Function:** at a follow-up of 3 years, OAT had statistically significant improved IKDC scores compared to MF and debridement. **Return to activity:** At a 3 year follow-up, OAT had slightly higher Tegner scores than both MF and debridement but statistical significance was unclear.	Included in Pareek 2016 SR Authors declare no conflicts of interest

Author (Year)	Demographics	Results	Conclusions	Comments
Knee: Autograft vs. MF or Debridement alone				
	medial and/or lateral margin of femoral trochlea were used MF (n=34) Awls used to make perforations 3 to 4 mm apart Debridement (n=34) Unstable cartilage was debrided and the calcified cartilage layer was removed	Return to activity: OAT vs debridement Tegner score, mean (SD): •Preoperative: 2.5 vs 2.5 •3 years: 7.1 vs 6.2		
Lim 2012†	N= 109 patients (120 knees) randomized, 69 patients (70 knees) evaluated Age, mean (range): 28.5 (18-42) years % Male: 57% F/U, mean (range): 5.7 (3 to 10.5) years Lesion size: 2.74 cm^2 Lesion description: single symptomatic articular cartilage lesion of the knee Area, n/N (%) medial vs n/N (%) lateral: 55/70 (79%) vs 15/70 (21%) OAT (Autograft mosaicplasty) (n=22 knees) Performed after arthroscopic examination and debridement of fibrillated cartilage. Plugs of 4, 6, and 8 mm were inserted using press-fit method MF (n=30 knees) After arthroscopic examination, tapered awls were used to make 0.5 to 1 mm diameter holes 4 mm deep and placed 3 to 4 mm apart	Function: OAT vs MF ‡ Lysholm, mean (SD), (MD, 95% CI): •Preoperative: 53.2 (7.2) vs 51.2 (6.2) •5 years: 84.8 (5.5) vs 85.6 (6.8), (-0.8, 95% CI -4.5 to 2.9), p = 0.66 HSS, mean (SD), (MD, 95% CI): •Preoperative: 78.66 (7.23) vs 78.22 (9.12) •5 years: 88.12 (4.15) vs 87.60 (4.56), (MD 0.52, 95% CI -2.06 to 3.09), p = 0.69 Return to activity Tegner, mean (SD), (MD, 95% CI): •Preoperative: 2.7 (1.5) vs 2.8 (1.4) •5 years: 5.3 (1.2) vs 5.1 (1.5), (MD 0.2, 95% CI -0.6 to 1.0), p = 0.62 Reoperation: OAT vs MF§ Reoperation, n/N (%), (RR, 95% CI), p: 1/22 (5%) vs 3/30 (10%), (RR 0.45, 95% CI 0.1 to 4.1), p = 0.47	**Function:** There were no differences in functional scores measured with Lysholm or HHS at a five year follow-up **Return to activity:** At a five year follow-up, there were no differences in return to activity as measured by Tegner **Reoperation:** There was no difference in number of reoperations at a five year follow-up	109 patients enrolled, only 69 underwent procedures Authors declare no conflicts of interest Funding NR

Author (Year)	Demographics	Results	Conclusions	Comments
Knee: Autograft vs. MF or Debridement alone				
Gudas 2012 (follow-up to Gudas 2005)**	N= 60 Age, mean (range): 24.3 (15 to 40) years % Male: 63% F/U: 3 and 10 years Lesion size: 2.8 (1.4) cm^2 Lesion description: articular cartilage defect or osteochondral defect of the knee Area, % medial vs % lateral: 84% vs 16% OAT (Autologous) Residual cartilage and calcified layers of subchondral bone were removed. 5.5 mm plugs from lateral and/or medial margin of the femoral trochlea were used and transplanted into defect using "press-fit" method. MF Debridement of unstable cartilage was done and calcified layer was removed. Arthroscopic awl made multiple holes 2 to 4 mm apart.	Function: OAT vs MF ICRS, mean (SD): • 10 years: $p < 0.005$ Return to activity: OAT vs MF Return to preinjury sports activities: • $p < 0.001$ Average duration of return to previous sports activities: • $p < 0.005$ Failure: OAT vs MF Reoperation during 10 year follow-up, n/N (%), p: 4/28 (14%) vs 11/29 (38%). $p < 0.05$ Authors report subanalysis based on lesion type ACD and OCD as well as on age less than 25 and greater than 25 but do not provide formal test for interaction.	**Function:** At a 10 year follow-up, a statistically significant difference in ICRS scores was found in favor of OAT **Return to activity:** In terms of return to preinjury sports activities and duration of continuation of sports after surgery was statistically significantly better in OAT at a 10 year follow-up. **Failure:** Over 10 year period, OAT had a statistically significant lower rate of reoperation.	Authors declare no conflict of interest Funding NR
Ankle/Talus: Autograft vs. MF or drilling or allograft				
Sun 2016	N= 153 Age, mean (SD): 33.6 (6.9) % Male: 59% F/U: mean 27.4 months Lesion size: NR Lesion description: osteochondral lesions of the talus Area: NR A. OAT (Autograft) (n=52) Follow debridement, 4-9 mm holes 5.0 mm in depth were drilled into	Function: OAT vs. MF AOFAS, mean (SD), p: • 2.3 (1.7 to 3.0) years: 79.6 (6.5) vs 76.7 (8.4), p = NS AOFAS, mean change (SD), p: • 2.3 (1.7 to 3.0) years: 25.1 (1.3) vs 24.3 (1.6), p = NS Mazur ankle scoring system, mean (SD), p: • 2.3 (1.7 to 3.0) years: 95.2 (8.8) vs 92.3 (7.4), p = NS Mazur ankle scoring system, mean change (SD), p:	**Function:** Authors report no difference in changes scores of AOFAS or Mazur ankle scoring system values between OAT and MF; both OAT and MF resulted in improved AOFAS and Mazur ankle scoring compared with drilling. **Return to activity:** No difference in Tegner scores were reported between OAT	Authors declare no conflict of interest Funding NR No description of how randomization was done and no description of concealed allocation

Author (Year)	Demographics	Results	Conclusions	Comments
Knee: Autograft vs. MF or Debridement alone				
	the cartilage surface under arthroscopy. Grafts were taken from the outside of the ipsilateral patellofemoral joint and transplanted into the defects. Ankle fracture fixation and/or ligament repair was done and the joint capsule was sutured. B. MF (n=53) Following debridement, holes of depth 3 to 4 mm were made 3 to 4 mm apart under arthroscopy. C. Drilling (n=48) Following debridement, the fracture surface was trimmed under arthroscopy. Holes with depths 1.0 to 1.5 cm were drilled.	• 2.3 (1.7 to 3.0) years: 41.8 (3.2) vs 40.5 (4.1), p = NS Return to activity: OAT vs MF Tegner, mean (SD), p: • 2.3 (1.7 to 3.0) years: 4.7 (2.1) vs 4.6 (1.3), p = NS Tegner, mean change (SD), p: • 2.3 (1.3 to 3.0) years: 2.8 (0.3) vs 2.8 (0.7), p = NS Pain: OAT vs MF VAS (0-10), mean (SD), p: • 2.3 (1.7 to 3.0) years: 2.4 (0.4) vs 2.7 (0.3), p = NS VAS (0-10), mean change (SD), p: • 2.3 (1.7 to 3.0) years: 5.1 (1.2) vs 4.9 (0.7), p = NS, Function: OAT vs. drilling AOFAS, mean (SD), p: • 2.3 (1.7 to 3.0) years: 79.6 (6.5) vs 64.9 (9.8), AOFAS, mean change (SD), p: • 2.3 (1.7 to 3.0) years: 25.1 (1.3) vs 11.2 (0.7), p < 0.05 Mazur ankle scoring system, mean (SD), p: • 2.3 (1.7 to 3.0) years: 95.2 (8.8) vs 80.1 (9.8), p < 0.05 Mazur ankle scoring system, mean change (SD), p: • 2.3 (1.7 to 3.0) years: 41.8 (3.2) vs 28.0 (1.7), p < 0.05 Return to activity Tegner, mean (SD), p: • 2.3 (1.7 to 3.0) years: 4.7 (2.1) vs 3.6 (1.1), Tegner, mean change (SD), p:	and MF, but both OAT and MF resulted in improved Tegner scores compared to drilling. **Pain:** Authors report no difference in pain VAS between OAT and MF; both OAT and MF resulted in improved pain VAS compared with drilling.	

Author (Year)	Demographics	Results	Conclusions	Comments
Knee: Autograft vs. MF or Debridement alone				
		• 2.3 (1.3 to 3.0) years: 2.8 (0.3) vs 1.8 (0.2), p < 0.05, <u>Pain: OAT vs drilling</u> VAS (0-10), mean (SD), p: • 2.3 (1.7 to 3.0) years: 2.4 (0.4) vs 5.2 (0.8), p < 0.05 VAS (0-10), mean change (SD), p: • 2.3 (1.7 to 3.0) years: 5.1 (1.2) vs 2.3 (0.4), p < 0.05		
Ahmad 2016	N= 40 Age, mean (range): 40.5 (14-63) years Male: 58% F/U: 2 weeks, 6 weeks, 3 months, 6 months, 1 year, final F/U mean of 3.2 years Lesion size, mean: 1.6 cm^2 Lesion description: recurrent or large osteochondral lesions of the talar dome (OLT) Area, n/N (%) anterior or central, n/N (%) posteromedial: 19/36 (53%) vs 17/36 (47%) <u>OAT (Autograft)</u> Open ankle arthrotomy with or without malleolar osteotomy was done. Osteochondral autografts from the extra-articular superolateral distal femoral condyle were transplanted into defects using "press-fit" method <u>Allograft</u> Open ankle arthrotomy with or without malleolar osteotomy was done. Fresh talar allografts were	<u>Function: OAT vs allograft</u> FAAM, mean (range): • 3.2 (1 to 6.4) years: 85.5 (56 to 97.6) vs 80.7 (56 to 95.2), p = 0.25 <u>Pain: OAT vs allograft</u> VAS Pain: • 3.2 (1 to 6.4) years: 2.2 (0 to 8) vs 2.7 (1 to 8), p = 0.15 <u>Safety/complications: OAT vs allograft§</u> Revision operative procedure, n/N (%), (RR, 95% CI), p: • 3.2 (1 to 6.4) years: 2/20 (10%) vs 2/16 (13%), (RR 0.80, 95% CI 0.13 to 5.1), p = 0.81 Graft nonunion, n/N (%), (RR, 95% CI), p: • 3.2 (1 to 6.4) years: 2/20 (10%) vs 3/16 (19%) (RR 0.53, 95% CI 0.10 to 2.8), p = 0.46	**Function:** There was no difference between OAT and allograft in FAAM at a mean follow-up of 3.2 years **Pain:** There was no difference between OAT and allograft in pain VAS at a mean follow-up of 3.2 years **Safety and complications:** There was no difference between OAT and allograft in graft nonunion or in revision operative procedures at a mean follow-up of 3.2 years	Authors declare no conflict of interest No external funding reported 4 patients in allograft group were excluded after randomization for having OLTs with significant involvement of either medial or lateral shoulder of the talar dome. Patients were treated with hemi-talus allograft.

Author (Year)	Demographics	Results	Conclusions	Comments
Knee: Autograft vs. MF or Debridement alone				
	harvested and transplanted into defects using "press-fit" method			
KNEE: Autograft vs. ACI				
Lim 2012	N= 109 patients (120 knees) randomized, 69 patients (70 knees) evaluated Age, mean (range): 28.5 (18-42) years Male: 57% F/U, mean (range): 5.7 (3 to 10.5) years Lesion size: 2.74 cm^2 Lesion description: single symptomatic articular cartilage lesion of the knee Area, n/N (%) medial, n/N (%) lateral: 55/70 (79%) vs 15/70 (21%) <u>OAT (Autograft mosaicplasty) (n=22 knees)</u> Performed after arthroscopic examination and debridement of fibrillated cartilage. Plugs of 4, 6, and 8 mm were inserted using press-fit method <u>ACI (n=18 knees)</u> First stage was arthroscopic harvest of 1 cm by 1 cm fragments from the margin of the trochlea. Fragment underwent enzymic digestion to release cells for culture. Six weeks later, arthrotomy procedure was done to place periosteal flap, harvested from the tibia, over defect, fixed with sutures, and sealed with fibrin glue. Solution of expanded chondrocytes was injected underneath flap.	<u>Function: OAT vs ACI</u> Lysholm, mean (SD), p: • 5 years: 84.8 (5.5) vs 84.6 (6.1), p NS HSS, mean (SD), p: • 5 years: 88.12 (4.15) vs 82.51 (4.58), p NS <u>Return to activity: OAT vs ACI</u> Tenger, mean (SD), p: • 5 years: 5.3 (1.2) vs 5.2 (1.3), p NS <u>Reoperation: OAT vs ACI§</u> Reoperation, n/N (%), (RR, 95% CI), p: 1/22 (4%) vs 2/18 (11%), (RR 0.41, 95% CI 0.04 to 4.2), p = 0.44	**Function:** The authors reported no differences in Lysholm or HSS scores at a follow-up up to a mean of 5.7 years **Return to activity:** The authors reported no differences Tegner at a follow-up up to a mean of 5.7 years **Reoperation:** There was no difference in rates of reoperation between OAT and ACI	109 patients enrolled, only 69 underwent procedures Authors declare no conflicts of interest Funding NR

Author (Year)	Demographics	Results	Conclusions	Comments
Knee: Autograft vs. MF or Debridement alone				
Bentley 2012 (follow-up to Bentley 2003)	N= 100 Age, mean (range): 31.3 (16 to 49) years % Male: 57% F/U: ≥10 years Lesion size, mean: 4.2 cm^2 Lesion description: symptomatic articular cartilage defect of the knee Area, n/N (%) medial, n/N (%) lateral, n/N (%) patella, n/N (%) other/unknown: 17/100 (17%), 50/100 (50%), 24/100 (24%), 9/100 (9%) <u>OAT (Autograft, mosaicplasty)</u> Parapatelar arthrotomy was done. After defect was debrided, 4.5 mm grafts were harvested from the margin of the trochlea and transplanted into the defect. <u>ACI</u> Biopsy of articular cartilage was harvested from the margin of the trochlear. Three to five weeks after enzymatic digestion, parapatellar arthrotomy was performed. The defect was debrided and covered with the cells at 3 to 4 mm intervals.	<u>Function: OAT vs ACI</u> Modified Cincinnati score, n/N (%): • Excellent (80-100): 4/15 (27%) vs 28/48 (58%) • Good (55-79): 5/15 (33%) vs 7/48 (15%) • Fair (30-54): 4/15 (27%) vs 6/48 (13%) • Poor (<30): 2/15 (13%) vs 2/48 (4%) • p-value: 0.02 Stanmore Bentley score, n/N (%): • Score of 0: 2/15 (13%) vs 7/48 (15%) • Score of 1: 4/15 (27%) vs 23/48 (48%) • Score of 2: 5/15 (33%) vs 3/48 (6%) • Score of 3: 2/15 (13%) vs 6/48 (13%) • Score of 4: 2/15 (13%) vs 4/48 (8%) • p-value: 0.27 <u>Failure of operation: OAT vs ACI</u> Failed cartilage repairs, n/N (%), p: 23/42 (55%) vs 10/58 (17%), p < 0.0001	**Function:** At a minimum of a 10 year follow-up, ACI demonstrated statistically significant better results than OAT in the modified Cincinnati score, while results of the Stanmore-Bentley functional rating showed no difference. **Failure of operation:** ACI showed statistically significant lower rates of failed cartilage repair at a minimum of 10 years follow-up.	Only 15 of 42 patients in the OAT group were evaluated for functional outcomes at the 10 year follow-up, compared to 48 of 58 patients in the ACI group. Authors declare no conflict of interest Funding NR

ACI: Autologous Chondrocyte Implantation; ACD: articular cartilage defect; AOFAS: American Orthopedic Foot and Ankle Society score; CI: confidence interval; FAAM: Foot and Ankle Ability Measure Sports scoring system; F/U: follow-up; HSS: Hospital for Special Surgery score; ICRS: International Cartilage Repair Society score; IQR: interquartile range; KOOS: Knee Injury and Osteoarthritis Outcome Score; MD: mean difference; MF: microfracture; NR: not reported; OAT: osteochondral autograft transplantation; OCD: osteochondral defect; RR: risk ratio; VAS: visual analog scale

*Estimated from graph

†Trial population included three groups; OAT, MF, and ACI. Comparison between OAT and ACI is included in corresponding section

‡MDs, CIs, and p values calculated by AAI

§RRs, CIs, and p values calculated by AAI

**Population was exclusively athletes

APPENDIX C. SYSTEMATIC REVIEWS EXCLUDED AT FULL TEST REVIEW

Excluded systematic reviews

Citation	Reason for exclusion
Bexkens R, Ogink PT, Doornberg JN, et al. Donor-site morbidity after osteochondral autologous transplantation for osteochondritis dissecans of the capitellum: a systematic review and meta-analysis. Knee Surg Sports Traumatol Arthrosc 2017;25:2237-46.	Evaluated chondral lesions of the elbow; elbow was not a region of interest
Camp CL, Stuart MJ, Krych AJ. Current concepts of articular cartilage restoration techniques in the knee. Sports Health 2014;6:265-73.	No quantitative synthesis
Chalmers PN, Vigneswaran H, Harris JD, Cole BJ. Activity-Related Outcomes of Articular Cartilage Surgery: A Systematic Review. Cartilage 2013;4:193-203.	No new RCTs included
Chawla A, Twycross-Lewis R, Maffulli N. Microfracture produces inferior outcomes to other cartilage repair techniques in chondral injuries in the paediatric knee. Br Med Bull 2015;116:93-103.	No RCTs included
Devitt BM, Bell SW, Webster KE, Feller JA, Whitehead TS. Surgical treatments of cartilage defects of the knee: Systematic review of randomised controlled trials. Knee 2017;24:508-17.	No quantitative synthesis
Haien Z, Jiachang W, Qiang L, Yufeng M, Zhenwei J. Osteochondral Autologous Transplantation Compared to Microfracture for Treating Osteochondral Defect: An Updated Meta-analysis of Randomized Controlled Trials. J Knee Surg 2017.	Lower quality review and substantial cross-over with included trials
Li Z, Zhu T, Fan W. Osteochondral autograft transplantation or autologous chondrocyte implantation for large cartilage defects of the knee: a meta-analysis. Cell Tissue Bank 2016;17:59-67.	Lower quality review and substantial cross-over with included trials
Lynch TS, Patel RM, Benedick A, Amin NH, Jones MH, Miniaci A. Systematic review of autogenous osteochondral transplant outcomes. Arthroscopy 2015;31:746-54.	Lower quality review and substantial cross-over with included trials
Mundi R, Bedi A, Chow L, et al. Cartilage Restoration of the Knee: A Systematic Review and Meta-analysis of Level 1 Studies. Am J Sports Med 2016;44:1888-95.	No new RCTs included
Naveen S, Robson N, Kamarul T. Comparative analysis of autologous chondrocyte implantation and other treatment modalities: A systematic review. European Journal of Orthopaedic Surgery and Traumatology 2012;22:89-96.	No new RCTs included
Riboh JC, Cvetanovich GL, Cole BJ, Yanke AB. Comparative efficacy of cartilage repair procedures in the knee: a network meta-analysis. Knee Surg Sports Traumatol Arthrosc 2017;25:3786-99.	Not a meta-analysis head-to-head of trials, network meta-analysis
Richter DL, Schenck RC, Jr., Wascher DC, Treme G. Knee Articular Cartilage Repair and Restoration Techniques: A Review of the Literature. Sports Health 2016;8:153-60.	Lower quality review and substantial cross-over with included trials
Smith MV, Bedi A, Chen NC. Surgical treatment for osteochondritis dissecans of the capitellum. Sports Health 2012;4:425-32.	No RCTs included
Westermann RW, Hancock KJ, Buckwalter JA, Kopp B, Glass N, Wolf BR. Return to Sport After Operative Management of Osteochondritis Dissecans of the Capitellum: A Systematic Review and Meta-analysis. Orthop J Sports Med 2016;4:2325967116654651.	No RCTs included

Excluded randomized controlled trials

Citation	Reason for exclusion
Clave A, Potel JF, Servien E, Neyret P, Dubrana F, Stindel E. Third-generation autologous chondrocyte implantation versus mosaicplasty for knee cartilage injury: 2-year randomized trial. Journal of orthopaedic research : official publication of the Orthopaedic Research Society 2016;34:658-65.	Product used ACI intervention not FDA approved

ACI: Autologous Chondrocyte Implantation; FDA: Food and Drug Administration