RACGP

Royal Australian College *of* General Practitioners

Guideline for the management of knee and hip osteoarthritis

Technical document

racgp.org.au

Healthy Profession.
Healthy Australia.

Guideline for the management of knee and hip osteoarthritis: Technical document

Disclaimer

We acknowledge the Traditional Custodians of the lands and seas on which we work and live, and pay our respects to Elders, past, present and future.

RACGP

Royal Australian College *of* General Practitioners

Guideline for the management of knee and hip osteoarthritis

Technical document

Contents

Appendix 1. Search strategy

1.1 PubMed

Systematic review OR meta-analysis of OA of hip OR knee	(("Osteoarthritis, Hip/diet therapy"[Mesh] OR "Osteoarthritis, Hip/drug therapy"[Mesh] OR "Osteoarthritis, Hip/nursing"[Mesh] OR "Osteoarthritis, Hip/psychology"[Mesh] OR "Osteoarthritis, Hip/radiotherapy"[Mesh] OR "Osteoarthritis, Hip/rehabilitation"[Mesh] OR "Osteoarthritis, Hip/therapy"[Mesh])) OR ("Osteoarthritis, Knee/diet therapy"[Mesh] OR "Osteoarthritis, Knee/drug therapy"[Mesh] OR "Osteoarthritis, Knee/nursing"[Mesh] OR "Osteoarthritis, Knee/psychology"[Mesh] OR "Osteoarthritis, Knee/radiotherapy"[Mesh] OR "Osteoarthritis, Knee/rehabilitation"[Mesh] OR "Osteoarthritis, Knee/therapy"[Mesh]) Filters: Meta-Analysis; Systematic Reviews; Publication date from 2005/01/01 to 2016/12/31; Humans
RCT OA hip or knee	Search (("Osteoarthritis, Hip/diet therapy"[Mesh] OR "Osteoarthritis, Hip/drug therapy"[Mesh] OR "Osteoarthritis, Hip/nursing"[Mesh] OR "Osteoarthritis, Hip/psychology"[Mesh] OR "Osteoarthritis, Hip/radiotherapy"[Mesh] OR "Osteoarthritis, Hip/rehabilitation"[Mesh] OR "Osteoarthritis, Hip/therapy"[Mesh])) OR ("Osteoarthritis, Knee/diet therapy"[Mesh] OR "Osteoarthritis, Knee/drug therapy"[Mesh] OR "Osteoarthritis, Knee/nursing"[Mesh] OR "Osteoarthritis, Knee/psychology"[Mesh] OR "Osteoarthritis, Knee/radiotherapy"[Mesh] OR "Osteoarthritis, Knee/rehabilitation"[Mesh] OR "Osteoarthritis, Knee/therapy"[Mesh]) Filters: Randomized Controlled Trial; Publication date from 2005/01/01 to 2016/12/31; Humans
Systematic review OR meta-analysis for arthroscopy in OA of hip or knee	Search ((((("Osteoarthritis, Knee"[Majr] OR "Osteoarthritis, Hip"[Majr])) OR ((("Osteoarthritis, Hip/diet therapy"[Mesh] OR "Osteoarthritis, Hip/drug therapy"[Mesh] OR "Osteoarthritis, Hip/nursing"[Mesh] OR "Osteoarthritis, Hip/psychology"[Mesh] OR "Osteoarthritis, Hip/radiotherapy"[Mesh] OR "Osteoarthritis, Hip/rehabilitation"[Mesh] OR "Osteoarthritis, Hip/therapy"[Mesh])) OR ("Osteoarthritis, Knee/diet therapy"[Mesh] OR "Osteoarthritis, Knee/drug therapy"[Mesh] OR "Osteoarthritis, Knee/nursing"[Mesh] OR "Osteoarthritis, Knee/psychology"[Mesh] OR "Osteoarthritis, Knee/radiotherapy"[Mesh] OR "Osteoarthritis, Knee/rehabilitation"[Mesh] OR "Osteoarthritis, Knee/therapy"[Mesh])))) AND "Arthroscopy"[Mesh] Filters: Meta-Analysis; Systematic Reviews; Publication date from 2005/01/01 to 2016/12/31; Humans;
RCT for arthroscopy of OA hip or knee	Search (((((("Osteoarthritis, Knee"[Majr] OR "Osteoarthritis, Hip"[Majr])) OR ((("Osteoarthritis, Hip/diet therapy"[Mesh] OR "Osteoarthritis, Hip/drug therapy"[Mesh] OR "Osteoarthritis, Hip/nursing"[Mesh] OR "Osteoarthritis, Hip/psychology"[Mesh] OR "Osteoarthritis, Hip/radiotherapy"[Mesh] OR "Osteoarthritis, Hip/rehabilitation"[Mesh] OR "Osteoarthritis, Hip/therapy"[Mesh])) OR ("Osteoarthritis, Knee/diet therapy"[Mesh] OR "Osteoarthritis, Knee/drug therapy"[Mesh] OR "Osteoarthritis, Knee/nursing"[Mesh] OR "Osteoarthritis, Knee/psychology"[Mesh] OR "Osteoarthritis, Knee/radiotherapy"[Mesh] OR "Osteoarthritis, Knee/rehabilitation"[Mesh] OR "Osteoarthritis, Knee/therapy"[Mesh])))) AND "Arthroscopy"[Mesh]) Filters: Randomized Controlled Trial; Publication

	date from 2015/01/01 to 2016/12/31; Humans;

1.2 CINAHL

Systematic review OR meta-analysis of OA of hip OR knee	SU osteoarthritis, hip OR SU osteoarthritis, knee Limiters – Published Date: 20050101-20161231; Peer Reviewed; Clinical Queries: Therapy – Best Balance; Human; Publication Type: Meta Analysis, Systematic Review
RCT OA hip or knee	SU osteoarthritis, hip OR SU osteoarthritis, knee Limiters Published Date: 20150101-20161231; Peer Reviewed; Clinical Queries: Therapy – Best Balance; Randomized Controlled Trials; Publication Type: Randomized Controlled Trial
Systematic review OR meta-analysis for arthroscopy in OA of hip or knee	SU osteoarthritis, hip OR SU osteoarthritis, knee AND SU arthroscopy OR TI arthoscop* Limiters – Published Date: 20050101-20171231; Peer Reviewed; English Language; Clinical Queries: Therapy – Best Balance; Human; Publication Type: Meta Analysis, Systematic Review
RCT for Arthroscopy of OA hip or knee	SU osteoarthritis, hip OR SU osteoarthritis, knee AND SU arthroscopy OR TI arthoscop* Limiters – Published Date: 20150101-20171231; Peer Reviewed; English Language; Clinical Queries: Therapy – Best Balance; Human; Randomized Controlled Trials; Publication Type: Randomized Controlled Trial

1.3 Cochrane Library

Systematic review OR meta-analysis of OA of hip OR knee	MeSH descriptor: [Osteoarthritis, Hip] explode all trees OR MeSH descriptor: [Osteoarthritis, Knee] explode all trees and limited to 2005 – 2016
Systematic review OR meta-analysis for arthroscopy in OA of hip or knee	Above search AND MeSH descriptor: [Arthroscopy] explode all trees

Appendix 2. Study selection and criteria

2.1 Inclusion/exclusion criteria

To be included, an article had to meet the following selection criteria:

- Study was of osteoarthritis of the hip or knee
- Study was performed in humans
- Study design was a systematic review, a meta-analysis, or a randomized controlled trial.
- At least 80% of the enrolled study population were 45 years of age or older
- Study results included outcomes of interest (Table 2.1)

The following publications were excluded:

- Animal or laboratory studies
- Studies performed on cadavers
- Non-randomised trials, observational studies, case series and case reports, analyses of medical records, narrative reviews, editorials, letters, and commentaries
- Studies that did not report on any outcomes of interest

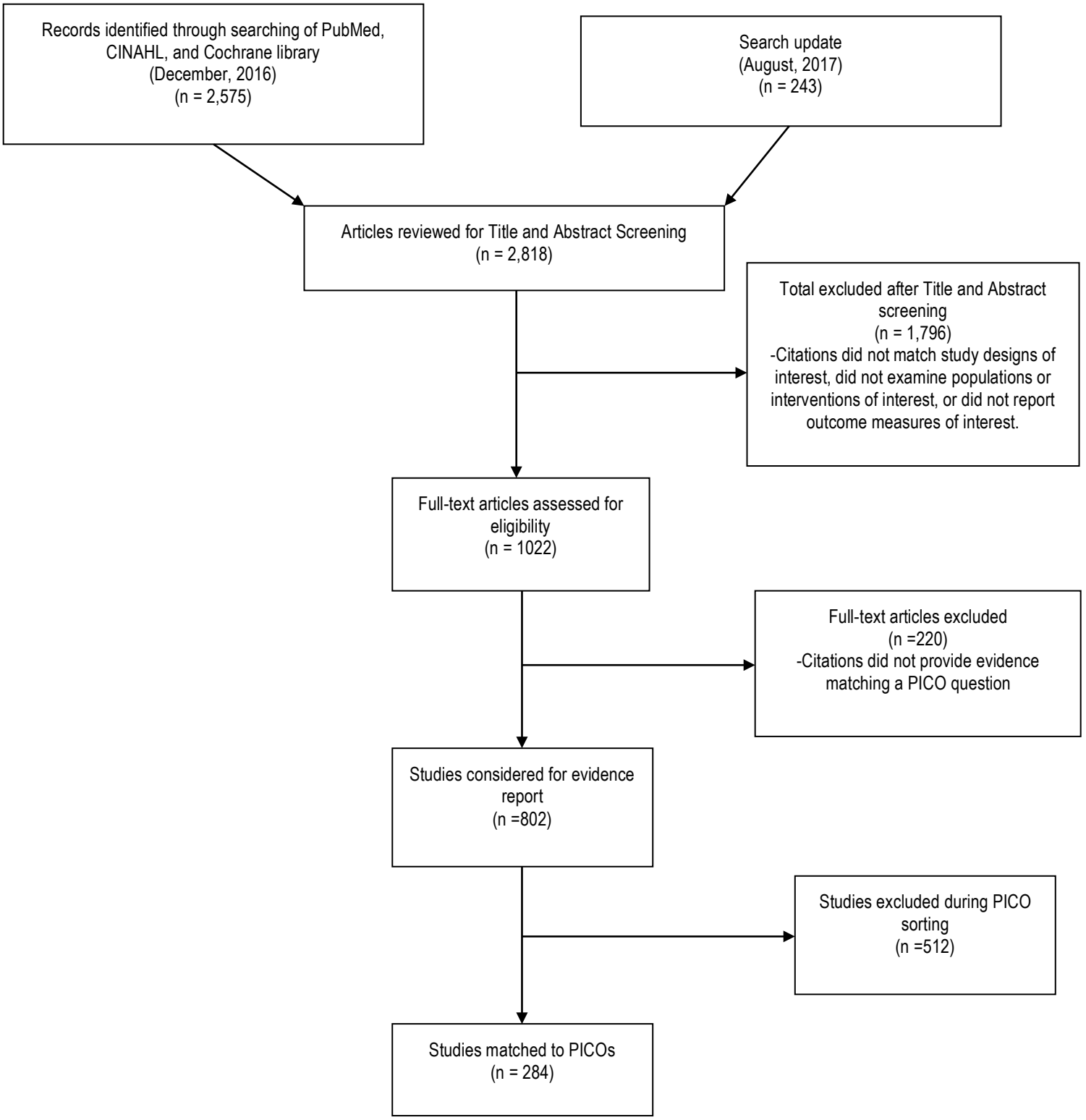
Table 2.1 Outcomes of Interest

Outcome Name	Outcome Type*	Acceptable Measures, Definitions	GRADE Designation
Pain	Primary: Continuous	WOMAC Pain, NRS Pain, VAS Pain, KOOS Pain, HOOS Pain, AIMS Pain, BPI, any other validated scale	CRITICAL
Function	Primary: Continuous	WOMAC Function, Lequesne Index, KOOS-ADL, HOOS-ADL, AIMS Function, any other validated scale	CRITICAL
Quality of Life	Primary: Continuous	SF-36, SF-12, EQ-5D, KOOS-QoL, HOOS-QoL, HAQ, any other validated scale	IMPORTANT
Structural Progression	Primary: Continuous	JSN (mm), JSW (mm), cartilage thickness (mm), cartilage volume (mm^3)	IMPORTANT
Structural Progression	Primary: Dichotomous	N experiencing Radiographic Progression	IMPORTANT
Withdrawals due to Adverse Events	Primary: Dichotomous	N withdrawing from study due to one or more Adverse Event(s)	CRITICAL
Total Adverse Events	Primary: Dichotomous	N experiencing one or more Adverse Event(s)	CRITICAL
Serious Adverse Events	Primary: Dichotomous	N experiencing one or more Serious Adverse Event(s)	CRITICAL
Treatment-specified Harms	Primary: Dichotomous	N experiencing one or more Adverse Event(s) specific to the intervention of interest	IMPORTANT
Self-Efficacy	Secondary: Continuous	ASES, any other validated scale	IMPORTANT
Depression	Secondary: Continuous	BDI, AIMS Psychological Disability, any other validated scale	IMPORTANT
Percentage Weight Loss	Secondary: Continuous	Mean percentage of baseline weight lost	IMPORTANT
Non-Adherence	Secondary: Dichotomous	N who left study or declined to participate due to lack of interest or who were classified as “non-adherent” or “non-compliant”	IMPORTANT
Opioid Withdrawal	Secondary: Dichotomous	N classified by practitioner as experiencing opioid withdrawal as measured by any validated instrument	IMPORTANT

*Primary outcomes were those that were selected *a priori* by the Working Group. Secondary outcomes were selected during the data extraction process to accommodate specific PICO questions.

WOMAC: Western Ontario & McMaster universities Osteoarthritis Index; NRS: Numerical Rating Scale; VAS: Visual Analog Scale; KOOS: Knee Injury and Osteoarthritis Outcome Score; HOOS: Hip Disability and Osteoarthritis Outcome Score; AIMS: Arthritis Impact Measurement Scale; BPI: Brief Pain Inventory; ADL: Activities of Daily Living; SF-36/SF-12: Medical Outcomes Study Short-Form Health Survey; EQ-5D: European 5 Dimensional Quality of Life measure; QoL: Quality of Life; HAQ: Health Assessment Questionnaire; JSN: Joint Space Narrowing; JSW: Joint Space Widening; N: Number of patients; ASES: Arthritis Self-Efficacy Scale; BDI: Beck Depression Inventory

Figure 1: Flowchart of the trial selection process



Appendix 3. Quality assessment and evidence report formulation

Table 3.1 GRADE Quality Assessment Rubric

Quality Assessment Domain	Standard Downgrade	*A priori* Domain Cutoffs
Risk of Bias	**"Serious"**= -1 **"Very Serious"**= -2	**"Serious":** ≥50% of trials received overall (≥1 out of 7 dimensions in the Cochrane Risk of Bias tool) "High" risk of bias ratings (-1); 25%-50% of trials received overall (≥4 out of 7 dimensions in the Cochrane Risk of Bias tool) "Unclear" risk of bias ratings (-1) **"Very Serious":** >25% of trials received multiple (≥2 out of 7 dimensions in the Cochrane Risk of Bias tool) "High" risk of bias ratings (-2); >50% of trials received overall (≥4 out of 7 dimensions in the Cochrane Risk of Bias tool) "Unclear" risk of bias ratings (-2)
Inconsistency	**"Serious"**= -1 **"Very Serious"**= -2	**"Serious":** I^2 >50% and ≤75%; "moderate heterogeneity" (-1) **"Very Serious":** I^2 >75%; "high heterogeneity" (-2)
Indirectness	**"Serious"**= -1 **"Very Serious"**= -2	**"Serious":** Indirectness present in one of the four key extraction categories- Population, Intervention, Comparator, Outcome (-1) **"Very Serious":** Indirectness present in more than one of the four key extraction categories- Population, Intervention, Comparator, Outcome (-2)
Imprecision	**"Serious"**= -1 **"Very Serious"**= -2	**"Serious":** 95% Confidence Interval crosses null (-1); sample size in one study arm <50 (-1) **"Very Serious":** Total sample size ≤30 (-2)
I^2= measure of heterogeneity, with 100% being the maximum possible heterogeneity		

Table 3.2 Recommendations considered but excluded

(relating to PICO questions that were considered and excluded)

PICO (knee): What are the benefits and harms of **decision aids** in the management of patients with knee OA?
PICO (hip): What are the benefits and harms of **decision aids** in the management of patients with hip OA?
PICO (knee): What are the benefits and harms of **balneotherapy** in the management of patients with knee OA?
PICO (hip): What are the benefits and harms of **balneotherapy** in the management of patients with hip OA?
PICO (hip): What are the benefits and harms of **foot orthotics** in the management of patients with hip OA?
PICO (hip): What are the benefits and harms of **hip orthotics** in the management of patients with hip OA?
PICO (hip): What are the benefits and harms of **taping** in the management of patients with hip OA?
PICO (knee): What are the benefits and harms of **radiotherapy** in the management of patients with knee OA?
PICO (hip): What are the benefits and harms of **radiotherapy** in the management of patients with hip OA?
PICO (knee): What are the benefits and harms of **statins** in the management of patients with knee OA?
PICO (hip): What are the benefits and harms of **statins** in the management of patients with hip OA?
PICO (knee): What are the benefits and harms of **bone marrow aspirate concentrate** in the management of patients with knee OA?
PICO (hip): What are the benefits and harms of **arthroscopic lavage and debridement** interventions in the management of patients with hip OA?
PICO (hip): What are the benefits and harms of **arthroscopic procedures** for cartilage repair interventions in the management of patients with hip OA?

Appendix 4. Recommendation support surveys

Working Group members indicated their extent of support for the 128 initial recommendations through an online survey voting process.

Initial voting process
Completed by ten members
Rated on an 11-point numerical scale
70% consensus agreement threshold set for accepting recommendation
Recommendations with consensus agreement
≥70% – 21 recommendations
≥80% – 41 recommendations
100% – 41 recommendations
25 recommendations did not achieve consensus agreement
57% – 9 recommendations
67% – 4 recommendations
43% – 6 recommendations
29% – 4 recommendations
14% – 2 recommendations
Further discussion was convened and recommendations which did not achieve consensus agreement were re-drafted
Second (Final) voting process
Completed by ten members
Rated with an agree/disagree response
70% consensus agreement threshold set for accepting recommendation
All recommendations received ≥70% consensus agreement
Complete voting details on specific recommendations are available on request.

Appendix 5. GRADE evidence report

Evidence Report: Pharmacologic and Non-Pharmacologic Management of Knee and Hip Osteoarthritis

June, 2017

Prepared for:
The Royal Australian College of General Practitioners
Melbourne, Victoria, Australia

Prepared by:
Center for Treatment Comparison and Integrative Analysis (CTCIA)
Tufts Medical Center
Boston, MA, USA

Investigators:
Mikala C. Osani
Mia-Cara Musetti
Elizaveta E. Vaysbrot
Timothy E. McAlindon
Raveendhara R. Bannuru

Table of Contents

PICO questions for knee osteoarthritis (79 PICOs)

1. Non-Pharmacological Interventions for Knee Osteoarthritis (37 PICOs)

1.1 What are the benefits and harms of **self-management education programmes** in the management of patients with knee OA?

1.2 What are the benefits and harms of **decision aids** in the management of patients with knee OA? – excluded

1.3 What are the benefits and harms of **cognitive behavioural therapy** in the management of patients with knee OA?

1.4 What are the benefits and harms of **all land-based exercise** in the management of patients with knee OA?

1.5 **Specific form of land-based exercise**

1.5.1 What are the benefits and harms of **muscle strengthening** in the management of patients with knee OA?

1.5.2 What are the benefits and harms of **walking** in the management of patients with knee OA?

1.5.3 What are the benefits and harms of **stationary cycling** in the management of patients with knee OA?

1.5.4 What are the benefits and harms of **Tai Chi** in the management of patients with knee OA?

1.5.5 What are the benefits and harms of **Hatha yoga** in the management of patients with knee OA?

1.6 What are the benefits and harms of **aquatic exercise/hydrotherapy** in the management of patients with knee OA?

1.7 **Manual therapy**

1.7.1 What are the benefits and harms of **massage** in the management of patients with knee OA?

1.7.2 What are the benefits and harms of **manipulation and mobilisation** in the management of patients with knee OA?

1.8 What are the benefits and harms of **weight management** in the management of patients with knee OA?

1.9 **Thermotherapy**

1.9.1 What are the benefits and harms of **local hot application** in the management of patients with knee OA?

1.9.2 What are the benefits and harms of **local cold application** in the management of patients with knee OA?

1.10 **Orthotic braces**

1.10.1 What are the benefits and harms of **varus unloading/re-alignment braces** in the management of patients with knee OA?

1.10.2 What are the benefits and harms of **valgus unloading/re-alignment braces** in the management of patients with knee OA?

1.10.3 What are the benefits and harms of **realigning patellofemoral braces** in the management of patients with knee OA?

1.11 **Insoles**

1.11.1 What are the benefits and harms of **medial wedged insoles** in the management of patients with knee OA?

1.11.2 What are the benefits and harms of **lateral wedge insoles** in the management of patients with knee OA?

1.11.3 What are the benefits and harms of **shock absorbing insoles** in the management of patients with knee OA?

1.11.4 What are the benefits and harms of **arch supports** in the management of patients with knee OA?

1.12 **Shoes**
1.12.1 What are the benefits and harms of **unloading shoes** in the management of patients with knee OA?
1.12.2 What are the benefits and harms of **minimalist footwear** in the management of patients with knee OA?
1.12.3 What are the benefits and harms of **rocker sole shoes** in the management of patients with knee OA?

1.13 **Taping**
1.13.1 What are the benefits and harms of **kinesio taping** in the management of patients with knee OA?
1.13.2 What are the benefits and harms of **patellar taping** in the management of patients with knee OA?

1.14 What are the benefits and harms of **walking cane/stick** in the management of patients with knee OA?

1.15 **Electromagnetic therapy**
1.15.1 What are the benefits and harms of **pulsed electromagnetic/shortwave therapy** in the management of patients with knee OA?
1.15.2 What are the benefits and harms of **shockwave therapy** in the management of patients with knee OA?

1.16 **Electrical stimulation**
1.16.1 What are the benefits and harms of **transcutaneous electrical nerve stimulation (TENS)** in the management of patients with knee OA?
1.16.2 What are the benefits and harms of **inferential currents** in the management of patients with knee OA?

1.17 What are the benefits and harms of **ultrasound** in the management of patients with knee OA?

1.18 What are the benefits and harms of **laser** in the management of patients with knee OA?

1.19 **Acupuncture**
1.19.1 What are the benefits and harms of **traditional acupuncture with manual stimulation** in the management of patients with knee OA?
1.19.2 What are the benefits and harms of **laser acupuncture** in the management of patients with knee OA?
1.19.3 What are the benefits and harms of **electroacupuncture** in the management of patients with knee OA?

2. Pharmacological Interventions for Knee Osteoarthritis (35 PICOs)

2.1 **Oral analgesics**
2.1.1 What are the benefits and harms of **oral paracetamol** in the management of patients with knee OA?
2.1.2 What are the benefits and harms of **oral NSAIDs including COX-2 inhibitors** in the management of patients with knee OA?
2.1.3 What are the benefits and harms of **oral opioids** in the management of patients with knee OA?

2.2 **Topical analgesics**
2.2.1 What are the benefits and harms of **topical NSAIDs** in the management of patients with knee OA?
2.2.2 What are the benefits and harms of **transdermal opioids** in the management of patients with knee OA?
2.2.3 What are the benefits and harms of **topical capsaicin** in the management of patients with knee OA?

2.3 **Herbal therapies**
2.3.1 What are the benefits and harms of **avocado soybean unsaponifiable** in the management of patients with knee OA?
2.3.2 What are the benefits and harms of **boswellia serrata** in the management of patients with knee OA?
2.3.3 What are the benefits and harms of **curcuma** in the management of patients with knee OA?
2.3.4 What are the benefits and harms of **pine bark extract** in the management of patients with knee OA?

2.4 **Nutraceuticals**

2.4.1 What are the benefits and harms of **glucosamine** in the management of patients with knee OA?

2.4.2 What are the benefits and harms of **chondroitin** in the management of patients with knee OA?

2.4.3 What are the benefits and harms of **glucosamine and chondroitin in compound form** in the management of patients with knee OA?

2.4.4 What are the benefits and harms of **vitamin D** in the management of patients with knee OA?

2.4.5 What are the benefits and harms of **(omega-3/6) poly-unsaturated fatty acids** in the management of patients with knee OA?

2.4.6 What are the benefits and harms of **collagen preparations** in the management of patients with knee OA?

2.4.7 What are the benefits and harms of **methylsulfonylmethane** in the management of patients with knee OA?

2.4.8 What are the benefits and harms of **diacerein** in the management of patients with knee OA?

2.5 What are the benefits and harms of **duloxetine** in the management of patients with knee OA?

2.6 What are the benefits and harms of **doxycycline** in the management of patients with knee OA?

2.7 **Anti-osteoporosis (anti-resorptive bone-acting) drugs**

2.7.1 What are the benefits and harms of **bisphosphonates** in the management of patients with knee OA?

2.7.2 What are the benefits and harms of **calcitonin** in the management of patients with knee OA?

2.7.3 What are the benefits and harms of **strontium ranelate** in the management of patients with knee OA?

2.8 **Investigational DMOADs (symptomatic or inflammatory modification)**

2.8.1 What are the benefits and harms of **IL-1 inhibitors** in the management of patients with knee OA?

2.8.2 What are the benefits and harms of **TNF-alpha inhibitors** in the management of patients with knee OA?

2.8.3 What are the benefits and harms of **anti-nerve growth factor (NGF) therapy** in the management of patients with knee OA?

2.8.4 What are the benefits and harms of **fibroblast growth factor (FGF) therapy** in the management of patients with knee OA?

2.8.5 What are the benefits and harms of **colchicine** in the management of patients with knee OA?

2.8.6 What are the benefits and harms of **methotrexate** in the management of patients with knee OA?

2.8.7 What are the benefits and harms of **statins** in the management of patients with knee OA? – excluded

2.9 **Intra-articular injections**

2.9.1 What are the benefits and harms of **corticosteroids** in the management of patients with knee OA?

2.9.2 What are the benefits and harms of **viscosupplementation** in the management of patients with knee OA?

2.9.3 What are the benefits and harms of **platelet-rich plasma** in the management of patients with knee OA?

2.9.4 What are the benefits and harms of **stem cell therapy** in the management of patients with knee OA?

2.9.5 What are the benefits and harms of **dextrose prolotherapy** in the management of patients with knee OA?

3. Surgical Interventions (non-arthroplasty) (3 PICOs)

3.1 What are the benefits and harms of **arthroscopic lavage and debridement** interventions in the management of patients with knee OA?

3.2 What are the benefits and harms of **arthroscopic meniscectomy** interventions in the management of patients with knee OA?

3.3 What are the benefits and harms of **arthroscopic procedures for cartilage repair** interventions in the management of patients with knee OA?

4. Combination Therapies (4 PICOs)

4.1 What are the benefits and harms of **combination weight management and exercise interventions** compared to **exercise** in patients with knee OA?

4.2 What are the benefits and harms of **combination weight management and exercise interventions** compared to **weight management** in patients with knee OA?

4.3 What are the benefits and harms of **combination exercise and cognitive behavioural interventions** compared to **exercise** in patients with knee OA?

4.4 What are the benefits and harms of **combination exercise and cognitive behavioural interventions** compared to **cognitive behavioural interventions** in patients with knee OA?

PICO questions for hip osteoarthritis (66 PICOs)

1. Non-Pharmacological Interventions for Hip Osteoarthritis (25 PICOs)

1.1 What are the benefits and harms of **self-management education programmes** in the management of patients with hip OA?

1.2 What are the benefits and harms of **decision aids** in the management of patients with hip OA? – excluded

1.3 What are the benefits and harms of **cognitive behavioural therapy** in the management of patients with hip OA?

1.4 What are the benefits and harms of **all land-based exercise** in the management of patients with hip OA?

1.5 **Specific form of land-based exercise**

1.5.1 What are the benefits and harms of **muscle strengthening** in the management of patients with hip OA?

1.5.2 What are the benefits and harms of **walking** in the management of patients with hip OA?

1.5.3 What are the benefits and harms of **stationary cycling** in the management of patients with hip OA?

1.5.4 What are the benefits and harms of **Tai Chi** in the management of patients with hip OA?

1.5.5 What are the benefits and harms of **Hatha yoga** in the management of patients with hip OA?

1.6 What are the benefits and harms of **aquatic exercise/hydrotherapy** in the management of patients with hip OA?

1.7 **Manual therapy**

1.7.1 What are the benefits and harms of **massage** in the management of patients with hip OA?

1.7.2 What are the benefits and harms of **manipulation and mobilisation** in the management of patients with hip OA?

1.8 What are the benefits and harms of **weight management** in the management of patients with hip OA?

1.9 **Thermotherapy**

1.9.1 What are the benefits and harms of **local hot application** in the management of patients with hip OA?

1.9.2 What are the benefits and harms of **local cold application** in the management of patients with hip OA?

1.10 What are the benefits and harms of **hip orthotics** in the management of patients with hip OA?

1.11 What are the benefits and harms of **kinesio taping** in the management of patients with hip OA?

1.12 What are the benefits and harms of **walking cane/stick** in the management of patients with hip OA?

1.13 **Electromagnetic therapy**

1.13.1 What are the benefits and harms of **pulsed electromagnetic/shortwave therapy** in the management of patients with hip OA?

1.13.2 What are the benefits and harms of **shockwave therapy** in the management of patients with hip OA?

1.14 **Electrical stimulation**

1.14.1 What are the benefits and harms of **transcutaneous electrical nerve stimulation (TENS)** in the management of patients with hip OA?

1.14.2 What are the benefits and harms of **inferential currents** in the management of patients with hip OA?

1.15 What are the benefits and harms of **therapeutic ultrasound** in the management of patients with hip OA?

1.16 What are the benefits and harms of **laser** in the management of patients with hip OA?

1.17 What are the benefits and harms of **acupuncture** in the management of patients with hip OA?

2. Pharmacological Interventions for Hip Osteoarthritis (35 PICOs)

2.1 **Oral analgesics**
2.1.1 What are the benefits and harms of **paracetamol** in the management of patients with hip OA?
2.1.2 What are the benefits and harms of **oral NSAIDs including COX-2 inhibitors** in the management of patients with hip OA?
2.1.3 What are the benefits and harms of **oral opioids** in the management of patients with hip OA?

2.2 **Topical analgesics**
2.2.1 What are the benefits and harms of **topical NSAIDs** in the management of patients with hip OA?
2.2.2 What are the benefits and harms of **transdermal opioids** in the management of patients with hip OA?
2.2.3 What are the benefits and harms of **topical capsaicin** in the management of patients with hip OA?

2.3 **Herbal therapies**
2.3.1 What are the benefits and harms of **avocado soybean unsaponifiable** in the management of patients with hip OA?
2.3.2 What are the benefits and harms of **boswellia serrata** in the management of patients with hip OA?
2.3.3 What are the benefits and harms of **curcuma** in the management of patients with hip OA?
2.3.4 What are the benefits and harms of **pine bark extract** in the management of patients with hip OA?

2.4 **Nutraceuticals**
2.4.1 What are the benefits and harms of **glucosamine** in the management of patients with hip OA?
2.4.2 What are the benefits and harms of **chondroitin** in the management of patients with hip OA?
2.4.3 What are the benefits and harms of **glucosamine and chondroitin in compound form** in the management of patients with hip OA?
2.4.4 What are the benefits and harms of **vitamin D** in the management of patients with hip OA?
2.4.5 What are the benefits and harms of **(omega-3/6) poly-unsaturated fatty acids** in the management of patients with hip OA?
2.4.6 What are the benefits and harms of **collagen preparations** in the management of patients with hip OA?
2.4.7 What are the benefits and harms of **methylsulfonylmethane** in the management of patients with hip OA?
2.4.8 What are the benefits and harms of **diacerein** in the management of patients with hip OA?

2.5 What are the benefits and harms of **duloxetine** in the management of patients with hip OA?

2.6 What are the benefits and harms of **doxycycline** in the management of patients with hip OA?

2.7 **Anti-osteoporosis (anti-resorptive bone-acting) drugs**
2.7.1 What are the benefits and harms of **bisphosphonates** in the management of patients with hip OA?
2.7.2 What are the benefits and harms of **calcitonin** in the management of patients with hip OA?
2.7.3 What are the benefits and harms of **strontium ranelate** in the management of patients with hip OA?

2.8 **Investigational DMOADs (symptomatic or inflammatory modification)**
2.8.1 What are the benefits and harms of **IL-1 inhibitors** in the management of patients with hip OA?
2.8.2 What are the benefits and harms of **TNF-alpha inhibitors** in the management of patients with hip OA?

2.8.3 What are the benefits and harms of **anti-nerve growth factor (NGF) therapy** in the management of patients with hip OA?
2.8.4 What are the benefits and harms of **fibroblast growth factor (FGF) therapy** in the management of patients with hip OA?
2.8.5 What are the benefits and harms of **colchicine** in the management of patients with hip OA?
2.8.6 What are the benefits and harms of **methotrexate** in the management of patients with hip OA?
2.8.7 What are the benefits and harms of **statins** in the management of patients with hip OA? – excluded

2.9 **Intra-articular injections**
2.9.1 What are the benefits and harms of **corticosteroids** in the management of patients with hip OA?
2.9.2 What are the benefits and harms of **viscosupplementation** in the management of patients with hip OA?
2.9.3 What are the benefits and harms of **platelet-rich plasma** in the management of patients with hip OA?
2.9.4 What are the benefits and harms of **stem cell therapy** in the management of patients with hip OA?
2.9.5 What are the benefits and harms of **dextrose prolotherapy** in the management of patients with hip OA?

3. Surgical Interventions (non-arthroplasty) (2 PICOs)

3.1 What are the benefits and harms of **arthroscopic lavage and debridement** interventions in the management of patients with hip OA? – excluded
3.2 What are the benefits and harms of **arthroscopic procedures for cartilage repair** interventions in the management of patients with hip OA? – excluded

4. Combination Therapies (4 PICOs)

4.1 What are the benefits and harms of **combination weight management and exercise interventions** compared to **exercise** in patients with hip OA?
4.2 What are the benefits and harms of **combination weight management and exercise interventions** compared to **weight management** in patients with hip OA?
4.3 What are the benefits and harms of **combination exercise and cognitive behavioural interventions** compared to **exercise** in patients with hip OA?
4.4 What are the benefits and harms of **combination exercise and cognitive behavioural interventions** compared to **cognitive behavioural interventions** in patients with hip OA?

GRADE Tables for Knee Osteoarthritis

Section 1: Non-Pharmacologic Interventions

PICO 1.1 (knee): What are the benefits and harms of self-management education programmes in the management of patients with knee OA?

SUMMARY
Very low-quality evidence shows this intervention has no significant effect on pain and function.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Coleman, et al. Arthritis Res Ther. 2012 Jan 27; 14(1): R21; 2. Keefe, et al. Behavior Therapy. 1990 Winter; 21(1): 49–62 ; 3. Mazzuca, et al. Arthritis Rheum. 1997 Aug; 40(8): 1466-74; 4. Ravaud, et al. BMJ. 2009 Feb 23; 338: b421; 5. Victor, et al. Clin Rheumatol. 2005 Aug; 24(4): 358-64.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Self-Management Education	Control	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 8 weeks to 12 months)												
5	randomised trials [1,2,3,4,5]	very serious [a]	serious [b]	not serious	serious [c]	none	N=405	N=418	SMD 0.16 lower (0.39 lower to 0.06 higher)		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 8 weeks to 12 months)												
5	randomised trials [1,2,3,4,5]	very serious [a]	very serious [d]	not serious	serious [c]	none	N=405	N=418	SMD 0.19 lower (0.52 lower to 0.14 higher)		⊕□□□ VERY LOW	CRITICAL
Quality of Life (Higher scores indicate better quality of life) (follow up: 4 months)												
2	randomised trials [3,4]	very serious [a]	not serious	not serious	serious [c]	none	N=228	N=264	SMD 0.04 lower (0.21 lower to 0.14 higher)		⊕□□□ VERY LOW	IMPORTANT
Depression (Higher scores indicate more severe depression) (follow up: range 8 weeks to 12 months)												
3	randomised trials [1,2,5]	very serious [a]	not serious	not serious	serious [c]	none	N=177	N=154	SMD 0.19 lower (0.5 lower to 0.11 higher)		⊕□□□ VERY LOW	IMPORTANT
Non-Adherence to Study Regimen [outcome does not necessitate, but does not exclude, withdrawal from study] (Risk ratios less than one favor Self-Management Education) (follow up: range 10 weeks to 12 months)												
4	randomised trials [1,2,3,5]	very serious [a]	not serious	not serious	serious [c]	none	45/284 (15.8%)	73/331 (22.1%)	RR 0.80 (0.47 to 1.35)	44 fewer per 1,000 (from 77 more to 117 fewer)	⊕□□□ VERY LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. All of the included trials received at least one High risk of bias rating; all trials were either single blind or unblinded. In addition some studies were rated High risk due to reporting errors or due to potential attrition bias.
b. I^2= 59%; moderate heterogeneity.
c. 95% CI crosses null.
d. I^2= 80%

PICO 1.3 (knee): What are the benefits and harms of cognitive behavioural therapy (CBT) in the management of patients with knee OA?

SUMMARY
Low-quality evidence in people with knee OA shows that CBT programs can have small benefits for pain, and can also improve self-efficacy, pain coping, depression and anxiety (O'Moore KA, et al. 2017;70(1):61–70). CBT for knee OA is associated with a low risk of adverse events. Very low-quality evidence from a limited number of studies indicates that combining CBT with exercise is more effective than either alone. While there is no evidence of the effects of CBT, specifically in people with hip OA. Benefits seen in people with knee OA and mixed samples of hip/knee OA are likely to be generalisable to those with hip OA.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. Helminen, et al. Clin Rehabil. 2015 Sep; 29(9): 868-81; 2. Keefe, et al. Behavior Therapy. 1991 Feb; 21(1): 49-62; 3. Smith, et al. Arthritis Rheumatol. 2015 May; 67(5): 1221-33; 4. Somers, et al. Pain. 2012 Jun; 153(6): 1199-209.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cognitive Behavioral Therapy†	Control	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate more severe pain) (follow up: range 6 weeks to 24 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	not serious	not serious	not serious	none	N=189	N=177	**SMD 0.21 lower (0.42 lower to 0.01 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 6 weeks to 24 weeks)												
3	randomised trials [1,2,4]	serious [a]	not serious	not serious	serious [b]	none	N=146	N=129	SMD 0.05 lower (0.28 lower to 0.19 higher)		⊕⊕□□ LOW	CRITICAL
Health-Related Quality of Life 15D (scale range 0-1, with higher scores indicating better quality of life) (follow up: 6 weeks)												
1	randomised trial [1]	serious [c]	not assessable	not serious	very serious [b,d]	none	N=55	N=48	MD 0.02 lower (0.06 lower to 0.02 higher)		⊕□□□ VERY LOW	IMPORTANT
Self-Efficacy (Higher scores indicate higher self-efficacy) (follow up: range 6 weeks to 24 weeks)												
2	randomised trials [1,4]	serious [a]	very serious [e]	not serious	serious [b]	none	N=115	N=99	SMD 0.06 lower (0.62 lower to 0.5 higher)		⊕□□□ VERY LOW	IMPORTANT
Depression (Higher scores indicate more severe depression) (follow up: range 6 weeks to 24 weeks)												
3	randomised trials [1,2,4]	serious [a]	not serious	not serious	serious [b]	none	N=146	N=129	SMD 0.07 lower (0.33 lower to 0.18 higher)		⊕⊕□□ LOW	IMPORTANT

No Participation due to Lack of Interest [defined as withdrawal due to "no response" or "dissatisfaction"] (Risk Ratios less than one favor Cognitive Behavioral Therapy) (follow up: range 6 months to 12 months)												
2	randomised trials [1,4]	serious [a]	not serious	not serious	serious [b]	none	5/115 (4.3%)	13/107 (12.1%)	RR 0.38 (0.14 to 1.09)	75 fewer per 1,000 (from 11 more to 104 fewer)	⊕⊕◯◯ LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

† **Different Cognitive Behavioral Therapeutic Techniques were analyzed together.** The following were included in the analysis: Helminen 2015- cognitive-behavioral group intervention focused on pain management; Keefe 1991- Pain coping skills training; Smith 2015- cognitive behavioral therapy focused on insomnia; Somers 2012- Pain coping skills training

a. The majority of studies received at least one High risk of bias rating due to unblinded study design.
b. 95% CI crosses null.
c. Study received High risk of bias rating due to unblinded study design.
d. Sample size in one study arm <50.
e. I^2= 77%

PICO 1.4 (knee): What are the benefits and harms of all land-based exercise in the management of patients with knee OA? *(efficacy estimates only)*

SUMMARY (1.4 all land based exercise, 1.5.1 muscle strengthening, 1.5.2 walking, 1.5.3 stationary cycling, 1.5.4 Tai Chi, 1.5.5 Hatha yoga)

There is low-quality evidence from a large number of randomised controlled trials (RCTs) that found that land-based exercise overall has significant and clinically relevant benefits for pain, function and quality of life in the short-to-medium term in people with knee OA.The benefits for pain and function are moderate in size, and are seen irrespective of the patients' age, structural disease severity, pain levels and functional status. There is evidence that long-term therapeutic exercise is safe and not associated with an increased risk of structural disease progression.(Quicke JG, et al. Osteoarth Cartilage 2015;23(9):1445–56) There are various forms of land-based exercise that may be adopted for knee OA. There is very low-quality evidence from a limited number of RCTs (1 for cycling and 2 for Hatha yoga), with small samples sizes showing benefits for pain (stationary cycling and Hatha yoga) and function (Hatha yoga).

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□ LOW

Bibliography: 1. Fransen, et al. Cochrane Database Syst Rev. 2015 Jan 9; 1: CD004376 (meta analysis).

Quality assessment							№ of events/ № of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Land-based Exercise**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 6 weeks to 24 months)												
44	randomised trials [1]	serious [a]	not serious	not serious	not serious	none	N=1992	N=1545	**SMD 0.49 lower (0.59 lower to 0.39 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 6 weeks to 24 months)												
44	randomised trials [1]	serious [a]	serious [b]	not serious	not serious	none	N=2260	N=1653	**SMD 0.52 lower (0.64 lower to 0.39 lower)**		⊕⊕□□ LOW	CRITICAL
Quality of Life (Higher scores indicate better quality of life) (follow up: range 8 weeks to 12 months)												
13	randomised trials [1]	serious [a]	not serious	not serious	not serious	none	N=581	N=492	**SMD 0.28 higher (0.15 higher to 0.40 higher)**		⊕⊕⊕□ MODERATE	IMPORTANT
Study Withdrawals (Risk ratios less than one favor Land-based Exercise) (follow up: range 6 weeks to 24 months)												
45	randomised trials [1]	serious [a]	not serious	not serious	serious [c]	none	343/2512 (13.7%)	320/2095 (15.3%)	RR 0.89 (0.78 to 1.03)	17 fewer per 1,000 (from 5 more to 34 fewer)	⊕⊕□□ LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference

a. All studies received at least one High risk of bias rating due to inadequate blinding or potential attrition bias.
b. I^2= 68%; moderate heterogeneity.
c. 95% CI crosses null.

PICO 1.5: SPECIFIC FORMS OF LAND-BASED EXERCISE

PICO 1.5.1 (knee): What are the benefits and harms of muscle strengthening in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Fransen, et al. Cochrane Database Syst Rev. 2015 Jan 9; 1: CD004376; 2. Bruce-Brand, et al. BMC Musculoskelet Disord. 2012 Jul 3; 13: 118; 3. Doi, et al. Am J Phys Med Rehabil. 2008 Apr; 87(4): 258-69; 4. Foley, et al. Ann Rheum Dis. 2003 Dec ;62(12): 1162-7; 5. Lund, et al. J Rehabil Med. 2008 Feb; 40(2): 137-44; 6. Maurer, et al. Arch Phys Med Rehabil. 1999 Oct; 80(10): 1293-9; 7. Salli, et al. Int J Rheum Dis. 2012 Apr; 15(2): 197-206; 8. Thorstensson, et al. BMC Musculoskelet Disord. 2005 May 30; 6: 27; 9. Baker, et al. J Rheumatol. 2001 Jul; 28(7): 1655-65; 10. Bennell, et al. Osteoarthritis Cartilage. 2010 May; 18(5): 621-8; 11. Bezalel, et al. Physiotherapy. 2010 Jun; 96(2): 137-43; 12. Chang, et al. Disabil Rehabil. 2012; 34(20): 1727-35; 13. Foroughi, et al. Clin Biomech (Bristol, Avon). 2011 Feb; 26(2): 167-74; 14. Gür, et al. Arch Phys Med Rehabil. 2002 Mar; 83(3): 308-16; 15. Huang, et al. Semin Arthritis Rheum. 2003 Jun; 32(6): 398-406; 16. Jan, et al. Phys Ther. 2008 Apr; 88(4): 427-36; 17. Lim, et al. Arthritis Rheum. 2008 Jul 15; 59(7): 943-51; 18. Lin, et al. J Orthop Sports Phys Ther. 2009 Jun; 39(6): 450-7; 19. Ettinger, et al. JAMA. 1997 Jan 1; 277(1): 25-31.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Muscle Strengthening Exercise	Control	Relative (95% CI)	Absolute (95% CI)		
***QUADRICEPS STRENGTHENING ONLY:* Pain** (Higher scores indicate higher pain severity) (follow up: range 6 weeks to 20 weeks)												
9	randomised trials [1]	serious [a]	serious [b]	not serious	not serious	none	N=327	N=293	**SMD 0.64 lower (0.95 lower to 0.33 lower)**		⊕⊕□□ LOW	CRITICAL
***QUADRICEPS STRENGTHENING ONLY:* WOMAC Function** (Higher scores indicate poorer functional outcome) (follow up: range 6 weeks to 20 weeks)												
10	randomised trials [1]	serious [a]	very serious [c]	not serious	not serious	none	N=398	N=328	**SMD 0.74 lower (1.07 lower to 0.41 lower)**		⊕□□□ VERY LOW	CRITICAL
***LOWER LIMB STRENGTHENING ONLY:* Pain** (Higher scores indicate higher pain severity) (follow up: range 8 weeks to 18 months)												
12	randomised trials [1]	serious [a]	serious [d]	not serious	not serious	none	N=522	N=341	**SMD 0.53 lower (0.78 lower to 0.28 lower)**		⊕⊕□□ LOW	CRITICAL
***LOWER LIMB STRENGTHENING ONLY:* Function** (Higher scores indicate poorer functional outcome) (follow up: range 8 weeks to 18 months)												
13	randomised trials [1]	serious [a]	very serious [e]	not serious	not serious	none	N=634	N=432	**SMD 0.54 lower (0.83 lower to 0.26 lower)**		⊕□□□ VERY LOW	CRITICAL
Quality of Life (Higher scores indicate better quality of life) (follow up: range 6 weeks to 8 weeks)												

7	randomised trials [2,3,4,5,6,7,8]	serious [a]	not serious	not serious	not serious	none	N=259	N=230	**SMD 0.26 higher (0.07 higher to 0.46 higher)**	⊕⊕⊕◻ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Muscle strengthening Exercise) (follow up: range 6 weeks to 12 months)											
16	randomised trials [2,3,4,5,7,8,9,10,11,12,13,14,15,16,17,18]	serious [a]	not serious	not serious	serious [f]	none	42/640 (6.6%)	22/502 (4.4%)	RR 1.33 (0.76 to 2.31)	14 more per 1,000 (from 11 fewer to 57 more)	⊕⊕◻◻ LOW	CRITICAL
Treatment-Related Adverse Events (Risk ratios less than one favor Muscle strengthening Exercise) (follow up: range 6 weeks to 12 months)											
10	randomised trials [3,4,5,8,10,13,15,16,17,18]	serious [a]	not serious	not serious	not serious	none	36/489 (7.4%)	3/392 (0.8%)	**RR 3.90 (1.55 to 9.81)**	**22 more per 1,000 (from 4 more to 67 more)**	⊕⊕⊕◻ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Muscle strengthening Exercise) (follow up: 20 weeks)											
1	randomised trial [19]	serious [a]	not assessable	not serious	serious [f]	none	3/146 (2.1%)	1/149 (0.7%)	RR 3.06 (0.32 to 29.10)	14 more per 1,000 (from 5 fewer to 189 more)	⊕⊕◻◻ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. All studies received at least one High risk of bias rating due to inadequate blinding or potential attrition bias.
b. I^2= 70%; moderate heterogeneity
c. I^2= 77%
d. I^2= 61%; moderate heterogeneity
e. I^2= 76%
f. 95% CI crosses null.

PICO 1.5.2 (knee): What are the benefits and harms of walking in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Fransen, et al. Cochrane Database Syst Rev. 2015 Jan 9; 1: CD004376; 2. Talbot, et al. J Am Geriatr Soc. 2003 Mar; 51(3): 387-92; 3. Ettinger, et al. JAMA. 1997 Jan 1; 277(1): 25-31.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Walking Programs**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 12 weeks to 18 months)												
4	randomised trials [1]	serious [a]	not serious	not serious	not serious	none	N=225	N=126	**SMD 0.48 lower (0.83 lower to 0.13 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 12 weeks to 18 months)												
3	randomised trials [1]	serious [a]	not serious	not serious	not serious	none	N=208	N=109	**SMD 0.35 lower (0.58 lower to 0.11 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Walking programs) (follow up: 12 weeks)												
1	randomised trial [2]	serious [a]	not assessable	not serious	very serious [b,c]	none	0/17 (0.0%)	1/17 (5.9%)	RR 0.33 (0.01 to 7.65)	39 fewer per 1,000 (from 58 fewer to 391 more)	⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Walking programs) (follow up: 18 months)												
1	randomised trial [3]	serious [a]	not assessable	not serious	serious [b]	none	2/144 (1.4%)	1/149 (0.7%)	RR 2.07 (0.19 to 22.57)	7 more per 1,000 (from 5 fewer to 145 more)	⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. All studies received at least one High risk of bias rating for inadequate blinding or unblinded design.
b. 95% CI crosses null.
c. Sample size <50 in each study arm.

PICO 1.5.3 (knee): What are the benefits and harms of cycling in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Salacinski, et al. J Orthop Sports Phys Ther. 2012 Dec; 42(12): 985-95.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Cycling**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (Scale range 0-100, with higher scores indicating higher pain severity) (follow up: 12 weeks)												
1	randomised trial [1]	very serious [a]	not assessable	not serious	very serious [b]	none	N=13	N=15	**MD 14.9 lower (25.3 lower to 4.5 lower)**		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (Scale range 0-100, with higher scores indicating poorer functional outcome) (follow up: 12 weeks)												
1	randomised trial [1]	very serious [a]	not assessable	not serious	very serious [b,c]	none	N=13	N=15	MD 11.1 lower (23.74 lower to 1.54 higher)		⊕□□□ VERY LOW	CRITICAL
KOOS Knee-related Quality of Life (Scale range 0-100, with higher scores indicating better quality of life) (follow up: 12 weeks)												
1	randomised trial [1]	very serious [a]	not assessable	not serious	very serious [b,c]	none	N=13	N=15	MD 6.8 higher (7.48 lower to 21.08 higher)		⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Cycling) (follow up: 12 weeks)												
1	randomised trial [1]	very serious [a]	not assessable	not serious	very serious [c,d]	none	4/19 (21.1%)	0/18 (0.0%)	RR 8.55 (0.49 to 148.33)	NA [e]	⊕□□□ VERY LOW	CRITICAL
Treatment-Related Adverse Events (Risk ratios less than one favor Cycling) (follow up: 12 weeks)												
1	randomised trial [1]	very serious [a]	not assessable	not serious	very serious [c,d]	none	3/19 (15.8%)	0/18 (0.0%)	RR 6.65 (0.37 to 120.36)	NA [e]	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Study received multiple High risk of bias ratings due to unblinded study design and potential attrition bias.
b. Total sample size <30.
c. 95% CI crosses null.
d. Sample size in each study arm <50.
e. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.

PICO 1.5.4 (knee): What are the benefits and harms of Tai Chi in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. Adler, P.A., 2007. *The effects of tai chi on pain and function in older adults with osteoarthritis* (Doctoral dissertation, Case Western Reserve University); 2. Brismée, et al. Clin Rehabil. 2007 Feb; 21(2): 99-111; 3. Fransen, et al. Arthritis Rheum. 2007 Apr 15; 57(3): 407-14; 4. Hartman, et al. J Am Geriatr Soc. 2000 Dec; 48(12): 1553-9; 5. Lee, et al. Clin Rehabil. 2009 Jun; 23(6): 504-11; 6. Song, et al. J Rheumatol. 2003 Sep; 30(9): 2039-44 ; 7. Song, et al. Journal of muscle and joint health. 2009; 16(1): 46-54; 8. Tsai, et al. J Pain Symptom Manage. 2013 Apr; 45(4): 660-9; 9. Wang, et al. Arthritis Rheum. 2009 Nov 15; 61(11): 1545-53.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Tai Chi**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 6 weeks to 12 weeks)												
9	randomised trials [1-9]	serious [a]	not serious	not serious	not serious	none	N=231	N=197	**SMD 0.57 lower (0.76 lower to 0.37 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 6 weeks to 12 weeks)												
7	randomised trials [2,3,5-9]	serious [a]	not serious	not serious	not serious	none	N=205	N=176	**SMD 0.67 lower (0.88 lower to 0.46 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Quality of Life (Higher scores indicate better quality of life) (follow up: 12 weeks)												
3	randomised trials [3,5,9]	serious [a]	not serious	not serious	not serious	none	N=105	N=76	**SMD 0.55 higher (0.11 higher to 0.99 higher)**		⊕⊕⊕□ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Tai Chi) (follow up: range 12 weeks to 48 weeks)												
6	randomised trials [1-3,5,8,9]	serious [a]	not serious	not serious	serious [b]	none	5/163 (3.1%)	2/128 (1.6%)	RR 1.90 (0.43 to 8.37)	14 more per 1,000 (from 9 fewer to 115 more)	⊕⊕□□ LOW	CRITICAL
Treatment-related Adverse Events (Risk ratios less than one favor Tai Chi) (follow up: range 12 weeks to 48 weeks)												
3	randomised trials [1,8,9]	serious [a]	not serious	not serious	serious [b]	none	1/56 (1.8%)	0/53 (0.0%)	RR 3.00 (0.13 to 69.52)	NA [c]	⊕⊕□□ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Tai Chi) (follow up: 48 weeks)												
1	randomised trials [9]	serious [a]	not assessable	not serious	serious [b,d]	none	1/20 (5.0%)	1/20 (5.0%)	RR 1.00 (0.07 to 14.90)	0 fewer per 1,000 (from 47 fewer to 695 more)	⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. The majority of studies received at least one High risk of bias rating due to single-blind design or inadequate blinding.
b. 95% CI crosses null.
c. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.
d. Sample size in each study arm <50.

PICO 1.5.5 (knee): What are the benefits and harms of yoga in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Cheung, et al. BMC Complement Altern Med. 2014 May 18; 14: 160; 2. Cheung, et al. Rheumatol Int. 2017 Mar; 37(3): 389-398.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Hatha Yoga**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 8 weeks)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	serious [b]	none	N=50	N=41	**MD 3.49 lower (5.06 lower to 1.91 lower)**		⊕⊕□□ LOW	CRITICAL
WOMAC Function (scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 8 weeks)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	serious [b]	none	N=50	N=41	**MD 10.58 lower (15.24 lower to 5.93 lower)**		⊕⊕□□ LOW	CRITICAL
SF-12 Physical Component Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 8 weeks)												
2	randomised trials [1,2]	serious [a]	very serious [c]	not serious	very serious [b,d]	none	N=50	N=41	MD 2.01 lower (10.82 lower to 6.8 higher)		⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Hatha Yoga) (follow up: 8 weeks)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	very serious [b,d]	none	1/50 (2.0%)	0/41 (0.0%)	RR 2.18 (0.09 to 51.28)	0 fewer per 1,000 (from 0 fewer to 0 fewer)	⊕□□□ VERY LOW	CRITICAL
Treatment-related Adverse Events (Risk ratios less than one favor Hatha Yoga) (follow up: 8 weeks)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	serious [b]	none	0/50 (0.0%)	0/41 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕□□ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Hatha Yoga) (follow up: 8 weeks)												
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [e]	none	0/32 (0.0%)	0/23 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Both trials received at least one High risk of bias rating due to single-blind study design.

b. Total sample size ≤50 in each study arm.
c. I^2= 89%
d. 95% CI crosses null.
e. Total sample size <50 in each study arm.
.

PICO 1.6 (knee): What are the benefits and harms of aquatic exercise/hydrotherapy in the management of patients with knee OA?

SUMMARY
There is low-quality evidence that aquatic exercise lead to small statistically significant improvements in pain, physical function and quality of life in people with knee and/or hip OA. There is a low risk of harm with aquatic exercise. Benefits in pain reduction and function from aquatic exercise therapy in the treatment of hip and/or knee OA are smaller than the effects from land-based exercise therapy.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. Bartels, et al. Cochrane Database Syst Rev. 2016 Mar 23; 3: CD005523 (meta-analysis).

Quality assessment							№ of events/ № of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Aquatic Exercise**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 6 weeks to 18 months)												
12	randomised trials [1]	not serious	not serious	serious [a]	not serious	none	N=539	N=537	**SMD 0.31 lower (0.47 lower to 0.15 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 6 weeks to 18 months)												
12	randomised trials [1]	not serious	not serious	serious [a]	not serious	none	N=529	N=530	**SMD 0.32 lower (0.47 lower to 0.17 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Quality of Life (Higher scores indicate better quality of life) (follow up: range 6 weeks to 18 months)												
10	randomised trials [1]	not serious	serious [b]	serious [a]	not serious	none	N=493	N=478	**SMD 0.25 lower (0.49 lower to 0.01 lower)**		⊕⊕□□ LOW	IMPORTANT
Total Adverse Events (Risk ratios less than one favor Aquatic Exercise) (follow up: range 6 weeks to 18 months)												
13	randomised trials [1]	not serious	not serious	serious [a]	serious [c]	none	112/605 (18.5%)	89/585 (15.2%)	RR 1.25 (0.98 to 1.60)	38 more per 1,000 (from 3 fewer to 91 more)	⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. Mixed population; includes Knee only studies, Hip only studies, and Hip/Knee studies.
b. I^2= 65%; moderate heterogeneity.
c. 95% CI crosses null.

PICO 1.7 MANUAL THERAPY
PICO 1.7.1 (knee): What are the benefits and harms of massage in the management of patients with knee OA?

SUMMARY (1.7.1 massage, 1.7.2 manipulation and mobilisation)
The evidence is from very low-quality or low-quality data. For some people with knee and/or hip OA, these therapies may have a positive effect on pain and/or function over a short term (low-quality to very low-quality evidence), and there is a very low risk of harm.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1.Atkins, et al. Int J Ther Massage Bodywork. 2013; 6(1): 4-14; 2. Perlman, et al. Arch Intern Med. 2006 Dec 11-25; 166(22): 2533-8; 3. Perlman, et al. PLoS One. 2012; 7(2): e30248; 4. Yip, YB., and Tam, AC. Complement Ther Med. 2008 Jun; 16(3): 131-8.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage Therapy†	Usual Care	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 4 weeks to 12 weeks)												
4	randomised trials [1,2,3,4]	very serious [a]	not serious	not serious	not serious	none	N=183	N=93	**SMD 0.7 lower (0.97 lower to 0.43 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 4 weeks to 8 weeks)												
3	randomised trials [2,3,4]	serious [b]	not serious	not serious	not serious	none	N=165	N=75	**SMD 0.58 lower (0.87 lower to 0.29 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Massage) (follow up: range 4 weeks to 24 weeks)												
3	randomised trials [2,3,4]	serious [b]	not serious	not serious	serious [c]	none	4/175 (2.3%)	1/77 (1.3%)	RR 1.17 (0.23 to 5.89)	2 more per 1,000 (from 10 fewer to 64 more)	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Massage) (follow up: range 4 weeks to 24 weeks)												
3	randomised trials [2,3,4]	serious [b]	not serious	not serious	serious [c]	none	2/175 (1.1%)	0/77 (0.0%)	RR 2.01 (0.22 to 18.82)	NA [d]	⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

† Interventions and comparators assessed included the following: Atkins 2013- Supervised self-massage (4 weeks), then unsupervised self-massage (4 weeks) vs. Wait-list Control; Perlman 2006- Swedish massage vs. Wait-list Control; Perlman 2012- Swedish massage (various doses) vs. Usual Care; Yip and Tam 2008- Aroma massage with orange and ginger vs. Olive Oil massage vs. Usual Care.

a. 3 of 4 trials which reported pain received at least one High risk of bias rating due to unblinded study design, potentially inadequate blinding, or potential for attrition bias or reporting bias.
b. 2 of 3 trials which reported this outcome received at least one High risk of bias rating due to unblinded study design, potentially inadequate blinding, or potential attrition bias.
c. 95% CI crosses null.
d. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.

PICO 1.7.2 (knee): What are the benefits and harms of manipulation and mobilisation in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Moss, et al. Man Ther. 2007 May; 12(2): 109-18; 2. Pollard, et al. J Can Chiropr Assoc. 2008 Dec; 52(4): 229-42.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Manipulation & Mobilisation**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 48 hours to 2 weeks)												
2	randomised trials [1,2]	very serious [a]	not serious	not serious	serious [b]	none	N=64	N=55	SMD 0.16 lower (0.52 lower to 0.21 higher)		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 48 hours to 2 weeks)												
2	randomised trials [1,2]	very serious [a]	not serious	not serious	not serious	none	N=64	N=55	**SMD 0.56 lower (0.93 lower to 0.19 lower)**		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Manipulation & Mobilisation) (follow up: 2 weeks)												
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [c]	none	0/26 (0.0%)	0/17 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was inestimable.		⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Manipulation & Mobilisation) (follow up: 2 weeks)												
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [c]	none	0/26 (0.0%)	0/17 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was inestimable.		⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. One of two studies which report this outcome received multiple High risk of bias ratings due to a crossover design without phase segregation, as well as pre- post- follow up time of merely 48 hours. The other study received a High risk of bias rating due to potentially inadequate blinding.
b. 95% CI crosses null.
c. Sample size <50 in each study arm.

PICO 1.8 (knee): What are the benefits and harms of weight management in the management of patients with knee OA?

SUMMARY

There is limited evidence of very low quality that weight loss alone (achieved through diet and exercise) has no significant effect on either pain or function in people with knee OA although benefits appear to be more significant with higher amounts of weight loss – starting at a minimum of 5–7.5% body weight loss. Ideally, it is beneficial to achieve a greater amount of weight loss. (Atukorala I, et al. Arthritis Care Res 2016;68(8):1106–14; Messier SP, et al. JAMA, 2013;310(12):1263–73). There is no randomised controlled trial (RCT) on the effects of bariatric surgery in people with hip and/or knee OA. There are no RCTs investigating weight management specifically in people with hip OA, although the numerous other systemic health benefits of weight loss and maintaining a healthy body weight are most likely transferable to people with hip OA. There are limitations of the available RCT evidence in OA. Clinicians are advised to refer to the National Health and Medical Research Council's (NHMRC's) guidelines for the most effective strategies for managing overweight/obesity in primary care. National Health and Medical Research Council. Clinical practice guidelines for the management of overweight and obesity in adults, adolescents and children in Australia. Melbourne: NHMRC, 2013).

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Bliddal, et al. Ann Rheum Dis. 2011 Oct; 70(10): 1798-803; 2. Messier, et al. Arthritis Rheum. 2004 May; 50(5): 1501-10 (ADAPT); 3. Miller, et al. Obesity (Silver Spring). 2006 Jul; 14(7): 1219-30; 4. Focht, et al. Arthritis Rheum. 2005 Oct 15; 53(5): 659-65 (ADAPT).

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Weight Management**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 6 months to 18 months)												
3	randomised trials [1,2,3]	serious [a]	very serious [b]	not serious	serious [c]	none	N=165	N=158	SMD 0.38 lower (0.88 lower to 0.11 higher)		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 6 months to 18 months)												
3	randomised trials [1,2,3]	serious [a]	serious [d]	not serious	serious [c]	none	N=165	N=158	SMD 0.29 lower (0.62 lower to 0.04 higher)		⊕□□□ VERY LOW	CRITICAL
Percentage Weight Loss (Percentage of weight at baseline lost by follow up time, with more loss indicating positive outcome) (follow up: range 6 months to 18 months)												
3	randomised trials [1,2,3]	serious [e]	not serious	not serious	not serious	none	N=165	N=156	**MD 6.48 % more lost (8.48 % more lost to 4.48 % more lost)**		⊕⊕⊕□ MODERATE	IMPORTANT
Lateral Joint Space Narrowing [mm] (Higher scores indicate poorer structural outcome) (follow up: 18 months)												
1	randomised trials [2]	serious [e]	not assessable	not serious	serious [c]	none	N=82	N=78	MD 0.09 mm higher (0.41 mm lower to 0.59 mm higher)		⊕⊕□□ LOW	IMPORTANT
Medial Joint Space Narrowing [mm] (Higher scores indicate poorer structural outcome) (follow up: 18 months)												

1	randomised trials [2]	serious [e]	not assessable	not serious	serious [c]	none	N=82	N=78	MD 0.05 mm higher (0.45 mm lower to 0.55 mm higher)	⊕⊕□□ LOW	IMPORTANT
Walking Self-Efficacy (Patient confidence in walking around a gymnasium twice without stopping; Scale range 0-100 with higher scores indicating more confidence) (follow up: 18 months)											
1	randomised trials [4]	serious [a]	not assessable	not serious	serious [c]	none	N=82	N=78	MD 0.1 higher (9.72 lower to 9.92 higher)	⊕⊕□□ LOW	IMPORTANT
Withdrawal due to Adverse Events (Risk ratios less than one favor Weight Management) (follow up: 12 months)											
1	randomised trials [1]	serious [a]	not assessable	not serious	very serious [c,g]	none	2/44 (4.5%)	2/45 (4.4%)	RR 1.02 (0.15 to 6.94)	1 more per 1,000 (from 38 fewer to 264 more)	⊕□□□ VERY LOW	CRITICAL
Non-Compliance with Regimen (Outcome includes, but does not necessitate, withdrawal from study. Defined as "non-compliance", "lack of motivation", "non-adherence") (Risk ratios less than one favor Weight Management) (follow up: range 12 months to 18 months)											
2	randomised trials [1,4]	serious [a]	serious [f]	not serious	serious [c]	none	32/126 (25.4%)	38/123 (30.9%)	RR 0.77 (0.37 to 1.63)	71 fewer per 1,000 (from 195 fewer to 195 more)	⊕□□□ VERY LOW	IMPORTANT
Serious Adverse Events (Risk ratios less than one favor Weight Management) (follow up: 6 months)											
1	randomised trials [3]	serious [e]	not assessable	not serious	serious [g]	none	0/44 (0.0%)	0/43 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. All studies received at least one High risk of bias rating due to single-blind or unblinded study design.
b. I^2= 78%
c. 95% CI crosses null.
d. I^2= 52%; moderate heterogeneity.
e. All studies received at least one High risk of bias rating due to single-blind or unblinded study design, but this outcome may not be susceptible to bias due to objective reporting.
f. I^2= 67%; moderate heterogeneity.
g. Sample size <50 in each study arm.

PICO 1.9 THERMOTHERAPY

PICO 1.9.1 (knee): What are the benefits and harms of local hot application in the management of patients with knee OA?

SUMMARY

Heat therapy may be effective in reducing pain for some people with knee and/or hip OA, but the quality of evidence is very low . Heat therapy is cheap and generally feasible for people to undertake independently as a self-management strategy.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Denegar, et al. Clin Interv Aging. 2010 Aug 9; 5: 199-206; 2. Mazzuca, et al. Arthritis Rheum. 2004 Oct 15; 51(5): 716-21; 3. Yildirim, et al. J Clin Nurs. 2010 Apr; 19(7-8): 1113-20.

Quality assessment							№ of events/ № of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Heat Therapy†**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 1 week to 4 weeks)												
3	randomised trials [1,2,3]	very serious [a]	not serious	not serious	not serious	none	N=83	N=82	**SMD 0.38 lower (0.69 lower to 0.07 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 1 week to 4 weeks)												
3	randomised trials [1,2,3]	very serious [a]	not serious	not serious	serious [b]	none	N=83	N=82	SMD 0.27 lower (0.68 lower to 0.15 higher)		⊕□□□ VERY LOW	CRITICAL
KOOS Quality of Life (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 1 week)												
1	randomised trial [2]	very serious [a]	not assessable	not serious	very serious [b,c]	none	N=34	N=34	MD 5.7 higher (0.63 lower to 12.03 higher)		⊕□□□ VERY LOW	IMPORTANT
Total Adverse Events (Risk ratios less than one favor Heat Therapy) (follow up: range 1 week to 4 weeks)												
2	randomised trials [1,2]	very serious [a]	not serious	not serious	not serious	none	0/60 (0.0%)	0/59 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

†Interventions and comparators assessed included the following: Denegar 2010- Heat Pad vs. No treatment ("Rest"); Mazzuca 2004- Heat-retaining knee sleeve vs. Placebo knee sleeve; Yildirim 2010- Heat Pad vs. Usual Care.

a. All studies received at least one High risk of bias rating.

b. 95% CI crosses null.

c. Sample size <50 in each study arm.

PICO 1.9.2 (knee): What are the benefits and harms of local cold application in the management of patients with knee OA?

SUMMARY
There is very low-quality evidence suggesting that the use of cold therapy is not effective in improving pain, function or quality of life in people with knee and/or hip OA.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Clark, et al. Rheumatol Rehabil. 1974 Nov; 13(4): 190-7; 2. Denegar, et al. Clin Interv Aging. 2010 Aug 9; 5: 199-206.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Cold Therapy†**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 1 week to 3 weeks)												
2	randomised trials [1,2]	very serious [a]	not serious	not serious	very serious [b,c]	none	N=49	N=47	SMD 0.5 lower (1.07 lower to 0.07 higher)		⊕□□□ VERY LOW	CRITICAL
KOOS Function in Daily Living (scale range 0 to 100, with higher scores indicating poorer functional outcome) (follow up: 1 week)												
1	randomised trial [2]	very serious [a]	not assessable	not serious	very serious [b,c]	none	N=34	N=34	MD 6.20 lower (13.71 lower to 1.31 higher)		⊕□□□ VERY LOW	CRITICAL
KOOS Quality of Life (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 1 week)												
1	randomised trial [2]	very serious [a]	not assessable	not serious	very serious [b,c]	none	N=34	N=34	MD 4.7 higher (1.86 lower to 11.26 higher)		⊕□□□ VERY LOW	IMPORTANT
Total Adverse Events (Risk ratios less than one favor Cold Therapy) (follow up: 1 week)												
1	randomised trial [2]	very serious [a]	not assessable	not serious	serious [c]	none	0/34 (0.0%)	0/34 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

†Interventions and comparators assessed included the following: Clark 1974- Ice application vs. "Untuned" short-wave diathermy (study received quality downgrade for inappropriate comparison); Denegar 2010- Cold treatment facilitated by circulation of water through a wrap-around garment vs. No treatment ("Rest").

a. Both studies received more than two High risk of bias ratings due to faults in study design and reporting.
b. 95% CI crosses null.
c. Sample size <50 in each study arm.

PICO 1.10 ORTHOTIC BRACES

PICO 1.10.1 (knee): What are the benefits and harms of varus unloading/re-alignment braces in the management of patients with knee OA?

SUMMARY (1.10.1 varus unloading/realignment braces, 1.10.2 valgus unloading/realignment braces, 1.10.3 realigning patellofemoral braces)
There is no RCT evidence on the effects of varus unloading braces on pain or physical function in people with knee OA. There is limited low-quality evidence that valgus unloading braces have no significant effect on pain or physical function. Similarly, there is limited, very low quality evidence that patellofemoral realigning braces have no significant effect on pain or function.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

No RCT data was found which related to this question.

PICO 1.10.2 (knee): What are the benefits and harms of valgus unloading/re-alignment braces in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□ LOW

Bibliography: 1. Duivenvoorden, et al. Cochrane Database Syst Rev. 2015 Mar 16; (3): CD004020; 2. Brouwer, et al. Osteoarthritis Cartilage. 2006 Aug; 14(8): 777-83.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Valgus Unloading Brace	Control	Relative (95% CI)	Absolute (95% CI)		
VAS Pain (scale range 0-10, with higher scores indicating more severe pain) (follow up: 6 months)												
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [b]	none	N=58	N=57	MD 0.1 lower (0.91 lower to 0.71 higher)		⊕⊕□□ LOW	CRITICAL
VAS Pain (scale range 0-10, with higher scores indicating more severe pain) (follow up: 12 months)												
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [b]	none	N=58	N=57	MD 0 (0.84 lower to 0.84 higher)		⊕⊕□□ LOW	CRITICAL
HSS Knee Function (scale range 0-100, with higher scores indicating better functional outcome) (follow up: 6 months)												
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [b]	none	N=56	N=54	MD 1.4 higher (2.36 lower to 5.16 higher)		⊕⊕□□ LOW	CRITICAL
HSS Knee Function (scale range 0-100, with higher scores indicating better functional outcome) (follow up: 12 months)												
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [b]	none	N=56	N=54	MD 1 higher (2.98 lower to 4.98 higher)		⊕⊕□□ LOW	CRITICAL
EQ-5D (QoL measure; scale range 0-1, with higher scores indicating better quality of life) (follow up: 6 months)												
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [b]	none	N=60	N=57	MD 0.05 lower (0.14 lower to 0.04 higher)		⊕⊕□□ LOW	IMPORTANT
EQ-5D (QoL measure; scale range 0-1, with higher scores indicating better quality of life) (follow up: 12 months)												
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [b]	none	N=60	N=57	MD 0.04 lower (0.12 lower to 0.04 higher)		⊕⊕□□ LOW	IMPORTANT
Total Adverse Events (Risk ratios less than one favor Valgus Unloading Brace) (follow up: 12 months)												

1	randomised trial [2]	serious [a]	not assessable	not serious	serious [b]	none	24/60 (40.0%)	14/57 (24.6%)	RR 1.63 (0.94 to 2.82)	155 more per 1,000 (from 15 fewer to 447 more)	⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Study received High risk of bias rating due to unblinded study design.
b. 95% CI crosses null.

PICO 1.10.3 (knee): What are the benefits and harms of realigning patellofemoral braces in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Callaghan, et al. Ann Rheum Dis. 2015 Jun; 74(6): 1164-70; 2. Hunter, et al. Osteoarthritis Cartilage. 2011 Jul; 19(7): 792-800.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Realigning Patellofemoral Brace**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: 6 weeks)												
2	randomised trials [1,2]	serious [a]	serious [b]	not serious	serious [c]	none	N=104	N=102	SMD 0.3 lower (0.75 lower to 0.14 higher)		⊕□□□ VERY LOW	CRITICAL
KOOS Activities of Daily Living subscale (scale range 0 to 100, with higher scores indicating better functional outcome) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	N=61	N=62	MD 8.8 higher (1.29 lower to 18.89 higher)		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Realigning Patellofemoral Brace) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	3/63 (4.8%)	1/63 (1.6%)	RR 3.00 (0.32 to 28.07)	32 more per 1,000 (from 11 fewer to 430 more)	⊕⊕□□ LOW	CRITICAL
Treatment-Related Adverse Events (Risk ratios less than one favor Realigning Patellofemoral Brace) (follow up: range 6 weeks to 18 weeks)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	serious [c]	none	15/143 (10.5%)	12/143 (8.4%)	RR 1.25 (0.63 to 2.50)	21 more per 1,000 (from 31 fewer to 126 more)	⊕⊕□□ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Realigning Patellofemoral Brace) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	1/63 (1.6%)	0/63 (0.0%)	RR 3.00 (0.12 to 72.27)	NA [d]	⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. One of 2 studies a High risk of bias rating, due to unblinded study design.

b. I^2= 60%; moderate heterogeneity.

c. 95% CI crosses null.

d. Due to zero events in the comparator arm, an absolute risk reduction could not be calculated.

PICO 1.11 INSOLES

PICO 1.11.1 (knee): What are the benefits and harms of medial wedged insoles in the management of patients with knee OA?

SUMMARY (1.11.1 medial wedge insoles, 1.11.2 lateral wedge insoles, 1.11.3 shock absorbing insoles, 1.11.4 arch supports)
Very low-quality evidence from a single, small RCT investigating medial wedge insoles found significant benefits of clinically relevant magnitude for pain and function in people with lateral tibiofemoral compartment knee OA and valgus deformity. This study provides preliminary evidence that would need to be confirmed in larger trials. Conversely, for lateral wedge insoles, very low-quality evidence from a number of RCTs found no significant benefits for pain, function, quality of life or structural disease progression in people with medial knee OA. As there is no RCT data available in people with either knee or hip OA for either shock-absorbing insoles or arch supports.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Rodrigues, et al. Arthritis Rheum. 2008 May 15; 59(5): 603-8.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Medial Wedge Insoles	Control	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (scale range 0-100, with higher scores indicating higher pain severity) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b]	none	N=16	N=14	**MD 15.5 lower (25.81 lower to 5.19 lower)**		⊕□□□ VERY LOW	CRITICAL
Lequesne Index (scale range 0-24, with higher scores indicating poorer functional outcome) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b]	none	N=16	N=14	**MD 4.5 lower (7.29 lower to 1.71 lower)**		⊕□□□ VERY LOW	CRITICAL
Withdrawal due to Adverse Events (Risk ratios less than one favor Medial Wedge Insoles) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b]	none	0/16 (0.0%)	0/14 (0.0%)	Due to zero events in both groups, an absolute risk reduction was not estimable.		⊕⊕□□ LOW	CRITICAL
Treatment-Related Adverse Events (Risk ratios less than one favor Medial Wedge Insoles) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	0/16 (0.0%)	1/14 (7.1%)	RR 0.29 (0.01 to 6.69)	51 fewer per 1,000 (from 71 fewer to 406 more)	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Baseline pain and function scores were higher in the intervention group than in the control group.
b. Total sample size for both study arms ≤30.
c. 95% CI crosses null.

PICO 1.11.2 (knee): What are the benefits and harms of lateral wedge insoles in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Barrios, et al. Knee. 2009 Mar; 16(2): 136-42; 2. Bennell, et al. BMJ. 2011 May 18; 342: d2912; 3. Hatef, et al. Int J Rheum Dis. 2014 Jan; 17(1): 84-8; 4. Pham, et al. Osteoarthritis Cartilage. 2004 Jan; 12(1): 46-55; 5. Toda, et al. Osteoarthritis Cartilage. 2008 Feb; 16(2): 244-53; 6. Wallace, D.A. *Efficacy of lateral heel wedge orthotics for the treatment of patients with knee osteoarthritis* (Doctoral dissertation); 7. Baker, et al. Arthritis Rheum. 2007 Apr; 56(4): 1198-203.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Lateral Wedge Insole	Control	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 8 weeks to 12 months)												
6	randomised trials [1,2,3,4,5,6]	serious [a]	very serious [b]	not serious	serious [c]	none	N=338	N=319	SMD 0.36 lower (0.92 lower to 0.2 higher)		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 8 weeks to 12 months)												
6	randomised trials [1,2,3,4,5,6]	serious [a]	very serious [b]	not serious	serious [c]	none	N=338	N=319	SMD 0.31 lower (0.85 lower to 0.24 higher)		⊕□□□ VERY LOW	CRITICAL
Health-Related Quality of Life (scale range -0.04 to 1, with higher scores indicating better quality of life) (follow up: 12 months)												
1	randomised trial [2]	not serious	not assessable	not serious	serious [c]	none	N=103	N=97	MD 0.01 lower (0.04 lower to 0.02 higher)		⊕⊕⊕□ MODERATE	IMPORTANT
Proportion of Patients with Radiographic Progression [JSN ≥0.5 mm] (Risk ratios less than one favor Lateral Wedge Insole) (follow up: 24 months)												
1	randomised trial [4]	serious [d]	not assessable	not serious	serious [c]	none	24/55 (43.6%)	20/55 (36.4%)	RR 1.20 (0.76 to 1.90)	73 more per 1,000 (from 87 fewer to 327 more)	⊕⊕□□ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Lateral Wedge Insole) (follow up: range 12 weeks to 12 months)												
4	randomised trials [1,2,4,5]	serious [e]	not serious	not serious	serious [c]	none	11/265 (4.2%)	4/247 (1.6%)	RR 2.06 (0.64 to 6.59)	17 more per 1,000 (from 6 fewer to 91 more)	⊕⊕□□ LOW	CRITICAL

Treatment-Related Adverse Events (Risk ratios less than one favor Lateral Wedge Insole) (follow up: range 12 weeks to 12 months)												
4	randomised trials [1,2,5,7]	serious [e]	serious [f]	not serious	serious [c]	none	57/255 (22.4%)	39/252 (15.5%)	RR 1.45 (0.55 to 3.80)	70 more per 1,000 (from 70 fewer to 433 more)	⨁◯◯◯ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. 5 of 6 trials received at least one High risk of bias rating, most commonly due to single blind or unblinded study design.
b. I^2= 91%
c. 95% CI crosses null.
d. Study received a High risk of bias rating due to single blind study design. Patients in this study were allowed to receive intra-articular injections of hyaluronate and/or corticosteroid, and joint lavage was permitted.
e. 3 of 4 trials received at least one High risk of bias rating.
f. I^2= 68%; moderate heterogeneity.

PICO 1.11.3 (knee): What are the benefits and harms of shock absorbing insoles in the management of patients with knee OA?

No RCT data was found which related to this question.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

PICO 1.11.4 (knee): What are the benefits and harms of arch supports in the management of patients with knee OA?

No RCT data was found which related to this question.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

PICO 1.12: SHOES
PICO 1.12.1 (knee): What are the benefits and harms of unloading shoes in the management of patients with knee OA?

SUMMARY (1.12.1 unloading shoes, 1.12.2 minimalist footwear, 1.12.3 rocker-sole shoes)
While unloading and minimalist shoes reduce medial tibiofemoral compartment knee joint loading (Erhart JC, et al. J Orthop Res, 2010;28(12):1548–53; Trombini-Souza F, et al. Gait Posture 2011;34(1):126–30; Bennell KL, et al. Arthritis Rheum 201;65(3):701–9), there is limited evidence of very low quality that these shoes offer no additional benefit on pain or clinically relevant effects on function, compared with conventional walking shoes There is limited evidence of low quality that rocker-sole shoes offer no significant benefit on pain or function, compared with conventional walking shoes. Clinicians may consider advising people to consider wearing footwear with shock-absorbing properties,and to advise avoidance of high-heeled shoes, given they increase knee joint loads (Radzimski AO, et al. Knee 2012;19(3):163–75), albeit in the absence of RCT data about which individual footwear features are beneficial and/or harmful.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Erhart, et al. J Orthop Res. 2010 Jul; 28(7): 873-9; 2. Hinman, et al. Ann Intern Med. 2016 Sep 20; 165(6): 381-9.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Unloading Shoes**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: 6 months)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	serious [b]	none	N=120	N=119	SMD 0.11 lower (0.37 lower to 0.14 higher)		⊕⊕□□ LOW	CRITICAL
WOMAC Function (scale range 0-68, with higher scores indicating poorer functional outcome) (follow up: 6 months)												
1	randomised trial [2]	not serious	not assessable	not serious	serious [b]	none	N=80	N=80	MD 0.5 lower (4.34 lower to 3.34 higher)		⊕⊕⊕□ MODERATE	CRITICAL
Assessment of Quality of Life 6D scale (scale range -0.04 to 1.00, with higher scores indicating higher quality of life) (follow up: 6 months)												
1	randomised trial [2]	not serious	not assessable	not serious	serious [b]	none	N=79	N=79	MD 0 (0.03 lower to 0.03 higher)		⊕⊕⊕□ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Unloading Shoes) (follow up: 6 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	3/40 (7.5%)	11/39 (28.2%)	**RR 0.27 (0.08 to 0.88)**	**206 fewer per 1,000 (from 34 fewer to 259 fewer)**	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Unloading Shoes) (follow up: 6 months)												

1	randomised trial [2]	not serious	not assessable	not serious	serious [b]	none	26/83 (31.3%)	20/81 (24.7%)	RR 1.27 (0.77 to 2.08)	67 more per 1,000 (from 57 fewer to 267 more)	⊕⊕⊕□ MODERATE	CRITICAL
Withdrawal due to Potentially Treatment-Related Reasons ["Too small shoe size", "Shoe discomfort", "Meniscectomy", "Hip pain", "Foot pain", "Total Knee Replacement"] (Risk ratios less than one favor Unloading Shoes) (follow up 6 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	4/40 (10.0%)	10/39 (25.6%)	RR 0.39 (0.13 to 1.14)	156 fewer per 1,000 (from 36 more to 223 fewer)	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. Erhart 2010 was given a High risk of bias rating due to single blind study design and potential attrition bias.
b. 95% CI crosses null.
c. Sample size <50 in each study arm.

PICO 1.12.2 (knee): What are the benefits and harms of minimalist footwear in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Trombini-Souza, et al. Clin Biomech (Bristol, Avon). 2015 Dec; 30(10): 1194-201.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Minimalist Footwear**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 6 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	N=28	N=28	MD 3.25 lower (6.78 lower to 0.28 higher)		⊕□□□ VERY LOW	CRITICAL
Lequesne Index (scale range 0 to 24, with higher scores indicating poorer functional outcome) (follow up: 6 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	N=28	N=28	**MD 2.4 lower (4.38 lower to 0.42 lower)**		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Minimalist Footwear) (follow up: 6 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	1/28 (3.6%)	1/28 (3.6%)	RR 1.00 (0.07 to 15.21)	0 fewer per 1,000 (from 33 fewer to 507 more)	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Study received a High risk of bias rating due to a single blind design.
b. 95% CI crosses null.
c. Total sample size in each study arm is <50.

PICO 1.12.3 (knee): What are the benefits and harms of rocker sole shoes in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. Nigg, et al. Med Sci Sports Exerc. 2006 Oct; 38(10): 1701-8.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Masai Barefoot Technology Footwear**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0-500, with higher scores indicating higher pain severity) (follow up: 12 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=53	N=66	MD 4.2 higher (27.8 lower to 36.2 higher)		⊕⊕□□ LOW	CRITICAL
WOMAC Function (scale range 0-1700, with higher scores indicating poorer functional outcome) (follow up: 12 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=53	N=66	MD 18.7 higher (94.31 lower to 131.71 higher)		⊕⊕□□ LOW	CRITICAL
Withdrawal due to Treatment-Related Adverse Events (Risk ratios less than one favor Masai Barefoot Technology footwear) (follow up: 12 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	1/58 (1.7%)	1/67 (1.5%)	RR 1.16 (0.07 to 18.06)	2 more per 1,000 (from 14 fewer to 255 more)	⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Patients are not blinded to allocation.
b. 95% CI crosses null

PICO 1.13: TAPING
PICO 1.13.1 (knee): What are the benefits and harms of kinesio taping in the management of patients with knee OA?

SUMMARY (11.13.1 kinesio taping, 1.13.2 patellar taping)
There is some evidence that patellar taping can immediately change patellar alignment measured on imaging and reduce pain.(Crossley KM, et al. Arthritis Rheum 2009;61(12):1719–25; Hinman RS, Bennell KL, Rheumatology, 2003;42(7):865–69). However, very low-quality evidence from a single RCT did not find any significant effect of taping when worn continuously for three weeks on pain and function, compared with sham tape in people with knee OA not specifically selected for patellofemoral pain. A limited number of trials of very low quality and with small sample sizes have evaluated kinesio taping for knee OA. These trials used different configurations of kinesio tape that was reapplied by the clinician after various intervals. There is no trial of kinesio taping for hip OA. The evidence shows no significant benefits of kinesio taping for pain or function.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Anandkumar, et al. Physiother Theory Pract. 2014 Aug; 30(6): 375-83; 2. Cho, et al. Am J Phys Med Rehabil. 2015 Mar; 94(3): 192-200; 3. Kocyigit, et al. Complement Ther Clin Pract. 2015 Nov; 21(4): 262-7.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Kinesio Taping**	**Sham Kinesio Taping**	**Relative (95% CI)**	**Absolute (95% CI)**		
VAS Pain (scale range 0 to 10, with higher scores indicating higher pain severity) (follow up: range 0 days [two pre- and post-test trials] to 12 days)												
3	randomised trials [1,2,3]	serious [a]	very serious [b]	not serious	very serious [c,d]	none	N=64	N=63	MD 1.28 lower (2.93 lower to 0.37 higher)		⊕□□□ VERY LOW	CRITICAL
Lequesne Index (scale range 0 to 24, with higher scores indicating poorer functional outcome) (follow up: 12 days)												
1	randomised trial [3]	serious [a]	not assessable	not serious	serious [e]	none	N=21	N=20	**MD 2.9 higher (0.09 higher to 5.71 higher)**		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Kinesio Taping) (follow up: 12 days)												
1	randomised trial [3]	serious [a]	not assessable	not serious	very serious [c,e]	none	0/22 (0.0%)	1/21 (4.8%)	RR 0.32 (0.01 to 7.42)	32 fewer per 1,000 (from 47 fewer to 306 more)	⊕□□□ VERY LOW	CRITICAL
Skin irritation (Risk ratios less than one favor Kinesio Taping) (follow up: 12 days)												
1	randomised trial [3]	serious [a]	not assessable	not serious	very serious [c,e]	none	0/22 (0.0%)	1/21 (4.8%)	RR 0.32 (0.01 to 7.42)	32 fewer per 1,000 (from 47 fewer to 306 more)	⊕□□□ VERY LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. All studies received either High or Unclear risk of bias ratings due to single-blind study design or potentially inadequate blinding measures.

b. I^2= 89%
c. 95% CI crosses null.
d. Insufficient follow up time to provide an adequate estimate.
e. Sample size <50 in each study arm.

PICO 1.13.2 (knee): What are the benefits and harms of patellar taping in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Hinman, et al. BMJ. 2003 Jul 19; 327(7407): 135; 2. Cushnaghan, et al. BMJ. 1994 Mar 19; 308(6931): 753-5.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Patellar Taping**	**Sham Patellar Taping**	**Relative (95% CI)**	**Absolute (95% CI)**		
VAS Pain (Scale range 0 to 10, with higher scores indicating higher pain severity) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	N=29	N=29	MD 0.8 lower (1.97 lower to 0.37 higher)		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (Scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	N=29	N=29	MD 2.6 higher (3.32 lower to 8.52 higher)		⊕□□□ VERY LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Patellar Taping) (follow up: 7 days)												
1	randomised trial [2]	serious [a]	not assessable	not serious	very serious [d]	none	0/14 (0.0%)	0/14 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was inestimable.		⊕□□□ VERY LOW	CRITICAL
Skin Irritation (Risk ratios less than one favor Therapeutic Taping) (follow up: 3 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	8/29 (27.6%)	1/29 (3.4%)	**RR 8.00 (1.07 to 59.95)**	**241 more per 1,000 (from 2 more to 1,000 more)**	⊕⊕□□ LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Both trials received High risk of bias ratings due to single-blind design, in which patients were not blinded.
b. 95% CI crosses null.
c. Sample size <50 in each study arm.
d. Total sample size <30.

PICO 1.14 (knee): What are the benefits and harms of walking cane/stick in the management of patients with knee OA?

SUMMARY
People with knee and/or hip OA often adopt an abnormal gait pattern because of pain, muscle weakness, joint mobility restrictions or other pain conditions. The use of an assistive walking device may be useful to improve gait pattern and posture to normalise musculoskeletal loads. There is low-quality evidence from one trial that the use of a walking aid (eg single point stick) is effective in improving pain and function in people with knee OA. These data could be reasonably transferred to people with hip OA (very low-quality evidence).

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. Jones, et al. Ann Rheum Dis. 2012 Feb; 71(2): 172-9.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Walking cane/stick**	**No Cane**	**Relative (95% CI)**	**Absolute (95% CI)**		
VAS Pain (Scale range 0 to 10 cm, with higher scores indicating higher pain severity) (follow up: 60 days)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=32	N=32	**MD 2.26 cm lower (2.9 lower to 1.62 lower)**		⊕⊕□□ LOW	CRITICAL
Lequesne Index (Scale range 0 to 24, with higher scores indicating poorer functional outcome) (follow up: 60 days)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=32	N=32	**MD 3.64 cm lower (5.35 lower to 1.93 lower)**		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Walking cane/stick) (follow up: 60 days)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious [b,c]	none	1/32 (3.1%)	2/32 (6.3%)	RR 0.50 (0.05 to 5.24)	31 fewer per 1,000 (from 59 fewer to 265 more)	⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Study received a High risk of bias rating due to single-blind study design; investigators are not blinded, and patient blinding may have been inadequate.
b. Sample size <50 in each study arm
c. 95% CI crosses null.

PICO 1.15: ELECTROMAGNETIC THERAPY

PICO 1.15.1 (knee): What are the benefits and harms of pulsed electromagnetic/shortwave therapy in the management of patients with knee OA?

SUMMARY

There is low-quality evidence that pulsed electromagnetic therapy significantly improves pain and function in people with knee OA by clinically relevant amounts. There is very low-quality evidence that pulsed electromagnetic therapy has no statistically significant effect on pain or function in people with hip OA. Most studies involved clinician-delivered treatments, at high frequency of servicing, ranging from three to five times per week. A minority of studies used portable devices that individuals applied themselves at home, with treatment dosage ranging from two to 12 hours per day. Although the evidence suggests moderate effect sizes and a low risk of harms for pulsed electromagnetic fields in peoplewith knee OA, it was noted current evidence is restricted to short-term (two to 10 weeks) follow-up only, so maintenance of a therapeutic effect remains uncertain. The available evidence suggests that three to five treatment sessions per week are required for benefits when this treatment is administered by clinicians. Promising data from a limited number of small trials investigating portable devices, but felt further research is required regarding effectiveness, acceptability and adherence.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. Bagnato, et al. Rheumatology (Oxford). 2016 Apr; 55(4): 755-62; 2. Callaghan, et al. Joint Bone Spine. 2005 Mar; 72(2): 150-5; 3. Fukuda, et al. Journal of Applied Research. 2008 Sep, 8(3),189-199; 4. Fukuda, et al. Phys Ther. 2011 Jul; 91(7): 1009-17; 5. Külcü, et al. Turk J Rheumatol 2009; 24: 144-8; 6. Lee, et al. J Korean Rheum Assoc. 2004, 11: 143-50; 7. Nelson, et al. Rheumatol Int. 2013 Aug; 33(8): 2169-73; 8. Perrot, et al. Arthritis & Rheumatism 41.9 (1998): S357; 9. Tejero Sánchez, et al. Patología del aparato locomotor. 2003; 1(3): 190-5; 10. Thamsborg, et al. Osteoarthritis Cartilage. 2005 Jul; 13(7): 575-81; 11. Trock, et al. J Rheumatol. 1994 Oct; 21(10): 1903-11; 12. Wuschech, et al. Bioelectromagnetics. 2015 Dec; 36(8): 576-85.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Pulsed Electromagnetic/ Shortwave Therapy**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 2 weeks to 10 weeks)												
12	randomised trials [1-12]	serious [a]	serious [b]	not serious	not serious	none	N=316	N=271	**SMD 0.53 lower (0.84 lower to 0.21 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 18 days to 10 weeks)												
9	randomised trials [1,3-6,8,10-12]	serious [c]	not serious	not serious	not serious	none	N=248	N=220	**SMD 0.39 lower (0.58 lower to 0.21 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Quality of Life (Higher scores indicate better quality of life) (follow up: range 3 weeks to 4 weeks)												
2	randomised trials [1,4]	not serious	not serious	not serious	serious [d]	none	N=60	N=51	SMD 0.36 higher (0.01 lower to 0.74 higher)		⊕⊕⊕□ MODERATE	IMPORTANT

Withdrawals due to Adverse Events (Risk ratios less than one favor Pulsed Electromagnetic/Shortwave Therapy) (follow up: range 1 month to 12 months)												
3	randomised trials [4,10,11]	not serious	not serious	not serious	serious [d]	none	1/116 (0.9%)	1/108 (0.9%)	RR 0.72 (0.05 to 10.91)	3 fewer per 1,000 (from 9 fewer to 92 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Treatment-related Adverse Events (Risk ratios less than one favor Pulsed Electromagnetic/Shortwave Therapy) (follow up: range 3 weeks to 12 weeks)												
8	randomised trials [1,5,7,9-12]	serious [e]	not serious	not serious	serious [d]	none	12/250 (4.8%)	6/205 (2.9%)	RR 1.95 (0.81 to 4.71)	28 more per 1,000 (from 6 fewer to 109 more)	⊕⊕◻◻ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Pulsed Electromagnetic/Shortwave Therapy) (follow up: 12 weeks)												
1	randomised trial [10]	not serious	not assessable	not serious	serious [f]	none	0/42 (0.0%)	0/41 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕◻ MODERATE	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. 6 of 12 trials reporting pain received at least one High risk of bias rating; one trial could not be adequately assessed for bias due to foreign language.
b. I^2= 73% moderate heterogeneity
c. 5 of 9 trials reporting function received at least one High risk of bias rating; one trial could not be adequately assessed for bias due to foreign language.
d. 95% CI crosses null.
e. 4 of 8 trials reporting this outcome received at least one High risk of bias rating; one trial could not be adequately assessed for bias due to foreign language.
f. Sample size in each study arm <50.

PICO 1.15: ELECTROMAGNETIC THERAPY
PICO 1.15.2 (knee): What are the benefits and harms of shockwave therapy in the management of patients with knee OA?

SUMMARY (1.15.2 shockwave, 1.16.2 interferential, 1.18 laser)
While very low-quality evidence suggests some possible benefits from shockwave and interferential current modalities on pain and function, these findings were limited trials (one for shockwave and two for interferential), with a limited sample size and serious or very serious risk of bias. Seven trials using laser therapy among people with knee OA suggest clinically meaningful benefits in short-term pain and function (up to three weeks); however, the quality of evidence is low to very low.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Zhao, et al. J Surg Res. 2013 Dec; 185(2): 661-6.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Extracorporeal Shockwave Therapy**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (Scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 12 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=34	N=36	**MD 2.3 lower (3.6 lower to 1 lower)**		⊕⊕□□ LOW	CRITICAL
WOMAC Function (Scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 12 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=34	N=36	**MD 7.9 lower (12.33 lower to 3.47 lower)**		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Extracorporeal Shockwave Therapy) (follow up: 12 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	0/34 (0.0%)	1/36 (2.8%)	RR 0.35 (0.01 to 8.36)	18 fewer per 1,000 (from 27 fewer to 204 more)	⊕⊕□□ LOW	CRITICAL
Treatment-related Adverse Events (Risk ratios less than one favor Extracorporeal Shockwave Therapy) (follow up: 12 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	0/34 (0.0%)	0/36 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio
a. Study received a High risk of bias rating due to single-blind study design.
b. Sample size <50 in each study arm.
c. 95% CI crosses null.

PICO 1.16: ELECTRICAL STIMULATION

PICO 1.16.1 (knee): What are the benefits and harms of transcutaneous electrical nerve stimulation (TENS) in the management of patients with knee OA?

SUMMARY

Very low-quality evidence from four trials in people with knee OA suggests that TENS has a clinically meaningful effect on pain and function. While no direct evidence is available from trials in people with hip OA, the working group felt that the mode of action with TENS could be transferable to the hip. Trials were limited to four weeks follow-up, so it remains uncertain whether treatment effects are maintained beyond this period.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Cheing, et al. Clin Rehabil. 2002 Nov; 16(7): 749-60; 2. Cheing, et al. J Rehabil Med. 2003 Mar; 35(2): 62-8; 3. Ng, et al. J Altern Complement Med. 2003 Oct; 9(5): 641-9; 4. Yurtkuran, M., and Kocagil, T. Am J Acupunct. 1999; 27(3-4): 133-40 ; 5. Law, et al. J Clin Rheumatol. 2004 Dec; 10(6): 295-9; 6. Smith, et al. Physiotherapy. 1983 Aug 10; 69(8): 266-8.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**TENS**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 10 days to 4 weeks)												
4	randomised trials [1,2,3,4]	very serious [a]	not serious	not serious	not serious	none	N=79	N=57	**SMD 0.76 lower (1.13 lower to 0.39 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: 2 weeks)												
3	randomised trials [3,4,5]	very serious [a]	not serious	not serious	serious [b]	none	N=54	N=48	**SMD 0.48 lower (0.88 lower to 0.08 lower)**		⊕□□□ VERY LOW	CRITICAL
Treatment-related Withdrawals (Risk ratios less than one favor TENS) (follow up: range 2 weeks to 8 weeks)												
3	randomised trials [4,5,6]	very serious [a]	not serious	not serious	serious [c]	none	1/63 (1.6%)	0/58 (0.0%)	RR 3.00 (0.13 to 68.57)	NA [d]	⊕□□□ VERY LOW	CRITICAL
Treatment-related Adverse Events (Risk ratios less than one favor TENS) (follow up: 2 weeks)												
1	randomised trial [5]	very serious [a]	not assessable	not serious	serious [e]	none	0/22 (0.0%)	0/17 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor TENS) (follow up: 2 weeks)												
1	randomised trial [2]	very serious [a]	not assessable	not serious	very serious [f]	none	0/10 (0.0%)	0/10 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. All studies received at least one High risk of bias rating; most received more than one High risk of bias rating, due to inadequate blinding/single blind or unblinded design, potential reporting bias, and other issues
b. Sample size<50 in one study arm.
c. 95% CI crosses null.
d. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.
e. Sample size<50 in one study arm.
f. Total sample size <30.

PICO 1.16.2 (knee): What are the benefits and harms of interferential currents in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Defrin, et al. Pain. 2005 May; 115(1-2): 152-60; 2. Gundog, et al. Am J Phys Med Rehabil. 2012 Feb; 91(2): 107-13.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Interferential Current**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: 4 weeks)												
2	randomised trials [1,2]	very serious [a]	not serious	not serious	serious [b]	none	N=90	N=24	**SMD 0.79 lower (1.27 lower to 0.32 lower)**		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (Scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 4 weeks)												
1	randomised trial [2]	very serious [a]	not assessable	not serious	serious [c]	none	N=45	N=15	**MD 30.93 lower (50.77 lower to 11.08 lower)**		⊕□□□ VERY LOW	CRITICAL
Treatment-related Withdrawals (Risk ratios less than one favor Interferential Current Therapy) (follow up: 4 weeks)												
1	randomised trial [2]	very serious [a]	not assessable	not serious	serious [c]	none	0/45 (0.0%)	0/15 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. Both studies received more than one High risk of bias rating, due to inadequate blinding/single blind or unblinded design, potential reporting bias, and other issues
b. Sample size <50 in one study arm.
c. Sample size <50 in each study arm.

PICO 1.17 (knee): What are the benefits and harms of ultrasound in the management of patients with knee OA?

SUMMARY
There is moderate-quality evidence that therapeutic ultrasound has statistically significant effects on pain and physical function in people with knee OA. There is no randomised controlled trial (RCT) involving participants with hip OA, thus the evidence level for this population group was downgraded to low quality because of concerns about indirectness. Although the evidence suggests moderate effect sizes and a low risk of harms for therapeutic ultrasound in people with knee OA, the working group noted that current evidence is restricted to short-term (two to eight weeks) follow-up only. There is concern about whether benefits are sustained once treatment finished. The available evidence suggests that three to five treatment sessions per week are required for benefits.

OVERALL QUALITY OF EVIDENCE: ⊕⊕⊕□MODERATE

Bibliography: 1. Külcü, et al., 2009. "Short-term efficacy of pulsed electromagnetic field therapy on pain and functional level in knee osteoarthritis: a randomised controlled study"; 2. Loyola-Sánchez, et al. Arch Phys Med Rehabil. 2012 Jan; 93(1): 35-42; 3. Tascioglu, et al. J Int Med Res. 2010 Jul-Aug; 38(4): 1233-42; 4. Özgönenel, et al. Ultrasound Med Biol. 2009 Jan; 35(1): 44-9.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Ultrasound**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 2 weeks to 8 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	not serious	not serious	not serious	none	N=116	N=86	**SMD 0.55 lower (0.88 lower to 0.22 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up range 2 weeks to 8 weeks)												
3	randomised trials [1,2,4]	serious [a]	not serious	not serious	not serious	none	N=61	N=59	**SMD 0.57 lower (1.03 lower to 0.10 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Change in Central Medial Femoral Cartilage Volume [mm³] (Higher values indicate better structural outcome) (follow up: 8 weeks)												
1	randomised trial [2]	not serious	not assessable	not serious	very serious [b,c]	none	N=11	N=12	MD 16.70 mm³ lower (136.32 mm³ lower to 102.92 mm³ higher)		⊕⊕□□ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Therapeutic Ultrasound) (follow up: range 2 weeks to 8 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	not serious	not serious	not serious	none	0/118 (0.0%)	0/88 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕□ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Therapeutic Ultrasound) (follow up: range 2 weeks to 8 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	not serious	not serious	not serious	none	0/118 (0.0%)	0/88 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕□ MODERATE	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. 3 of 4 trials received at least one High risk of bias rating due to unblinded design or due to potential attrition bias.
b. 95% CI crosses null.
c. Total sample size <30.

PICO 1.18 (knee): What are the benefits and harms of laser in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1.Alfredo, et al. Clin Rehabil. 2012 Jun; 26(6): 523-33; 2. Fukuda, et al. Rev Bras Ortop. 2015 Dec 6; 46(5): 526-33; 3. Gworys, et al. Ortop Traumatol Rehabil. 2012 May-Jun; 14(3): 269-77; 4. Hegedüs, et al. Photomed Laser Surg. 2009 Aug; 27(4): 577-84; 5. Hsieh, et al. Arch Phys Med Rehabil. 2012 May; 93(5): 757-64; 6. Stelian, et al. J Am Geriatr Soc. 1992 Jan; 40(1): 23-6; 7. Tascioglu, et al. Swiss Med Wkly. 2004 May 1; 134(17-18): 254-8; 8. Bülow, et al. Scand J Rehabil Med. 1994 Sep; 26(3): 155-9.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Laser Therapy**	**Sham Laser Therapy**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 2 weeks to 3 weeks)												
7	randomised trials [1,2,3,4,5,6,7]	serious [a]	serious [b]	not serious	not serious	none	N=252	N=154	**SMD 0.49 lower (0.82 lower to 0.17 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 2 weeks to 3 weeks)												
6	randomised trials [1,2,3,5,6,7]	serious [c]	serious [d]	not serious	not serious	none	N=234	N=145	**SMD 0.67 lower (1.04 lower to 0.31 lower)**		⊕⊕□□ LOW	CRITICAL
Patients Reporting Improvement [N values are patients who reported "Treatment did help" over (N "Treatment did help" + "Treatment did not help")] (Risk ratios greater than one favor Laser Therapy) (follow up: 9 weeks)												
1	randomised trial [8]	serious [e]	not assessable	not serious	very serious [f,g]	none	4/14 (28.6%)	6/15 (40.0%)	RR 0.71 (0.25 to 2.01)	116 fewer per 1,000 (from 300 fewer to 404 more)	⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Laser Therapy) (follow up: range 3 weeks to 6 months)												
4	randomised trials [2,5,7,8]	serious [h]	not serious	not serious	serious [f]	none	0/116 (0.0%)	1/92 (1.1%)	RR 0.32 (0.01 to 7.50)	7 fewer per 1,000 (from 11 fewer to 71 more)	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Laser Therapy) (follow up: range 4 weeks to 6 months)												
3	randomised trials [5,7,8]	serious [i]	not serious	not serious	not serious	none	0/71 (0.0%)	0/70 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕□ MODERATE	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. 5 of 7 studies reporting pain received at least one High risk of bias rating.
b. I^2= 56% moderate heterogeneity.
c. 4 of 6 studies reporting function received at least one High risk of bias rating.
d. I^2= 62% moderate heterogeneity.
e. Study received High risk of bias rating due to potential selective reporting.
f. 95% CI crosses null.
g. Total sample size <30.
h. 3 of 4 studies received at least one High risk of bias rating.
i. All studies received at least one High risk of bias rating.

PICO 1.19.1 (knee): What are the benefits and harms of traditional acupuncture with manual stimulation in the management of patients with knee OA?

Summary: Four randomized controlled trials (RCTs) assessing the effects of traditional acupuncture with manual stimulation on patients with clinically confirmed knee osteoarthritis were included. Only included articles which involved sham acupuncture comparator groups; RCTs with comparator groups that involved no intervention or an incompatible sham (e.g. not a sham for traditional acupuncture with manual stimulation) were excluded. Overall, the RCTs were assessed to be of moderate quality. The majority of trials were double-blind, and one trial was single-blind; two trials reported funding from insurance companies, and the funding relationships of two trials were unclear. Two RCTs randomized less than 100 participants into groups of interest, but only one of these RCTs randomized less than 50 participants into groups of interest. Traditional acupuncture with manual stimulation demonstrated moderate, statistically significant benefits on pain within 8 to 13 weeks, but demonstrated no significant benefits on subjective function, quality of life, or depression within this time period. Acupuncture demonstrated no significant benefits on any efficacy outcome within 26 weeks. There were no significant differences observed between acupuncture and sham groups with regard to safety outcomes.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Huang, et al. J Am Geriatr Soc. 2010 Jun; 58(6): 1218-20; 2. Miller, et al. Evid Based Complement Alternat Med. 2011; 2011: 792975; 3. Scharf, et al. Ann Intern Med. 2006 Jul 4; 145(1): 12-20; 4. Witt, et al. Lancet. 2005 Jul 9-15; 366(9480): 136-43.

Certainty assessment							№ of events/№ of patients		Effect		Certainty	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Acupuncture (Manual)**	**Sham**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Short term) *[Decreasing values indicate improvement]*: Follow-up time ranged from 8 weeks to 13 weeks												
4	randomized trials	serious [a]	serious [b]	not serious	serious [c]	none [d]	N= 509	N= 473	**SMD 0.3 lower (0.61 lower to 0)**		⊕□□□ VERY LOW	CRITICAL
Pain (Long term) *[Decreasing values indicate improvement]*: Follow-up at 26 weeks												
2	randomized trials	serious [a]	not serious	not serious	not serious	none [d]	N= 475	N= 440	SMD 0.11 lower (0.25 lower to 0.02 higher)		⊕⊕⊕□ MODERATE	CRITICAL
Function (Short term) *[Decreasing values indicate improvement]*: Follow-up time ranged from 8 weeks to 13 weeks												
3	randomized trials	not serious	serious [e]	not serious	serious [c]	none [d]	N= 503	N= 467	SMD 0.2 lower (0.49 lower to 0.09 higher)		⊕⊕□□ LOW	CRITICAL
Function (Long term) *[Decreasing values indicate improvement]*: Follow-up at 26 weeks												
2	randomized trials	serious [a]	not serious	not serious	not serious	none [d]	N= 475	N= 440	SMD 0.17 lower (0.37 lower to 0.04 higher)		⊕⊕⊕□ MODERATE	CRITICAL
Quality of Life (short term) *[Increasing values indicate improvement]*: Follow-up time ranged from 8 weeks to 13 weeks												
2	randomized trials	serious [a]	not serious	not serious	serious [c]	none [d]	N= 475	N= 440	SMD 0.16 higher (0.05 lower to 0.38 higher)		⊕⊕□□ LOW	IMPORTANT
Quality of Life (Long term) *[Increasing values indicate improvement]*: Follow-up at 26 weeks												

Quality of Life (Long term) *[Increasing values indicate improvement]*: Follow-up at 26 weeks											
2	randomized trials	serious [a]	not serious	not serious	not serious	none [d]	N= 475	N= 440	SMD 0.12 higher (0.01 lower to 0.25 higher)	⊕⊕⊕◯ MODERATE	IMPORTANT
Depression (Short term) *[Decreasing values indicate improvement]*: Follow-up at 8 weeks											
1	randomized trial	serious [a]	not assessable	not serious	serious [c]	none [d]	N= 149	N= 75	SMD 0.03 higher (0.25 lower to 0.31 higher)	⊕⊕◯◯ LOW	IMPORTANT
Depression (Long term) *[Decreasing values indicate improvement]*: Follow-up at 26 weeks											
1	randomized trial	serious [a]	not assessable	not serious	serious [c]	none [d]	N= 149	N= 75	SMD 0.04 higher (0.24 lower to 0.32 higher)	⊕⊕◯◯ LOW	IMPORTANT
Withdrawals due to Adverse Events *[Risk ratios less than 1 favor Manual Acupuncture]*: Follow-up at 8 weeks											
2	randomized trials	serious [a]	not serious	not serious	serious [f]	none [d]	4/177 (2.3%)	4/102 (3.9%)	RR 0.82 (0.23 to 2.94)	7 fewer per 1,000 (from 30 fewer to 76 more)	⊕⊕◯◯ LOW	CRITICAL
Treatment-related Adverse Events *[Risk ratios less than 1 favor Manual Acupuncture]*: Follow-up at 8 weeks											
2	randomized trials	serious [a]	not serious	not serious	serious [f]	none [d]	20/177 (11.3%)	13/102 (12.7%)	RR 0.77 (0.41 to 1.47)	29 fewer per 1,000 (from 60 more to 75 fewer)	⊕⊕◯◯ LOW	CRITICAL
Serious Adverse Events *[Risk ratios less than 1 favor Manual Acupuncture]*: Follow-up at 26 weeks											
2	randomized trials	serious [a]	not serious	not serious	serious [f]	none [d]	23/475 (4.8%)	11/440 (2.5%)	RR 1.81 (0.64 to 5.09)	20 more per 1,000 (from 9 fewer to 102 more)	⊕⊕◯◯ LOW	IMPORTANT
Total Knee Replacement *[Risk ratios less than 1 favor Manual Acupuncture]*: Follow-up at 26 weeks											
1	randomized trial	not serious	not assessable	not serious	serious [f]	none [d]	0/326 (0.0%)	1/265 (0.4%)	RR 0.27 (0.01 to 6.63)	3 fewer per 1,000 (from 4 fewer to 21 more)	⊕⊕⊕◯ MODERATE	IMPORTANT

CI: Confidence interval; SMD: Standardized mean difference; RR: Risk ratio

Explanations

a. ≥50% of trials received “High” risk of bias ratings (≥1 out of 6 dimensions in the Cochrane Risk of Bias tool)
b. I^2= 63%, T^2= 0.05; moderate heterogeneity
c. 95% Confidence Interval of an SMD extends between >0.2-≤0.5 points in either direction (Cohen 1988*)
d. See Supplementary Table for funding information
e. I^2= 67%, T^2= 0.04; moderate heterogeneity
f. 95% Confidence Interval of a Risk Ratio crosses null

PICO 1.19.2 (knee): What are the benefits and harms of laser acupuncture in the management of patients with knee OA?

Summary: Three randomized controlled trials (RCTs) assessing the effects of laser acupuncture on patients with clinically confirmed knee osteoarthritis were included. Only included articles which involved sham acupuncture comparator groups; RCTs with comparator groups that involved no intervention or an incompatible sham (e.g. not a sham for laser acupuncture) were excluded. The RCTs were assessed to be of moderate to high quality; since none of the trials was assessed to be of Very Low Quality*, sensitivity analyses limiting by study quality were not warranted. All of the trials utilized appropriate double-blinding techniques; two of the trials confirmed no industry sponsorship, and one trial did not report funding sources. Two of the RCTs randomized less than 100 participants into groups of interest, but only one of these RCTs randomized less than 50 participants into groups of interest. Laser acupuncture demonstrated no significant benefits on pain, objective or subjective function, or quality of life at short-term or longer-term follow-up times. There were no significant differences observed between laser acupuncture and sham groups with regard to safety outcomes. No treatment-related adverse events occurred.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

A: All trials

Bibliography: 1. Al Rashoud, et al. Physiotherapy. 2014 Sep; 100(3): 242-8; 2. Hinman, et al. JAMA. 2014 Oct 1; 312(13): 1313-22; 3. Yurtkuran, et al. Photomed Laser Surg. 2007 Feb; 25(1): 14-20.

Certainty assessment							№ of events/№ of patients		Effect		Certainty	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Acupuncture (Laser)**	**Sham**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Short term) *[Decreasing values indicate improvement]*: Follow-up at 12 weeks												
3	randomized trials	not serious	not serious	not serious	serious [a]	none [b]	N= 118	N= 106	SMD 0.08 lower (0.4 lower to 0.24 higher)		⊕⊕⊕□ MODERATE	CRITICAL
Pain (Long term) *[Decreasing values indicate improvement]*: Follow-up time ranged from 26 weeks to 12 months												
2	randomized trials	serious [c]	very serious [d]	not serious	very serious [e]	none [b]	N= 84	N= 74	SMD 0.25 lower (1.06 lower to 0.56 higher)		⊕□□□ VERY LOW	CRITICAL
Function (Short term) *[Decreasing values indicate improvement]*: Follow-up at 12 weeks												
2	randomized trials	not serious	not serious	not serious	serious [a]	none [b]	N= 92	N= 83	SMD 0.02 lower (0.32 lower to 0.27 higher)		⊕⊕⊕□ MODERATE	CRITICAL
Function (Long term) *[Decreasing values indicate improvement]*: Follow-up at 12 months												
1	randomized trial	not serious	not assessable	not serious	serious [a]	none [b]	N= 58	N= 51	SMD 0.12 higher (0.26 lower to 0.49 higher)		⊕⊕⊕□ MODERATE	CRITICAL
Time to walk a certain distance *[Decreasing values indicate improvement]*: Follow-up at 12 weeks												
1	randomized trial	not serious	not assessable	not serious	very serious [e]	none [b]	N= 27	N= 25	SMD 0.12 higher (0.42 lower to 0.67 higher)		⊕⊕□□ LOW	IMPORTANT
Quality of Life *[Increasing values indicate improvement]*: Follow-up at 12 weeks												

2	randomized trials	not serious	not serious	not serious	serious [a]	none [b]	N= 92	N= 83	SMD 0.06 lower (0.36 lower to 0.24 higher)		⊕⊕⊕◻ MODERATE	IMPORTANT
Withdrawals due to Adverse Events *[Risk ratios less than 1 favor Laser Acupuncture]*: Follow-up at 12 weeks												
2	randomized trials	not serious	not serious	not serious	serious [f]	none [b]	1/87 (1.1%)	0/88 (0.0%)	RR 3.10 (0.13 to 74.61)	NA [g]	⊕⊕⊕◻ MODERATE	CRITICAL
Total Adverse Events *[Risk ratios less than 1 favor Laser Acupuncture]*: Follow-up time ranged from 12 weeks to 26 weeks												
3	randomized trials	not serious	not serious	not serious	serious [f]	none [b]	11/117 (9.4%)	7/116 (6.0%)	RR 1.62 (0.68 to 3.91)	37 more per 1,000 (from 19 fewer to 176 more)	⊕⊕⊕◻ MODERATE	IMPORTANT
Treatment-related Adverse Events *[Risk ratios less than 1 favor Laser Acupuncture]*: Follow-up time ranged from 12 weeks to 26 weeks												
2	randomized trials	serious [c]	not serious	not serious	not serious	none [b]	0/58 (0.0%)	0/55 (0.0%)	Due to zero events in both study arms, a relative risk could not be estimated		⊕⊕⊕◻ MODERATE	CRITICAL
Total Knee Replacement *[Risk ratios less than 1 favor Laser Acupuncture]*: Follow-up at 12 months												
1	randomized trial	not serious	not assessable	not serious	serious [f]	none [b]	1/59 (1.7%)	1/61 (1.6%)	RR 1.03 (0.07 to 16.15)	0 fewer per 1,000 (from 15 fewer to 248 more)	⊕⊕⊕◻ MODERATE	IMPORTANT

***"VERY LOW QUALITY" was defined as studies that received ≥2 High Risk of Bias ratings OR one specific High Risk Rating in the "Other" category in addition to ≥2 Unclear Risk ratings OR ≥3 Unclear Risk of Bias ratings in dimensions other than blinding or the "Other" category.**

CI: Confidence interval; **SMD:** Standardized mean difference; **RR:** Risk ratio

Explanations

a. 95% Confidence Interval of an SMD extends between >0.2-≤0.5 points in either direction (Cohen 1988*)
b. See Supplemental Table for funding information
c. ≥50% of trials received "High" risk of bias ratings (≥1 out of 6 dimensions in the Cochrane Risk of Bias tool)
d. I^2= 82%
e. 95% Confidence Interval of an SMD extends >0.5 points in either direction (Cohen 1988*)
f. 95% Confidence Interval of a Risk Ratio crosses null
g. Due to zero events in one study arm, an absolute risk reduction could not be estimated

Supplementary Table: funding information

Author, Year	Funding Statement	Industry?	Positive Result Pain?
Al Rashoud, 2014	"The authors would like to thank the Physiotherapy Department at the Security Forces Hospital, Riyadh, Saudi Arabia who granted approval to conduct this study without any expenses being incurred."	No	Yes (not within 6 months; only over 6 months)

Hinman, 2014	"This trial was funded by the National Health and Medical Research Council (project 566783)."	No	Negative result
Yurtkuran, 2007	Not reported	Unclear	No effect

PICO 1.19.3 (knee): What are the benefits and harms of electroacupuncture in the management of patients with knee OA?

Summary: Two randomized controlled trials (RCTs) assessing the effects of electroacupuncture on patients with clinically confirmed knee osteoarthritis were included. Only included articles which involved sham acupuncture comparator groups; RCTs with comparator groups that involved no intervention or an incompatible sham (e.g. not a sham for electroacupuncture) were excluded. Both RCTs were assessed to be of high quality and utilized appropriate double-blinding techniques. Both RCTs confirmed no industry funding. Both RCTs randomized more than 100 participants into groups of interest. Electroacupuncture had little to no positive effect on pain, subjective function, objective function, or quality of life within 8 to 12 weeks. Electroacupuncture demonstrated a small, statistically significant benefit on pain by 26 weeks, but did not demonstrate any significant benefits on subjective or objective function or quality of life within this time period. Acupuncture demonstrated no significant benefits on any efficacy outcome within 26 weeks. Participants receiving electroacupuncture were not significantly more likely to report treatment-related adverse events, but were significantly more likely to report at least one serious adverse event.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

A: All trials

Bibliography: 1. Berman, et al. Ann Intern Med. 2004 Dec 21; 141(12): 901-10; 2. Suarez-Almazor, et al. Arthritis Care Res (Hoboken). 2010 Sep; 62(9): 1229-36.

Certainty assessment							№ of events/№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Acupuncture (Electric)	Sham	Relative (95% CI)	Absolute (95% CI)		
Pain (Short term) *[Decreasing values indicate improvement]*: Follow-up time ranged from 8 weeks to 12 weeks												
2	randomized trials	not serious	not serious	not serious	not serious	none [a]	N= 322	N= 463	SMD 0.05 lower (0.2 lower to 0.1 higher)		⊕⊕⊕⊕ HIGH	CRITICAL
Pain (Long term) *[Decreasing values indicate improvement]*: Follow-up at 26 weeks												
1	randomized trial	not serious	not assessable	not serious	serious [b]	none [a]	N= 142	N= 141	**SMD 0.23 lower (0.47 lower to 0)**		⊕⊕⊕□ MODERATE	CRITICAL
Function (Short term) *[Decreasing values indicate improvement]*: Follow-up time ranged from 8 weeks to 12 weeks												
2	randomized trials	not serious	very serious [c]	not serious	serious [b]	none [a]	N= 322	N= 463	SMD 0.11 lower (0.42 lower to 0.2 higher)		⊕□□□ VERY LOW	CRITICAL
Function (Long term) *[Decreasing values indicate improvement]*: Follow-up at 26 weeks												
1	randomized trial	not serious	not assessable	not serious	serious [b]	none [a]	N= 142	N= 141	SMD 0.21 lower (0.44 lower to 0.03 higher)		⊕⊕⊕□ MODERATE	CRITICAL
6-minute walk distance (short term) *[Increasing values indicate improvement]*: Follow-up at 8 weeks												
1	randomized trial	not serious	not assessable	not serious	serious [b]	none [a]	N= 163	N= 156	SMD 0.02 lower (0.24 lower to 0.2 higher)		⊕⊕⊕□ MODERATE	IMPORTANT
6-minute walk distance (short term) *[Increasing values indicate improvement]*: Follow-up at 26 weeks												

1	randomized trial	not serious	not assessable	not serious	serious [b]	none [a]	N= 136	N= 129	SMD 0.13 lower (0.37 lower to 0.11 higher)	⊕⊕⊕◯ MODERATE	IMPORTANT
Timed get up and go test *[Decreasing values indicate improvement]*: Follow-up at 12 weeks											
1	randomized trial	not serious	not assessable	not serious	not serious	none [a]	N= 153	N= 302	SMD 0.12 lower (0.32 lower to 0.07 higher)	⊕⊕⊕⊕ HIGH	IMPORTANT
Quality of Life (short term) *[Increasing values indicate improvement]*: Follow-up time ranged from 8 weeks to 12 weeks											
2	randomized trials	not serious	not serious	not serious	not serious	none [a]	N= 322	N= 471	SMD 0 (0.16 lower to 0.17 higher)	⊕⊕⊕⊕ HIGH	IMPORTANT
Quality of Life (Long term) *[Increasing values indicate improvement]*: Follow-up at 26 weeks											
1	randomized trial	not serious	not assessable	not serious	serious [b]	none [a]	N= 142	N= 141	SMD 0.14 higher (0.1 lower to 0.37 higher)	⊕⊕⊕◯ MODERATE	IMPORTANT
Treatment-related Adverse Events *[Risk ratios less than 1 favor Electroacupuncture]*: Follow-up time ranged from 12 weeks to 26 weeks											
2	randomized trials	not serious	not serious	not serious	serious [d]	none [a]	22/343 (6.4%)	31/493 (6.3%)	RR 1.40 (0.84 to 2.33)	25 more per 1,000 (from 10 fewer to 84 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Serious Adverse Events *[Risk ratios less than 1 favor Electroacupuncture]*: Follow-up at 26 weeks											
1	randomized trial	not serious	not assessable	not serious	not serious	none [a]	14/190 (7.4%)	5/191 (2.6%)	**RR 2.81 (1.03 to 7.66)**	**47 more per 1,000 (from 1 more to 174 more)**	⊕⊕⊕⊕ HIGH	IMPORTANT

CI: Confidence interval; **SMD:** Standardized mean difference; **RR:** Risk ratio

Explanations

a. See Supplementary Table for funding information

b. 95% Confidence Interval of an SMD extends between >0.2-≤0.5 points in either direction (Cohen 1988*)

c. I^2= 78%

d. 95% Confidence Interval of a Risk Ratio crosses null

Supplementary Table: funding information

Author, Year	Funding Statement	Industry?	Positive Result Pain?
Berman, 2004	"By the National Center for Complementary and Alternative Medicine (National Institutes of Health Cooperative Agreement U01 AT-00171), with advice and encouragement by the National Institute of Arthritis and Musculoskeletal and Skin Diseases" "Potential Financial Conflicts of Interest: None disclosed."	No	No
Suarez-Almazor, 2010	"Supported by the National Institute of Arthritis and Musculoskeletal and Skin Disorders (grant R01-AR49999)."	No	Negative result

GRADE tables for knee osteoarthritis

Section 2: Pharmacologic Interventions

PICO 2.1: ORAL ANALGESICS
PICO 2.1.1 (knee): What are the benefits and harms of paracetamol in the management of patients with knee OA?

SUMMARY
While paracetamol has long been considered first-line therapy for OA, this has mainly reflected its relative safety, availability and cost, compared with other pharmacological options (eg NSAIDs, opioids). Current evidence from a systematic review of randomised controlled trials (RCTs) suggests that, on average, the reduction in OA pain achieved with paracetamol is too small to be of clinical relevance.(Machado, et al. BMJ. 2015 Mar 31; 350: h1225). Moreover, paracetamol is associated with infrequent potential for significant harms, both short-term excess dosing and long-term regular use.(Roberts E, Delgado Nunes V, Buckner S, et al. Ann Rheum Dis 2016;75(3):552–59).

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Machado, et al. BMJ. 2015 Mar 31; 350: h1225.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Paracetamol	Placebo	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: range 2 weeks to 3 months)												
7	randomised trials [1]	not serious	not serious	serious [a]	not serious	none	N=1279	N=1076	**MD 3.7 lower (5.5 lower to 1.9 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
WOMAC Physical Function (scale range from 0 to 100, with higher score indicating poorer functional outcome) (follow up: range 2 weeks to 3 months)												
7	randomised trials [1]	not serious	not serious	serious [a]	not serious	none	N=1279	N=1076	**MD 2.9 lower (4.9 lower to 0.9 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Withdrawals due to Adverse Events (follow up: range 2 weeks to 3 months)												
7	randomised trials [1]	not serious	not serious	serious [a]	serious [b]	none	120/1630 (7.4%)	90/1393 (6.5%)	RR 1.2 (0.9 to 1.5)	13 more per 1,000 (from 6 fewer to 32 more)	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (follow up: range 2 weeks to 3 months)												

9	randomised trials [1]	not serious	serious [c]	serious [a]	serious [b]	none	735/2729 (26.9%)	613/2117 (29.0%)	RR 1.0 (0.9 to 1.1)	0 fewer per 1,000 (from 29 fewer to 29 more)	⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (follow up: range 2 weeks to 3 months)												
7	randomised trials 1	not serious	not serious	serious [a]	serious [b]	none	37/2825 (1.3%)	22/2027 (1.1%)	RR 1.2 (0.7 to 2.1)	2 more per 1,000 (from 3 fewer to 12 more)	⊕⊕□□ LOW	CRITICAL
Abnormal Liver Function (AST/ALN>1.5 ULN) (follow up: range 2 weeks to 3 months)												
3	randomised trials [1]	not serious	not serious	serious [a]	not serious	none	46/693 (6.6%)	10/544 (1.8%)	**RR 3.8 (1.9 to 7.4)**	**51 more per 1,000 (from 17 more to 118 more)**	⊕⊕⊕□ MODERATE	IMPORTANT

Pincus, 2004a; % knee: 84%, % hip: 16%
Pincus, 2004b; % knee: 84%, % hip: 16%
Altman, 2007; % knee: 81%, % hip: 19%
Prior, 2014; % knee: 82%, % hip: 18%
Zoppi, 1995; mixed population, no data

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Mixed population; all patients have osteoarthritis of the hip and/or knee
b. 95% CI crosses null.
c. I^2= 68%; moderate heterogeneity

PICO 2.1.2 (knee): What are the benefits and harms of oral NSAIDs including COX-2 inhibitors in the management of patients with knee OA?

SUMMARY
On average, the use of NSAIDs result in small but clinically relevant improvements in pain and function in individuals with knee and/or hip OA, and are likely to be more effective than paracetamol for most people. The direct costs of NSAIDs are relatively low. Evidence for effectiveness is derived from trials of relatively short duration, so the relative benefits versus harms of long-term NSAID therapy are unknown. It is likely that the risk of harms increases with duration of therapy; therefore, the balance of benefits and harms may become less favourable with time. Given the variable natural history of OA symptoms, periodic trials of drug withdrawal are recommended.

OVERALL QUALITY OF EVIDENCE: ⊕⊕⊕◻MODERATE

Bibliography: 1. Bensen, et al. Mayo Clin Proc. 1999 Nov; 74(11): 1095-105; 2. Boswell, et al. Medscape J Med. 2008; 10(11): 259; 3. Case, et al. Arch Intern Med. 2003 Jan 27; 163(2): 169-78; 4. Clegg, et al. N Engl J Med. 2006 Feb 23; 354(8): 795-808; 5. Conaghan, et al. Rheumatology (Oxford). 2013 Jul; 52(7): 1303-12; 6. Fleischmann, et al. Clin Rheumatol. 2006 Feb; 25(1): 42-53; 7. Hochberg, et al. Curr Med Res Opin. 2011 Jun; 27(6): 1243-53[1]; 8. Hochberg, et al. Curr Med Res Opin. 2011 Jun; 27(6): 1243-53[2]; 9. Kivitz, et al. J Fam Pract. 2002 Jun; 51(6): 530-7; 10. Lehmann, et al. Curr Med Res Opin. 2005 Apr; 21(4): 517-26; 11. Schnitzer, et al. Osteoarthritis Cartilage. 2010 May; 18(5): 629-39; 12. Schnitzer, et al. Semin Arthritis Rheum. 2011 Feb; 40(4): 285-97; 13. Sheldon, et al. Clin Ther. 2005 Jan; 27(1): 64-77; 14. Tannenbaum, et al. Ann Rheum Dis. 2004 Nov; 63(11): 1419-26.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Oral NSAID**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (Higher scores indicate higher pain severity) (follow up: range 12 weeks to 13 weeks)												
14	randomised trials [1-14]	not serious	not serious	not serious	not serious	none	N=6498	N=3145	**SMD 0.26 lower (0.31 lower to 0.22 lower)**		⊕⊕⊕⊕ HIGH	CRITICAL
WOMAC Function (Higher scores indicate poorer functional outcome) (follow up: range 12 weeks to 13 weeks)												
14	randomised trials [1-14]	not serious	not serious	not serious	not serious	none	N=6498	N=3145	**SMD 0.31 lower (0.36 lower to 0.26 lower)**		⊕⊕⊕⊕ HIGH	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Oral NSAID) (follow up: range 12 weeks to 13 weeks)												
13	randomised trials [1,3-14]	not serious	not serious	not serious	serious [a]	none	413/6298 (6.6%)	175/2958 (5.9%)	RR 1.01 (0.85 to 1.21)	1 more per 1,000 (from 9 fewer to 12 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Oral NSAID) (follow up: range 12 weeks to 13 weeks)												
12	randomised trials [1-2,5-14]	not serious	not serious	not serious	not serious	none	3327/6155 (54.1%)	1362/2805 (48.6%)	**RR 1.07 (1.01 to 1.13)**	**34 more per 1,000 (from 5 more to 63 more)**	⊕⊕⊕⊕ HIGH	CRITICAL

Serious Adverse Events (Risk ratios less than one favor Oral NSAID) (follow up: range 12 weeks to 13 weeks)												
10	randomised trials [2,5-8,10-14]	not serious	not serious	not serious	serious [a]	none	83/5398 (1.5%)	38/2424 (1.6%)	RR 0.85 (0.56 to 1.29)	2 fewer per 1,000 (from 5 more to 7 fewer)	⊕⊕⊕◻ MODERATE	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Oral NSAID) (follow up: range 12 weeks to 13 weeks)												
11	randomised trials [1-2,5-11,13-14]	not serious	not serious	not serious	not serious	none	1110/5898 (18.8%)	355/2544 (14.0%)	**RR 1.26 (1.13 to 1.40)**	**36 more per 1,000 (from 18 more to 56 more)**	⊕⊕⊕⊕ HIGH	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. 95% CI crosses null.

PICO 2.1.3 (knee): What are the benefits and harms of oral opioids in the management of patients with knee OA?

SUMMARY
Opioid prescription for chronic non-cancer pain (including OA) has increased in recent years despite a lack of high-quality evidence demonstrating benefit, particularly with long-term use. Evidence for the efficacy of opioids in the treatment of OA is mostly derived from short-term trials. There is moderate-quality evidence from trials that improvement in pain and function with opioids is of marginal clinical significance at best, and is offset by the risk of harms.
.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. Afilalo, et al. Clin Drug Investig. 2010; 30(8): 489-505; 2. Babul, et al. J Pain Symptom Manage. 2004 Jul; 28(1): 59-71; 3. Fleischmann, et al. Current Therapeutic Research 62(2) (2001): 113-128; 4. NCT00486811.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Oral opioid†**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: 12 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	serious [b]	not serious	not serious	none	N=1523	N=862	**SMD 0.21 lower (0.35 lower to 0.07 lower)**		⊕⊕□□ LOW	CRITICAL
WOMAC Function (scale range 0 to 100, with higher scores indicating poorer functional outcome) (follow up: 12 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	not serious	none	N=241	N=158	**MD 0.21 lower (0.35 lower to 0.07 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
EuroQoL-5 Health Status Index (scale range 0 to 1, with higher scores indicating better quality of life) (follow up: 12 weeks)												
2	randomised trials [1,4]	serious [a]	very serious [c]	not serious	serious [d]	none	N=1336	N=674	MD 0.01 lower (0.08 lower to 0.07 higher)		⊕□□□ VERY LOW	IMPORTANT
% Patients Experiencing Opioid Withdrawal ≥5 days after last medication intake (As evidenced by a score between 13-36 on the Clinical Opiate Withdrawal Scale [COWS; score >36 indicative of severe withdrawal) (Risk ratios less than one favor Oral Opioid) (follow up: 12 weeks)												

1	randomised trial [1]	not serious	not assessable	not serious	serious [d]	none	13/154 (8.4%)	5/59 (8.5%)	RR 1.00 (0.37 to 2.67)	0 fewer per 1,000 (from 53 fewer to 142 more)	⊕⊕⊕◻ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Oral Opioid) (follow up: 12 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	serious [e]	not serious	not serious	none	454/1526 (29.8%)	69/862 (8.0%)	**RR 3.31 (2.19 to 5.02)**	**185 more per 1,000 (from 95 more to 322 more)**	⊕⊕◻◻ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Oral Opioid) (follow up: 12 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	not serious	not serious	not serious	none	1153/1523 (75.7%)	471/862 (54.6%)	**RR 1.33 (1.25 to 1.42)**	**180 more per 1,000 (from 137 more to 229 more)**	⊕⊕⊕◻ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Oral Opioid) (follow up: 12 weeks)												
3	randomised trials [1,3,4]	not serious	not serious	not serious	serious [d]	none	29/1399 (2.1%)	12/740 (1.6%)	RR 1.30 (0.65 to 2.60)	5 more per 1,000 (from 6 fewer to 26 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Oral Opioid) (follow up: 12 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	serious [f]	not serious	not serious	none	1065/1523 (69.9%)	226/862 (26.2%)	**RR 2.67 (2.07 to 3.45)**	**438 more per 1,000 (from 281 more to 642 more)**	⊕⊕◻◻ LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio; **MD:** Mean difference

† **Different oral opioids on various dosing regimens were analyzed together.** The following were included in the analysis: Afilalo, 2010- Controlled, adjustable, oral doses of Tapentadol ER 100–250 mg BID or controlled, adjustable, oral doses of oxycodone HCl CR 20–50 mg BID; Babul, 2004- Oral Tramadol extended release (ER) 100 mg/day increased to 200 mg/day (between days 4 and 8; further increases to 300 or 400 mg/day were allowed after week 1); Fleischmann, 2001- Oral Tramadol 50 mg/day increased by 50 mg/day on 2 day increments to a target dose of 200 mg/day; once the target dose was reached, patients could take up to 400 mg/day for the remainder of the double-blind period; NCT00486811- Tapentadol (CG5503) extended-release (ER)100-250 mg BID or Oxycodone controlled-release (CR) 20-50 mg BID.

a. All trials received High risk of bias ratings for potential attrition bias; potential reporting bias in two trials
b. I^2= 63%; moderate heterogeneity
c. I^2= 100%
d. 95% CI crosses null

e. I^2=60%; moderate heterogeneity
f. I^2= 67%; moderate heterogeneity
.

PICO 2.2: TOPICAL ANALGESICS
PICO 2.2.1 (knee): What are the benefits and harms of topical NSAIDs in the management of patients with knee OA?

SUMMARY
The effectiveness of topical NSAID application in OA is variable. Generally, the benefit is small, but the risk of harm is also small.

OVERALL QUALITY OF EVIDENCE: ⊕⊕⊕◻MODERATE

Bibliography: 1. Baraf, et al. Phys Sportsmed. 2010 Jun; 38(2): 19-28; 2. Barthel, et al. Semin Arthritis Rheum. 2009 Dec; 39(3): 203-12; 3. Conaghan, et al. Rheumatology (Oxford). 2013 Jul; 52(7): 1303-12; 4. Kneer, et al. J Pain Res. 2013 Oct 25; 6: 743-53; 5. Roth and Shainhouse. Arch Intern Med. 2004 Oct 11; 164(18): 2017-23; 6. Rother, et al. J Rheumatol. 2013 Oct; 40(10): 1742-8; 7. Simon, et al. Pain. 2009 Jun; 143(3): 238-45.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Topical NSAID**	**Vehicle Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (Higher scores indicate higher pain severity) (follow up: 12 weeks)												
7	randomised trials [1,2,3,4,5,6,7]	not serious	not serious	not serious	not serious	none	N=2016	N=1559	**SMD 0.2 lower (0.29 lower to 0.11 lower)**		⊕⊕⊕⊕ HIGH	CRITICAL
WOMAC Function (Higher scores indicate poorer functional outcome) (follow up: 12 weeks)												
7	randomised trials [1,2,3,4,5,6,7]	not serious	not serious	not serious	not serious	none	N=2015	N=1559	**SMD 0.19 lower (0.28 lower to 0.1 lower)**		⊕⊕⊕⊕ HIGH	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Topical NSAID) (follow up: 12 weeks)												
7	randomised trials [1,2,3,4,5,6,7]	not serious	not serious	not serious	not serious	none	130/2048 (6.3%)	71/1571 (4.5%)	**RR 1.36 (1.02 to 1.82)**	**16 more per 1,000 (from 1 more to 37 more)**	⊕⊕⊕⊕ HIGH	CRITICAL
Total Adverse Events (Risk ratios less than one favor Topical NSAID) (follow up: 12 weeks)												
6	randomised trials [1,2,3,4,6,7]	not serious	not serious	not serious	serious [a]	none	963/1884 (51.1%)	685/1409 (48.6%)	RR 1.07 (0.98 to 1.16)	34 more per 1,000 (from 10 fewer to 78 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Topical NSAID) (follow up: 12 weeks)												

5	randomised trials [1,2,3,5,7]	not serious	not serious	not serious	serious [a]	none	12/1243 (1.0%)	10/1245 (0.8%)	RR 1.08 (0.26 to 4.47)	1 more per 1,000 (from 6 fewer to 28 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Topical NSAID) (follow up: 12 weeks)												
7	randomised trials [1,2,3,4,5,6,7]	not serious	not serious	not serious	serious [a]	none	96/2048 (4.7%)	84/1571 (5.3%)	RR 0.94 (0.70 to 1.25)	3 fewer per 1,000 (from 13 more to 16 fewer)	⊕⊕⊕◻ MODERATE	IMPORTANT
Local Reactions (local reactions included dermatitis, skin dryness, eczema, exanthema, erythema, papules, pruritus, itching, dermatosis, allergic reaction, parasthesia (Roth 2004 only), and/or rash near the application site of topical NSAID or Vehicle control) (Risk ratios less than one favor Topical NSAID) (follow up: 12 weeks)												
7	randomised trials [1,2,3,4,5,6,7]	not serious	not serious	not serious	not serious	none	354/2048 (17.3%)	193/1571 (12.3%)	**RR 1.32 (1.04 to 1.67)**	**39 more per 1,000 (from 5 more to 82 more)**	⊕⊕⊕⊕ HIGH	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. 95% CI crosses null

PICO 2.2.2 (knee): What are the benefits and harms of transdermal opioids in the management of patients with knee OA?

SUMMARY
Evidence is mostly derived from two short-term trials of transdermal opioids: buprenorphine and fentanyl. Similar to oral opioid, the low-quality evidence demonstrated the improvements in pain and function are of marginal clinical significance at best,and is offset by the risk of harms

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

TRANSDERMAL BUPRENORPHINE

Bibliography: 1. da Costa, Bruno R., et al. "Oral or transdermal opioids for osteoarthritis of the knee or hip." The Cochrane Library (2014).

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Transdermal Buprenorphine	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 4 weeks to 30 weeks)												
4	randomised trials [1]	serious [a]	not serious	serious [b]	not serious	none	N=691	N=710	**SMD 0.19 lower (0.3 lower to 0.09 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (multiple measures including WOMAC function, "participant global assessment") (follow up: range 4 weeks to 28 weeks)												
2	randomised trials [1]	serious [a]	not serious	serious [b]	not serious	none	N=243	N=258	**SMD 0.23 lower (0.4 lower to 0.05 lower)**		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Transdermal Buprenorphine) (follow up: range 4 weeks to 30 weeks)												
4	randomised trials [1]	serious [a]	serious [c]	not serious	not serious	none	119/698 (17.0%)	51/709 (7.2%)	**RR 3.10 (1.38 to 6.94)**	**151 more per 1,000 (from 27 more to 427 more)**	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Transdermal Buprenorphine) (follow up: 28 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	not serious	none	92/100 (92.0%)	73/99 (73.7%)	**RR 1.25 (1.09 to 1.42)**	**184 more per 1,000 (from 66 more to 310 more)**	⊕⊕⊕□ MODERATE	CRITICAL

Breivik, 2010; % knee: 63%, % hip: 37% Munera, 2010; mixed population, no data Shannon, 2005; mixed population, no data

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. All studies received at least one High risk of bias rating due to inadequate masking of interventions or non-ITT analysis

b. All studies involved both Knee and Hip OA patients; NCT00531427 was only in Knee OA patients

c. I^2= 74%; moderate heterogeneity

TRANSDERMAL FENTANYL

Bibliography: 1. da Costa, Bruno R., et al. "Oral or transdermal opioids for osteoarthritis of the knee or hip." The Cochrane Library (2014).

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Transdermal Fentanyl**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	not serious	none	N=202	N=197	**SMD 0.22 lower (0.42 lower to 0.03 lower)**		⊕⊕◻◻ LOW	CRITICAL
WOMAC Function (Higher scores indicate poorer functional outcome) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	not serious	none	N=202	N=197	**SMD 0.28 lower (0.48 lower to 0.09 lower)**		⊕⊕◻◻ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Transdermal Fentanyl) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	not serious	none	54/202 (26.7%)	20/197 (10.2%)	**RR 2.63 (1.64 to 4.23)**	**165 more per 1,000 (from 65 more to 328 more)**	⊕⊕⊕◻ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Transdermal Fentanyl) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	not serious	none	169/216 (78.2%)	101/200 (50.5%)	**RR 1.55 (1.33 to 1.81)**	**278 more per 1,000 (from 167 more to 409 more)**	⊕⊕⊕◻ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Transdermal Fentanyl) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	6/216 (2.8%)	2/200 (1.0%)	RR 2.78 (0.57 to 13.60)	18 more per 1,000 (from 4 fewer to 126 more)	⊕⊕◻◻ LOW	CRITICAL

Langford, 2006; % knee: 53%, % hip: 47%

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. Study received High risk of bias ratings for inadequate masking of interventions and non-ITT analyses

b. Study involves both knee and hip OA patients

c. 95% CI crosses null.

PICO 2.2.3 (knee): What are the benefits and harms of topical capsaicin in the management of patients with knee OA?

SUMMARY
Evidence from one trial demonstrated that 0.025% of topical capsaicin had small effects of pain relief in people with knee OA. It is uncertain whether individuals with multi-joint OA or with relevant comorbidities will benefit from capsaicin. The principle benefit of capsaicin is in neuropathic pain, which is not the major pain source in knee or hip OA. Similar to other topical analgesia, the topical application process is very regime orientated, and local irritation side effects can be detrimental. These issues often outweigh possible benefits to individuals.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: Kosuwon, et al. J Med Assoc Thai. 2010 Oct; 93(10): 1188-95.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Topical capsaicin**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
VAS Pain (Mean change from baseline, scale range 0 to 10, with higher scores indicating higher pain severity**)** (follow up: 4 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious [a,b]	none	N=65	N=34	MD 0.27 higher (0.33 lower to 0.87 higher)		⊕⊕□□ LOW	CRITICAL
Total WOMAC (scale range 0 to 100, with higher scores indicating poorer outcomes) (follow-up: 4 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [b]	none	N=65	N=34	**MD 6.75 lower (12.95 lower to 0.55 lower)**		⊕⊕⊕□ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Topical capsaicin) (follow up: 8 weeks, a.k.a two treatment periods in cross-over)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [b]	none	0/99 (0.0%)	0/99 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕□ MODERATE	CRITICAL
Burning sensation (N episodes reported per group/Total N episodes reported) (Risk ratios less than one favor Topical capsaicin) (follow up: 8 weeks, a.k.a two treatment periods in cross-over)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [c]	none	272/338 (80.5%)	66/338 (19.5%)	**RR 4.12 (3.30 to 5.15)**	**609 more per 1,000 (from 449 more to 810 more)**	⊕⊕⊕□ MODERATE	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. 95% CI crosses null
b. Sample size <50 in one study arm
c. Authors chose to report skin irritation as the proportion of N total burning sensation events in each group.

PICO 2.3: HERBAL THERAPIES

PICO 2.3.1 (knee): What are the benefits and harms of avocado soybean unsaponifiable in the management of patients with knee OA?

SUMMARY

The 2014 Cochrane review reports ASU 300 mg produced a small and clinically questionable improvement in symptoms, and probably no increased adverse events, compared with placebo after three to 12 months treatment. (Cameron M, Chrubasik S. Cochrane Database Syst Rev, 2014;22(5):CD002947). In the new evidence review for this guideline, short-term pain and function up to six months was improved by about 0.5 standard deviations, and there were no significant longer-term benefits in pain or function. Studies examining the use of supplements are often of low quality constrained by small sample sizes, industry publication bias and potential for positive publication bias. In the context of low-quality to very low-quality studies, despite some suggestion of beneficial effects, it is prudent to use caution and advocate that further research is needed before a firm recommendation on ASU can be made.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Appelboom, et al. Scand J Rheumatol. 2001; 30(4): 242-7; 2. Blotman, et al. Rev Rhum Engl Ed. 1997 Dec; 64(12): 825-34; 3. Maheu, et al. Arthritis Rheum. 1998 Jan; 41(1): 81-91; 4. Lequesne, et al. Arthritis Rheum. 2002 Feb; 47(1): 50-8; 5. Maheu, et al. Ann Rheum Dis. 2014 Feb; 73(2): 376-84; 6. Liu, et al. Osteoarthritis and Cartilage. 2017 Apr 1; 25: S292-3.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Avocado Soybean Unsaponifiable	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain [short term] (Higher scores indicate higher pain severity) (follow up: 90 days)												
2	randomised trials [1,2]	serious [a]	serious [b]	serious [c]	not serious	none	N=243	N=161	**SMD 0.57 lower (0.95 lower to 0.19 lower)**		⊕□□□ VERY LOW	CRITICAL
Pain [moderate term] (Higher scores indicate higher pain severity) (follow up: 6 months)												
1	randomised trial [3]	not serious	not assessable	serious [c]	not serious	none	N=84	N=78	**SMD 0.45 lower (0.77 lower to 0.14 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Pain [long term] (Higher scores indicate higher pain severity) (follow up: range 2 years to 3 years)												
2	randomised trials [4,5]	not serious	not serious	serious [d]	serious [e]	none	N=251	N=257	SMD 0.04 higher (0.14 lower to 0.21 higher)		⊕⊕□□ LOW	CRITICAL
Function [short term] (Higher scores indicate poorer functional outcome) (follow up: 90 days)												

2	randomised trials [1,2]	serious [a]	not serious	serious [c]	not serious	none	N=234	N=161	**SMD 0.48 lower (0.69 lower to 0.28 lower)**	⊕⊕□□ LOW	CRITICAL
Function [moderate term] (Higher scores indicate poorer functional outcome) (follow up: 6 months)											
1	randomised trial [3]	not serious	not assessable	serious [c]	not serious	none	N=84	N=78	**SMD 0.58 lower (0.94 lower to 0.23 lower)**	⊕⊕⊕□ MODERATE	CRITICAL
Function [long term] (Higher scores indicate poorer functional outcome) (follow up: range 2 years to 3 years)											
2	randomised trials [4,5]	not serious	not serious	serious [d]	serious [e]	none	N=251	N=257	SMD 0.03 lower (0.21 lower to 0.14 higher)	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Avocado Soybean Unsaponifiables) (follow up: range 90 days to 3 years)											
5	randomised trials [1,2,3,4,5]	not serious	not serious	serious [c]	not serious	none	291/610 (47.7%)	270/537 (50.3%)	**RR 1.0 (1.0 to 1.1)**	**0 fewer per 1,000 (from 0 fewer to 50 more)**	⊕⊕⊕□ MODERATE	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. One study received at least one High risk of bias rating.
b. I^2= 69%; moderate heterogeneity.
c. Mixed population; included trials involve patients with Knee and/or Hip Osteoarthritis.
d. All patients in both trials have Hip Osteoarthritis
e. 95% CI crosses null.

PICO 2.3.2 (knee): What are the benefits and harms of boswellia serrata in the management of patients with knee OA?

SUMMARY
Three small RCTs found significant short-term benefits in pain and function; however, these are all sponsored by the same company, raising concern about possible bias. In the context of low-quality to very low-quality studies, despite potential beneficial effects, it is prudent to use caution and advocate that further research is needed before a firm recommendation on *Boswellia serrata* can be made.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Sengupta, et al. Arthritis Res Ther. 2008; 10(4): R85; 2. Sengupta, et al. Int J Med Sci 2010; 7(6): 366-377; 3. Vishal, et al. Int J Med Sci. 2011; 8(7): 615-22.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Boswellia serrata extract**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 30 days to 90 days)												
3	randomised trials [1,2,3]	serious [a]	not serious	not serious	not serious	none	N=115	N=71	**SMD 1.61 lower (2.1 lower to 1.13 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 30 days to 90 days)												
3	randomised trials [1,2,3]	serious [a]	serious [b]	not serious	not serious	none	N=115	N=71	**SMD 1.15 lower (1.63 lower to 0.68 lower)**		⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Boswellia serrata extract) (follow up: range 30 days to 90 days)												
2	randomised trials [2,3]	serious [a]	not serious	not serious	serious [c]	none	2/68 (2.9%)	2/49 (4.1%)	RR 0.70 (0.10 to 4.80)	12 fewer per 1,000 (from 37 fewer to 155 more)	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. One study reporting this outcome received a High risk of bias rating; another study received primarily "Unclear" risk of bias ratings.
b. I^2= 51%; moderate heterogeneity.
c. 95% CI crosses null.
d. Sample size in one study arm <50.

PICO 2.3.3 (knee): What are the benefits and harms of curcuma in the management of patients with knee OA?

SUMMARY
Three small RCTs found significant short-term (ie six to eight weeks) benefits in pain and function; however, these are all industry-sponsored trials, raising concern about possible bias. Additionally, there were inconsistency in the results. All of the studies involved knee OA, so extrapolation to hip or other OA requires additional caution. In the context of low-quality to very low-quality studies, despite potential beneficial effects, it is prudent to use caution and advocate that further research is needed before a firm recommendation on curcuma can be made.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□ LOW

Bibliography: 1. Madhu, et al. Inflammopharmacology. 2013 Apr; 21(2): 129-36; 2. Nakagawa, et al. J Orthop Sci. 2014 Nov; 19(6): 933-9; 3. Panahi, et al. Phytother Res. 2014 Nov; 28(11): 1625-31.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Curcuminoid**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 6 weeks to 8 weeks)												
3	randomised trials [1,2,3]	serious [a]	serious [b]	not serious	not serious	none	N=63	N=70	**SMD 1.1 lower (1.66 lower to 0.54 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: 6 weeks)												
1	randomised trial [3]	serious [a]	not assessable	not serious	serious [c]	none	N=19	N=21	**SMD 0.81 lower (1.46 lower to 0.16 lower)**		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Curcuminoid) (follow up: range 6 weeks to 8 weeks)												
3	randomised trials [1,2,3]	serious [a]	not serious	not serious	serious [d]	none	4/75 (5.3%)	3/79 (3.8%)	RR 1.46 (0.34 to 6.31)	17 more per 1,000 (from 25 fewer to 202 more)	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Curcuminoid) (follow up: range 6 weeks to 8 weeks)												
2	randomised trials [1,3]	serious [a]	not serious	not serious	serious [d]	none	9/57 (15.8%)	6/56 (10.7%)	RR 1.48 (0.57 to 3.83)	51 more per 1,000 (from 46 fewer to 303 more)	⊕⊕□□ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Curcuminoid) (follow up: range 6 weeks to 8 weeks)												
2	randomised trials [2,3]	serious [a]	not serious	not serious	not serious	none	0/52 (0.0%)	0/51 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕□ MODERATE	CRITICAL

Gastrointestinal Adverse Events (Risk ratios less than one favor Curcuminoid) (follow up: range 6 weeks to 8 weeks)												
2	randomised trials [1,3]	serious [a]	not serious	not serious	serious [d]	none	9/57 (15.8%)	4/56 (7.1%)	RR 1.92 (0.68 to 5.41)	66 more per 1,000 (from 23 fewer to 315 more)	⊕⊕◻◻ LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. All trials received at least one High risk of bias rating due to single blind study design, potential attrition bias, or potential reporting bias.
b. I^2= 56%; moderate heterogeneity.
c. Sample size <50 in each study arm.
d. 95% CI crosses null.

PICO 2.3.4 (knee): What are the benefits and harms of pine bark extract in the management of patients with knee OA?

SUMMARY
Three small RCTs found short-term benefits in pain and function; however, these could not be pooled because of heterogeneity and reporting weaknesses. All three trials were industry-sponsored, with the larger trial at very high risk of bias. Evidence is based on studies of knee OA, so extrapolation to hip or other OA requires additional caution. In the context of low-quality to very low-quality studies, despite potential beneficial effects, it is prudent to use caution and advocate that further research is needed before a firm recommendation on the use of pine bark extract.

OVERALL QUALITY OF EVIDENCE: ⊕⊕◻◻ LOW

Bibliography: 1.Belcaro, et al. Phytotherapy Research. 2008 Apr 1; 22(4): 518-23; 2. Farid, et al. Nutrition Research. 2007 Nov 30; 27(11): 692-7; 3. Cisár, et al. Phytother Res. 2008 Aug; 22(8): 1087-92.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Pycnogenol	Placebo	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (Scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 3 months)												
1	randomised trial [1]	very serious [a]	not assessable [b]	not serious	not serious	none	N=71	N=74	**MD 7.7 lower (8.24 lower to 7.16 lower)**		⊕⊕◻◻ LOW	CRITICAL
WOMAC Pain VAS (Scale range 0 to 500 mm, with higher scores indicating higher pain severity) (follow up: 3 months)												
1	randomised trial [2]	not serious	not assessable [b]	not serious	serious [c]	none	N=19	N=18	**MD 133 mm lower (198.66 mm lower to 67.34 mm lower)**		⊕⊕⊕◻ MODERATE	CRITICAL
WOMAC Function (Scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 3 months)												
1	randomised trial [1]	very serious [a]	not assessable [b]	not serious	not serious	none	N=71	N=74	**MD 26 lower (26.49 lower to 25.51 lower)**		⊕⊕◻◻ LOW	CRITICAL
WOMAC Function VAS (Scale range 0 to 1700 mm, with higher scores indicating poorer functional outcome) (follow up: 3 months)												
1	randomised trial [2]	not serious	not assessable [b]	not serious	serious [c]	none	N=19	N=18	**MD 484 mm lower (718.42 mm lower to 249.58 mm lower)**		⊕⊕⊕◻ MODERATE	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Pycnogenol) (follow up: 3 months)												
2	randomised trials [2,3]	serious [d]	not serious	not serious	serious [e]	none	2/69 (2.9%)	4/68 (5.9%)	RR 0.50 (0.10 to 2.61)	29 fewer per 1,000 (from 53 fewer to 95 more)	⊕⊕◻◻ LOW	CRITICAL

Total Adverse Events (Risk ratios less than one favor Pycnogenol) (follow up: 3 months)											
2	randomised trials [2,3]	serious [d]	not serious	not serious	serious [e]	none	2/69 (2.9%)	4/68 (5.9%)	RR 0.50 (0.10 to 2.61)	29 fewer per 1,000 (from 53 fewer to 95 more)	⊕⊕◻◻ LOW	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Pycnogenol) (follow up: 3 months)											
2	randomised trials [2,3]	serious [d]	not serious	not serious	serious [e]	none	1/69 (1.4%)	1/68 (1.5%)	RR 1.00 (0.06 to 15.55)	0 fewer per 1,000 (from 14 fewer to 214 more)	⊕⊕◻◻ LOW	IMPORTANT
Serious Adverse Events (Risk ratios less than one favor Pycnogenol) (follow up: 3 months)											
1	randomised trial [3]	serious [d]	not assessable	not serious	not serious	none	0/50 (0.0%)	0/50 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕◻ MODERATE	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Study received High risk of bias rating due to potential reporting bias; inadequate reporting of critical efficacy outcomes and no Adverse Event reporting, despite description of collection in study methods.
b. A standardised mean difference could not be provided due to excessive heterogeneity in efficacy reporting between two eligible studies.
c. Sample size<50 in each study arm.
d. One study received a High risk of bias rating, as well as more than one Unclear risk of bias rating, in important categories.
e. 95% CI crosses null.

PICO 2.4: NUTRACEUTICALS
PICO 2.4.1 (knee): What are the benefits and harms of glucosamine in the management of patients with knee OA?

SUMMARY
Overall, there is very low-quality evidence from a large number of randomised controlled trials (RCTs) that found that glucosamine provides some benefits to pain in the short term, but no apparent benefits to function, quality of life or joint space narrowing. When the studies are restricted to higher quality trials, no benefit is demonstrated. There are concerns in the literature of publication bias, effects being driven by small industry-sponsored trials, and the overall poor quality of the positive trials. Larger publicly funded trials generally demonstrate no effect over placebo. High-quality trial data suggest no effect.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Frestedt, et al. Nutr J. 2008 Feb 17; 7: 9; 2. Giordano, et al. Curr Ther Res Clin Exp. 2009 Jun; 70(3): 185-96;3. Kwoh, et al. Arthritis Rheumatol. 2014 Apr; 66(4): 930-9; 4. McAlindon, et al. ; 5. Usha, P.R., Naidu, M.U.R. Clin Drug Investig. 2004; 24(6): 353-63; 6. Clegg, et al. N Engl J Med. 2006 Feb 23; 354(8): 795-808 (GAIT); 7. Fransen, et al. Ann Rheum Dis. 2015 May; 74(5): 851-8; 8. Herrero-Beaumont, et al. Arthritis Rheum. 2007 Feb; 56(2): 555-67; 9. Hughes, R., Carr, A. Rheumatology (Oxford). 2002 Mar; 41(3): 279-84; 10. Pavelká, et al. Arch Intern Med. 2002 Oct 14; 162(18): 2113-23; 11. Reginster, et al. Lancet. 2001 Jan 27; 357(9252): 251-6; 12. Sawitzke, et al. Annals of the rheumatic diseases. 2010 Aug 1; 69(8): 1459-64 (GAIT); 13. Sawitzke, et al. Arthritis & Rheumatology. 2008 Oct 1; 58(10): 3183-91.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Glucosamine	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain [short term] (Higher scores indicate higher pain severity) (follow up: 3 months)												
5	randomised trials [1,2,3,4,5]	serious [a]	very serious [b]	not serious	not serious	none	N=274	N=281	**SMD 0.83 lower (1.55 lower to 0.11 lower)**		⊕□□□ VERY LOW	CRITICAL
Pain [moderate term] (Higher scores indicate higher pain severity) (follow up: range 6 months to 12 months)												
5	randomised trials [3,6,7,8,9]	not serious	not serious	not serious	serious [c]	none	N=712	N=710	SMD 0 (0.12 lower to 0.12 higher)		⊕⊕⊕□ MODERATE	CRITICAL
Pain [long term] (Higher scores indicate higher pain severity) (follow up: range 24 months to 36 months)												
4	randomised trials [7,10,11,12]	not serious	not serious	not serious	serious [c]	none	N=493	N=489	SMD 0.14 lower (0.3 lower to 0.03 higher)		⊕⊕⊕□ MODERATE	CRITICAL
Function [short term] (Higher scores indicate poorer functional outcome) (follow up: 3 months)												
5	randomised trials [1,2,3,4,5]	serious [a]	very serious [d]	not serious	serious [c]	none	N=274	N=281	SMD 0.57 lower (1.18 lower to 0.04 higher)		⊕□□□ VERY LOW	CRITICAL
Function [moderate term] (Higher scores indicate poorer functional outcome) (follow up: range 6 months to 12 months)												

5	randomised trials [3,6,7,8,9]	not serious	not serious	not serious	serious [c]	none	N=712	N=710	SMD 0.02 lower (0.18 lower to 0.14 higher)		⊕⊕⊕◻ MODERATE	CRITICAL
Function [long term] (Higher scores indicate poorer functional outcome) (follow up: range 24 months to 36 months)												
4	randomised trials [7,10,11,12]	not serious	serious [e]	not serious	serious [c]	none	N=493	N=489	SMD 0.14 lower (0.32 lower to 0.05 higher)		⊕⊕◻◻ LOW	CRITICAL
Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: range 24 months to 36 months)												
4	randomised trials [7,10,11,13]	not serious	serious [f]	not serious	serious [c]	none	N=409	N=398	MD 0.13 higher (0.02 lower to 0.29 higher)		⊕⊕◻◻ LOW	IMPORTANT
SF-12 Physical Component Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 24 months)												
1	randomised trial [7]	not serious	not assessable	not serious	serious [c]	none	N=152	N=151	MD 0.9 higher (1.26 lower to 3.06 higher)		⊕⊕⊕◻ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Glucosamine) (follow up: range 3 months to 36 months)												
11	randomised trials [1-11]	not serious	not serious	not serious	serious [c]	none	60/1096 (5.5%)	69/1096 (6.3%)	RR 0.89 (0.64 to 1.24)	7 fewer per 1,000 (from 15 more to 23 fewer)	⊕⊕⊕◻ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Glucosamine) (follow up: range 3 months to 36 months)												
8	randomised trials [1,2,3,4,7,9,10,11]	not serious	not serious	not serious	serious [c]	none	254/643 (39.5%)	255/651 (39.2%)	RR 1.01 (0.95 to 1.07)	4 more per 1,000 (from 20 fewer to 27 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Glucosamine) (follow up: range 3 months to 24 months)												
5	randomised trials [2,4,5,9,12]	serious [g]	not serious	not serious	serious [c]	none	4/335 (1.2%)	5/333 (1.5%)	RR 0.82 (0.23 to 3.00)	3 fewer per 1,000 (from 12 fewer to 30 more)	⊕⊕◻◻ LOW	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Glucosamine) (follow up: range 3 months to 36 months)												
8	randomised trials [1,2,4,7,8,9,10,11]	not serious	not serious	not serious	serious [c]	none	85/651 (13.1%)	104/652 (16.0%)	RR 0.82 (0.64 to 1.05)	29 fewer per 1,000 (from 8 more to 57 fewer)	⊕⊕⊕◻ MODERATE	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. 3 of 5 trials received at least one High risk of bias rating.

b. I^2= 93%; excluding studies which received quality downgrades reduces I^2 to 0%, with an SMD (95% CI) of -0.07 (-0.26, 0.13).

c. 95% CI crosses null.
d. I^2= 91%; excluding studies which received quality downgrades reduces I^2 to 0%, with an SMD (95% CI) of 0.05 (-0.15, 0.24).
e. I^2= 53%; moderate heterogeneity.
f. I^2= 74%; moderate heterogeneity.
g. 4 of 5 trials received at least one High risk of bias rating.

PICO 2.4.2 (knee): What are the benefits and harms of chondroitin in the management of patients with knee OA?

SUMMARY
There are a large number of trials on the use of chondroitin where at least seven are industry sponsored. When all 16 studies are pooled, there is a clinically and statistically significant effect on pain and function in the short-term (up to three months), which lessens to clinically not significant by six to 12 months, and no effect is demonstrated at 24 months. However, when the analysis is restricted to studies of higher quality or free of industry sponsorship, no benefit is demonstrated. There are some moderate-term to long-term (12–24 months) benefits on joint space narrowing, but these are not clinically meaningful. The studies are all on participants with knee OA, so extrapolation to OA of hip or other joints requires further caution.
There are concerns in the literature of publication bias, effects being driven by small industry-sponsored trials, and the overall poor quality of the positive trials. High-quality trial data suggest no effect.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1.Bourgeois, et al. Osteoarthritis Cartilage. 1998 May; 6 Suppl A: 25-30; 2. Mazieres, et al. J Rheumatol. 2001 Jan; 28(1): 173-81; 3. Pavelká, et al. Litera Rheumatologica. 1999; 24: 21-30; 4. Zegels, et al. Osteoarthritis Cartilage. 2013 Jan; 21(1): 22-7; 5. Reginster, et al. Ann Rheum Dis. 2017 May 22. pii: annrheumdis-2016-210860; 6. Bucsi, L. and Poór, G. Osteoarthritis Cartilage. 1998 May; 6 Suppl A: 31-6 ; 7. Clegg, et al. N Engl J Med. 2006 Feb 23; 354(8): 795-808 (GAIT); 8. Fransen, et al. Ann Rheum Dis. 2015 May; 74(5): 851-8; 9. Kahan, et al. Arthritis Rheum. 2009 Feb; 60(2): 524-33; 10. Railhac, et al. Clin Rheumatol. 2012 Sep; 31(9): 1347-57; 11. Uebelhart, et al. Osteoarthritis Cartilage. 1998 May; 6 Suppl A: 39-46; 12. Uebelhart, et al. Osteoarthritis Cartilage. 2004 Apr; 12(4): 269-76 ; 13. Wildi, et al. Ann Rheum Dis. 2011 Jun; 70(6): 982-9; 14. Michel, et al. Arthritis Rheum. 2005 Mar; 52(3): 779-86; 15. Sawitzke, et al. Annals of the rheumatic diseases. 2010 Aug 1; 69(8): 1459-64 (GAIT); 16. Sawitzke, et al. Arthritis & Rheumatology. 2008 Oct 1; 58(10): 3183-91.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Chondroitin**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain [short term] (Higher scores indicate higher pain severity) (follow up: 3 months)												
5	randomised trials 1,2,3,4,5	serious [a]	serious [b]	not serious	not serious	none	N=666	N=451	**SMD 0.63 lower (0.91 lower to 0.36 lower)**		⊕⊕□□ LOW	CRITICAL
Pain [moderate term] (Higher scores indicate higher pain severity) (follow up: range 6 months to 12 months)												
9	randomised trials 5,6,7,8,9,10,11,12,13	not serious	very serious [c]	not serious	not serious	none	N=1109	N=1127	**SMD 0.28 lower (0.49 lower to 0.06 lower)**		⊕⊕□□ LOW	CRITICAL
Pain [long term] (Higher scores indicate higher pain severity) (follow up: 24 months)												
4	randomised trials 8,9,14,15	not serious	not serious	not serious	serious [d]	none	N=736	N=745	SMD 0.03 lower (0.13 lower to 0.07 higher)		⊕⊕⊕□ MODERATE	CRITICAL
Function [short term] (Higher scores indicate poorer functional outcome) (follow up: 3 months)												
5	randomised trials 1,2,3,4,5	serious [a]	not serious	not serious	not serious	none	N=666	N=451	**SMD 0.55 lower (0.78 lower to 0.33 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Function [moderate term] (Higher scores indicate poorer functional outcome) (follow up: range 6 months to 12 months)												
6	randomised trials 5,6,7,8,10,11	not serious	very serious [e]	not serious	serious [d]	none	N=711	N=724	**SMD 0.33 lower (0.62 lower to 0.04 lower)**		⊕□□□ VERY LOW	CRITICAL

Function [long term] (Higher scores indicate poorer functional outcome) (follow up: 24 months)											
3	randomised trials [8,14,15]	not serious	not serious	not serious	serious [d]	none	N=427	N=432	SMD 0.04 lower (0.17 lower to 0.1 higher)	⊕⊕⊕◻ MODERATE	CRITICAL
Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: range 12 months to 2 years)											
6	randomised trials [8,9,11,12,14,16]	not serious	serious [f]	not serious	not serious	none	N=635	N=646	**MD 0.16 mm higher (0.03 mm higher to 0.28 mm higher)**	⊕⊕⊕◻ MODERATE	IMPORTANT
SF-12 Physical Component Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 2 years)											
1	randomised trials [8]	not serious	not assessable	not serious	serious [d]	none	N=151	N=151	MD 1 higher (1.12 lower to 3.12 higher)	⊕⊕⊕◻ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Chondroitin) (follow up: range 3 months to 24 months)											
13	randomised trials [1,2,4,5,6,7,8,9,10,11,12,13,14]	not serious	not serious	not serious	serious [d]	none	85/1691 (5.0%)	67/1547 (4.3%)	RR 1.16 (0.85 to 1.59)	7 more per 1,000 (from 6 fewer to 26 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Chondroitin) (follow up: range 3 months to 24 months)											
5	randomised trials [1,2,3,8,10]	not serious	not serious	not serious	serious [d]	none	74/322 (23.0%)	57/286 (19.9%)	RR 1.21 (0.90 to 1.61)	42 more per 1,000 (from 20 fewer to 122 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Chondroitin) (follow up: range 3 months to 24 months)											
6	randomised trials [2,4,11,12,13,15]	serious [g]	not serious	not serious	serious [d]	none	10/543 (1.8%)	6/434 (1.4%)	RR 1.32 (0.45 to 3.87)	4 more per 1,000 (from 8 fewer to 40 more)	⊕⊕◻◻ LOW	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Chondroitin) (follow up: range 3 months to 24 months)											
10	randomised trials [1,3,5,6,8,9,10,12,13,14]	not serious	not serious	not serious	not serious	none	62/1156 (5.4%)	79/1062 (7.4%)	RR 0.72 (0.52 to 0.99)	21 fewer per 1,000 (from 1 fewer to 36 fewer)	⊕⊕⊕⊕ HIGH	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. Two of four studies received High risk of bias ratings. One study did not provide sufficient information to adequately assess bias.

b. I^2= 65%; moderate heterogeneity.
c. I^2= 77%
d. 95% CI crosses null.
e. I^2= 86%
f. I^2= 68%; moderate heterogeneity.
g. Five of six studies received at least one High risk of bias rating.

PICO 2.4.3 (knee): What are the benefits and harms of glucosamine and chondroitin in compound form in the management of patients with knee OA?

SUMMARY
With pooling (where possible) of results from the nine available RCTs, no benefit for pain, function or joint space narrowing was demonstrated. Participants in all trials had knee OA, so extrapolation to hip OA needs additional caution.There are concerns in the literature of publication bias, effects being driven by small industry-sponsored trials, and the overall poor quality of the positive trials. High-quality trial data suggest no effect.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Clegg, et al. N Engl J Med. 2006 Feb 23; 354(8): 795-808 (GAIT); 2. Fransen, et al. Ann Rheum Dis. 2015 May; 74(5): 851-8; 3. Kanzaki, et al. J Sci Food Agric. 2012 Mar 15; 92(4): 862-9; 4. Lugo, et al. Nutr J. 2016 Jan 29; 15: 14; 5. Messier, et al. Osteoarthritis Cartilage. 2007 Nov; 15(11): 1256-66; 6. Roman-Blas, et al. Arthritis Rheumatol. 2017 Jan; 69(1): 77-85; 7. Tsuji, et al. Aging Clin Exp Res. 2016 Apr; 28(2): 197-205; 8. Sawitzke, et al. Annals of the rheumatic diseases. 2010 Aug 1; 69(8): 1459-64 (GAIT); 9. Sawitzke, et al. Arthritis & Rheumatology. 2008 Oct 1; 58(10): 3183-91.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Chondroitin + Glucosamine	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 16 weeks to 12 months)												
7	randomised trials [1,2,3,4,5,6,7]	not serious	very serious [a]	not serious	serious [b]	none	N=694	N=682	SMD 0.28 higher (0.4 lower to 0.97 higher)		⊕□□□ VERY LOW	CRITICAL
Pain [long term] (Higher scores indicate higher pain severity) (follow up: 24 months)												
2	randomised trials [2,8]	serious [c]	not serious	not serious	serious [b]	none	N=280	N=282	SMD 0.02 higher (0.15 lower to 0.19 higher)		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 6 months to 12 months)												
6	randomised trials [1,2,4,5,6,7]	not serious	very serious [d]	not serious	serious [b]	none	N=675	N=664	SMD 0.26 higher (0.43 lower to 0.96 higher)		⊕□□□ VERY LOW	CRITICAL
Function [long term] (Higher scores indicate poorer functional outcome) (follow up: 24 months)												
2	randomised trials [2,8]	serious [c]	not serious	not serious	serious [b]	none	N=280	N=282	SMD 0.03 higher (0.19 lower to 0.24 higher)		⊕⊕□□ LOW	CRITICAL
Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: 24 months)												

2	randomised trials [2,9]	serious [c]	serious [e]	not serious	serious [b]	none	N=180	N=191	MD 0.04 lower (0.14 lower to 0.22 higher)	⊕□□□ VERY LOW	IMPORTANT
SF-12 Physical Component Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 24 months)											
1	randomised trial [2]	not serious	not assessable	not serious	serious [b]	none	N=151	N=151	MD 0.7 lower (2.92 lower to 1.52 higher)	⊕⊕⊕□ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Glucosamine + Chondroitin) (follow up: range 16 weeks to 24 months)											
5	randomised trials [1,2,3,4,5]	not serious	not serious	not serious	serious [b]	none	28/633 (4.4%)	23/620 (3.7%)	RR 1.18 (0.68 to 2.04)	7 more per 1,000 (from 12 fewer to 39 more)	⊕⊕⊕□ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Glucosamine + Chondroitin) (follow up: 24 months)											
3	randomised trials [2,4,5]	not serious	serious [f]	not serious	serious [b]	none	82/296 (27.7%)	56/287 (19.5%)	RR 1.45 (0.77 to 2.73)	88 more per 1,000 (from 45 fewer to 338 more)	⊕⊕□□ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Glucosamine + Chondroitin) (follow up: range 16 weeks to 24 months)											
4	randomised trials [3,4,5,8]	not serious	not serious	not serious	serious [b]	none	6/294 (2.0%)	5/287 (1.7%)	RR 1.16 (0.36 to 3.75)	3 more per 1,000 (from 11 fewer to 48 more)	⊕⊕⊕□ MODERATE	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Glucosamine + Chondroitin) (follow up: 24 months)											
2	randomised trials [2,4]	not serious	serious [g]	not serious	serious [b]	none	11/216 (5.1%)	9/209 (4.3%)	RR 1.00 (0.21 to 4.81)	0 fewer per 1,000 (from 34 fewer to 164 more)	⊕⊕□□ LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. I^2= 97%; removal of Roman-Blas 2017 reduces I^2 to 50% with an effect size (SMD and 95% CI) of -0.14 (-0.33, 0.05).

b. 95% CI crosses null.

c. One study was a 2 year follow up study of a subset of patients who underwent a "departure from randomization." Randomization for structural outcomes was adequate. However, for all outcomes, maintenance of blinding from 24 weeks to 24 months is not adequately described. Not enough information is supplied to adequately assess other dimensions of the risk of bias tool.

d. I^2= 97%; removal of Roman-Blas 2017 reduces I^2 to 29% with an effect size (SMD and 95% CI) of -0.16 (-0.31, 0.00).

e.I^2= 56%; moderate heterogeneity.

f. I^2= 69%; moderate heterogeneity.

g. I^2= 61%; moderate heterogeneity.

PICO 2.4.4 (knee): What are the benefits and harms of Vitamin D supplementation in the management of patients with knee OA?

SUMMARY
There were four RCTs (one to three years' duration), all without serious risk of bias. However, there was very serious inconsistent results, leading to low-quality evidence. When combined, there is a suggestion of a favourable effect, statistically significant but not reaching what was regarded as clinical significance (standardised mean difference [SMD] >0.4). The results for function were similar to the effect estimates. There was no evidence of favourable structural effects on cartilage volume or joint space narrowing. Notably, one study in participants who were vitamin D deficient also failed to show clinically meaningful beneficial effects. Participants in all studies had knee OA, so extrapolation to OA of hip or other joints requires additional caution.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□ LOW

Bibliography: 1. Arden, et al. Osteoarthritis Cartilage. 2016 Nov; 24(11): 1858-1866; 2. Jin, et al. JAMA. 2016 Mar 8; 315(10): 1005-13; 3. McAlindon, et al. JAMA. 2013 Jan 9; 309(2): 155-62; 4. Sanghi, et al. Clin Orthop Relat Res. 2013 Nov; 471(11): 3556-62.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Vitamin D**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 12 months to 3 years)												
4	randomised trials [1,2,3,4]	not serious	very serious [a]	not serious	not serious	none	N=571	N=565	**SMD 0.36 lower (0.7 lower to 0.02 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 12 months to 3 years)												
4	randomised trials [1,2,3,4]	not serious	very serious [b]	not serious	not serious	none	N=571	N=565	**SMD 0.34 lower (0.62 lower to 0.07 lower)**		⊕⊕□□ LOW	CRITICAL
Tibial cartilage volume [mm^3] (Higher values indicate better structural outcome) (follow up: 24 months)												
2	randomised trials [2,3]	not serious	not serious	not serious	serious [c]	none	N=282	N=277	MD 35.44 mm^3 higher (13.66 mm^3 lower to 84.54 mm^3 higher)		⊕⊕⊕□ MODERATE	IMPORTANT
Radiographic Progression [JSN >0.5 mm] (Risk ratios less than one favor Vitamin D) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [c]	none	92/237 (38.8%)	88/237 (37.1%)	RR 1.05 (0.83 to 1.32)	19 more per 1,000 (from 63 fewer to 119 more)	⊕⊕⊕□ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Vitamin D) (follow up: range 12 months to 3 years)												

4	randomised trials [1,2,3,4]	not serious	not serious	not serious	serious [c]	none	14/572 (2.4%)	14/567 (2.5%)	RR 0.99 (0.48 to 2.05)	0 fewer per 1,000 (from 13 fewer to 26 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Vitamin D) (follow up: 24 months)												
2	randomised trials [2,3]	not serious	serious [d]	not serious	serious [c]	none	103/282 (36.5%)	83/277 (30.0%)	RR 1.20 (0.82 to 1.77)	60 more per 1,000 (from 54 fewer to 231 more)	⊕⊕◯◯ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Vitamin D) (follow up: range 24 months to 3 years)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	serious [c]	none	86/519 (16.6%)	87/514 (16.9%)	RR 0.97 (0.75 to 1.27)	5 fewer per 1,000 (from 42 fewer to 46 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Hypercalcemia (Risk ratios less than one favor Vitamin D) (follow up: range 24 months to 3 years)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	serious [c]	none	7/519 (1.3%)	7/514 (1.4%)	RR 0.99 (0.32 to 3.10)	0 fewer per 1,000 (from 9 fewer to 29 more)	⊕⊕⊕◯ MODERATE	IMPORTANT
Hypercalciuria (Risk ratios less than one favor Vitamin D) (follow up: range 24 months to 3 years)												
2	randomised trials [1,3]	not serious	not serious	not serious	serious [c]	none	52/310 (16.8%)	38/310 (12.3%)	RR 1.37 (0.93 to 2.01)	45 more per 1,000 (from 9 fewer to 124 more)	⊕⊕⊕◯ MODERATE	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. I^2= 86%; with the exclusion of Sanghi, 2013, I^2 drops to 0%. Sanghi specifically selected osteoarthritis patients with Vitamin D deficiency at baseline, which may have contributed to discordant results (Clin Orthop Relat Res. 2013 Nov;471(11):3714-5)
b. I^2= 78%; with the exclusion of Sanghi, 2013, I^2 drops to 15%. Sanghi specifically selected osteoarthritis patients with Vitamin D deficiency at baseline, which may have contributed to discordant results (Clin Orthop Relat Res. 2013 Nov;471(11):3714-5)
c. 95% CI crosses null.
d. I^2= 68%; moderate heterogeneity.

PICO 2.4.5 (knee): What are the benefits and harms of (omega-3/6) poly-unsaturated fatty acids in the management of patients with knee OA?

SUMMARY
Pooled data from five RCTs (15–26 weeks) demonstrated no benefits on pain and function in people with hip and knee OA. Three studies received at least one high risk of bias. Most trials used marine oil from whole fish, but some used cod liver oil and mussel extracts. High heterogeneity was expected from pooling different sources of omega-3 fatty acids, and measures within each outcome. The optimal type of omega-3 fatty acids could not be established in OA because only a few trials included marine oil from sources other than whole fish. There are high variations in doses of eicosapentaenoic acid (EPA; 0.01–1.7 g/day), and doses of docosahexaenoic acid (DHA; 0.01–1.10 g/day). A controlled trial that was not included found no additional benefit of high dose fish oil (4.5 g/day), compared with low dose fishoil (0.45 g/day).

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Gruenwald, et al. Adv Ther. 2009 Sep; 26(9): 858-71; 2. Lau, et al. Progress in Nutrition. 2004; 3. Stammers, et al. Lancet. 1989 Aug 26; 2(8661): 503; 4. Stammers, et al. Ann Rheum Dis. 1992 Jan; 51(1): 128-9; 5. Stebbings, et al. Annals of the Rheumatic Diseases. 2014 Jun 1; 73(Suppl 2): 755.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	(omega-3/6) poly-unsaturated fatty acids	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 15 weeks to 26 weeks)												
5	randomised trials [1,2,3,4,5]	serious [a]	very serious [b]	serious [c]	serious [d]	none	N=201	N=207	SMD 0.16 lower (0.57 lower to 0.24 higher)		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 15 weeks to 26 weeks)												
5	randomised trials [1,2,3,4,5]	serious [a]	very serious [b]	serious [c]	serious [d]	none	N=201	N=207	SMD 0.11 higher (0.13 lower to 0.35 higher)		⊕□□□ VERY LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Fish oil) (follow up: range 24 weeks to 26 weeks)												
3	randomised trials [1,2,4]	serious [a]	not serious	serious [c]	serious [d]	none	10/174 (5.7%)	7/170 (4.1%)	RR 1.33 (0.52 to 3.39)	14 more per 1,000 (from 20 fewer to 98 more)	⊕□□□ VERY LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Fish oil) (follow up: range 24 weeks to 26 weeks)												
3	randomised trials [1,2,4]	serious [a]	not serious	serious [c]	serious [d]	none	29/174 (16.7%)	21/170 (12.4%)	RR 1.31 (0.79 to 2.18)	38 more per 1,000 (from 26 fewer to 146 more)	⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Fish oil) (follow up: range 15 weeks to 26 weeks)												

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	(omega-3/6) poly-unsaturated fatty acids	Placebo	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials [1,5]	serious [a]	not serious	serious [c]	not serious	none	0/129 (0.0%)	0/129 (0.0%)	Due to zero events in both groups, an absolute risk reduction was not estimable.		⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. 3 of 5 studies received at least one High risk of bias rating; 1 of 5 studies received all Unclear risk of bias ratings due to insufficient information.
b. I^2 >75%
c. Mixed populations of Hip, Knee, Hip/Knee Osteoarthritis patients.
d. 95% CI crosses null.

PICO 2.4.6 (knee): What are the benefits and harms of collagen preparations in the management of patients with knee OA?

SUMMARY
Pooled results from six studies found short-term (13–26 weeks) clinical benefits in pain; however, there have been very serious inconsistent results across the studies. Available data from four studies found no effect in function. All of the studies were conducted in knee OA, so extrapolation to hip or other OA requires additional caution. There are concerns in the literature of publication bias, effects being mostly driven by industrysponsored trials, and the overall poor quality of the positive trials.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□ LOW

Bibliography: 1. Benito-Ruiz, et al. Int J Food Sci Nutr. 2009; 60 Suppl 2: 99-113; 2. Kumar, et al.[a] J Sci Food Agric. 2015 Mar 15; 95(4): 702-7; 3. Kumar, et al.[b] J Sci Food Agric. 2015 Mar 15; 95(4): 702-7; 4. Lugo, et al. Nutr J. 2016 Jan 29; 15: 14; 5. McAlindon, et al. Osteoarthritis Cartilage. 2011 Apr; 19(4): 399-405; 6. Moskowitz, RW. Semin Arthritis Rheum. 2000 Oct; 30(2): 87-99.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Collagen**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 13 weeks to 6 months)												
6	randomised trials [1,2,3,4,5,6]	not serious	very serious [a]	not serious	not serious	none	N=404	N=371	**SMD 0.58 lower (0.98 lower to 0.17 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: 6 months)												
4	randomised trials [1,4,5,6]	not serious	serious [b]	not serious	serious [c]	none	N=367	N=350	SMD 0.18 lower (0.42 lower to 0.07 higher)		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Collagen) (follow up: 6 months)												
3	randomised trials [1,4,6]	not serious	not serious	not serious	not serious	none	4/376 (1.1%)	16/368 (4.3%)	**RR 0.26 (0.09 to 0.78)**	**32 fewer per 1,000 (from 10 fewer to 40 fewer)**	⊕⊕⊕⊕ HIGH	CRITICAL
Total Adverse Events (Risk ratios less than one favor Collagen) (follow up: 6 months)												
6	randomised trials [1,2,3,4,5,6]	not serious	not serious	not serious	serious [c]	none	212/431 (49.2%)	194/403 (48.1%)	RR 1.08 (0.87 to 1.36)	39 more per 1,000 (from 63 fewer to 173 more)	⊕⊕⊕□ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Collagen) (follow up: 6 months)												

2	randomised trials [4,5]	not serious	not serious	not serious	serious [c]	none	1/78 (1.3%)	0/73 (0.0%)	RR 3.00 (0.13 to 68.26)	NA [d]	⊕⊕⊕◯ MODERATE	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Collagen) (follow up: 6 months)												
5	randomised trials [1,2,3,4,6]	not serious	not serious	not serious	serious [c]	none	42/416 (10.1%)	28/388 (7.2%)	RR 1.43 (0.90 to 2.25)	31 more per 1,000 (from 7 fewer to 90 more)	⊕⊕⊕◯ MODERATE	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. I^2= 81%
b. I^2= 52%
c. 95% CI crosses null.
d. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.

PICO 2.4.7 (knee): What are the benefits and harms of methylsulfonylmethane in the management of patients with knee OA?

SUMMARY

There are three trials with short study durations (12–13 weeks), and pooled data found statistically and clinically significant benefits in pain. Even larger effects were found in function, but with very serious inconsistent results and high heterogeneity across studies. One trial had a high risk of bias because of inappropriate randomisation technique; while the other had potential reporting bias. The doses in the trials ranged from 1.5–6 g/day for 12 weeks. All of the studies were conducted in knee OA, so extrapolation to hip or other OA sites requires additional caution. There are concerns in the literature of publication bias, effects being mostly driven by industrysponsored trials, and the overall poor quality of the positive trials.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Debbi, et al. BMC Complement Altern Med. 2011 Jun 27; 11: 50; 2. Kim, et al. Osteoarthritis Cartilage. 2006 Mar; 14(3): 286-94; 3. Usha, P.R., Naidu, M.U.R. Clin Drug Investig. 2004; 24(6): 353-63.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Methylsulfonyl-methane**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 12 weeks to 13 weeks)												
3	randomised trials [1,2,3]	serious [a]	not serious	not serious	not serious	none	N=76	N=72	**SMD 0.47 lower (0.8 lower to 0.14 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 12 weeks to 13 weeks)												
3	randomised trials [1,2,3]	serious [a]	very serious [b]	not serious	not serious	none	N=76	N=72	**SMD 1.1 lower (1.81 lower to 0.38 lower)**		⊕□□□ VERY LOW	CRITICAL
Quality of Life (Higher scores indicate better quality of life) (follow up: range 12 weeks to 13 weeks)												
2	randomised trials [1,2]	serious [c]	not serious	not serious	very serious [d,e]	none	N=46	N=44	SMD 0.42 lower (0.86 lower to 0.01 higher)		⊕□□□ VERY LOW	IMPORTANT
Total Adverse Events (Risk ratios less than one favor Methylsulfonylmethane) (follow up: 13 weeks)												
1	randomised trial [2]	not serious	not assessable	not serious	very serious [d,e]	none	21/25 (84.0%)	19/25 (76.0%)	RR 1.1 (0.8 to 1.5)	76 more per 1,000 (from 152 fewer to 380 more)	⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. Two studies received at least one High risk of bias rating due to inappropriate randomization technique and potential reporting bias, respectively.
b. I^2= 76%
c. One study received at least one High risk of bias rating due to inappropriate randomization technique.
d. 95% CI crosses null.
e. Sample size in each study arm <50.

PICO 2.4.8 (knee): What are the benefits and harms of Diacerein in the management of patients with knee OA?

SUMMARY
Five trials were included, with time durations ranging from eight weeks to 12 months, all receiving high risk of bias because of weak allocation concealment and random sequence generation. Very low-quality evidence from four trials indicated a small clinical benefit on pain reduction. Data from five trials indicated statistically significant effects on function, but this did not reach the clinically meaningful threshold. Analysis of one study demonstrated no benefit in reducing joint space narrowing. There are concerns in the literature of publication bias, effects being mostly driven by industrysponsored trials, and the overall poor quality of the positive trials. A search of regulatory websites found a recommendation from the European Medicines Agency (EMA) Pharmacovigilance Risk Assessment Committee (PRAC) that the marketing authorisation of diacerein should be suspended across Europe because of harms (particularly the risk of severe diarrhoea and potentially harmful effects on the liver) outweighing benefits. However, this guidance is not final as the PRAC recommendation will be re-examined.
All of the studies were conducted in knee OA, so extrapolation to hip or other OA sites requires additional caution.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Brahmachari, et al. Clin Rheumatol. 2009 Oct; 28(10): 1193-8; 2. Pavelka, et al. Arthritis Rheum. 2007 Dec; 56(12): 4055-64; 3. Pelletier, et al. Arthritis Rheum. 2000 Oct; 43(10): 2339-48; 4. Pham, et al. Ann Rheum Dis. 2004 Dec; 63(12): 1611-7; 5. Lequesne, et al. Rev Prat. 1998 Nov 1; 48(17 Suppl): S31-5.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Diacerein	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 8 weeks to 1 year)												
4	randomised trials [1,2,3,4]	serious [a]	very serious [b]	not serious	not serious	none	N=305	N=319	**SMD 0.45 lower (0.91 lower to 0.01 lower)**		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 8 weeks to 1 year)												
5	randomised trials [1,2,3,4,5]	serious [a]	not serious	not serious	not serious	none	N=364	N=373	**SMD 0.23 lower (0.38 lower to 0.09 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Proportion of Patients with Radiographic Progression (JSN ≥0.5mm) (Risk ratios less than one favor Diacerein) (follow up: 1 year)												
1	randomised trial [4]	serious [a]	not assessable	not serious	serious [c]	none	16/85 (18.8%)	17/85 (20.0%)	RR 0.94 (0.51 to 1.74)	12 fewer per 1,000 (from 98 fewer to 148 more)	⊕⊕□□ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Diacerein) (follow up: range 8 weeks to 1 year)												

5	randomised trials [1,2,3,4,5]	serious [a]	not serious	not serious	serious [c]	none	50/396 (12.6%)	55/413 (13.3%)	RR 0.94 (0.67 to 1.30)	8 fewer per 1,000 (from 40 more to 44 fewer)	⊕⊕◻◻ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Diacerein) (follow up: range 8 weeks to 1 year)												
5	randomised trials [1,2,3,4,5]	serious [a]	very serious [d]	not serious	not serious	none	169/311 (54.3%)	117/328 (35.7%)	**RR 1.95 (1.08 to 3.54)**	**339 more per 1,000 (from 29 more to 906 more)**	⊕◻◻◻ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Diacerein) (follow up: 16 weeks)												
1	randomised trial [3]	serious [a]	not assessable	not serious	not serious	none	0/111 (0.0%)	0/125 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable		⊕⊕⊕◻ MODERATE	CRITICAL
Rash/Pruritus (Risk ratios less than one favor Diacerein) (follow up: range 8 weeks to 1 year)												
2	randomised trials [1,4]	serious [a]	not serious	not serious	serious [c]	none	7/113 (6.2%)	2/112 (1.8%)	RR 2.97 (0.51 to 17.29)	35 more per 1,000 (from 9 fewer to 291 more)	⊕⊕◻◻ LOW	IMPORTANT
Diarrhea (Risk ratios less than one favor Diacerein) (follow up: range 8 weeks to 1 year)												
5	randomised trials [1,2,3,4,5]	serious [a]	not serious	not serious	not serious	none	90/285 (31.6%)	24/288 (8.3%)	**RR 3.50 (1.95 to 6.27)**	**208 more per 1,000 (from 79 more to 439 more)**	⊕⊕⊕◻ MODERATE	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. All studies received at least one High risk of bias rating due to attrition bias or inadequate blinding.
b. I^2=83%
c. 95% CI crosses null.
d. I^2=88%

PICO 2.5 (knee): What are the benefits and harms of duloxetine in the management of patients with knee OA?

SUMMARY
In the three trials reviewed, significant response and moderate effects in knee pain (standardised mean difference [SMD] 0.43) and function (SMD 0.45) were found over 13–16 weeks at doses of 60/120 mg. However, most study participants were also already using NSAIDs and paracetamol. The use of duloxetine for knee OA adjunctively with NSAIDs, thus reducing the usage of NSAIDs and paracetamol, would be clinically useful to reduce adverse events. (Brown JP, Boulay LJ. Ther Adv Musculoskelet Dis 2013;5(6):291–304). In addition, results differed as to whether significant reduction in depression symptoms was needed for analgesic impact. There is no direct randomised controlled trial (RCT) evidence for hip OA; thus, using knee OA data to extrapolate to hip or other OA requires additional caution. Duloxetine currently does not have an indication via the Therapeutic Goods Administration (TGA) for OA, and should be considered as an investigational medication only.

OVERALL QUALITY OF EVIDENCE: ⊕⊕⊕◻MODERATE

Bibliography: 1. Abou-Raya, et al. Age Ageing. 2012 Sep; 41(5): 646-52; 2. Chappell, et al. Pain. 2009 Dec; 146(3): 253-60; 3. Chappell, et al. Pain Pract. 2011 Jan-Feb; 11(1): 33-41.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Duloxetine**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 13 weeks to 16 weeks)												
3	randomised trials [1,2,3]	serious [a]	not serious	not serious	not serious	none	N=379	N=389	**SMD 0.43 lower (0.58 lower to 0.29 lower)**		⊕⊕⊕◻ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 13 weeks to 16 weeks)												
2	randomised trials [1,3]	serious [a]	not serious	not serious	not serious	none	N=272	N=272	**SMD 0.45 lower (0.81 lower to 0.08 lower)**		⊕⊕⊕◻ MODERATE	CRITICAL
EQ-5D UK index (Higher scores indicate better quality of life) (follow up: 13 weeks)												
1	randomised trial [2]	not serious	not assessable	not serious	not serious	none	N=103	N=114	**MD 0.1 higher (0.04 higher to 0.16 higher)**		⊕⊕⊕⊕ HIGH	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Duloxetine) (follow-up: range 13 weeks to 16 weeks)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	not serious	none	48/383 (12.5%)	20/392 (5.1%)	**RR 2.42 (1.46 to 4.03)**	**72 more per 1,000 (from 23 more to 155 more)**	⊕⊕⊕⊕ HIGH	CRITICAL
Treatment- Related Adverse Events (Risk ratios less than one favor Duloxetine) (follow-up: 13 weeks)												

1	randomised trial [3]	not serious	not assessable	not serious	not serious	none	65/128 (50.8%)	42/128 (32.8%)	**RR 1.55 (1.15 to 2.09)**	**180 more per 1,000 (from 49 more to 358 more)**	⊕⊕⊕⊕ HIGH	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Duloxetine) (follow-up: range 13 weeks to 16 weeks)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	serious [b]	none	4/383 (1.0%)	4/392 (1.0%)	RR 1.04 (0.25 to 4.33)	0 fewer per 1,000 (from 8 fewer to 34 more)	⊕⊕⊕□ MODERATE	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. Chappell 2011 received a High risk of bias rating due to potential attrition bias b. 95% CI crosses null

PICO 2.6 (knee): What are the benefits and harms of doxycycline in the management of patients with knee OA?

SUMMARY
Preclinical research and earlier human studies indicated doxycycline might be useful in managing symptomatic knee OA. However, current evidence found that doxycycline did not reduce the mean severity of joint pain, although pain scores in both treatment groups were low at baseline and remained low throughout the trial, which may suggest the presence of a floor effect. Brown JP, Boulay LJ. Ther Adv Musculoskelet Dis 2013;5(6):291–304). Despite the small benefit (SMD 0.15 mm) in joint space narrowing, it is outweighed by medication harms. There is no RCT of doxycycline for hip OA, thus using knee OA data to extrapolate to hip or other OA requires additional caution. Doxycycline currently does not have an indication via the TGA for OA, and should be considered as an investigational medication only.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. da Costa, et al. Cochrane Database Syst Rev. 2012 Nov 14; 11: CD007323.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Doxycycline**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (multiple measures including 50-ft walking pain, WOMAC pain) (Higher scores indicate higher pain severity) (follow up: range 24 weeks to 30 months)												
2	randomised trials [1]	serious [a]	not serious	not serious	serious [b]	none	N=256	N=268	SMD 0.05 lower (0.2 2 lower to 0.13 higher)		⊕⊕□□ LOW	CRITICAL
WOMAC function (Higher scores indicate poorer functional outcome) (follow up: range 24 weeks to 30 months)												
2	randomised trials [1]	serious [a]	not serious	not serious	serious [b]	none	N=252	N=265	SMD 0.07 lower (0.25 lower to 0.1 higher)		⊕⊕□□ LOW	CRITICAL
Minimum Joint Space Width [mm] (Higher values indicate a better structural outcome) (follow up: 30 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	not serious	none	N=181	N=180	**SMD 0.15 mm lower (0.28 mm lower to 0.02 mm lower)**		⊕⊕⊕□ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Doxycycline) (follow up: range 24 weeks to 30 months)												
2	randomised trials [1]	not serious	serious [c]	not serious	not serious	none	57/334 (17.1%)	27/329 (8.2%)	**RR 2.28 (1.06 to 4.90)**	**105 more per 1,000 (from 5 more to 320 more)**	⊕⊕⊕□ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Doxycycline) (follow up: 24 weeks)												

1	randomised trial [1]	not serious	not assessable	not serious	not serious	none	75/116 (64.7%)	55/116 (47.4%)	**RR 1.36 (1.08 to 1.72)**	**171 more per 1,000 (from 38 more to 341 more)**	⊕⊕⊕⊕ HIGH	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Doxycycline) (follow up: range 24 weeks to 30 months)												
2	randomised trials [1]	not serious	not serious	not serious	serious [b]	none	34/334 (10.2%)	31/329 (9.4%)	RR 1.07 (0.68 to 1.68)	7 more per 1,000 (from 30 fewer to 64 more)	⊕⊕⊕◻ MODERATE	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. Both studies received at least one High risk bias rating due to non-ITT analyses and possible selective reporting
b. 95% CI crosses null
c. I^2=55%; moderate heterogeneity

PICO 2.7: ANTI-OSTEOPOROSIS DRUGS
PICO 2.7.1 (knee): What are the benefits and harms of bisphosphonates in the management of patients with knee OA?

SUMMARY
Evidence from six trials found no statistically significant benefits in symptom relief, and structural and functional improvement. The quality of evidence has varied from moderate to very low, with inconsistent results. A meta-analysis of the two largest knee studies using risedronate 15 mg found the odds ratios (ORs) favouring placebos for Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain (1.73), function (2.03) and stiffness (1.82). However, eight trials (61.5%) reported that bisphosphonates improved pain assessed by Visual Analogue Scale (VAS) scores, and two (38.5%) reported significant improvement in WOMAC pain scores, compared with control groups. (Davis AJ, et al. PLoS One 2013;8(9):e72714). There were no statistically significant differences or trends noted for any dose of risedronate. Similarly, there was no difference between the five groups with respect to radiographic joint space narrowing, joint space width or osteophyte formation at 24-month follow-up. Davis AJ, et al. PLoS One 2013;8(9):e72714). There is one very low-quality trial conducted in 42 participants with hip OA, demonstrating no effect over 24 months. Biphosphonates currently do not have an indication via the Therapeutic Goods Administration (TGA) for OA, and should be considered as investigational medications only.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Bingham, et al.[EU] Arthritis Rheum. 2006 Nov; 54(11): 3494-507 (KOSTAR); 2. Bingham, et al. [US] Arthritis Rheum. 2006 Nov; 54(11): 3494-507 (KOSTAR); 3. Laslett, et al. Ann Rheum Dis. 2012 Aug; 71(8): 1322-8; 4. Rossini, et al. Rheumatol Int. 2015 Feb; 35(2):255-63; 5. Spector, et al. Arthritis Res Ther. 2005; 7(3): R625-33 (BRISK); 6. Varenna, et al. Rheumatology (Oxford). 2015 Oct; 54(10): 1826-32; 7. Adami, et al.[EU+US] Mayo Clin Proc. 2005 Oct; 80(10): 1278-85 (KOSTAR); 8. Jokar, et al. Iran J Med Sci. 2013 Sep; 38(3): 221-6.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Bisphosphonate†	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain [all] (follow up: range 60 days to 24 months)												
6	randomised trials [1,2,3,4,5,6]	not serious [a]	very serious [b]	not serious	serious [c]	none	N=2113	N=794	SMD 0.16 lower (0.34 lower to 0.02 higher)		⊕□□□ VERY LOW	CRITICAL
Pain [short term] (Higher scores indicate higher pain severity) (follow up: range 60 days to 24 weeks)												
3	randomised trials [3,4,6]	serious [d]	very serious [e]	not serious	serious [c]	none	N=101	N=92	SMD 0.86 lower (2.09 lower to 0.37 higher)		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 16 weeks to 24 months)												
4	randomised trials [1,2,4,5]	not serious	not serious	not serious	serious [c]	none	N=2051	N=741	SMD 0.02 lower (0.1 lower to 0.07 higher)		⊕⊕⊕□ MODERATE	CRITICAL
Joint Space Widening [mm] (Positive values indicate better structural outcome) (follow up: 12 months)												

1	randomised trial [5]	not serious	not assessable	not serious	serious [c]	none	N=151	N=80	MD 0.05 higher (0.9 lower to 1 higher)		⊕⊕⊕□ MODERATE	IMPORTANT
Proportion of Patients Experiencing Radiographic Progression [≥0.6 mm JSN] (follow up: 24 months)												
2	randomised trials [1,2]	not serious	not serious	not serious	serious [c]	none	222/1676 (13.2%)	72/533 (13.5%)	RR 0.98 (0.77 to 1.26)	3 fewer per 1,000 (from 31 fewer to 35 more)	⊕⊕⊕□ MODERATE	IMPORTANT
SF-36 Physical Component Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 60 days)												
1	randomised trial [6]	very serious [f]	not assessable	not serious	serious [g]	none	N=31	N=25	**MD 8.6 higher (1.21 higher to 15.99 higher)**		⊕□□□ VERY LOW	IMPORTANT
SF-36 Mental Component Score (scale range 0 to 100, with higher scores indicating better outcome) (follow up: 60 days)												
1	randomised trial [6]	very serious [f]	not assessable	not serious	serious [g]	none	N=31	N=25	**MD 7.5 higher (0.77 higher to 14.23 higher)**		⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Bisphosphonate) (follow up: range 60 days to 24 months)												
6	randomised trials [3,4,5,6,7,8]	not serious	not serious	not serious	serious [c]	none	214/2172 (9.9%)	88/841 (10.5%)	RR 0.88 (0.70 to 1.12)	13 fewer per 1,000 (from 13 more to 31 fewer)	⊕⊕⊕□ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Bisphosphonate) (follow up: range 60 days to 12 months)												
4	randomised trials [3,4,5,8]	not serious	serious [h]	not serious	serious [c]	none	226/277 (81.6%)	132/185 (71.4%)	RR 1.10 (0.80 to 1.51)	71 more per 1,000 (from 143 fewer to 364 more)	⊕⊕□□ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Bisphosphonate) (follow up: range 16 weeks to 12 months)												
3	randomised trials [3,4,5]	not serious	not serious	not serious	serious [c]	none	9/257 (3.5%)	3/166 (1.8%)	RR 1.83 (0.38 to 8.75)	15 more per 1,000 (from 11 fewer to 140 more)	⊕⊕⊕□ MODERATE	CRITICAL

Gastrointestinal Adverse Events (Risk ratios less than one favor Bisphosphonate) (follow up: range 24 weeks to 24 months)												
3	randomised trials [5,7,8]	not serious	not serious	not serious	serious [c]	none	387/2067 (18.7%)	247/1401 (17.6%)	RR 1.07 (0.93 to 1.24)	12 more per 1,000 (from 12 fewer to 42 more)	⊕⊕⊕◻ MODERATE	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

† Bisphosphonates of varying routes of administration were analyzed together. The following Bisphosphonates were included in the analysis: Bingham 2006/Adami 2005 (Europe + N. America)- Oral risedronate, Jokar, 2010- Oral alendronate, Laslett, 2012- Intravenous Zoledronic Acid, Rossini, 2015- Intra-articular Clodronate, Spector, 2005- Oral risedronate, Varenna, 2015- Intravenous Neridronate

a. Only one of 6 studies received a High risk of bias rating in at least one dimension (Varenna, 2015)
b. I^2=77%
c. 95% CI crosses null.
d. One of 3 studies received a High risk of bias rating due to potential attrition bias.
e. I^2=94%
f. Study received High risk of bias rating due to potential attrition bias and due to potential unblinding.
g. Sample size <50 in each study arm.
h. I^2=72% moderate heterogeneity

PICO 2.7.2 (knee): What are the benefits and harms of calcitonin in the management of patients with knee OA?

SUMMARY
The two phase III studies found no significant effect of salmon calcitonin on total WOMAC, WOMAC subscores and joint space narrowing. There is a potentially small effect on markers of bone and cartilage degradation, CTX-I and CTX-II respectively, and no positive balance between bone formation and bone resorption. There is no randomised controlled trial (RCT) of calcitonin for hip OA; thus, using knee OA data to extrapolate to hip or other OA requires additional caution.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Karsdal, et al.[1] Osteoarthritis and Cartilage 23 (2015) 532-543; 2. Karsdal, et al.[2] Osteoarthritis and Cartilage 23 (2015) 532-543; 3. Manicourt, et al. Arthritis Rheum. 2006 Oct; 54(10):3205-11

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Calcitonin**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0 to 500, with higher scores indicate higher pain severity) (follow up: 12 months)												
2	randomised trials [1,2]	serious [a]	very serious [b]	not serious	serious [c]	none	N=803	N=885	MD 6.65 lower (30.15 lower to 16.85 higher)		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (scale range 0 to 1700, with higher scores indicating poorer functional outcome) (follow up: 12 months)												
2	randomised trials [1,2]	serious [a]	serious [d]	not serious	serious [c]	none	N=791	N=874	MD 18.42 lower (76.86 lower to 40.01 higher)		⊕□□□ VERY LOW	CRITICAL
Joint Space Widening [mm] (positive values indicate better structural outcome) (follow up: 12 months)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	serious [c]	none	N=804	N=885	MD 0.02 higher (0.04 lower to 0.08 higher)		⊕⊕□□ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Calcitonin) (follow-up: range 85 days to 24 months)												
3	randomised trials [1,2,3]	serious [a]	not serious	not serious	not serious	none	196/1140 (17.2%)	71/1110 (6.4%)	**RR 2.68 (2.07 to 3.47)**	**107 more per 1,000 (from 68 more to 158 more)**	⊕⊕⊕□ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Calcitonin) (follow-up: range 85 days to 24 months)												

3	randomised trials [1,2,3]	not serious	serious [e]	not serious	serious [c]	none	1034/1140 (90.7%)	981/1110 (88.4%)	RR 1.03 (0.98 to 1.09)	27 more per 1,000 (from 18 fewer to 80 more)	⊕⊕◯◯ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Calcitonin) (follow-up: 24 months)												
2	randomised trials [1,2]	not serious	not serious	not serious	serious [c]	none	18/1105 (1.6%)	11/1092 (1.0%)	RR 1.59 (0.75 to 3.40)	6 more per 1,000 (from 3 fewer to 24 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Calcitonin) (follow-up: range 85 days to 24 months)												
3	randomised trials [1,2,3]	not serious	serious [f]	not serious	not serious	none	480/1140 (42.1%)	303/1110 (27.3%)	**RR 1.55 (1.20 to 2.00)**	**150 more per 1,000 (from 55 more to 273 more)**	⊕⊕⊕◯ MODERATE	IMPORTANT
Hot Flush (Risk ratios less than one favor Calcitonin) (follow-up: range 85 days to 24 months)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	not serious	none	199/1140 (17.5%)	47/1110 (4.2%)	**RR 4.11 (3.02 to 5.59)**	**132 more per 1,000 (from 86 more to 194 more)**	⊕⊕⊕⊕ HIGH	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. All trials received at least one High risk of bias rating due to poor attrition without ITT and/or differential dropout
b. I^2=85%
c. 95% CI crosses null
d. I^2=72% moderate heterogeneity
e. I^2=59%; moderate heterogeneity
f. I^2=66%; moderate heterogeneity

PICO 2.7.3 (knee): What are the benefits and harms of strontium ranelate in the management of patients with knee OA?

SUMMARY
Data from one moderate-quality trial found no effect of strontium ranelate in altering OA symptoms. However, strontium ranelate reatment had a beneficial effect on joint space widening, with a mean difference (MD) of 0.12 mm over three years. Similarly, the risk ratio of radiographic progression (joint space narrowing ≥0.5 mm) favoured strontium ranelate over three years. As strontium ranelate is not accessible for people in Australia, the working group considered this treatment as unfeasible for use.

OVERALL QUALITY OF EVIDENCE: ⊕⊕⊕◻MODERATE

Bibliography: 1. Reginster, et al. Ann Rheum Dis. 2013 Feb; 72(2): 179-86.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Strontium ranelate**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	N=899	N=472	MD 2.14 lower (6.55 lower to 2.27 higher)		⊕⊕⊕◻ MODERATE	CRITICAL
WOMAC Function (scale range 0 to 100, with higher scores indicating poorer functional outcome) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	N=899	N=472	MD 1.88 lower (4.63 lower to 0.86 higher)		⊕⊕⊕◻ MODERATE	CRITICAL
Joint Space Widening [mm] (Higher values indicate better structural outcome) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	not serious	not serious	none	N=899	N=472	**MD 0.12 higher (0.05 higher to 0.19 higher)**		⊕⊕⊕⊕ HIGH	IMPORTANT
Patients experiencing Radiographic Progression [JSN ≥0.5 mm] (Risk ratios less than one favor Strontium ranelate) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	not serious	not serious	none	215/899 (23.9%)	156/472 (33.1%)	**RR 0.72 (0.61 to 0.86)**	**93 fewer per 1,000 (from 46 fewer to 129 fewer)**	⊕⊕⊕⊕ HIGH	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Strontium ranelate) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	45/1112 (4.0%)	27/556 (4.9%)	RR 0.83 (0.52 to 1.33)	8 fewer per 1,000 (from 16 more to 23 fewer)	⊕⊕⊕◻ MODERATE	CRITICAL

Gastrointestinal Adverse Events (Risk ratios less than one favor Strontium ranelate) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	136/1112 (12.2%)	52/556 (9.4%)	RR 1.31 (0.97 to 1.77)	29 more per 1,000 (from 3 fewer to 72 more)	⊕⊕⊕◻ MODERATE	IMPORTANT
Skin and Subcutaneous Disorders (Sum of N patients experiencing "Dermatitis", "Allergic Dermatitis", "Eczema", and "Rash") (Risk ratios less than one favor Strontium ranelate) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	44/1112 (4.0%)	20/556 (3.6%)	RR 1.10 (0.65 to 1.85)	4 more per 1,000 (from 13 fewer to 31 more)	⊕⊕⊕◻ MODERATE	IMPORTANT
Venous Thromboembolism Events "Sum of N patients experiencing "Deep venous thrombosis" and "Pulmonary embolism") (Risk ratios less than one favor Strontium ranelate) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	9/1112 (0.8%)	2/556 (0.4%)	RR 2.25 (0.49 to 10.38)	4 more per 1,000 (from 2 fewer to 34 more)	⊕⊕⊕◻ MODERATE	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. 95% CI crosses null

PICO 2.8: INVESTIGATIONAL DMOADs
PICO 2.8.1 (knee): What are the benefits and harms of IL-1 inhibitors in the management of patients with knee OA?

SUMMARY
Results from a three-arm trial of a single intra-articular injection of anakinra at a dose of 50 mg (n = 34) and 150 mg (n = 67) were vailable. The mean improvement from baseline at week 12 on the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) score was not statistically different between the anakinra and placebo groups. A placebo-controlled randomised controlled trial (RCT) of AMG-108, (not included in this review) found non-statistically significant improvement on WOMAC pain after subcutaneous administration of AMG-108. (Cohen SB, et al. Arthritis Res Ther 2011;13(4):R125). Due to the limitations in current efficacy, safety,access and costs, it is considered that IL-1 inhibitors are not a feasible nor cost-effective treatment.There are currently no trials that have investigated the benefits and safety of IL-1 inhibitors in people with hip OA. Using knee OA data to extrapolate to hip or other OA requires additional caution.

OVERALL QUALITY OF EVIDENCE: ⊕⊕⊕◻MODERATE

Bibliography: 1. Chevalier, et al. Arthritis Rheum. 2009 Mar 15; 61(3): 344-52

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	IA Anakinra (single 50 mg/150mg dose)	Placebo	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (Scale range 0 to 500, with higher scores indicating higher pain severity) (follow up: 12 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	N=101	N=69	MD 2.71 lower (33.78 lower to 28.36 higher)		⊕⊕⊕◻ MODERATE	CRITICAL
WOMAC Function (Scale range 0 to 1,700, with higher scores indicating poorer functional outcome) (follow up: 12 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	N=101	N=69	MD 21.12 lower (132.53 lower to 90.29 higher)		⊕⊕⊕◻ MODERATE	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Anakinra) (follow up: 12 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	not serious	none	0/101 (0.0%)	0/69 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕⊕ HIGH	CRITICAL
Total Adverse Events (Risk ratios less than one favor Anakinra) (follow up: 12 weeks)												

1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	53/101 (52.5%)	41/69 (59.4%)	RR 0.88 (0.67 to 1.16)	71 fewer per 1,000 (from 95 more to 196 fewer)	⊕⊕⊕◯ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Anakinra) (follow up: 12 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	1/101 (1.0%)	1/69 (1.4%)	RR 0.68 (0.04 to 10.74)	5 fewer per 1,000 (from 14 fewer to 141 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Infections (Risk ratios less than one favor Anakinra) (follow up: 12 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	13/101 (12.9%)	4/69 (5.8%)	RR 2.22 (0.76 to 6.53)	71 more per 1,000 (from 14 fewer to 321 more)	⊕⊕⊕◯ MODERATE	IMPORTANT
Injection Site Reactions (Risk ratios less than one favor Anakinra) (follow up: 12 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	4/101 (4.0%)	4/69 (5.8%)	RR 0.68 (0.18 to 2.64)	19 fewer per 1,000 (from 48 fewer to 95 more)	⊕⊕⊕◯ MODERATE	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. 95% CI crosses null

PICO 2.8.2 (knee): What are the benefits and harms of TNF-alpha inhibitors in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

No RCT data was found for this question.

PICO 2.8.3 (knee): What are the benefits and harms of anti-nerve growth factor (NGF) therapy in the management of patients with knee OA?

SUMMARY
Results from five trials of tanezumab and one of fasimumab found a statistically significant lower WOMAC pain and function score, compared with placebos with a pooled standardised mean differences (SMDs) of 0.6 and 0.64 respectively The dosage of tanezumab differed between phase II and phase III studies included in the systematic review. There were two phase II studies of tanezumab (References 3 and 5) which demonstrated SMD ranging from –0.31 to 0.94 with five different dose groups (10 µg/kg, 25 µg/kg, 50 µg/kg, 100 µg/kg, 200 µg/kg).109 The other phase III studies evaluated a narrower dose range (2.5 mg, 5 mg, 10 mg), and reported a correspondingly narrower range of SMD from 0.26 to 0.61, all of which are statistically significantly from placebo. In the study of fasinumab, all three doses of fasinumab were associated with significant improvements, compared with placebo in walking knee pain and WOMAC total and subscale scores. Anti-NGF requires off-label prescribing and is expensive, which limited its accessibility and affordability.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□ LOW

Bibliography: 1. Brown, et al. J Pain. 2012 Aug; 13(8): 790-8; 2. Brown, et al. Arthritis Rheum. 2013 Jul; 65(7): 1795-803; 3. Lane, et al. N Engl J Med. 2010 Oct 14; 363(16): 1521-31; 4. Mayorga, et al. Int J Clin Pract. 2016 Jun; 70(6): 493-505; 5. Nagashima, et al. Osteoarthritis Cartilage. 2011 Dec; 19(12): 1405-12; 6. Tiseo, et al. Pain. 2014 Jul; 155(7): 1245-52.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Anti-Nerve Growth Factor	Placebo	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (Higher scores indicate higher pain severity) (follow-up: range 8 weeks to 16 weeks)												
6	randomised trials [1,2,3,4,5,6]	not serious	serious [a]	not serious	not serious	none	N=1613	N=498	**SMD 0.6 lower (0.83 lower to 0.38 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
WOMAC Function (Higher scores indicate poorer functional outcome) (follow-up: range 8 weeks to 16 weeks)												
6	randomised trials [1,2,3,4,5,6]	not serious	serious [b]	not serious	not serious	none	N=1613	N=498	**SMD 0.64 lower (0.84 lower to 0.44 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
SF-36 Composite score (scale range 0 to 100, with higher scores indicating better quality of life) (follow-up: 24 weeks)												
1	randomised trials [1]	not serious	not assessable	not serious	serious [c]	none	N=464	N=154	MD 0.19 higher (2 lower to 2.37 higher)		⊕⊕⊕□ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Anti-Nerve Growth Factor) (follow-up: range 8 weeks to 16 weeks)												

6	randomised trials [1,2,3,4,5,6]	not serious	not serious	not serious	serious [c]	none	80/1679 (4.8%)	12/520 (2.3%)	RR 1.78 (0.97 to 3.25)	18 more per 1,000 (from 1 fewer to 52 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Anti-Nerve Growth Factor) (follow-up: range 8 weeks to 16 weeks)												
6	randomised trials [1,2,3,4,5,6]	not serious	serious [d]	not serious	serious [c]	none	1037/1676 (61.9%)	277/520 (53.3%)	RR 1.08 (0.91 to 1.27)	43 more per 1,000 (from 48 fewer to 144 more)	⊕⊕◻◻ LOW	CRITICAL
Treatment-related Adverse Events (Risk ratios less than one favor Anti-Nerve Growth Factor) (follow-up: range 8 weeks to 16 weeks)												
4	randomised trials [3,4,5,6]	not serious	very serious [e]	not serious	serious [c]	none	279/692 (40.3%)	85/193 (44.0%)	RR 1.04 (0.72 to 1.51)	18 more per 1,000 (from 123 fewer to 225 more)	⊕◻◻◻ VERY LOW	IMPORTANT
Serious Adverse Events (Risk ratios less than one favor Anti-Nerve Growth Factor) (follow-up: range 8 weeks to 16 weeks)												
6	randomised trials [1,2,3,4,5,6]	not serious	not serious	not serious	serious [c]	none	41/1676 (2.4%)	14/519 (2.7%)	RR 0.89 (0.49 to 1.62)	3 fewer per 1,000 (from 14 fewer to 17 more)	⊕⊕⊕◻ MODERATE	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio; **MD:** Mean difference

a. 1^2=75% moderate heterogeneity
b. 1^2= 68%; moderate heterogeneity
c. 95% CI crosses null
d. I^2= 72%; moderate heterogeneity
e. I^2= 77%

PICO 2.8.4 (knee): What are the benefits and harms of fibroblast growth factor (FGF) therapy in the management of patients with knee OA?

SUMMARY
There is one trial of 190 participants with knee OA evaluating the effects of intra-articular injection of sprifermin as a single treatment and multiple-dose regimen (three doses of either 10, 30 or 100 µg). Results found that all groups had improved WOMAC pain scores, with statistically significantly less improvement at 12 months in participants receiving the 100 µg dose of sprifermin, compared with participants receiving placebo. No statistically significant relationship between treatment group and reduction in central medial femorotibial compartment cartilage thickness was observed. However, sprifermin was associated with statistically significant, dose-dependent reductions in the loss of total and lateral femorotibial cartilage thickness and volume, and in joint space widening in the lateral femorotibial compartment. The reasons for the seemingly preferential effect on the lateral knee compartment in the present and previous studies are not clear. In OA, the status of cartilage differs between the medial and lateral femorotibial compartments, with the medial compartment more commonly severely affected. An anabolic agent acting on cartilage may be less effective in tissue that is severely damaged. Currently, sprifermin is expensive and mainly available in phase II trials. No trial has investigated the benefits and safety of sprifermin in people with hip OA. Using knee OA data to extrapolate to hip or other OA requires additional caution.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Lohmander, et al. Arthritis Rheumatol. 2014 Jul; 66(7): 1820-31; 2. Dahlberg, et al. Clin Exp Rheumatol. 2016 May-Jun; 34(3): 445-50.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Fibroblast Growth Factor	Placebo	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (Scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	N=122	N=41	**MD 2.14 higher (0.61 higher to 3.67 higher)**		⊕⊕⊕□ MODERATE	CRITICAL
WOMAC Function (Scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious [a,b]	none	N=122	N=41	MD 4.23 higher (0.64 lower to 9.1 higher)		⊕⊕□□ LOW	CRITICAL
Medial Joint Space Widening [mm] (Positive values indicate better structural outcome) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious [a,b]	none	N=120	N=38	MD 0.03 higher (0.3 lower to 0.36 higher)		⊕⊕□□ LOW	IMPORTANT
Lateral Joint Space Widening [mm] (Positive values indicate better structural outcome) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [b]	none	N=120	N=38	**MD 0.31 higher (0.03 higher to 0.59 higher)**		⊕⊕⊕□ MODERATE	IMPORTANT
Medial Femorotibial Compartment Cartilage Thickness [mm] (Higher values indicate better structural outcome) (follow up: 12 months)												

1	randomised trial [1]	not serious	not assessable	not serious	very serious [a,b]	none	N=115	N=37	MD 0.05 higher (0.03 lower to 0.12 higher)		⊕⊕◯◯ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Fibroblast Growth Factor) (follow-up: range 24 weeks to 12 months)												
2	randomised trials [1,2]	serious [c]	not serious	not serious	serious [a]	none	3/199 (1.5%)	1/66 (1.5%)	RR 1.00 (0.11 to 9.39)	0 fewer per 1,000 (from 13 fewer to 127 more)	⊕⊕◯◯ LOW	CRITICAL
Treatment-emergent Adverse Events (Risk ratios less than one favor Fibroblast Growth Factor) (follow-up: range 24 weeks to 12 months)												
2	randomised trials [1,2]	serious [c]	very serious [d]	not serious	serious [a]	none	131/199 (65.8%)	42/66 (63.6%)	RR 0.94 (0.48 to 1.82)	38 fewer per 1,000 (from 331 fewer to 522 more)	⊕◯◯◯ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Fibroblast Growth Factor) (follow-up: 24 weeks)												
1	randomised trial [2]	serious [c]	not assessable	not serious	serious [b]	none	5/55 (9.1%)	5/18 (27.8%)	**RR 0.33 (0.11 to 1.00)**	**186 fewer per 1,000 (from 0 fewer to 247 fewer)**	⊕⊕◯◯ LOW	CRITICAL
Local Treatment-Emergent Adverse Events (Risk ratios less than one favor Fibroblast Growth Factor) (follow-up: range 24 weeks to 12 months)												
2	randomised trials [1,2]	serious [c]	not serious	not serious	serious [a]	none	50/199 (25.1%)	12/66 (18.2%)	RR 1.37 (0.78 to 2.41)	67 more per 1,000 (from 40 fewer to 256 more)	⊕⊕◯◯ LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. 95% CI crosses null
b. Sample size in one study arm <50.
c. Dahlberg, 2016: 25% more patients in the Placebo group were Kellgren Lawrence Grade IV than in the Sprifermin groups at baseline
d. I^2=92%

PICO 2.8.5 (knee): What are the benefits and harms of colchicine in the management of patients with knee OA?

SUMMARY

There is currently a lack of high-quality evidence supporting the use of colchicine for symptomatic relief for people with knee OA. While two small trials (one comparing colchicine to placebo; one comparing the combination of colchicine and an anti-inflammatory medication to the anti-inflammatory medication alone) indicate colchicine may provide symptomatic relief, its efficacy and safety remains unproven. In the trials, participants who received colchicine reported more gastrointestinal adverse effects, and the benefit to risk profile needs to be investigated in larger studies. One randomised placebo-controlled trial for people with knee OA that commenced enrolment of 120 participants in June 2014 in Singapore is reported to have been completed (Identifier: NCT02176460; ClinicalTrials.gov), but the results have not been published. One additional trial was identified in a search of the World Health Organization's (WHO's) International Clinical Trials Registry Platform (ICTRP). This trial is reported to have recruited 81 participants between March and September 2012 in Iran, and was retrospectively registered in September 2015 (IRCT2015071623240N1). These results have also not been published. There are currently no trials investigating the benefits and safety of colchicine in people with OA of the hip. Colchicine currently does not have an indication via the Therapeutic Goods Administration (TGA) for OA, and should be considered as an investigational medication only.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Das, et al. Arthritis Rheum. 2002 Jun 15; 47(3): 280-4 [concomitant nimesulide]; 2. Das, et al. Osteoarthritis Cartilage. 2002 Apr; 10(4): 247-52 [concomitant piroxicam]; 3. Aran, et al. Clin Exp Rheumatol. 2011 May-Jun; 29(3): 513-8 [concomitant OA treatment, various]; 4. Ediz, et al. Journal of Clinical and Analytical Medicine 3, no. 1 (2012): 63-67[concomitant acetaminophen].

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Colchicine**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
VAS Pain (scale range 0 to 15 cm, with higher scores indicating higher pain severity) (follow up: 20 weeks)												
2	randomised trials [1,2]	serious [a]	very serious [b]	not serious	very serious [c,d]	none	N=38	N=37	MD 1.24 lower (3.27 lower to 0.79 higher)		⊕□□□ VERY LOW	CRITICAL
Modified HAQ (Quality of Life) (range unclear, unvalidated measure) (Higher scores indicate better quality of life) (follow up: 20 weeks)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	serious [c]	none	N=38	N=37	**MD 3.13 lower (4.43 lower to 1.83 lower)**		⊕⊕□□ LOW	IMPORTANT
Patients' Global Assessment of Disease Severity (scale range 0 to 15 cm, with higher scores indicating higher disease severity) (follow up: range 3 months to 20 weeks)												
2	randomised trials [1,3]	serious [e]	very serious [f]	not serious	serious [c]	none	N=48	N=46	**MD 5.76 lower (10.17 lower to 1.35 lower)**		⊕□□□ VERY LOW	IMPORTANT
Physician's Global Assessment of Disease Severity (scale range 0 to 15 cm, with higher scores indicating higher disease severity) (follow up: range 3 months to 20 weeks)												

2	randomised trials [1,3]	serious [e]	very serious [g]	not serious	serious [c]	none	N=48	N=46	**MD 4.77 lower (7.32 lower to 2.21 lower)**		⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Colchicine) (follow up: range 3 months to 6 months)												
4	randomised trials [1,2,3,4]	not serious	not serious	not serious	serious [d]	none	2/101 (2.0%)	0/100 (0.0%)	RR 2.89 (0.31 to 26.79)	NA [h]	⊕⊕⊕□ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Colchicine) (follow up: 3 months)												
1	randomised trial [3]	not serious	not assessable	not serious	very serious [c,d]	none	1/30 (3.3%)	0/31 (0.0%)	RR 3.10 (0.13 to 73.16)	NA [h]	⊕⊕□□ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Colchicine) (follow up: 20 weeks)												
1	randomised trial [2]	not serious	not assessable	not serious	serious [c]	none	0/19 (0.0%)	0/20 (0.0%)	Due to zero events in both groups, an absolute risk reduction was not estimable.		⊕⊕⊕□ MODERATE	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Colchicine) (follow up: range 20 weeks to 6 months)												
2	randomised trials [2,4]	serious [i]	not serious	not serious	serious [d]	none	19/52 (36.5%)	15/52 (28.8%)	RR 1.26 (0.83 to 1.93)	75 more per 1,000 (from 49 fewer to 268 more)	⊕⊕□□ LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Both studies received High risk of bias ratings due to potential attrition bias
b. I^2=79%
c. Sample size <50 in each study arm.
d. 95% CI crosses null.
e. One study received a High risk of bias rating due to potential attrition bias
f. I^2=93%
g. I^2=77%
h. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.
i. One study received High risk of bias rating due to unblinded design.

PICO 2.8.6 (knee): What are the benefits and harms of methotrexate in the management of patients with knee OA?

SUMMARY

There is very low-quality evidence from one small trial of 56 participants who used 7.5 mg of methotrexate weekly versus placebo for painful knee OA, which did not find a reduction in pain at four months. Another open-label study evaluated the effects of methotrexate for pain relief in participants with knee OA. At 24 weeks, 13/30 participants (43%) achieved ≥30% reduction in Visual Analogue Scale (VAS) pain, of whom, seven (23%) had achieved ≥50% reduction. Conversely, four participants (13%) experienced a flare. Thirteen of 30 (43%) participants achieved Osteoarthritis Research Society International's responder criteria.(Wenham CY, et al. Rheumatology, 2013;52(5):888–92). An ongoing pragmatic phase III trial (ISRCTN77854383) has been designed to confirm these inconsistent findings. In terms of cost and access, methotrexate is a relatively cheap and widely available. Currently, there is no direct evidence for hip OA.Methotrexate does not currently have an indication via the TGA for OA, and should be considered as an investigational medication only.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. De Holanda, et al. Revista Brasileira de Reumatologia 47, no. 5 (2007): 334-340.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Methotrexate	Placebo	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (Scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 4 months)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious [a,b]	none	N=28	N=28	MD 0.62 higher (1.28 lower to 2.52 higher)		⊕⊕□□ LOW	CRITICAL
WOMAC Function (Scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 4 months)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious [a,b]	none	N=28	N=28	MD 3.18 lower (10.1 lower to 3.74 higher)		⊕⊕□□ LOW	CRITICAL
Lequesne Index (Scale range 0 to 24, with higher scores indicating poorer functional outcome) (follow up: 4 months)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious [a,b]	none	N=28	N=28	MD 0.1 lower (2.43 lower to 2.23 higher)		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Methotrexate) (follow up: 4 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [b]	none	0/29 (0.0%)	0/29 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕□ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Methotrexate) (follow up: 4 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [b]	none	0/29 (0.0%)	0/29 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕□ MODERATE	CRITICAL

Gastrointestinal Adverse Events (Risk ratios less than one favor Methotrexate) (follow up: 4 months)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious [a,b]	none	6/29 (20.7%)	6/29 (20.7%)	RR 1.00 (0.37 to 2.74)	0 fewer per 1,000 (from 130 fewer to 360 more)	⊕⊕◻◻ LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. 95% CI crosses null

b. Sample size <50 in each study arm

PICO 2.9: INTRA-ARTICULAR INJECTIONS
PICO 2.9.1 (knee): What are the benefits and harms of corticosteroids in the management of patients with knee OA?

SUMMARY
The studies upon which the recommendation is based were at serious risk of bias and generally small in size. The overall quality of the evidence was judged to be low to very low. Beneficial effects on knee pain and function were demonstrated at up to six weeks. These findings were not present when follow-up was extended to three months. For hip pain, the clinical benefits were demonstrated for up to 12 weeks; however, there is lack of long-term data. In addition, considering the complexity of the hip joint, image guidance would be required, which would further add to the costs.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Jüni, et al. Cochrane Database Syst Rev. 2015 Oct 22; (10):CD005328.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intra-articular corticosteroid	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain [immediate] (Higher scores indicate higher pain severity) (follow up: range 4 weeks to 6 weeks)												
26	randomised trials [1]	serious [a]	serious [b]	not serious	not serious	none	N=922	N=827	**SMD 0.4 lower (0.58 lower to 0.22 lower)**		⊕⊕□□ LOW	CRITICAL
Pain [short term] (Higher scores indicate higher pain severity) (follow up: mean 3 months)												
18	randomised trials [1]	serious [a]	serious [c]	not serious	not serious	none	N=646	N=587	SMD 0.22 lower (0.44 lower to 0)		⊕⊕□□ LOW	CRITICAL
Pain [moderate term] (Higher Scores indicate higher pain severity) (follow up: mean 6 months)												
7	randomised trials [1]	serious [a]	not serious	not serious	serious [d]	none	N=267	N=259	SMD 0.7 lower (0.25 lower to 0.11 higher)		⊕⊕□□ LOW	CRITICAL
Function [immediate] (Higher scores indicate poorer functional outcome) (follow up: range 4 weeks to 6 weeks)												
15	randomised trials [1]	serious [a]	serious [c]	not serious	not serious	none	N=546	N=468	**SMD 0.33 lower (0.56 lower to 0.09 lower)**		⊕⊕□□ LOW	CRITICAL
Function [short term] (Higher scores indicate poorer functional outcome) (follow up: mean 3 months)												
11	randomised trials [1]	serious [a]	serious [e]	not serious	serious [d]	none	N=433	N=367	SMD 0.13 lower (0.37 lower to 0.1 higher)		⊕□□□ VERY LOW	CRITICAL
Function [moderate term] (Higher scores indicate poorer functional outcome) (follow up: mean 6 months)												

4	randomised trials [1]	serious [a]	not serious	not serious	serious [d]	none	N=168	N=160	**SMD 0.06 higher (0.16 lower to 0.28 higher)**		⊕⊕□□ LOW	CRITICAL
Quality of Life (Higher scores indicate better quality of life) (follow up: range 4 weeks to 6 weeks)												
2	randomised trials [1]	serious [a]	not serious	not serious	serious [d]	none	N=92	N=92	SMD 0.01 lower (0.3 lower to 0.28 higher)		⊕⊕□□ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Intra-articular corticosteroids) (follow up: range 20 weeks to 26 weeks)												
2	randomised trials [1]	serious [a]	not serious	not serious	serious [d]	none	1/102 (1.0%)	4/102 (3.9%)	RR 0.33 (0.05 to 2.07)	26 fewer per 1,000 (from 37 fewer to 42 more)	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Intra-articular corticosteroids) (follow up: range 24 weeks to 26 weeks)												
2	randomised trials [1]	serious [a]	not serious	not serious	very serious [d,f]	none	22/42 (52.4%)	24/42 (57.1%)	RR 0.89 (0.64 to 1.23)	63 fewer per 1,000 (from 131 more to 206 fewer)	⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Intra-articular corticosteroids) (follow up: range 9 weeks to 26 weeks)												
5	randomised trials [1]	serious [a]	not serious	not serious	serious [d]	none	4/165 (2.4%)	5/166 (3.0%)	RR 0.80 (0.22 to 2.94)	6 fewer per 1,000 (from 23 fewer to 58 more)	⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. Most trials had high or unclear risk of bias overall
b. I^2= 68%; moderate heterogeneity
c. I^2=69%; moderate heterogeneity
d. 95% CI crosses null
e. I^2=62%; moderate heterogeneity
f. Sample size in each study arm <50.

PICO 2.9.2 (knee): What are the benefits and harms of viscosupplementation in the management of patients with knee OA?

SUMMARY
The major analyses upon which the recommendation is based were considered to be at serious risk of bias, but the large number of studies analysed involved, in total, a large number of participants. For knee pain, function and adverse events, the overall quality of the evidence was judged to be moderate. Despite some inconsistency on the conclusions among the analyses, a positive effect, albeit small and not clinically relevant, was demonstrated for pain and function. The recommendation for hip OA is based on three small randomised controlled trials (RCTs), which were judged to not be at serious risk of bias.The overall quality of evidence was judged to low. No effect on pain nor function was demonstrated, and the risk of total and serious adverse events and local reactions was greater in the viscosupplementation group. In addition, for a hip injection,image guidance would be required, further adding to complexity and cost. The increased risk of total and serious adverse events are of concern.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. Bannuru, et al. Ann Intern Med. 2015 Jan 6; 162(1): 46-54; 2. Bannuru, et al. Osteoarthritis Cartilage. 2016 Dec;24(12):2022-2041.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	IA Hyaluronic Acid	IA Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: 12 weeks)												
52	randomised trials [1]	serious [a]	not serious	not serious	not serious	none	N=3852	N=3413	**SMD 0.34 lower (0.42 lower to 0.26 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: 12 weeks)												
23	randomised trials [1]	serious [a]	not serious	not serious	not serious	none	N=1546	N=1831	**SMD 0.3 lower (0.4 lower to 0.2 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Intra-articular Hyaluronic Acid) (follow up range: 4 weeks to 12 months)												
41	randomised trials [2]	serious [a]	not serious	not serious	serious [b]	none	153/4247 (3.6%)	111/3676 (3.0%)	RR 1.19 (0.94 to 1.52)	6 more per 1,000 (from 2 fewer to 16 more)	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Intra-articular Hyaluronic Acid) (follow up range: 4 weeks to 12 months)												
40	randomised trials [2]	serious [a]	not serious	not serious	not serious	none	1699/3935 (43.2%)	1333/3362 (39.6%)	**RR 1.09 (1.03 to 1.15)**	**36 more per 1,000 (from 12 more to 59 more)**	⊕⊕⊕□ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Intra-articular Hyaluronic Acid) (follow up range: 4 weeks to 12 months)												

35	randomised trials [2]	serious [a]	not serious	not serious	not serious	none	103/3530 (2.9%)	60/2874 (2.1%)	**RR 1.40 (1.02 to 1.91)**	**8 more per 1,000 (from 0 fewer to 19 more)**	⊕⊕⊕◻ MODERATE	CRITICAL
Local Reactions (Risk ratios less than one favor Intra-articular Hyaluronic Acid) (follow up range: 4 weeks to 12 months)												
43	randomised trials [2]	serious [a]	not serious	not serious	not serious	none	546/4152 (13.2%)	367/3615 (10.2%)	**RR 1.30 (1.14 to 1.47)**	**30 more per 1,000 (from 14 more to 48 more)**	⊕⊕⊕◻ MODERATE	IMPORTANT
Septic Joint (Risk ratios less than one favor Intra-articular Hyaluronic Acid) (follow up range: 4 weeks to 12 months)												
15	randomised trials [2]	serious [a]	not serious	not serious	serious [b]	none	0/1120 (0.0%)	1/931 (0.1%)	RR 0.28 (0.01 to 6.79)	1 fewer per 1,000 (from 1 fewer to 6 more)	⊕⊕◻◻ LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. Studies were overall low to moderate quality
b. 95% CI crosses null

PICO 2.9.3 (knee): What are the benefits and harms of platelet-rich plasma in the management of patients with knee OA?

SUMMARY

The studies upon which the recommendation is based were at serious risk of bias and inconsistency, and were generally small in size. The overall quality of the evidence was judged to be very low. Beneficial effects on both knee pain and Western Ontario and McMaster Universities (WOMAC) function were demonstrated at six months. With the concern of potential reporting bias and low-quality data, the beneficial effects are likely to be overinflated. In addition, there is no consensus on eligible participant selection, number and frequency of injections, preparation technique, or appropriate platelet concentration,(Chang KV, et al. Arch Phys Med Rehabil, 2014;95(3):562–75) leading to large variationsin the design of PRP trials. No RCT was conducted in hip OA. However, during working group discussions, it was suggested that the mechanism of action should be no different in hip OA. Therefore, the findings might be transferrable to hip OA, but with a particular caution in terms of the complexity of the hip joint. The cost of PRP treatment is high, and additional equipment might be required for the preparation and administration.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Patel, et al. Am J Sports Med. 2013 Feb; 41(2): 356-64; 2. Rayegani, et al. Orthop Rev (Pavia). 2014 Sep 18; 6(3): 5405; 3. Smith, Patrick A. Am J Sports Med. 2016 Apr; 44(4): 884-91 4. Görmeli, et al. Knee Surg Sports Traumatol Arthrosc. 2017 Mar;25(3):958-965.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Platelet-rich Plasma**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow-up: 6 months)												
4	randomised trials [1,2,3,4]	serious [a]	very serious [b]	not serious	not serious	none	N=231	N=132	**SMD 1.87 lower (2.47 lower to 1.27 lower)**		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (Scale range 0 to 68, wither higher scores indicating poorer functional outcome) (follow-up: 6 months)												
3	randomised trials [1,2,3]	serious [a]	very serious [c]	not serious	not serious	none	N=148	N=92	**MD 16.07 lower (22.76 lower to 9.37 lower)**		⊕□□□ VERY LOW	CRITICAL
SF-36 Physical Component Score (Scale range 0 to 100, with higher scores indicating better quality of life) (follow-up: 6 months)												
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [d]	none	N=31	N=31	**MD 9.54 higher (0.22 higher to 18.86 higher)**		⊕⊕□□ LOW	IMPORTANT
SF-36 Mental Component Score (Scale range 0 to 100, with higher scores indicating better quality of life) (follow-up: 6 months)												
1	randomised trial [2]	serious [a]	not assessable	not serious	very serious [d,e]	none	N=31	N=31	MD 10.17 higher (0.67 lower to 21.01 higher)		⊕□□□ VERY LOW	IMPORTANT
Treatment-related Adverse Events (Risk Ratios less than one favor Platelet-rich Plasma) (follow-up: range 6 months to 12 months)												

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Platelet-rich Plasma	Placebo	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials [1,3]	not serious	not serious	not serious	serious [e]	none	17/117 (14.5%)	1/61 (1.6%)	RR 3.06 (0.39 to 23.87)	34 more per 1,000 (from 10 fewer to 375 more)	⊕⊕⊕◻ MODERATE	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. 2 studies received at least one High risk of bias rating; 1 due to unblinded design and 1 due to potential reporting bias

b. I^2=81%

c. I^2=77%

d. Sample size in each study arm <50.

e. 95% CI crosses null

PICO 2.9.4 (knee): What are the benefits and harms of stem cell therapy in the management of patients with knee OA?

SUMMARY
The two studies upon which the recommendation is based were at very serious risk of bias and were small in size. The overall quality of the evidence was judged to be low to very low. Beneficial effects on pain and function were demonstrated at up to six months. The between-group differences reported for pain and function appeared to be remarkably good. As they deviate significantly from those of other successful interventions, replication is required in high quality, large RCTs before a more favourable recommendation can be considered. onsistent with a recent position statement from the Australian College of Sports and Exercise Physicians, stem cell administration should be part of a rigorously designed study and the priority for individual health and welfare. (Osborne H, et al. Br J Sports Med 2016;50(20):1237–44).

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Varma HS, et al. J Indian Med Assoc 2010;108:583–58; 2. Tan, et al. J Tradit Chin Orthop Traumatol 2013;10:35–8.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**IA Mesenchymal Stem Cells**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
VAS Pain (scale range 0 to 10, with higher scores indicating higher pain severity) (follow up: 6 months)												
1	randomised trial [1]	very serious [a]	not assessable	not serious	serious [b]	none	N=25	N=25	**MD 3.4 lower (3.94 lower to 2.86 lower)**		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 6 months to 12 months)												
2	randomised trials [1,2]	very serious [a]	very serious [c]	not serious	not serious	none	N=61	N=61	**SMD 5.05 lower (7.01 lower to 3.1 lower)**		⊕□□□ VERY LOW	CRITICAL
Change in Cartilage Thickness [mm] (Higher values indicate better structural outcome) (follow up: 12 months)												
1	randomised trial [2]	very serious [a]	not assessable	not serious	serious [b]	none	N=36	N=36	**MD 0.6 higher (0.54 higher to 0.66 higher)**		⊕□□□ VERY LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **SMD:** Standardised mean difference

a. Both trials received High risk of bias ratings, and/or "Low" quality ratings. Adverse Events are not reported for either study.
b. Sample size <50 in each study arm.
c. I^2= 83%

PICO 2.9.5 (knee): What are the benefits and harms of dextrose prolotherapy in the management of patients with knee OA?

SUMMARY
The recommendation is based on the evidence of only one small RCT of low quality. The risk of bias in this study was not determined to be serious. No clinically significant effects were found for pain at 24 and 52 weeks follow-up. In terms of function, no clinically significant effects were found for pain at 24 weeks, but a marginally significant effect was recorded at 52 weeks. Furthermore, high-quality RCTs with low risk of bias and specifically for hip OA are required. As prolotherapy is relatively cheap and accessible, it is likely to be injudiciously used.
OVERALL QUALITY OF EVIDENCE: ⊕⊕□□ LOW

Bibliography: 1.Rabago, et al. Ann Fam Med. 2013 May-Jun; 11(3): 229-37.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Dextrose Prolotherapy	Placebo	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (Scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 24 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious [a,b]	none	N=28	N=25	MD 9.1 lower (19.07 lower to 0.87 higher)		⊕⊕□□ LOW	CRITICAL
WOMAC Pain (Scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 52 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious [a,b]	none	N=26	N=25	MD 6.8 lower (16.9 lower to 3.3 higher)		⊕⊕□□ LOW	CRITICAL
WOMAC Function (Scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 24 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious [a,b]	none	N=28	N=25	MD 9.57 lower (19.43 lower to 0.29 higher)		⊕⊕□□ LOW	CRITICAL
WOMAC Function (Scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 52 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [b]	none	N=26	N=25	**MD 10.79 lower (20.26 lower to 1.32 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Dextrose Prolotherapy) (follow up: 52 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [b]	none	0/30 (0.0%)	0/29 (0.0%)	Due to zero events in both study arms, an effect was not estimable.		⊕⊕⊕□ MODERATE	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio
a. 95% CI crosses null
b. Sample size <50 in each study arm. (Authors noted the following for adverse event collection: "The study was not large enough to detect uncommon adverse events, such as intolerance to study medication or rare injection-related sequelae")

GRADE tables for knee osteoarthritis

Section 3: Surgical Interventions

PICO 3.1 (knee): What are the benefits and harms of arthroscopic lavage and debridement interventions in the management of patients with knee OA?

SUMMARY (3.1 Arthroscopic, lavage and debridement,3.2 meniscectomy and 3.3 cartilage repair)
There is very low-quality evidence that there is no apparent benefit in terms of pain, function or quality of life for joint lavage, debridement and meniscectomy in the setting of knee OA. Arthroscopy occurs more commonly in the private hospital setting than public hospitals. It is important to note that arthroscopy rates in knee OA have been declining in the past few years. In the context of an intervention where there is a debatable benefit, measurable costs and potentially serious harms, the working group strongly recommends against the use of arthroscopy for lavage and debridement in the setting of knee OA.The Australian Orthopaedic Association and the Knee Society position statement (www.kneesociety.org.au/resources/aksarthroscopy-position-statement.pdf) strongly states that arthroscopy is not indicated for the treatment of knee OA. In the infrequent instance where exercise fails to release the locked knee, arthroscopy could be indicated.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Moseley, et al. N Engl J Med. 2002 Jul 11; 347(2): 81-8.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Arthroscopic Lavage and Debridement	Sham Surgery	Relative (95% CI)	Absolute (95% CI)		
AIMS Pain (Scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 6 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=55	N=57	MD 2.2 higher (5.45 lower to 9.85 higher)		⊕⊕□□ LOW	CRITICAL
AIMS Pain (Scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 12 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=51	N=54	MD 0.1 lower (8.66 lower to 8.46 higher)		⊕⊕□□ LOW	CRITICAL
AIMS Pain (Scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 24 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=53	N=55	MD 1.7 higher (6.85 lower to 10.25 higher)		⊕⊕□□ LOW	CRITICAL
"Physical Functioning Scale" (Number of seconds to walk 30 m, climb and descend a flight of stairs; score values always positive and technically infinite, and higher scores mean worse outcome) (follow up: 6 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=54	N=54	MD 0.8 lower (7.06 lower to 5.46 higher)		⊕⊕□□ LOW	CRITICAL
"Physical Functioning Scale" (Number of seconds to walk 30 m, climb and descend a flight of stairs; score values always positive and technically infinite, and higher scores mean worse outcome) (follow up: 12 months)												

1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	N=47	N=49	MD 3.3 higher (3.52 lower to 10.12 higher)	⨁◯◯◯ VERY LOW	CRITICAL
"Physical Functioning Scale" (Number of seconds to walk 30 m, climb and descend a flight of stairs; score values always positive and technically infinite, and higher scores mean worse outcome) (follow up: 24 months)											
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	N=44	N=44	MD 1.3 higher (5.48 lower to 8.08 higher)	⨁◯◯◯ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference

a. Study received High risk of bias rating because authors discuss potential selection bias and non-representative cohort.
b. 95% CI crosses null.
c. Sample size <50 in each study arm.

PICO 3.2 (knee): What are the benefits and harms of arthroscopic meniscectomy interventions in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. Sihvonen, et al. N Engl J Med. 2013 Dec 26; 369(26): 2515-24.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Arthroscopic Partial Meniscectomy	Sham Surgery	Relative (95% CI)	Absolute (95% CI)		
Knee Pain after Exercise (Scale range 0 to 10, with higher scores indicating more severe pain) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	N=70	N=76	MD 0.2 higher (0.57 lower to 0.97 higher)		⊕⊕□□ LOW	CRITICAL
Lysholm Knee Score (Designed to evaluate knee function and symptoms in activities of daily living) (Scale range 0-100, with higher scores indicating better outcome) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	N=70	N=76	MD 1.6 lower (7.16 lower to 3.96 higher)		⊕⊕□□ LOW	CRITICAL
WOMET [The Western Ontario Meniscal Evaluation Tool] (Scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	N=70	N=76	MD 2.5 lower (9.1 lower to 4.1 higher)		⊕⊕□□ LOW	NOT IMPORTANT
15D Quality of Life (Scale range 0 to 1, with higher scores indicating better quality of life) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	N=70	N=76	MD 0 (0.01 lower to 0.01 higher)		⊕⊕□□ LOW	IMPORTANT
Subsequent High Tibial Osteotomy or Total Knee Replacement (Risk ratios less than one favor Arthroscopic Partial Meniscectomy) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	1/70 (1.4%)	1/76 (1.3%)	RR 1.09 (0.07 to 17.03)	1 more per 1,000 (from 12 fewer to 211 more)	⊕⊕□□ LOW	IMPORTANT

Patients Reporting Improvement (Risk ratios greater than one favor Arthroscopic Partial Meniscectomy) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	62/70 (88.6%)	63/76 (82.9%)	RR 1.07 (0.94 to 1.22)	58 more per 1,000 (from 50 fewer to 182 more)	⊕⊕◻◻ LOW	IMPORTANT
Serious Adverse Events (Risk ratios less than one favor Arthroscopic Partial Meniscectomy) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	1/70 (1.4%)	0/76 (0.0%)	RR 3.25 (0.13 to 78.58)	NA [c]	⊕⊕◻◻ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Patients in this study have "degenerative meniscal tear and no knee osteoarthritis"; Kellgren Lawrence grade >I was an exclusion criterion. 54.3% of patients in the arthroscopic partial meniscectomy group and 39.5% of patients in the sham group had chondral degeneration observed by arthroscopy that was classified as "Degenerative". 25.7% and 27.6% of patients in the partial meniscectomy and sham groups, respectively, had chondral degeneration which was classified as "Osteoarthritic". Per the authors: "...increasing evidence suggests that a degenerative meniscal tear may be an early sign of knee osteoarthritis rather than a separate clinical problem requiring meniscal intervention."

b. 95% CI crosses null.

c. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.

PICO 3.3 (knee): What are the benefits and harms of arthroscopic procedures for cartilage repair interventions in the management of patients with knee OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

No RCT data was found for this question.

GRADE tables for knee osteoarthritis

Section 4: Combination Therapy

PICO 4.1 (knee): What are the benefits and harms of combination weight management and exercise interventions compared to mono-therapy in patients with knee OA?

SUMMARY
There is limited evidence of very low quality that weight loss alone (achieved through diet and exercise) has no significant effect on either pain or function in people with knee OA although benefits appear to be more significant with higher amounts of weight loss – starting at a minimum of 5–7.5% body weight loss. Ideally, it is beneficial to achieve a greater amount of weight loss.(Atukorala I, et al. Arthritis Care Res 2016;68(8):1106–14; Messier SP, et al. JAMA, 2013;310(12):1263–73). There is no randomised controlled trial (RCT) on the effects of bariatric surgery in people with hip and/or knee OA. There are no RCTs investigating weight management specifically in people with hip OA, although the numerous other systemic health benefits of weight loss and maintaining a healthy body weight are most likely transferable to people with hip OA. There are limitations of the available RCT evidence in OA. Clinicians are advised to refer to the National Health and Medical Research Council's (NHMRC's) guidelines for the most effective strategies for managing overweight/obesity in primary care. National Health and Medical Research Council. Clinical practice guidelines for the management of overweight and obesity in adults, adolescents and children in Australia. Melbourne: NHMRC, 2013).

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

WEIGHT MANAGEMENT+EXERCISE VS. EXERCISE ALONE

Bibliography: 1. Messier, et al. Arthritis Rheum. 2004 May; 50(5): 1501-10 (ADAPT); 2. Messier, et al. JAMA. 2013 Sep 25; 310(12): 1263-73 (IDEA); 3. Focht, et al. Arthritis Rheum. 2005 Oct 15; 53(5): 659-65 (ADAPT).

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Weight Management + Exercise**	**Exercise Alone**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 18 months)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	not serious	none	N=205	N=208	**MD 1.43 lower (2.09 lower to 0.77 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
WOMAC Function (scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 18 months)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	not serious	none	N=205	N=208	**MD 4.21 lower (6.46 lower to 1.95 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Walking Self-Efficacy (Patient confidence in walking around a gymnasium twice without stopping; score range 0-100 with higher scores indicating more confidence) (follow up: 18 months)												
1	randomised trial [3]	serious [a]	not assessable	not serious	serious [b]	none	N=76	N=80	MD 3.97 higher (5.52 lower to 13.46 higher)		⊕⊕□□ LOW	IMPORTANT
SF-36 Physical Component Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 18 months)												

1	randomised trial [2]	serious [a]	not assessable	not serious	not serious	none	N=129	N=128	**MD 3 higher (0.76 higher to 5.24 higher)**	⊕⊕⊕◯ MODERATE	IMPORTANT
Percentage Weight Loss (Percentage of weight at baseline lost by follow up time, with more loss indicating positive outcome) (follow up: 18 months)											
2	randomised trials [1,2]	serious [a]	very serious [c]	not serious	serious [b]	none	N=197	N=195	MD 5.42% more lost (1.86% less lost to 12.69% more lost)	⊕◯◯◯ VERY LOW	IMPORTANT
Lateral Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: 18 months)											
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=76	N=80	MD 0.16 mm lower (0.66 mm lower to 0.34 mm higher)	⊕⊕◯◯ LOW	IMPORTANT
Medial Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: 18 months)											
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=76	N=80	MD 0.11 mm lower (0.68 mm lower to 0.46 mm higher)	⊕⊕◯◯ LOW	IMPORTANT
Non-Compliance with Regimen [defined as "non-adherence"] (Risk ratios less than one favor Weight Management + Exercise) (follow up: 18 months)											
2	randomised trials [1,2]	serious [a]	not serious	not serious	serious [b]	none	91/228 (39.9%)	101/230 (43.9%)	RR 0.91 (0.73 to 1.13)	40 fewer per 1,000 (from 57 more to 119 fewer)	⊕⊕◯◯ LOW	IMPORTANT
Withdrawal due to Lack of Interest (Risk ratios less than one favor Weight Management + Exercise) (follow up: 18 months)											
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [b]	none	2/152 (1.3%)	2/150 (1.3%)	RR 0.99 (0.14 to 6.92)	0 fewer per 1,000 (from 11 fewer to 79 more)	⊕⊕◯◯ LOW	IMPORTANT
Serious Adverse Events (Risk ratios less than one favor Weight Management + Exercise) (follow up: 18 months)											
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [b]	none	6/152 (3.9%)	3/150 (2.0%)	RR 1.97 (0.50 to 7.75)	19 more per 1,000 (from 10 fewer to 135 more)	⊕⊕◯◯ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Both studies received High risk of bias ratings due to single blinded design, and due to potential for attrition bias.
b. 95% CI crosses null.
c. I^2= 79%

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

WEIGHT MANAGEMENT+EXERCISE VS. WEIGHT MANAGEMENT ALONE

Bibliography: 1. Messier, et al. Arthritis Rheum. 2004 May; 50(5): 1501-10 (ADAPT); 2. Messier, et al. JAMA. 2013 Sep 25; 310(12): 1263-73 (IDEA); 3. Focht, et al. Arthritis Rheum. 2005 Oct 15; 53(5): 659-65 (ADAPT).

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Weight management + Exercise**	**Weight management alone**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 18 months)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	not serious	none	N=205	N=206	**MD 1.18 lower (1.85 lower to 0.51 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
WOMAC Function (scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 18 months)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	not serious	none	N=205	N=206	**MD 2.7 lower (4.99 lower to 0.4 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Walking Self-Efficacy (Patient confidence in walking around a gymnasium twice without stopping; score range 0-100 with higher scores indicating more confidence) (follow up: 18 months)												
1	randomised trial [3]	serious [a]	not assessable	not serious	not serious	none	N=76	N=82	**MD 13.4 higher (3.84 higher to 22.96 higher)**		⊕⊕⊕□ MODERATE	IMPORTANT
SF-36 Physical Component Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 18 months)												
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [b]	none	N=129	N=124	MD 2 higher (0.27 lower to 4.27 higher)		⊕⊕□□ LOW	IMPORTANT
Percentage Weight Loss (Percentage of weight at baseline lost by follow up time, with more loss indicating positive outcome) (follow up: 18 months)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	serious [b]	none	N=197	N=202	MD 1.14% more lost (1.88% less lost to 4.17% more lost)		⊕⊕□□ LOW	IMPORTANT
Lateral Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: 18 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=76	N=82	MD 0.16 mm lower (0.65 mm lower to 0.33 mm higher)		⊕⊕□□ LOW	IMPORTANT

Medial Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: 18 months)											
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=76	N=82	MD 0.1 mm lower (0.62 mm lower to 0.42 mm higher)	⊕⊕◯◯ LOW	IMPORTANT
Non-Compliance with Regimen [defined as "non-adherence"] (Risk ratios less than one favor Weight Management + Exercise) (follow up: 18 months)											
2	randomised trials [1,2]	serious [a]	not serious	not serious	serious [b]	none	83/228 (36.4%)	82/234 (35.0%)	RR 1.03 (0.80 to 1.34)	11 more per 1,000 (from 70 fewer to 119 more)	⊕⊕◯◯ LOW	IMPORTANT
Withdrawal due to Lack of Interest (Risk ratios less than one favor Weight Management + Exercise) (follow up: 18 months)											
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [b]	none	2/152 (1.3%)	3/152 (2.0%)	RR 0.67 (0.11 to 3.93)	7 fewer per 1,000 (from 18 fewer to 58 more)	⊕⊕◯◯ LOW	IMPORTANT
Serious Adverse Events (Risk ratios less than one favor Weight Management + Exercise) (follow up: 18 months)											
1	randomised trial [2]	serious [a]	not assessable	not serious	serious [b]	none	6/152 (3.9%)	1/152 (0.7%)	RR 6.00 (0.73 to 49.24)	33 more per 1,000 (from 2 fewer to 317 more)	⊕⊕◯◯ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Both studies received High risk of bias ratings due to single blinded design, and due to potential for attrition bias.
b. 95% CI crosses null.

PICO 4.2 (knee): What are the benefits and harms of combination exercise and cognitive behavioural interventions compared to mono-therapy in patients with knee OA?

SUMMARY
Low-quality evidence in people with knee OA shows that CBT programs can have small benefits for pain, and can also improve self-efficacy, pain coping, depression and anxiety (O'Moore KA, et al. 2017;70(1):61–70). CBT for knee OA is associated with a low risk of adverse events. Very low-quality evidence from a limited number of studies indicates that combining CBT with exercise is more effective than either alone. While there is no evidence of the effects of CBT, specifically in people with hip OA. Benefits seen in people with knee OA and mixed samples of hip/knee OA are likely to be generalisable to those with hip OA.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

COGNITIVE BEHAVIORAL THERAPY + EXERCISE VS. EXERCISE ALONE

Bibliography: 1. Bennell, et al. Arthritis Care Res (Hoboken). 2016 May; 68(5): 590-602; 2. Somers, et al. Pain. 2012 Jun; 153(6): 1199-209.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cognitive Behavioral Therapy + Exercise†	Exercise Alone	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 12 weeks to 24 weeks)												
2	randomised trials [1,2]	very serious [a]	serious [b]	not serious	not serious	none	N=130	N=126	**SMD 0.52 lower (1 lower to 0.03 lower)**		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 12 weeks to 24 weeks)												
2	randomised trials [1,2]	very serious [a]	not serious	not serious	not serious	none	N=130	N=126	**SMD 0.43 lower (0.68 lower to 0.18 lower)**		⊕⊕□□ LOW	CRITICAL
Self-Efficacy (Higher scores indicate higher self-efficacy) (follow up: range 12 weeks to 24 weeks)												
2	randomised trials [1,2]	very serious [a]	not serious	not serious	not serious	none	N=130	N=126	**SMD 0.28 higher (0.03 higher to 0.53 higher)**		⊕⊕□□ LOW	IMPORTANT
Depression (Higher scores indicate more severe depression) (follow up: range 12 weeks to 24 weeks)												

2	randomised trials [1,2]	very serious [a]	not serious	not serious	serious [c]	none	N=130	N=126	SMD 0.04 lower (0.29 lower to 0.21 higher)		⊕□□□ VERY LOW	IMPORTANT
Treatment-related Adverse Events (Risk ratios less than one favor Cognitive Behavioral Therapy + Exercise) (follow up: range 12 weeks to 24 weeks)												
2	randomised trials [1,2]	very serious [a]	not serious	not serious	serious [c]	none	25/135 (18.5%)	28/134 (20.9%)	RR 0.90 (0.58 to 1.39)	21 fewer per 1,000 (from 81 more to 88 fewer)	⊕□□□ VERY LOW	CRITICAL
No Participation due to Lack of Interest [defined as withdrawal due to "no response" or "dissatisfaction" or "no longer interested"] (Risk ratios less than one favor Cognitive Behavioral Therapy + Exercise) (follow up: range 24 weeks to 52 weeks)												
2	randomised trials [1,2]	very serious [a]	serious [d]	not serious	serious [c]	none	4/135 (3.0%)	5/134 (3.7%)	RR 1.02 (0.08 to 12.39)	1 more per 1,000 (from 34 fewer to 425 more)	⊕□□□ VERY LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

† **The following Cognitive Behavioral Therapeutic Technique was included in the analysis:** Bennell 2016- Pain coping skills training + Strengthening exercise vs. Strengthening exercise; Somers 2012- Pain coping skills training + Diet and Exercise vs. Diet and Exercise

a. All studies received High risk of bias ratings due to single blind study design or inadequate description of blinding and potential attrition bias. Bennell 2016 received an additional High risk of bias for reporting bias due to errors in reporting.
b. I^2= 74%; moderate heterogeneity.
c. 95% CI crosses null.
d. I^2= 56%; moderate heterogeneity.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

COGNITIVE BEHAVIORAL THERAPY + EXERCISE VS. COGNITIVE BEHAVIORAL THERAPY ALONE

Bibliography: 1. Bennell, et al. Arthritis Care Res (Hoboken). 2016 May; 68(5): 590-602; 2. Keefe, et al. Pain. 2004 Aug; 110(3): 539-49; 3. Somers, et al. Pain. 2012 Jun; 153(6): 1199-209.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Cognitive Behavioral Therapy + Exercise†**	**Cognitive Behavioral Therapy Alone**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 12 weeks to 24 weeks)												
3	randomised trials [1,2,3]	very serious [a]	serious [b]	not serious	serious [c]	none	N=149	N=142	SMD 0.28 lower (0.64 lower to 0.08 higher)		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 12 weeks to 24 weeks)												
2	randomised trials [1,3]	very serious [a]	not serious	not serious	not serious	none	N=130	N=126	**SMD 0.77 lower (1.03 lower to 0.52 lower)**		⊕⊕□□ LOW	CRITICAL
Self-Efficacy (Higher scores indicate higher self-efficacy) (follow up: range 12 weeks to 24 weeks)												
3	randomised trials [1,2,3]	very serious [a]	not serious	not serious	not serious	none	N=149	N=142	**SMD 0.32 higher (0.09 higher to 0.56 higher)**		⊕⊕□□ LOW	IMPORTANT
Depression (Higher scores indicate more severe depression) (follow up: range 12 weeks to 24 weeks)												
3	randomised trials [1,2,3]	very serious [a]	not serious	not serious	serious [c]	none	N=149	N=142	SMD 0.02 lower (0.25 lower to 0.21 higher)		⊕□□□ VERY LOW	IMPORTANT
Treatment-related Adverse Events (Risk ratios less than one favor Cognitive Behavioral Therapy + Exercise) (follow up: range 12 weeks to 24 weeks)												
2	randomised trials [1,3]	very serious [a]	not serious	not serious	not serious	none	25/135 (18.5%)	4/134 (3.0%)	**RR 5.69 (2.18 to 14.86)**	**140 more per 1,000 (from 35 more to 414 more)**	⊕⊕□□ LOW	CRITICAL
No Participation due to Lack of Interest [defined as withdrawal due to "no response" or "dissatisfaction" or "no longer interested"] (Risk ratios less than one favor Cognitive Behavioral Therapy + Exercise) (follow up: range 24 weeks to 52 weeks)												
2	randomised trials [1,3]	very serious [a]	not serious	not serious	serious [c]	none	4/135 (3.0%)	7/134 (5.2%)	RR 0.57 (0.17 to 1.90)	22 fewer per 1,000 (from 43 fewer to 47 more)	⊕□□□ VERY LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

† **Different Cognitive Behavioral Therapeutic Techniques were analyzed together.** The following were included in the analysis: Bennell 2016- Pain coping skills training + Strengthening exercise vs. Strengthening exercise; Keefe 2004- Spouse-assisted pain coping skills training + Spouse-attended exercise vs. Spouse-assisted pain coping skills training; Somers 2012- Pain coping skills training + Diet and Exercise vs. Pain coping skills training

a. All studies received High risk of bias ratings due to single blind study design or inadequate description of blinding and potential attrition bias. Bennell 2016 received an additional High risk of bias for reporting bias due to errors in reporting.
b. I^2= 53%; moderate heterogeneity.
c. 95% CI crosses null.

GRADE tables for hip osteoarthritis

Section 1: Non-Pharmacologic Interventions

PICO 1.1 (hip): What are the benefits and harms of self-management education programmes in the management of patients with hip OA?

SUMMARY
Very low-quality evidence shows this intervention has no significant effect on pain and function.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1.Poulsen, et al. Osteoarthritis Cartilage. 2013 Oct; 21(10): 1494-503.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Self-Management Education**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
NRS Pain (scale range 0 to 10, with higher scores indicating higher pain severity) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	N=37	N=36	MD 0.6 higher (0.18 lower to 1.38 higher)		⊕□□□ VERY LOW	CRITICAL
HOOS Functioning in Daily Life (scale range 0 to 100, with higher scores indicating better function) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	N=37	N=36	MD 4 lower (9.33 lower to 1.33 higher)		⊕□□□ VERY LOW	CRITICAL
HOOS Hip-related Quality of Life (scale range 0 to 100, with higher scores indicating higher quality of life) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	N=37	N=36	**MD 6.0 lower (10.82 lower to 1.18 lower)**		⊕⊕□□ LOW	IMPORTANT
Total Adverse Events (Risk ratios less than one favor Self-Management Education) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	0/37 (0.0%)	2/36 (5.6%)	RR 0.19 (0.01 to 3.92)	45 fewer per 1,000 (from 55 fewer to 162 more)	⊕□□□ VERY LOW	CRITICAL
No Participation due to Lack of Interest [defined as withdrawal due to "lack of commitment" or "disappointed with group"] (Risk ratios less than one favor Self-Management Education) (follow up: 6 weeks)												

1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	1/39 (2.6%)	3/36 (8.3%)	RR 0.31 (0.03 to 2.83)	57 fewer per 1,000 (from 81 fewer to 153 more)	⨁◻◻◻ VERY LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Patients were not blinded to allocation. Blinding of outcome assessors could not be confirmed in all cases.
b. 95% CI crosses null.
c. Small sample size; <50 patients in each study arm.

PICO 1.3 (hip): What are the benefits and harms of cognitive behavioural therapy in the management of patients with hip OA?

SUMMARY

Low-quality evidence in people with knee OA shows that CBT programs can have small benefits for pain, and can also improve self-efficacy, pain coping, depression and anxiety (O'Moore KA, et al. 2017;70(1):61–70). CBT for knee OA is associated with a low risk of adverse events. Very low-quality evidence from a limited number of studies indicates that combining CBT with exercise is more effective than either alone. While there is no evidence of the effects of CBT, specifically in people with hip OA. Benefits seen in people with knee OA and mixed samples of hip/knee OA are likely to be generalisable to those with hip OA.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1.Broderick, et al. Pain. 2014 Sep; 155(9): 1743-54; 2. Gay, et al. Eur J Pain. 2002; 6(1): 1-16; 3. Murphy, et al. Pain. 2016 Jul; 157(7): 1563-73; 4. Rini, et al. Pain. 2015 May; 156(5): 837-48; 5. Allen, et al. Ann Intern Med. 2016 Jan 19; 164(2): 73-83.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Cognitive Behavioral Therapy†**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate more severe pain) (follow up: range 8 weeks to 10 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	not serious	serious [b]	not serious	none	N=330	N=286	**SMD 0.18 lower (0.35 lower to 0.02 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 8 weeks to 10 weeks)												
3	randomised trials [1,3,4]	serious [a]	not serious	serious [b]	serious [c]	none	N=317	N=245	SMD 0.1 lower (0.28 lower to 0.08 higher)		⊕□□□ VERY LOW	CRITICAL
Quality of Life [16 item scale] (scale range 16 to 112, with higher scores indicating better quality of life) (follow up: 10 weeks)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [c]	none	N=129	N=127	MD 1.44 higher (1.47 lower to 4.35 higher)		⊕□□□ VERY LOW	IMPORTANT
Self-Efficacy (Higher scores indicate higher self-efficacy) (follow up: range 8 weeks to 10 weeks)												

2	randomised trials [1,4]	serious [a]	not serious	serious [b]	not serious	none	N=187	N=182	**SMD 0.45 higher (0.19 higher to 0.71 higher)**		⊕⊕□□ LOW	IMPORTANT
Depression (Higher scores indicate more severe depression) (follow up: range 8 weeks to 12 months)												
3	randomised trials [1,4,5]	serious [a]	not serious	serious [b]	not serious	none	N=338	N=331	**SMD 0.23 lower (0.38 lower to 0.08 lower)**		⊕⊕□□ LOW	IMPORTANT
No Participation due to Lack of Interest [defined as any patient who did not return for first assessment due to "not liking the intervention", "unable to follow instructions", "could not contact"] (Risk Ratios less than one favor Cognitive Behavioral Therapy) (follow up: range 8 weeks to 10 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	not serious	serious [b]	serious [c]	none	10/330 (3.0%)	0/255 (0.0%)	RR 5.33 (0.96 to 29.51)	NA [d]	⊕□□□ VERY LOW	IMPORTANT
Daily Additional Use of Pain Medication (Risk ratios less than one favor Cognitive Behavioral Therapy) (follow up: 10 weeks)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	not serious	none	N=129	N=127	**MD 0.08 lower (0.16 lower to 0)**		⊕⊕□□ LOW	IMPORTANT

Allen, 2016: % knee: 79.3%, % hip: 10.7%, % both: 10.0%
Broderick, 2014: % knee: 77.4%, % hip: 22.6%
Gay, 2002: % knee: 66.7%, % hip: 69.4%
Murphy, 2016: % knee: 72.0%, % hip: 28.0%
Rini, 2015: % knee: 35.0%, % hip: 12.0%, % both: 52.0%

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

† **Different Cognitive Behavioral Therapeutic Techniques were analyzed together.** The following were included in the analysis: Broderick 2014- Nurse-delivered pain coping skills training; Gay 2002- Jacobson Relaxation technique; Murphy 2016- Tailored Activity Pacing and General Activity Pacing (2 arms, both included in this analysis); Rini 2015- Internet-based pain coping skills training; Allen 2016- Telephone-based intervention with educational element and cognitive behavioral pain management instruction.

a. All trials were single-blind. Due to self-reported outcomes, they were rated High risk of bias.
b. All trials involved mixed populations of patients with Hip and/or Knee Osteoarthritis.
c. 95% CI crosses null.
d. Due to zero events in the comparator arm, an absolute risk reduction could not be estimated.

PICO 1.4 (hip): What are the benefits and harms of all land-based exercise in the management of patients with hip OA? *(efficacy estimates only)*

Summary (1.4 all land based exercise, 1.5.1 muscle strengthening, 1.5.2 walking, 1.5.3 stationary cycling, 1.5.4 Tai Chi, 1.5.5 Hatha yoga)
Overall, there is moderate-quality evidence from a limited number of trials in people specifically with hip OA to support the short-term benefits of land-based exercise – conducted either at home or in groups – on pain and function . Exercise is also beneficial for other comorbidities and overall health. Most studies included multi-modal exercise programs comprising strengthening, range of motion and functional exercise. However, we are unable to recommend either for or against any specific type of exercise for hip OA because of limited or non-existent trials in people with hip OA that isolate the effects of different types of exercise. Clinicians should educate the individual about the benefits of regular exercise, and prescribe an individualised progressive exercise program, taking into account patient presentation, functional capacity, comorbidities, preferences and resource availability. Clinicians should emphasise that some discomfort may be experienced with exercise, but that this is not likely to be associated with harm. Some people may benefit from referral to an exercise professional (eg physiotherapist, exercise physiologist) to assist with exercise prescription and supervision. Attention should be paid to adherence strategies (eg written material, logbooks, SMS reminders).

OVERALL QUALITY OF EVIDENCE: ⊕⊕⊕□ MODERATE

Bibliography: 1. Fransen, et al. Cochrane Database Syst Rev. 2014 Apr 22; (4): CD007912 (meta analysis).

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Land-based exercise**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 6 weeks to 24 months)												
9	randomised trials [1]	serious [a]	not serious	not serious	not serious	none	N=282	N=267	**SMD 0.38 lower (0.55 lower to 0.20 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 6 weeks to 24 months)												
9	randomised trials [1]	serious [a]	not serious	not serious	not serious	none	N=269	N=252	**SMD 0.30 lower (0.54 lower to 0.05 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Quality of Life (Higher scores indicate better quality of life) (follow up: range 8 weeks to 12 weeks)												
3	randomised trials [1]	serious [a]	not serious	not serious	serious [b]	none	N=96	N=87	SMD 0.07 higher (0.23 lower to 0.36 higher)		⊕⊕□□ LOW	IMPORTANT
Study Withdrawals (Risk ratios less than one favor Land-based Exercise) (follow up: range 6 weeks to 24 months)												
7	randomised trials[1]	serious [a]	not serious	not serious	serious [b]	none	23/364 (6.3%)	12/351 (3.4%)	RR 1.85 (0.93 to 3.66)	29 more per 1,000 (from 2 fewer to 91 more)	⊕⊕□□ LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference
a. All studies received at least one High risk of bias rating due to inadequate blinding.
b. 95% CI crosses null.

PICO 1.5: SPECIFIC FORMS OF LAND-BASED EXERCISE

PICO 1.5.1 (hip): What are the benefits and harms of muscle strengthening in the management of patients with hip OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Fernandes, et al. Osteoarthritis Cartilage. 2010 Oct; 18(10): 1237-43; 2. Juhakoski, et al. Clin Rehabil. 2011 Apr; 25(4): 370-83; 3. Krauß, et al. Dtsch Arztebl Int. 2014 Sep 1; 111(35-36): 592-9..

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Muscle Strengthening Exercise**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 3 months to 4 months)												
3	randomised trials [1,2,3]	serious [a]	very serious [b]	not serious	serious [c]	none	N=186	N=180	SMD 0.02 lower (0.68 lower to 0.64 higher)		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional status) (follow up: range 3 months to 4 months)												
3	randomised trials [1,2,3]	serious [a]	very serious [d]	not serious	serious [c]	none	N=186	N=180	SMD 0.07 lower (0.58 lower to 0.44 higher)		⊕□□□ VERY LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Muscle strengthening Exercise) (follow up: range 3 months to 4 months)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	serious [c]	none	1/114 (0.9%)	0/112 (0.0%)	RR 2.95 (0.12 to 70.77)	NA [e]	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Muscle strengthening Exercise) (follow up: 4 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	1/55 (1.8%)	0/54 (0.0%)	RR 2.95 (0.12 to 70.77)	NA [e]	⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. All studies received at least one High risk of bias rating due to unblinded study design.
b. I^2= 90%
c. 95% CI crosses null.
d. I^2= 84%
e. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.

PICO 1.5.2 (hip): What are the benefits and harms of walking in the management of patients with hip OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Fransen, et al. Cochrane Database Syst Rev. 2015 Jan 9; 1: CD004376; 2. Talbot, et al. J Am Geriatr Soc. 2003 Mar; 51(3): 387-92; 3. Ettinger, et al. JAMA. 1997 Jan 1; 277(1): 25-31.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Walking Programs**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 12 weeks to 18 months)												
4	randomised trials [1]	serious [a]	not serious	serious [b]	not serious	none	N=225	N=126	**SMD 0.48 lower (0.83 lower to 0.13 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 12 weeks to 18 months)												
3	randomised trials [1]	serious [a]	not serious	serious [b]	not serious	none	N=208	N=109	**SMD 0.35 lower (0.58 lower to 0.11 lower)**		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Walking programs) (follow up: 12 weeks)												
1	randomised trial [2]	serious [a]	not assessable	serious [b]	very serious [c,d]	none	0/17 (0.0%)	1/17 (5.9%)	RR 0.33 (0.01 to 7.65)	39 fewer per 1,000 (from 58 fewer to 391 more)	⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Walking programs) (follow up: 18 months)												
1	randomised trial [3]	serious [a]	not assessable	serious [b]	serious [c]	none	2/144 (1.4%)	1/149 (0.7%)	RR 2.07 (0.19 to 22.57)	7 more per 1,000 (from 5 fewer to 145 more)	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. All studies received at least one High risk of bias rating for inadequate blinding or unblinded design.
b. In 3 of 4 trials, all patients have Knee Osteoarthritis. In Minor 1989, which is included in the Pain and Function analyses, the OA population is mixed and unclear.
c. 95% CI crosses null.
d. Sample size <50 in each study arm.

PICO 1.5.3 (hip): What are the benefits and harms of cycling in the management of patients with hip OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Salacinski, et al. J Orthop Sports Phys Ther. 2012 Dec; 42(12): 985-95.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Cycling**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 12 weeks)												
1	randomised trial [1]	very serious [a]	not assessable	serious [b]	very serious [c]	none	N=13	N=15	**MD 14.9 lower (25.3 lower to 4.5 lower)**		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (scale range 0 to 100, with higher scores indicating poorer functional outcome) (follow up: 12 weeks)												
1	randomised trial [1]	very serious [a]	not assessable	serious [b]	very serious [c,d]	none	N=13	N=15	MD 11.1 lower (23.74 lower to 1.54 higher)		⊕□□□ VERY LOW	CRITICAL
KOOS Knee-related Quality of Life (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 12 weeks)												
1	randomised trial [1]	very serious [a]	not assessable	serious [b]	very serious [c,d]	none	N=13	N=15	MD 6.8 higher (7.48 lower to 21.08 higher)		⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Cycling) (follow up: 12 weeks)												
1	randomised trial [1]	very serious [a]	not assessable	serious [b]	very serious [d,e]	none	4/19 (21.1%)	0/18 (0.0%)	RR 8.55 (0.49 to 148.33)	NA [f]	⊕□□□ VERY LOW	CRITICAL
Treatment-Related Adverse Events (Risk ratios less than one favor Cycling) (follow up: 12 weeks)												
1	randomised trial [1]	very serious [a]	not assessable	serious [b]	very serious [d,e]	none	3/19 (15.8%)	0/18 (0.0%)	RR 6.65 (0.37 to 120.36)	NA [f]	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Study received multiple High risk of bias ratings due to unblinded study design and potential attrition bias.

b. All patients in this trial have Knee Osteoarthritis.

c. Total sample size <30.

d. 95% CI crosses null.

e. Sample size <50 in each study arm.

f. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.

PICO 1.5.4 (hip): What are the benefits and harms of Tai Chi in the management of patients with hip OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Adler, P.A., 2007. *The effects of tai chi on pain and function in older adults with osteoarthritis* (Doctoral dissertation, Case Western Reserve University); 2. Brismée, et al. Clin Rehabil. 2007 Feb; 21(2): 99-111; 3. Fransen, et al. Arthritis Rheum. 2007 Apr 15; 57(3): 407-14; 4. Hartman, et al. J Am Geriatr Soc. 2000 Dec; 48(12): 1553-9; 5. Lee, et al. Clin Rehabil. 2009 Jun; 23(6): 504-11; 6. Song, et al. J Rheumatol. 2003 Sep; 30(9): 2039-44 ; 7. Song, et al. Journal of muscle and joint health. 2009; 16(1): 46-54; 8. Tsai, et al. J Pain Symptom Manage. 2013 Apr; 45(4): 660-9; 9. Wang, et al. Arthritis Rheum. 2009 Nov 15; 61(11): 1545-53.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Tai Chi**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 6 weeks to 12 weeks)												
9	randomised trials [1-9]	serious [a]	not serious	serious [b]	not serious	none	N=231	N=197	**SMD 0.57 lower (0.76 lower to 0.37 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 6 weeks to 12 weeks)												
7	randomised trials [2,3,5-9]	serious [a]	not serious	serious [b]	not serious	none	N=205	N=176	**SMD 0.67 lower (0.88 lower to 0.46 lower)**		⊕⊕□□ LOW	CRITICAL
Quality of Life (Higher scores indicate better quality of life) (follow up: 12 weeks)												
3	randomised trials [3,5,9]	serious [a]	not serious	serious [b]	not serious	none	N=105	N=76	**SMD 0.55 higher (0.11 higher to 0.99 higher)**		⊕⊕□□ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Tai Chi) (follow up: range 12 weeks to 48 weeks)												
6	randomised trials [1-3,5,8,9]	serious [a]	not serious	serious [b]	serious [c]	none	5/163 (3.1%)	2/128 (1.6%)	RR 1.90 (0.43 to 8.37)	14 more per 1,000 (from 9 fewer to 115 more)	⊕□□□ VERY LOW	CRITICAL
Treatment-related Adverse Events (Risk ratios less than one favor Tai Chi) (follow up: range 12 weeks to 48 weeks)												
3	randomised trials [1,8,9]	serious [a]	not serious	serious [b]	serious [c]	none	1/56 (1.8%)	0/53 (0.0%)	RR 3.00 (0.13 to 69.52)	NA [d]	⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Tai Chi) (follow up: 48 weeks)												
1	randomised trial [9]	serious [a]	not assessable	serious [b]	very serious [c,e]	none	1/20 (5.0%)	1/20 (5.0%)	RR 1.00 (0.07 to 14.90)	0 fewer per 1,000 (from 47 fewer to 695 more)	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. The majority of studies received at least one High risk of bias rating due to single-blind design or inadequate blinding.
b. All patients in all trials have Knee Osteoarthritis.
c. 95% CI crosses null.
d. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.
e. Sample size in each study arm <50.

PICO 1.5.5 (hip): What are the benefits and harms of yoga in the management of patients with hip OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Cheung, et al. BMC Complement Altern Med. 2014 May 18; 14: 160; 2. Cheung, et al. Rheumatol Int. 2017 Mar; 37(3): 389-398.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Hatha Yoga**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (Scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 8 weeks)												
2	randomised trials [1,2]	serious [a]	not serious	serious [b]	serious [c]	none	N=50	N=41	**MD 3.49 lower (5.06 lower to 1.91 lower)**		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (Scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 8 weeks)												
2	randomised trials [1,2]	serious [a]	not serious	serious [b]	serious [c]	none	N=50	N=41	**MD 10.58 lower (15.24 lower to 5.93 lower)**		⊕□□□ VERY LOW	CRITICAL
SF-12 Physical Component Score (Scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 8 weeks)												
2	randomised trials [1,2]	serious [a]	very serious [d]	serious [b]	very serious [c,e]	none	N=50	N=41	MD 2.01 lower (10.82 lower to 6.8 higher)		⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Hatha Yoga) (follow up: 8 weeks)												
2	randomised trials [1,2]	serious [a]	not serious	serious [b]	very serious [c,e]	none	1/50 (2.0%)	0/41 (0.0%)	RR 2.18 (0.09 to 51.28)	0 fewer per 1,000 (from 0 fewer to 0 fewer)	⊕□□□ VERY LOW	CRITICAL
Treatment-related Adverse Events (Risk ratios less than one favor Hatha Yoga) (follow up: 8 weeks)												
2	randomised trials [1,2]	serious [a]	not serious	serious [b]	serious [c]	none	0/50 (0.0%)	0/41 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Hatha Yoga) (follow up: 8 weeks)												
1	randomised trial [2]	serious [a]	not assessable	serious [b]	serious [f]	none	0/32 (0.0%)	0/23 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Both trials received at least one High risk of bias rating due to single-blind study design.

b. All patients in both trials have Knee Osteoarthritis.
c. Total sample size ≤50 in each study arm.
d. I^2= 89%
e. 95% CI crosses null.
f. Total sample size <50 in each study arm.

PICO 1.6 (hip): What are the benefits and harms of aquatic exercise/hydrotherapy in the management of patients with hip OA?

SUMMARY
There is low-quality evidence that aquatic exercise lead to small statistically significant improvements in pain, physical function and quality of life in people with knee and/or hip OA. There is a low risk of harm with aquatic exercise. Benefits in pain reduction and function from aquatic exercise therapy in the treatment of hip and/or knee OA are smaller than the effects from land-based exercise therapy.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□ LOW

Bibliography: 1. Bartels, et al. Cochrane Database Syst Rev. 2016 Mar 23; 3: CD005523 (meta-analysis).

Quality assessment							№ of events/ № of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Aquatic Exercise	Control	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 6 weeks to 18 months)												
12	randomised trials [1]	not serious	not serious	serious [a]	not serious	none	N=539	N=537	**SMD 0.31 lower (0.47 lower to 0.15 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 6 weeks to 18 months)												
12	randomised trials [1]	not serious	not serious	serious [a]	not serious	none	N=529	N=530	**SMD 0.32 lower (0.47 lower to 0.17 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Quality of Life (Higher scores indicate better quality of life) (follow up: range 6 weeks to 18 months)												
10	randomised trials [1]	not serious	serious [b]	serious [a]	not serious	none	N=493	N=478	**SMD 0.25 lower (0.49 lower to 0.01 lower)**		⊕⊕□□ LOW	IMPORTANT
Total Adverse Events (Risk ratios less than one favor Aquatic Exercise) (follow up: range 6 weeks to 18 months)												
13	randomised trials [1]	not serious	not serious	serious [a]	serious [c]	none	112/605 (18.5%)	89/585 (15.2%)	RR 1.25 (0.98 to 1.60)	38 more per 1,000 (from 3 fewer to 91 more)	⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. Mixed population; includes Knee only studies, Hip only studies, and Hip/Knee studies.
b. I^2= 65%; moderate heterogeneity.
c. 95% CI crosses null.

PICO 1.7 MANUAL THERAPY
PICO 1.7.1 (hip): What are the benefits and harms of massage in the management of patients with hip OA?

SUMMARY (1.7.1 massage, 1.7.2 manipulation and mobilisation)
The evidence is from very low-quality or low-quality data. For some people with knee and/or hip OA, these therapies may have a positive effect on pain and/or function over a short term (low-quality to very low-quality evidence), and there is a very low risk of harm.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1.Atkins, et al. Int J Ther Massage Bodywork. 2013; 6(1): 4-14; 2. Perlman, et al. Arch Intern Med. 2006 Dec 11-25; 166(22): 2533-8; 3. Perlman, et al. PLoS One. 2012; 7(2): e30248; 4. Yip, YB., and Tam, AC. Complement Ther Med. 2008 Jun; 16(3): 131-8.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Massage †	Usual Care	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 4 weeks to 12 weeks)												
4	randomised trials [1,2,3,4]	very serious [a]	not serious	serious [b]	not serious	none	N=183	N=93	**SMD 0.7 lower (0.97 lower to 0.43 lower)**		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 4 weeks to 8 weeks)												
3	randomised trials [2,3,4]	serious [c]	not serious	serious [b]	not serious	none	N=165	N=75	**SMD 0.58 lower (0.87 lower to 0.29 lower)**		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Massage) (follow up: range 4 weeks to 24 weeks)												
3	randomised trials [2,3,4]	serious [c]	not serious	serious [b]	serious [d]	none	4/175 (2.3%)	1/77 (1.3%)	RR 1.17 (0.23 to 5.89)	2 more per 1,000 (from 10 fewer to 64 more)	⊕□□□ VERY LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Massage) (follow up: range 4 weeks to 24 weeks)												
3	randomised trials [2,3,4]	serious [c]	not serious	serious [b]	serious [d]	none	2/175 (1.1%)	0/77 (0.0%)	RR 2.01 (0.22 to 18.82)	NA [e]	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

† Interventions and comparators assessed included the following: Atkins 2013- Supervised self-massage (4 weeks), then unsupervised self-massage (4 weeks) vs. Wait-list Control; Perlman 2006- Swedish massage vs. Wait-list Control; Perlman 2012- Swedish massage (various doses) vs. Usual Care; Yip and Tam 2008- Aroma massage with orange and ginger vs. Olive Oil massage vs. Usual Care.

a. 3 of 4 trials which reported pain received at least one High risk of bias rating due to unblinded study design, potentially inadequate blinding, or potential for attrition bias or reporting bias.
b. All patients in all trials have Knee Osteoarthritis.
c. 2 of 3 trials which reported this outcome received at least one High risk of bias rating due to unblinded study design, potentially inadequate blinding, or potential attrition bias.

d. 95% CI crosses null.
e. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.

PICO 1.7.2 (hip): What are the benefits and harms of manipulation and mobilisation in the management of patients with hip OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Poulsen, et al. Osteoarthritis Cartilage. 2013 Oct; 21(10): 1494-503; 2. Abbott, et al. Osteoarthritis Cartilage. 2013 Apr; 21(4): 525-34.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Mobilisation & Manipulation**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
NRS Pain (Scale range 0 to 10, with higher scores indicating higher pain severity) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=34	N=36	**MD 2.20 lower (3.19 lower to 1.21 lower)**		⊕⊕□□ LOW	CRITICAL
NRS Pain (Scale range 0 to 10, with higher scores indicating higher pain severity) (follow up: 12 months)												
2	randomised trials [1,2]	serious [a]	not serious	serious [c]	serious [d]	none	N=80	N=80	MD 0.99 lower (2.1 lower to 0.11 higher)		⊕□□□ VERY LOW	CRITICAL
HOOS Function in Daily Living (Scale range 0 to 100, with higher scores indicating poorer functional outcome) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=34	N=36	**MD 14.00 lower (20.29 lower to 7.71 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: 12 months)												
2	randomised trials [1,2]	serious [a]	not serious	serious [c]	serious [d]	none	N=80	N=80	SMD 0.18 lower (0.49 lower to 0.13 higher)		⊕□□□ VERY LOW	CRITICAL
HOOS Hip-related Quality of Life (Scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=34	N=36	**MD 14.00 higher (6.96 higher to 21.04 higher)**		⊕⊕□□ LOW	IMPORTANT
HOOS Hip-related Quality of Life (Scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 12 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,d]	none	N=38	N=37	MD 0.00 (10.78 lower to 10.78 higher)		⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Manipulation & Mobilisation) (follow up: range 6 weeks to 12 months)												
2	randomised trials [1,2]	serious [a]	serious [e]	serious [c]	serious [d]	none	3/92 (3.3%)	3/88 (3.4%)	RR 1.02 (0.07 to 14.83)	1 more per 1,000 (from 32 fewer to 471 more)	⊕□□□ VERY LOW	CRITICAL

Treatment-related Adverse Events (Risk ratios less than one favor Manipulation & Mobilisation) (follow up: range 6 weeks to 12 months)											
2	randomised trials [1,2]	serious [a]	not serious	serious [c]	not serious	none	9/92 (9.8%)	0/88 (0.0%)	**RR 18.51 (1.12 to 307.04)**	**NA [f]**	⊕⊕□□ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Manipulation & Mobilisation) (follow up: 12 months)											
1	randomised trial [2]	serious [a]	not assessable	serious [c]	not serious	none	0/54 (0.0%)	0/51 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was inestimable.		⊕⊕□□ LOW	CRITICAL

Abbott, et al., 2013: % Knee: 55%, % Hip: 45%, %
Both: 24%

CI: Confidence interval; **MD:** Mean difference; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. Both studies received a High risk of bias rating due to single-blind design, in which patients were not blinded.
b. Sample size <50 in each study arm.
c. In Abbott 2013, approximately 55% of the population has Knee Osteoarthritis, and 45% has Hip Osteoarthritis; 24% of the population has both types.
d. 95% CI crosses null.
e. I^2= 52%; moderate heterogeneity.
f. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.

PICO 1.8 (hip): What are the benefits and harms of weight management in the management of patients with hip OA?

SUMMARY

There is limited evidence of very low quality that weight loss alone (achieved through diet and exercise) has no significant effect on either pain or function in people with knee OA although benefits appear to be more significant with higher amounts of weight loss – starting at a minimum of 5–7.5% body weight loss. Ideally, it is beneficial to achieve a greater amount of weight loss.(Atukorala I, et al. Arthritis Care Res 2016;68(8):1106–14; Messier SP, et al. JAMA, 2013;310(12):1263–73). There is no randomised controlled trial (RCT) on the effects of bariatric surgery in people with hip and/or knee OA. There are no RCTs investigating weight management specifically in people with hip OA, although the numerous other systemic health benefits of weight loss and maintaining a healthy body weight are most likely transferable to people with hip OA. There are limitations of the available RCT evidence in OA. Clinicians are advised to refer to the National Health and Medical Research Council's (NHMRC's) guidelines for the most effective strategies for managing overweight/obesity in primary care. National Health and Medical Research Council. Clinical practice guidelines for the management of overweight and obesity in adults, adolescents and children in Australia. Melbourne: NHMRC, 2013).

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Bliddal, et al. Ann Rheum Dis. 2011 Oct; 70(10): 1798-803; 2. Messier, et al. Arthritis Rheum. 2004 May; 50(5): 1501-10 (ADAPT); 3. Miller, et al. Obesity (Silver Spring). 2006 Jul; 14(7): 1219-30; 4. Focht, et al. Arthritis Rheum. 2005 Oct 15; 53(5): 659-65 (ADAPT).

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Weight Management**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 6 months to 18 months)												
3	randomised trials [1,2,3]	serious [a]	very serious [b]	serious [c]	serious [d]	none	N=165	N=158	SMD 0.38 lower (0.88 lower to 0.11 higher)		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 6 months to 18 months)												
3	randomised trials [1,2,3]	serious [a]	serious [e]	serious [c]	serious [d]	none	N=165	N=158	SMD 0.29 lower (0.62 lower to 0.04 higher)		⊕□□□ VERY LOW	CRITICAL
Percentage Weight Loss (Percentage of weight at baseline lost by follow up time, with more loss indicating positive outcome) (follow up: range 6 months to 18 months)												
3	randomised trials [1,2,3]	serious [f]	not serious	serious [c]	not serious	none	N=165	N=156	**MD 6.48 % more lost (8.48 % lower to 4.48 % lower)**		⊕⊕□□ LOW	IMPORTANT
Lateral Joint Space Narrowing [mm] (Higher scores indicate poorer structural outcome) (follow up: 18 months)												
1	randomised trials [2]	serious [f]	not assessable	serious [c]	serious [d]	none	N=82	N=78	MD 0.09 mm higher (0.41 mm lower to 0.59 mm higher)		⊕□□□ VERY LOW	IMPORTANT
Medial Joint Space Narrowing [mm] (Higher scores indicate poorer structural outcome) (follow up: 18 months)												

1	randomised trials [2]	serious [f]	not assessable	serious [c]	serious [d]	none	N=82	N=78	MD 0.05 mm higher (0.45 mm lower to 0.55 mm higher)		⨁◯◯◯ VERY LOW	IMPORTANT
Walking Self-Efficacy (Patient confidence in walking around a gymnasium twice without stopping; Scale range 0-100 with higher scores indicating more confidence) (follow up: 18 months)												
1	randomised trials [4]	serious [a]	not assessable	serious [c]	serious [d]	none	N=82	N=78	MD 0.1 higher (9.72 lower to 9.92 higher)		⨁◯◯◯ VERY LOW	IMPORTANT
Withdrawal due to Adverse Events (Risk ratios less than one favor Weight Management) (follow up: 12 months)												
1	randomised trials [1]	serious [a]	not assessable	serious [c]	very serious [d,h]	none	2/44 (4.5%)	2/45 (4.4%)	RR 1.02 (0.15 to 6.94)	1 more per 1,000 (from 38 fewer to 264 more)	⨁◯◯◯ VERY LOW	CRITICAL
Non-Compliance with Regimen (Outcome includes, but does not necessitate, withdrawal from study. Defined as "non-compliance", "lack of motivation", "non-adherence") (Risk ratios less than one favor Weight Management) (follow up: range 12 months to 18 months)												
2	randomised trials [1,4]	serious [a]	serious [g]	serious [c]	serious [d]	none	32/126 (25.4%)	38/123 (30.9%)	RR 0.77 (0.37 to 1.63)	71 fewer per 1,000 (from 195 fewer to 195 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Serious Adverse Events (Risk ratios less than one favor Weight Management) (follow up: 6 months)												
1	randomised trials [3]	serious [f]	not assessable	serious [c]	serious [h]	none	0/44 (0.0%)	0/43 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⨁◯◯◯ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. All studies received at least one High risk of bias rating due to single-blind or unblinded study design.
b. I^2= 78%
c. All patients in all trials have Knee Osteoarthritis.
d. 95% CI crosses null.
e. I^2= 52%; moderate heterogeneity.
f. All studies received at least one High risk of bias rating due to single-blind or unblinded study design, but this outcome may not be susceptible to bias due to objective reporting.
g. I^2= 67%; moderate heterogeneity.
h. Sample size <50 in each study arm.

PICO 1.9: THERMOTHERAPY
PICO 1.9.1 (hip): What are the benefits and harms of local hot application in the management of patients with hip OA?

SUMMARY
Heat therapy may be effective in reducing pain for some people with knee and/or hip OA, but the quality of evidence is very low . Heat therapy is cheap and generally feasible for people to undertake independently as a self-management strategy.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Denegar, et al. Clin Interv Aging. 2010 Aug 9; 5: 199-206; 2. Mazzuca, et al. Arthritis Rheum. 2004 Oct 15; 51(5): 716-21; 3. Yildirim, et al. J Clin Nurs. 2010 Apr; 19(7-8): 1113-20.

Quality assessment							№ of events/ № of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Heat Therapy†**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 1 week to 4 weeks)												
3	randomised trials [1,2,3]	very serious [a]	not serious	serious [b]	not serious	none	N=83	N=82	**SMD 0.38 lower (0.69 lower to 0.07 lower)**		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 1 week to 4 weeks)												
3	randomised trials [1,2,3]	very serious [a]	not serious	serious [b]	serious [c]	none	N=83	N=82	SMD 0.27 lower (0.68 lower to 0.15 higher)		⊕□□□ VERY LOW	CRITICAL
KOOS Quality of Life (Scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 1 week)												
1	randomised trial [2]	very serious [a]	not assessable	serious [b]	very serious [c,d]	none	N=34	N=34	MD 5.7 higher (0.63 lower to 12.03 higher)		⊕□□□ VERY LOW	IMPORTANT
Total Adverse Events (Risk ratios less than one favor Heat Therapy) (follow up: range 1 week to 4 weeks)												
2	randomised trials [1,2]	very serious [a]	not serious	serious [b]	not serious	none	0/60 (0.0%)	0/59 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

†Interventions and comparators assessed included the following: Denegar 2010- Heat Pad vs. No treatment ("Rest"); Mazzuca 2004- Heat-retaining knee sleeve vs. Placebo knee sleeve; Yildirim 2010- Heat Pad vs. Usual Care.

a. All studies received at least one High risk of bias rating.
b. All patients in all studies have Knee Osteoarthritis.
c. 95% CI crosses null.
d. Sample size <50 in each study arm.

PICO 1.9.2 (hip): What are the benefits and harms of local cold application in the management of patients with hip OA?

SUMMARY
There is very low-quality evidence suggesting that the use of cold therapy is not effective in improving pain, function or quality of life in people with knee and/or hip OA.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Clark, et al. Rheumatol Rehabil. 1974 Nov; 13(4): 190-7; 2. Denegar, et al. Clin Interv Aging. 2010 Aug 9; 5: 199-206.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cold Therapy†	Control	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 1 week to 3 weeks)												
2	randomised trials [1,2]	very serious [a]	not serious	serious [b]	very serious [c,d]	none	N=49	N=47	SMD 0.5 lower (1.07 lower to 0.07 higher)		⊕□□□ VERY LOW	CRITICAL
KOOS Function in Daily Living (scale range 0 to 100, with higher scores indicating poorer functional outcome) (follow up: 1 week)												
1	randomised trial [2]	very serious [a]	not assessable	serious [b]	very serious [c,d]	none	N=34	N=34	MD 6.20 lower (13.71 lower to 1.31 higher)		⊕□□□ VERY LOW	CRITICAL
KOOS Quality of Life (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 1 week)												
1	randomised trial [2]	very serious [a]	not assessable	serious [b]	very serious [c,d]	none	N=34	N=34	MD 4.7 higher (1.86 lower to 11.26 higher)		⊕□□□ VERY LOW	IMPORTANT
Total Adverse Events (Risk ratios less than one favor Cold Therapy) (follow up: 1 week)												
1	randomised trial [2]	very serious [a]	not assessable	serious [b]	serious [d]	none	0/34 (0.0%)	0/34 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

†Interventions and comparators assessed included the following: Clark 1974- Ice application vs. "Untuned" short-wave diathermy (study received quality downgrade for inappropriate comparison); Denegar 2010- Cold treatment facilitated by circulation of water through a wrap-around garment vs. No treatment ("Rest").

a. Both studies received more than two High risk of bias ratings due to faults in study design and reporting.
b. All patients in both trials have Knee Osteoarthritis.
c. 95% CI crosses null.
d. Sample size <50 in each study arm.

PICO 1.10: HIP ORTHOTICS

PICO 1.10.1a (hip): What are the benefits and harms of foot orthotics (minimalist footwear) in the management of patients with hip OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Trombini-Souza, et al. Clin Biomech (Bristol, Avon). 2015 Dec; 30(10): 1194-201.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Minimalist Footwear**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 6 months)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	very serious [c,d]	none	N=28	N=28	MD 3.25 lower (6.78 lower to 0.28 higher)		⊕□□□ VERY LOW	CRITICAL
Lequesne Index (scale range 0 to 24, with higher scores indicating poorer functional outcome) (follow up: 6 months)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [d]	none	N=28	N=28	**MD 2.4 lower (4.38 lower to 0.42 lower)**		⊕□□□ VERY LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Minimalist Footwear) (follow up: 6 months)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	very serious [c,d]	none	1/28 (3.6%)	1/28 (3.6%)	RR 1.00 (0.07 to 15.21)	0 fewer per 1,000 (from 33 fewer to 507 more)	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Study received a High risk of bias rating due to a single blind design.
b. All patients in this study have Knee Osteoarthritis.
c. 95% CI crosses null.
d. Total sample size in each study arm is <50.

PICO 1.10.1b (hip): What are the benefits and harms of foot orthotics (unloading shoes) in the management of patients with hip OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Erhart, et al. J Orthop Res. 2010 Jul; 28(7): 873-9; 2. Hinman, et al. Ann Intern Med. 2016 Sep 20; 165(6): 381-9.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Unloading Shoes	Control	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: 6 months)												
2	randomised trials [1,2]	serious [a]	not serious	serious [b]	serious [c]	none	N=120	N=119	SMD 0.11 lower (0.37 lower to 0.14 higher)		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 6 months)												
1	randomised trial [2]	not serious	not assessable	serious [b]	serious [c]	none	N=80	N=80	MD 0.5 lower (4.34 lower to 3.34 higher)		⊕⊕□□ LOW	CRITICAL
Assessment of Quality of Life 6D scale (scale range -0.04 to 1.00, with higher scores indicating higher quality of life) (follow up: 6 months)												
1	randomised trial [2]	not serious	not assessable	serious [b]	serious [c]	none	N=79	N=79	MD 0 (0.03 lower to 0.03 higher)		⊕⊕□□ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Unloading Shoes) (follow up: 6 months)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [d]	none	3/40 (7.5%)	11/39 (28.2%)	**RR 0.27 (0.08 to 0.88)**	**206 fewer per 1,000 (from 34 fewer to 259 fewer)**	⊕□□□ VERY LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Unloading Shoes) (follow up: 6 months)												
1	randomised trial [2]	not serious	not assessable	serious [b]	serious [c]	none	26/83 (31.3%)	20/81 (24.7%)	RR 1.27 (0.77 to 2.08)	67 more per 1,000 (from 57 fewer to 267 more)	⊕⊕□□ LOW	CRITICAL

Withdrawal due to Potentially Treatment-Related Reasons ["Too small shoe size", "Shoe discomfort", "Meniscectomy", "Hip pain", "Foot pain", "Total Knee Replacement"] (Risk ratios less than one favor Unloading Shoes) (follow up 6 months)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	very serious [c,d]	none	4/40 (10.0%)	10/39 (25.6%)	RR 0.39 (0.13 to 1.14)	156 fewer per 1,000 (from 36 more to 223 fewer)	⨁◻◻◻ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. Erhart 2010 was given a High risk of bias rating due to single blind study design and potential attrition bias.
b. All participants in both trials have Knee Osteoarthritis.
c. 95% CI crosses null.
d. Sample size <50 in each study arm.

PICO 1.10.1c (hip): What are the benefits and harms of foot orthotics (unstable shoes) in the management of patients with hip OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Nigg, et al. Med Sci Sports Exerc. 2006 Oct; 38(10): 1701-8.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Masai Barefoot Technology Footwear	Control	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (scale range 0-500, with higher scores indicating higher pain severity) (follow up: 12 weeks)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [c]	none	N=53	N=66	MD 4.2 higher (27.8 lower to 36.2 higher)		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (scale range 0-1700, with higher scores indicating poorer functional outcome) (follow up: 12 weeks)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [c]	none	N=53	N=66	MD 18.7 higher (94.31 lower to 131.71 higher)		⊕□□□ VERY LOW	CRITICAL
Withdrawal due to Treatment-Related Adverse Events (Risk ratios less than one favor Masai Barefoot Technology footwear) (follow up: 12 weeks)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [c]	none	1/58 (1.7%)	1/67 (1.5%)	RR 1.16 (0.07 to 18.06)	2 more per 1,000 (from 14 fewer to 255 more)	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Patients are not blinded to allocation.
b. All participants in the included trial have Knee Osteoarthritis.
c. 95% CI crosses null.

PICO 1.11 (hip): What are the benefits and harms of kinesio taping in the management of patients with hip OA?

SUMMARY
A limited number of trials of very low quality and with small sample sizes have evaluated kinesio taping for knee OA. These trials used different configurations of kinesio tape that was reapplied by the clinician after various intervals. There is no trial of kinesio taping for hip OA. The evidence shows no significant benefits of kinesio taping for pain or function.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Anandkumar, et al. Physiother Theory Pract. 2014 Aug; 30(6): 375-83; 2. Cho, et al. Am J Phys Med Rehabil. 2015 Mar; 94(3): 192-200; 3. Kocyigit, et al. Complement Ther Clin Pract. 2015 Nov; 21(4): 262-7.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Kinesio Taping	Sham Kinesio Taping	Relative (95% CI)	Absolute (95% CI)		
VAS Pain (Scale range 0 to 10, with higher scores indicating higher pain severity) (follow up: range 0 days [two pre- and post-test trials] to 12 days)												
3	randomised trials [1,2,3]	serious [a]	very serious [b]	serious [c]	very serious [d,e]	none	N=64	N=63	MD 1.28 lower (2.93 lower to 0.37 higher)		⊕□□□ VERY LOW	CRITICAL
Lequesne Index (Scale range 0 to 24, with higher scores indicating poorer functional outcome) (follow up: 12 days)												
1	randomised trial [3]	serious [a]	not assessable	serious [c]	serious [f]	none	N=21	N=20	**MD 2.9 higher (0.09 higher to 5.71 higher)**		⊕□□□ VERY LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Kinesio Taping) (follow up: 12 days)												
1	randomised trial [3]	serious [a]	not assessable	serious [c]	very serious [d,f]	none	0/22 (0.0%)	1/21 (4.8%)	RR 0.32 (0.01 to 7.42)	32 fewer per 1,000 (from 47 fewer to 306 more)	⊕□□□ VERY LOW	CRITICAL
Skin irritation (Risk ratios less than one favor Kinesio Taping) (follow up: 12 days)												
1	randomised trial [3]	serious [a]	not assessable	serious [c]	very serious [d,f]	none	0/22 (0.0%)	1/21 (4.8%)	RR 0.32 (0.01 to 7.42)	32 fewer per 1,000 (from 47 fewer to 306 more)	⊕□□□ VERY LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio
a. All studies received either High or Unclear risk of bias ratings due to single-blind study design or potentially inadequate blinding measures.
b. I^2= 89%
c. All patients in all trials had Knee Osteoarthritis.
d. 95% CI crosses null.
e. Insufficient follow up time to provide an adequate estimate.
f. Sample size <50 in each study arm.

PICO 1.12 (hip): What are the benefits and harms of walking cane/stick in the management of patients with hip OA?

SUMMARY
People with knee and/or hip OA often adopt an abnormal gait pattern because of pain, muscle weakness, joint mobility restrictions or other pain conditions. The use of an assistive walking device may be useful to improve gait pattern and posture to normalise musculoskeletal loads. There is low-quality evidence from one trial that the use of a walking aid (eg single point stick) is effective in improving pain and function in people with knee OA. These data could be reasonably transferred to people with hip OA (very low-quality evidence).

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Jones, et al. Ann Rheum Dis. 2012 Feb; 71(2): 172-9.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Cane**	**No Cane**	**Relative (95% CI)**	**Absolute (95% CI)**		
VAS Pain (scale range 0 to 10 cm, with higher scores indicating higher pain severity) (follow up: 60 days)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [c]	none	N=32	N=32	**MD 2.26 cm lower (2.9 lower to 1.62 lower)**		⊕□□□ VERY LOW	CRITICAL
Lequesne Index (scale range 0 to 24, with higher scores indicate poorer functional outcome) (follow up: 60 days)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [c]	none	N=32	N=32	**MD 3.64 cm lower (5.35 lower to 1.93 lower)**		⊕□□□ VERY LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Walking cane/stick) (follow up: 60 days)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious [c,d]	none	1/32 (3.1%)	2/32 (6.3%)	RR 0.50 (0.05 to 5.24)	31 fewer per 1,000 (from 59 fewer to 265 more)	⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Study received a High risk of bias rating due to single-blind study design; investigators are not blinded, and patient blinding may have been inadequate.
b. All patients in this trial have Knee Osteoarthritis.
c. Sample size <50 in each study arm.
d. 95% CI crosses null.

PICO 1.13: ELECTROMAGNETIC THERAPY

PICO 1.13.1 (hip): What are the benefits and harms of pulsed electromagnetic/shortwave therapy in the management of patients with hip OA?

SUMMARY
There is very low-quality evidence that pulsed electromagnetic therapy has no statistically significant effect on pain or function in people with hip OA.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Moffett, et al. Pain. 1996 Sep; 67(1):121-7.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Electromagnetic Therapy**	**Sham Treatment**	**Relative (95% CI)**	**Absolute (95% CI)**		
NRS Pain (Scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 12 weeks)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	very serious [c,d]	none	N=26	N=22	MD 0.49 higher (12.26 lower to 13.24 higher)		⊕□□□ VERY LOW	CRITICAL
Moffett 1996: % knee: 50%, % hip: 50%												

CI: Confidence interval; **MD:** Mean difference

a. Study received a High risk of bias rating due to inadequate reporting of patient discontinuation through the full follow up period.
b. 50% of randomised patients have Knee Osteoarthritis.
c. 95% CI crosses null.
d. Total sample size <50 in each study arm.

PICO 1.13.2 (hip): What are the benefits and harms of shockwave therapy in the management of patients with hip OA?

SUMMARY
There is no direct evidence for the effects of 1.13.2 shockwave, 1.14.2 interferential, 1.16 laser therapy on people with hip OA but modes of physiologic actions of the knee interventions are transferable to hip.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

No RCT data was found for this question.

PICO 1.14: ELECTRICAL STIMULATION

PICO 1.14.1 (hip): What are the benefits and harms of transcutaneous electrical nerve stimulation (TENS) in the management of patients with hip OA?

SUMMARY

While no direct evidence is available from trials in people with hip OA, the working group felt that the mode of action with TENS could be transferable to the hip. Trials were limited to four weeks follow-up, so it remains uncertain whether treatment effects are maintained beyond this period.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

No RCT data were found for this question.

PICO 1.14.2 (hip): What are the benefits and harms of interferential currents in the management of patients with hip OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

No RCT data was found for this question.

PICO 1.15 (hip): What are the benefits and harms of ultrasound in the management of patients with hip OA?

SUMMARY

There is moderate-quality evidence that therapeutic ultrasound has statistically significant effects on pain and physical function in people with knee OA. There is no randomised controlled trial (RCT) involving participants with hip OA, thus the evidence level for this population group was downgraded to low quality because of concerns about indirectness. Although the evidence suggests moderate effect sizes and a low risk of harms for therapeutic ultrasound in people with knee OA, the working group noted that current evidence is restricted to short-term (two to eight weeks) follow-up only. There is concern about whether benefits are sustained once treatment finished. The available evidence suggests that three to five treatment sessions per week are required for benefits.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□ LOW

Bibliography: 1. Külcü, et al., 2009. "Short-term efficacy of pulsed electromagnetic field therapy on pain and functional level in knee osteoarthritis: a randomised controlled study"; 2. Loyola-Sánchez, et al. Arch Phys Med Rehabil. 2012 Jan; 93(1): 35-42; 3. Tascioglu, et al. J Int Med Res. 2010 Jul-Aug; 38(4): 1233-42; 4. Özgönenel, et al. Ultrasound Med Biol. 2009 Jan; 35(1): 44-9.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Ultrasound**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 2 weeks to 8 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	not serious	serious [b]	not serious	none	N=116	N=86	**SMD 0.55 lower (0.88 lower to 0.22 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up range 2 weeks to 8 weeks)												
3	randomised trials [1,2,4]	serious [a]	not serious	serious [b]	not serious	none	N=61	N=59	**SMD 0.57 lower (1.03 lower to 0.10 lower)**		⊕⊕□□ LOW	CRITICAL
Change in Central Medial Femoral Cartilage Volume [mm³] (Higher values indicate better structural outcome) (follow up: 8 weeks)												
1	randomised trial [2]	not serious	not assessable	serious [b]	very serious [c,d]	none	N=11	N=12	MD 16.70 mm³ lower (136.32 mm³ lower to 102.92 mm³ higher)		⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Therapeutic Ultrasound) (follow up: range 2 weeks to 8 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	not serious	serious [b]	not serious	none	0/118 (0.0%)	0/88 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Therapeutic Ultrasound) (follow up: range 2 weeks to 8 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	not serious	serious [b]	not serious	none	0/118 (0.0%)	0/88 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. 3 of 4 trials received at least one High risk of bias rating due to unblinded design or due to potential attrition bias.
b. All patients in all trials have Knee Osteoarthritis.
c. 95% CI crosses null.
d. Total sample size <30.

PICO 1.16 (hip): What are the benefits and harms of laser in the management of patients with hip OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1.Alfredo, et al. Clin Rehabil. 2012 Jun; 26(6): 523-33; 2. Fukuda, et al. Rev Bras Ortop. 2015 Dec 6; 46(5): 526-33; 3. Gworys, et al. Ortop Traumatol Rehabil. 2012 May-Jun; 14(3): 269-77; 4. Hegedüs, et al. Photomed Laser Surg. 2009 Aug; 27(4): 577-84; 5. Hsieh, et al. Arch Phys Med Rehabil. 2012 May; 93(5): 757-64; 6. Stelian, et al. J Am Geriatr Soc. 1992 Jan; 40(1): 23-6; 7. Tascioglu, et al. Swiss Med Wkly. 2004 May 1; 134(17-18): 254-8; 8. Bülow, et al. Scand J Rehabil Med. 1994 Sep; 26(3): 155-9.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Laser Therapy**	**Sham Laser Therapy**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 2 weeks to 3 weeks)												
7	randomised trials [1,2,3,4,5,6,7]	serious [a]	serious [b]	serious [c]	not serious	none	N=252	N=154	**SMD 0.49 lower (0.82 lower to 0.17 lower)**		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 2 weeks to 3 weeks)												
6	randomised trials [1,2,3,5,6,7]	serious [d]	serious [e]	serious [c]	not serious	none	N=234	N=145	**SMD 0.67 lower (1.04 lower to 0.31 lower)**		⊕□□□ VERY LOW	CRITICAL
Patients Reporting Improvement [N values are patients who reported "Treatment did help" over (N "Treatment did help" + "Treatment did not help")] (Risk ratios greater than one favor Laser Therapy) (follow up: 9 weeks)												
1	randomised trial [8]	serious [f]	not assessable	serious [c]	very serious [g,h]	none	4/14 (28.6%)	6/15 (40.0%)	RR 0.71 (0.25 to 2.01)	116 fewer per 1,000 (from 300 fewer to 404 more)	⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Laser Therapy) (follow up: range 3 weeks to 6 months)												
4	randomised trials [2,5,7,8]	serious [i]	not serious	serious [c]	serious [g]	none	0/116 (0.0%)	1/92 (1.1%)	RR 0.32 (0.01 to 7.50)	7 fewer per 1,000 (from 11 fewer to 71 more)	⊕□□□ VERY LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Laser Therapy) (follow up: range 4 weeks to 6 months)												
3	randomised trials [5,7,8]	serious [j]	not serious	serious [c]	not serious	none	0/71 (0.0%)	0/70 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. 5 of 7 studies reporting pain received at least one High risk of bias rating.
b. I^2= 56% moderate heterogeneity.
c. All patients in all trials have Knee Osteoarthritis.
d. 4 of 6 studies reporting function received at least one High risk of bias rating.
e. I^2= 62% moderate heterogeneity.
f. Study received High risk of bias rating due to potential selective reporting.
g. 95% CI crosses null.
h. Total sample size <30.
i. 3 of 4 studies received at least one High risk of bias rating.
j. All studies received at least one High risk of bias rating.

PICO 1.17 (hip): What are the benefits and harms of acupuncture in the management of patients with hip OA?

SUMMARY
There is low-quality evidence that traditional, laser and electro acupuncture have statistically significant benefits on pain and function, compared with sham acupuncture in people with knee OA However, these benefits are small and not of a clinically relevant magnitude. There is very low-quality evidence suggesting no statistically significant effect of laser acupuncture on either pain or function in people with knee OA. There is very low-quality evidence that traditional acupuncture has no statistically significant effect onpain or function, compared with sham in people with hip OA.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

TRADITIONAL ACUPUNCTURE:

Bibliography: 1. Fink, et al. Complement Ther Med. 2001 Jun; 9(2): 82-9.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Acupuncture**	**Sham Acupuncture**	**Relative (95% CI)**	**Absolute (95% CI)**		
VAS Pain [short term] (scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	N=32	N=30	MD 4.14 lower (16.1 lower to 7.83 higher)		⊕□□□ VERY LOW	CRITICAL
VAS Pain [long term] (scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 6 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	N=17	N=24	MD 8 lower (24.4 lower to 8.4 higher)		⊕□□□ VERY LOW	CRITICAL
Lequesne Index [short term] (scale range 0 to 24, with higher scores indicating poorer functional outcome) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	N=32	N=30	MD 0.45 lower (2.06 lower to 1.16 higher)		⊕□□□ VERY LOW	CRITICAL
Lequesne Index [long term] (scale range 0 to 24, with higher scores indicating poorer functional outcome) (follow up: 6 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	N=17	N=24	MD 0.57 lower (2.62 lower to 1.49 higher)		⊕□□□ VERY LOW	CRITICAL
Bullinger's "Everyday Life" questionnaire (Higher scores indicate better quality of life) (follow up: 6 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	N=32	N=30	**MD 9.54 higher (2.5 higher to 16.58 higher)**		⊕⊕□□ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Acupuncture) (follow up: 6 months)												

1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	0/33 (0.0%)	0/34 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.	⊕⊕◯◯ LOW	CRITICAL
Treatment-related Adverse Events (Risk ratios less than one favor Acupuncture) (follow up: 6 months)											
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	0/33 (0.0%)	0/34 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.	⊕⊕◯◯ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Study received High risk of bias ratings for potential attrition bias and potential reporting bias.
b. 95% CI crosses null.
c. Sample size <50 in each study arm.

GRADE tables for hip osteoarthritis

Section 2: Pharmacologic Interventions

PICO 2.1: ORAL ANALGESICS
PICO 2.1.1 (knee): What are the benefits and harms of paracetamol in the management of patients with knee OA?

SUMMARY
While paracetamol has long been considered first-line therapy for OA, this has mainly reflected its relative safety, availability and cost, compared with other pharmacological options (eg NSAIDs, opioids). Current evidence from a systematic review of randomised controlled trials (RCTs) suggests that, on average, the reduction in OA pain achieved with paracetamol is too small to be of clinical relevance.(Machado, et al. BMJ. 2015 Mar 31; 350: h1225). Moreover, paracetamol is associated with infrequent potential for significant harms, both short-term excess dosing and long-term regular use.(Roberts E, Delgado Nunes V, Buckner S, et al. Ann Rheum Dis 2016;75(3):552–59).

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Machado, et al. BMJ. 2015 Mar 31; 350: h1225.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Paracetamol	Placebo	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: range 2 weeks to 3 months)												
7	randomised trials [1]	not serious	not serious	serious [a]	not serious	none	N=1279	N=1076	**MD 3.7 lower (5.5 lower to 1.9 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
WOMAC Physical Function (scale range from 0 to 100, with higher score indicating poorer functional outcome) (follow up: range 2 weeks to 3 months)												
7	randomised trials [1]	not serious	not serious	serious [a]	not serious	none	N=1279	N=1076	**MD 2.9 lower (4.9 lower to 0.9 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Withdrawals due to Adverse Events (follow up: range 2 weeks to 3 months)												
7	randomised trials [1]	not serious	not serious	serious [a]	serious [b]	none	120/1630 (7.4%)	90/1393 (6.5%)	RR 1.2 (0.9 to 1.5)	13 more per 1,000 (from 6 fewer to 32 more)	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (follow up: range 2 weeks to 3 months)												
9	randomised trials [1]	not serious	serious [c]	serious [a]	serious [b]	none	735/2729 (26.9%)	613/2117 (29.0%)	RR 1.0 (0.9 to 1.1)	0 fewer per 1,000 (from 29 fewer to 29 more)	⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (follow up: range 2 weeks to 3 months)												

7	randomised trials [1]	not serious	not serious	serious [a]	serious [b]	none	37/2825 (1.3%)	22/2027 (1.1%)	RR 1.2 (0.7 to 2.1)	2 more per 1,000 (from 3 fewer to 12 more)	⊕⊕◻◻ LOW	CRITICAL
Abnormal Liver Function (AST/ALN>1.5 ULN) (follow up: range 2 weeks to 3 months)												
3	randomised trials [1]	not serious	not serious	serious [a]	not serious	none	46/693 (6.6%)	10/544 (1.8%)	**RR 3.8 (1.9 to 7.4)**	**51 more per 1,000 (from 17 more to 118 more)**	⊕⊕⊕◻ MODERATE	IMPORTANT

Pincus, 2004a; % knee: 84%, % hip: 16%
Pincus, 2004b; % knee: 84%, % hip: 16%
Altman, 2007; % knee: 81%, % hip: 19%
Prior, 2014; % knee: 82%, % hip: 18%
Zoppi, 1995; mixed population, no data

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Mixed population; all patients have osteoarthritis of the hip and/or knee
b. 95% CI crosses null.
c. I^2= 68%; moderate heterogeneity

PICO 2.1.2 (hip): What are the benefits and harms of oral NSAIDs including COX-2 inhibitors in the management of patients with hip OA?

SUMMARY
On average, the use of NSAIDs result in small but clinically relevant improvements in pain and function in individuals with knee and/or hip OA, and are likely to be more effective than paracetamol for most people. The direct costs of NSAIDs are relatively low. Evidence for effectiveness is derived from trials of relatively short duration, so the relative benefits versus harms of long-term NSAID therapy are unknown. It is likely that the risk of harms increases with duration of therapy; therefore, the balance of benefits and harms may become less favourable with time. Given the variable natural history of OA symptoms, periodic trials of drug withdrawal are recommended.

OVERALL QUALITY OF EVIDENCE: ⊕⊕⊕◻MODERATE

Bibliography: 1. Baerwald, et al. Arthritis Rheum. 2010 Dec; 62(12): 3635-44; 2. Makarowski, et al. Osteoarthritis Cartilage. 2002 Apr; 10(4): 290-6; 3. Schnitzer, et al. Clin Rheumatol. 2011 Nov; 30(11): 1433-46.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Oral NSAID**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (Higher scores indicate higher pain severity) (follow up: 12 weeks)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	not serious	none	N=1231	N=695	**SMD 0.32 lower (0.42 lower to 0.21 lower)**		⊕⊕⊕⊕ HIGH	CRITICAL
WOMAC Function (Higher scores indicate poorer functional outcome) (follow up: 12 weeks)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	not serious	none	N=1672	N=864	**SMD 0.37 lower (0.50 lower to 0.24 lower)**		⊕⊕⊕⊕ HIGH	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Oral NSAID) (follow up: 12 weeks)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	not serious	none	87/1231 (7.1%)	43/865 (5.0%)	**RR 1.53 (1.05 to 2.22)**	**26 more per 1,000 (from 2 more to 61 more)**	⊕⊕⊕⊕ HIGH	CRITICAL
Total Adverse Events (Risk ratios less than one favor Oral NSAID) (follow up: 12 weeks)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	not serious	none	664/1231 (53.9%)	384/863 (44.5%)	**RR 1.15 (1.05 to 1.26)**	**67 more per 1,000 (from 22 more to 116 more)**	⊕⊕⊕⊕ HIGH	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Oral NSAID) (follow up: 12 weeks)												

2	randomised trials [1,3]	not serious	not serious	not serious	serious [a]	none	16/1002 (1.6%)	14/746 (1.9%)	RR 0.92 (0.38 to 2.25)	2 fewer per 1,000 (from 12 fewer to 23 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Oral NSAID) (follow up: 12 weeks)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	serious [a]	none	119/1231 (9.7%)	91/863 (10.5%)	RR 1.24 (0.74 to 2.07)	25 more per 1,000 (from 27 fewer to 113 more)	⊕⊕⊕◻ MODERATE	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. 95% CI crosses null.

PICO 2.1.3 (hip): What are the benefits and harms of oral opioids in the management of patients with hip OA?

SUMMARY
Opioid prescription for chronic non-cancer pain (including OA) has increased in recent years despite a lack of high-quality evidence demonstrating benefit, particularly with long-term use. Evidence for the efficacy of opioids in the treatment of OA is mostly derived from short-term trials. There is moderate-quality evidence from trials that improvement in pain and function with opioids is of marginal clinical significance at best, and is offset by the risk of harms.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Friedmann, et al. J Opioid Manag. 2011 May-Jun; 7(3): 193-202; 2. Katz, et al. Postgrad Med. 2010 Jul; 122(4): 112-28 ; 3. NCT00980798; 4. Rauck, et al. Pain Pract. 2013 Jan; 13(1): 18-29; 5. Zautra, et al. Clin J Pain. 2005 Nov-Dec; 21(6): 471-7.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Oral opioid†**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (scale range 0 to 10, with higher scores indicating higher pain severity) (follow up: 12 weeks)												
5	randomised trials [1,2,3,4,5]	serious [a]	serious [b]	serious [c]	not serious	none	N=1209	N=903	**MD 0.43 lower (0.73 lower to 0.14 lower)**		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (Scale range 0 to 100, with higher scores indicating poorer functional outcome) (follow up: 12 weeks)												
2	randomised trials [2,4]	serious [d]	not serious	serious [c]	serious [e]	none	N=820	N=504	MD 0.39 lower (0.80 lower to 0.02 higher)		⊕□□□ VERY LOW	CRITICAL
% Patients Experiencing Opioid Withdrawal (As evidenced by a score between 13-36 on the Clinical Opiate Withdrawal Scale [COWS; score >36 indicative of severe withdrawal) (Risk ratios less than one favor Oral Opioid) (follow up: 12 weeks)												
1	randomised trial [2]	not serious	not assessable	not serious	serious [e]	none	5/171 (2.9%)	1/173 (0.6%)	RR 5.06 (0.60 to 42.85)	23 more per 1,000 (from 2 fewer to 242 more)	⊕⊕⊕□ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Oral Opioid) (follow up: 12 weeks)												

5	randomised trials [1,2,3,4,5]	serious [a]	very serious [f]	not serious	not serious	none	345/1219 (28.3%)	91/910 (10.0%)	**RR 2.63 (1.47 to 4.72)**	**163 more per 1,000 (from 47 more to 372 more)**	⊕□□□ VERY LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Oral Opioid) (follow up: 12 weeks)												
3	randomised trials [2,3,4]	serious [a]	very serious [g]	not serious	not serious	none	745/959 (77.7%)	346/654 (52.9%)	**RR 1.47 (1.05 to 2.06)**	**249 more per 1,000 (from 26 more to 561 more)**	⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Oral Opioid) (follow up: 12 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	not serious	not serious	serious [e]	none	36/1164 (3.1%)	17/861 (2.0%)	RR 1.57 (0.83 to 2.97)	11 more per 1,000 (from 3 fewer to 39 more)	⊕⊕□□ LOW	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Oral Opioid) (follow up: 12 weeks)												
4	randomised trials [1,2,3,4]	serious [a]	very serious [h]	not serious	not serious	none	886/1164 (76.1%)	201/861 (23.3%)	**RR 2.87 (1.80 to 4.57)**	**437 more per 1,000 (from 187 more to 833 more)**	⊕□□□ VERY LOW	IMPORTANT

Friedmann, 2011: % knee: 78.5%, % hip: 21.5%
Katz, 2010: % knee: 77.5%, % hip: 22.5%
NCT00980798: mixed population, no data
Rauck, 2013: % knee: 76.5%, % hip: 23.5%
Zautra, 2005: mixed population, no data

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

† Different oral opioids on various dosing regimens were analyzed together. The following were included in the analysis: Friedmann, 2011- Oxycodone 5 mg- 40 mg BID [up or down titration allowed for four weeks, then dose was fixed for 8 weeks]; Katz, 2010- Morphine Sulfate/Naltrexone started at an average daily dose of 25.3 ± 9.7 mg (range, 20–120 mg) and ended at an average daily dose of 43.5 ± 31.7 mg (range, 20–160 mg); NCT00980798- OROS Hydromorphone HCl 4-32 mg once/day; Rauck, 2013- OROS Hydromporphone ER 8 mg/day or 16 mg/day, FIXED DOSE, Zautra, 2005- Oxycodone CR 10 mg/12 hours, upward dose adjustments to a final dose of 12 tablets/day (120 mg CR oxycodone) were allowed to control pain; downward adjustments were available to control side effects.

a. The majority of studies received at least one High risk of bias rating due to potential attrition bias, or potential reporting bias.
b. I^2= 58%; moderate heterogeneity.
c. Mixed populations; the majority of patients in all trials have Knee Osteoarthritis.
d. 1 of 2 trials received a High risk of bias rating due to potential attrition bias.
e. 95% CI crosses null.

f. I^2= 91%
g. I^2= 82%
h. I^2= 92%

PICO 2.2: TOPICAL ANALGESICS
PICO 2.2.1 (hip): What are the benefits and harms of topical NSAIDs in the management of patients with hip OA?

SUMMARY
The effectiveness of topical NSAID application in OA is variable. Generally, the benefit is small, but the risk of harm is also small.

OVERALL QUALITY OF EVIDENCE: ⊕⊕⊕◻MODERATE

Bibliography: 1. Baraf, et al. Phys Sportsmed. 2010 Jun; 38(2): 19-28; 2. Barthel, et al. Semin Arthritis Rheum. 2009 Dec; 39(3): 203-12; 3. Conaghan, et al. Rheumatology (Oxford). 2013 Jul; 52(7): 1303-12; 4. Kneer, et al. J Pain Res. 2013 Oct 25; 6: 743-53; 5. Roth and Shainhouse. Arch Intern Med. 2004 Oct 11; 164(18): 2017-23; 6. Rother, et al. J Rheumatol. 2013 Oct; 40(10): 1742-8; 7. Simon, et al. Pain. 2009 Jun; 143(3): 238-45.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Topical NSAID	Vehicle Control	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (Higher scores indicate higher pain severity) (follow up: 12 weeks)												
7	randomised trials 1,2,3,4,5,6,7	not serious	not serious	serious [a]	not serious	none	N=2016	N=1559	**SMD 0.2 lower (0.29 lower to 0.11 lower)**		⊕⊕⊕◻ MODERATE	CRITICAL
WOMAC Function (Higher scores indicate poorer functional outcome) (follow up: 12 weeks)												
7	randomised trials 1,2,3,4,5,6,7	not serious	not serious	serious [a]	not serious	none	N=2015	N=1559	**SMD 0.19 lower (0.28 lower to 0.1 lower)**		⊕⊕⊕◻ MODERATE	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Topical NSAID) (follow up: 12 weeks)												
7	randomised trials 1,2,3,4,5,6,7	not serious	not serious	not serious	not serious	none	130/2048 (6.3%)	71/1571 (4.5%)	**RR 1.36 (1.02 to 1.82)**	**16 more per 1,000 (from 1 more to 37 more)**	⊕⊕⊕⊕ HIGH	CRITICAL
Total Adverse Events (Risk ratios less than one favor Topical NSAID) (follow up: 12 weeks)												
6	randomised trials 1,2,3,4,6,7	not serious	not serious	not serious	serious [b]	none	963/1884 (51.1%)	685/1409 (48.6%)	RR 1.07 (0.98 to 1.16)	34 more per 1,000 (from 10 fewer to 78 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Topical NSAID) (follow up: 12 weeks)												

5	randomised trials [1,2,3,5,7]	not serious	not serious	not serious	serious [b]	none	12/1243 (1.0%)	10/1245 (0.8%)	RR 1.08 (0.26 to 4.47)	1 more per 1,000 (from 6 fewer to 28 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Topical NSAID) (follow up: 12 weeks)												
7	randomised trials [1,2,3,4,5,6,7]	not serious	not serious	not serious	serious [b]	none	96/2048 (4.7%)	84/1571 (5.3%)	RR 0.94 (0.70 to 1.25)	3 fewer per 1,000 (from 13 more to 16 fewer)	⊕⊕⊕◻ MODERATE	IMPORTANT
Local Reactions (local reactions included dermatitis, skin dryness, eczema, exanthema, erythema, papules, pruritus, itching, dermatosis, allergic reaction, parasthesia (Roth 2004 only), and/or rash near the application site of topical NSAID or Vehicle control) (Risk ratios less than one favor Topical NSAID) (follow up: 12 weeks)												
7	randomised trials [1,2,3,4,5,6,7]	not serious	not serious	not serious	not serious	none	354/2048 (17.3%)	193/1571 (12.3%)	**RR 1.32 (1.04 to 1.67)**	**39 more per 1,000 (from 5 more to 82 more)**	⊕⊕⊕⊕ HIGH	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. All patients in all trials have Knee Osteoarthritis.
b. 95% CI crosses null

PICO 2.2.2 (hip): What are the benefits and harms of transdermal opioids in the management of patients with hip OA?

SUMMARY
Evidence is mostly derived from two short-term trials of transdermal opioids: buprenorphine and fentanyl. Similar to oral opioid, the low-quality evidence demonstrated the improvements in pain and function are of marginal clinical significance at best,and is offset by the risk of harms

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

TRANSDERMAL BUPRENORPHINE

Bibliography: 1. da Costa, Bruno R., et al. "Oral or transdermal opioids for osteoarthritis of the knee or hip." The Cochrane Library (2014).

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Transdermal Buprenorphine**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 4 weeks to 30 weeks)												
4	randomised trials [1]	serious [a]	not serious	serious [b]	not serious	none	N=691	N=710	**SMD 0.19 lower (0.3 lower to 0.09 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 4 weeks to 28 weeks)												
2	randomised trials [1]	serious [a]	not serious	serious [b]	not serious	none	N=243	N=258	**SMD 0.23 lower (0.4 lower to 0.05 lower)**		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Transdermal Buprenorphine) (follow up: range 4 weeks to 30 weeks)												
4	randomised trials [1]	serious [a]	serious [c]	not serious	not serious	none	119/698 (17.0%)	51/709 (7.2%)	**RR 3.10 (1.38 to 6.94)**	**151 more per 1,000 (from 27 more to 427 more)**	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Transdermal Buprenorphine) (follow up: 28 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	not serious	none	92/100 (92.0%)	73/99 (73.7%)	**RR 1.25 (1.09 to 1.42)**	**184 more per 1,000 (from 66 more to 310 more)**	⊕⊕⊕□ MODERATE	CRITICAL

Breivik, 2010; % knee: 63%, % hip: 37% Munera, 2010; mixed population, no data Shannon, 2005; mixed population, no data

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. All studies received at least one High risk of bias rating due to inadequate masking of interventions or non-ITT analysis
b. All studies involved both Knee and Hip OA patients; NCT00531427 was only in Knee OA patients
c. I^2= 74% moderate heterogeneity

TRANSDERMAL FENTANYL

Bibliography: 1. da Costa, Bruno R., et al. "Oral or transdermal opioids for osteoarthritis of the knee or hip." The Cochrane Library (2014).

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Transdermal Fentanyl**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	not serious	none	N=202	N=197	**SMD 0.22 lower (0.42 lower to 0.03 lower)**		⊕⊕□□ LOW	CRITICAL
WOMAC Function (Higher scores indicate poorer functional outcome) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	not serious	none	N=202	N=197	**SMD 0.28 lower (0.48 lower to 0.09 lower)**		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Transdermal Fentanyl) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	not serious	none	54/202 (26.7%)	20/197 (10.2%)	**RR 2.63 (1.64 to 4.23)**	**165 more per 1,000 (from 65 more to 328 more)**	⊕⊕⊕□ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Transdermal Fentanyl) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	not serious	none	169/216 (78.2%)	101/200 (50.5%)	**RR 1.55 (1.33 to 1.81)**	**278 more per 1,000 (from 167 more to 409 more)**	⊕⊕⊕□ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Transdermal Fentanyl) (follow up: 8 weeks)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	6/216 (2.8%)	2/200 (1.0%)	RR 2.78 (0.57 to 13.60)	18 more per 1,000 (from 4 fewer to 126 more)	⊕⊕□□ LOW	CRITICAL

Langford, 2006; % knee: 53%, % hip: 47%

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio
a. Study received High risk of bias ratings for inadequate masking of interventions and non-ITT analyses
b. Study involves both knee and hip OA patients
c. Wide 95% CI; crosses null

PICO 2.2.3 (hip): What are the benefits and harms of topical capsaicin in the management of patients with hip OA?

SUMMARY
Evidence from one trial demonstrated that 0.025% of topical capsaicin had small effects of pain relief in people with knee OA. It is uncertain whether individuals with multi-joint OA or with relevant comorbidities will benefit from capsaicin. The principle benefit of capsaicin is in neuropathic pain, which is not the major pain source in knee or hip OA. Similar to other topical analgesia, the topical application process is very regime orientated, and local irritation side effects can be detrimental. These issues often outweigh possible benefits to individuals.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: Kosuwon, et al. J Med Assoc Thai. 2010 Oct; 93(10): 1188-95.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Topical capsaicin**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
VAS Pain (Mean change from baseline, scale range 0 to 10, with higher scores indicating higher pain severity) (follow up: 4 weeks)												
1	randomised trial [1]	not serious	not assessable	serious [a]	very serious [b,c]	none	N=65	N=34	MD 0.27 higher (0.33 lower to 0.87 higher)		⊕□□□ VERY LOW	CRITICAL
Total WOMAC (Higher scores indicate poorer outcomes) (follow-up: 4 weeks)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [c]	none	N=65	N=34	**MD 6.75 lower (12.95 lower to 0.55 lower)**		⊕⊕□□ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Topical capsaicin) (follow up: 8 weeks, a.k.a two treatment periods in cross-over)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [c]	none	0/99 (0.0%)	0/99 (0.0%)	Due to zero events in both groups, an absolute risk reduction was not estimable.		⊕⊕⊕□ MODERATE	CRITICAL
Burning sensation (N episodes reported per group/Total N episodes reported) (Risk ratios less than one favor Topical capsaicin) (follow up: 8 weeks, a.k.a two treatment periods in cross-over)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious	none	272/338 (80.5%)	66/338 (19.5%)	**RR 4.12 (3.30 to 5.15)**	**609 more per 1,000 (from 449 more to 810 more)**	⊕⊕□□ LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. All patients in this trial have Knee Osteoarthritis.
b. 95% CI crosses null
c. Sample size <50 in one study arm
d. Authors chose to report skin irritation as the proportion of N total burning sensation events in each group.

PICO 2.3: HERBAL THERAPIES

PICO 2.3.1 (hip): What are the benefits and harms of avocado soybean unsaponifiable in the management of patients with hip OA?

SUMMARY

The 2014 Cochrane review reports ASU 300 mg produced a small and clinically questionable improvement in symptoms, and probably no increased adverse events, compared with placebo after three to 12 months treatment. (Cameron M, Chrubasik S. Cochrane Database Syst Rev, 2014;22(5):CD002947). In the new evidence review for this guideline, short-term pain and function up to six months was improved by about 0.5 standard deviations, and there were no significant longer-term benefits in pain or function. Studies examining the use of supplements are often of low quality constrained by small sample sizes, industry publication bias and potential for positive publication bias. In the context of low-quality to very low-quality studies, despite some suggestion of beneficial effects, it is prudent to use caution and advocate that further research is needed before a firm recommendation on ASU can be made.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Appelboom, et al. Scand J Rheumatol. 2001; 30(4): 242-7; 2. Blotman, et al. Rev Rhum Engl Ed. 1997 Dec; 64(12): 825-34; 3. Maheu, et al. Arthritis Rheum. 1998 Jan; 41(1): 81-91; 4. Lequesne, et al. Arthritis Rheum. 2002 Feb; 47(1): 50-8; 5. Maheu, et al. Ann Rheum Dis. 2014 Feb; 73(2): 376-84; 6. Liu, et al. Osteoarthritis and Cartilage. 2017 Apr 1; 25: S292-3.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Avocado Soybean Unsaponifiable	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: 90 days)												
2	randomised trials [1,2]	serious [a]	serious [b]	serious [c]	not serious	none	N=243	N=161	**SMD 0.57 lower (0.95 lower to 0.19 lower)**		⊕□□□ VERY LOW	CRITICAL
Pain (Higher scores indicate higher pain severity) (follow up: 6 months)												
1	randomised trial [3]	not serious	not assessable	serious [c]	not serious	none	N=84	N=78	**SMD 0.45 lower (0.77 lower to 0.14 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Pain (Higher scores indicate higher pain severity) (follow up: range 2 years to 3 years)												
2	randomised trials [4,5]	not serious	not serious	not serious	serious [d]	none	N=251	N=257	SMD 0.04 higher (0.14 lower to 0.21 higher)		⊕⊕⊕□ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: 90 days)												

2	randomised trials [1,2]	serious [a]	not serious	serious [c]	not serious	none	N=234	N=161	**SMD 0.48 lower (0.69 lower to 0.28 lower)**	⊕⊕◯◯ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: 6 months)											
1	randomised trial [3]	not serious	not assessable	serious [c]	not serious	none	N=84	N=78	**SMD 0.58 lower (0.94 lower to 0.23 lower)**	⊕⊕⊕◯ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 2 years to 3 years)											
2	randomised trials [4,5]	not serious	not serious	not serious	serious [d]	none	N=251	N=257	SMD 0.03 lower (0.21 lower to 0.14 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Avocado Soybean Unsaponifiable) (follow up: range 90 days to 3 years)											
5	randomised trials [1,2,3,4,5]	not serious	not serious	serious [c]	not serious	none	291/610 (47.7%)	270/537 (50.3%)	**RR 1.0 (1.0 to 1.1)**	**0 fewer per 1,000 (from 0 fewer to 50 more)**	⊕⊕⊕◯ MODERATE	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. One study received at least one High risk of bias rating.
b. I^2= 69%; moderate heterogeneity.
c. Mixed population; included trials involve patients with Knee and/or Hip Osteoarthritis.
d. 95% CI crosses null.

PICO 2.3.2 (hip): What are the benefits and harms of boswellia serrata in the management of patients with hip OA?

SUMMARY
Three small RCTs found significant short-term benefits in pain and function; however, these are all sponsored by the same company, raising concern about possible bias. In the context of low-quality to very low-quality studies, despite potential beneficial effects, it is prudent to use caution and advocate that further research is needed before a firm recommendation on *Boswellia serrata* can be made.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Sengupta, et al. Arthritis Res Ther. 2008; 10(4): R85; 2. Sengupta, et al. Int J Med Sci 2010; 7(6): 366-377; 3. Vishal, et al. Int J Med Sci. 2011; 8(7): 615-22.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Boswellia serrata extract**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 30 days to 90 days)												
3	randomised trials [1,2,3]	serious [a]	not serious	serious [b]	not serious	none	N=115	N=71	**SMD 1.61 lower (2.1 lower to 1.13 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 30 days to 90 days)												
3	randomised trials [1,2,3]	serious [a]	serious [c]	serious [b]	not serious	none	N=115	N=71	**SMD 1.15 lower (1.63 lower to 0.68 lower)**		⊕□□□ VERY LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Boswellia serrata extract) (follow up: range 30 days to 90 days)												
2	randomised trials [2,3]	serious [a]	not serious	serious [b]	very serious [d,e]	none	2/68 (2.9%)	2/49 (4.1%)	RR 0.70 (0.10 to 4.80)	12 fewer per 1,000 (from 37 fewer to 155 more)	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. One study reporting this outcome received a High risk of bias rating; another study received primarily "Unclear" risk of bias ratings.
b. All patients in all trials have Knee Osteoarthritis.
c. I^2= 51%; moderate heterogeneity.
d. 95% CI crosses null.
e. Sample size in one study arm <50.

PICO 2.3.3 (hip): What are the benefits and harms of curcuma in the management of patients with hip OA?

SUMMARY
Three small RCTs found significant short-term (ie six to eight weeks) benefits in pain and function; however, these are all industry-sponsored trials, raising concern about possible bias. Additionally, there were inconsistency in the results. All of the studies involved knee OA, so extrapolation to hip or other OA requires additional caution. In the context of low-quality to very low-quality studies, despite potential beneficial effects, it is prudent to use caution and advocate that further research is needed before a firm recommendation on curcuma can be made.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Madhu, et al. Inflammopharmacology. 2013 Apr; 21(2): 129-36; 2. Nakagawa, et al. J Orthop Sci. 2014 Nov; 19(6): 933-9; 3. Panahi, et al. Phytother Res. 2014 Nov; 28(11): 1625-31.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Curcuminoid**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 6 weeks to 8 weeks)												
3	randomised trials [1,2,3]	serious [a]	serious [b]	serious [c]	not serious	none	N=63	N=70	**SMD 1.1 lower (1.66 lower to 0.54 lower)**		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: 6 weeks)												
1	randomised trial [3]	serious [a]	not assessable	serious [c]	serious [d]	none	N=19	N=21	**SMD 0.81 lower (1.46 lower to 0.16 lower)**		⊕□□□ VERY LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Curcuminoid) (follow up: range 6 weeks to 8 weeks)												
3	randomised trials [1,2,3]	serious [a]	not serious	serious [c]	serious [e]	none	4/75 (5.3%)	3/79 (3.8%)	RR 1.46 (0.34 to 6.31)	17 more per 1,000 (from 25 fewer to 202 more)	⊕□□□ VERY LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Curcuminoid) (follow up: range 6 weeks to 8 weeks)												
2	randomised trials [1,3]	serious [a]	not serious	serious [c]	serious [e]	none	9/57 (15.8%)	6/56 (10.7%)	RR 1.48 (0.57 to 3.83)	51 more per 1,000 (from 46 fewer to 303 more)	⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Curcuminoid) (follow up: range 6 weeks to 8 weeks)												
2	randomised trials [2,3]	serious [a]	not serious	serious [c]	not serious	none	0/52 (0.0%)	0/51 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕□□ LOW	CRITICAL

Gastrointestinal Adverse Events (Risk ratios less than one favor Curcuminoid) (follow up: range 6 weeks to 8 weeks)												
2	randomised trials [1,3]	serious [a]	not serious	serious [c]	serious [e]	none	9/57 (15.8%)	4/56 (7.1%)	RR 1.92 (0.68 to 5.41)	66 more per 1,000 (from 23 fewer to 315 more)	⨁◯◯◯ VERY LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. All trials received at least one High risk of bias rating due to single blind study design, potential attrition bias, or potential reporting bias.
b. I^2= 56%; moderate heterogeneity.
c. All patients in all trials have Knee Osteoarthritis.
d. Sample size <50 in each study arm.
e. 95% CI crosses null.

PICO 2.3.4 (hip): What are the benefits and harms of pine bark extract in the management of patients with hip OA?

SUMMARY
Three small RCTs found short-term benefits in pain and function; however, these could not be pooled because of heterogeneity and reporting weaknesses. All three trials were industry-sponsored, with the larger trial at very high risk of bias. Evidence is based on studies of knee OA, so extrapolation to hip or other OA requires additional caution. In the context of low-quality to very low-quality studies, despite potential beneficial effects, it is prudent to use caution and advocate that further research is needed before a firm recommendation on the use of pine bark extract.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1.Belcaro, et al. Phytotherapy Research. 2008 Apr 1; 22(4): 518-23; 2. Farid, et al. Nutrition Research. 2007 Nov 30; 27(11): 692-7; 3. Cisár, et al. Phytother Res. 2008 Aug; 22(8): 1087-92.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Pycnogenol**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (Scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 3 months)												
1	randomised trial [1]	very serious [a]	not assessable [b]	serious [c]	not serious	none	N=71	N=74	**MD 7.7 lower (8.24 lower to 7.16 lower)**		⊕□□□ VERY LOW	CRITICAL
WOMAC Pain VAS (Scale range 0 to 500 mm, with higher scores indicating higher pain severity) (follow up: 3 months)												
1	randomised trial [2]	not serious	not assessable [b]	serious [c]	serious [d]	none	N=19	N=18	**MD 133 mm lower (198.66 mm lower to 67.34 mm lower)**		⊕⊕□□ LOW	CRITICAL
WOMAC Function (Scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 3 months)												
1	randomised trial [1]	very serious [a]	not assessable [b]	serious [c]	not serious	none	N=71	N=74	**MD 26 lower (26.49 lower to 25.51 lower)**		⊕□□□ VERY LOW	CRITICAL
WOMAC Function VAS (Scale range 0 to 1700 mm, with higher scores indicating poorer functional outcome) (follow up: 3 months)												
1	randomised trial [2]	not serious	not assessable [b]	serious [c]	serious [d]	none	N=19	N=18	**MD 484 mm lower (718.42 mm lower to 249.58 mm lower)**		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Pycnogenol) (follow up: 3 months)												
2	randomised trials [2,3]	serious [e]	not serious	serious [c]	serious [f]	none	2/69 (2.9%)	4/68 (5.9%)	RR 0.50 (0.10 to 2.61)	29 fewer per 1,000 (from 53 fewer to 95 more)	⊕□□□ VERY LOW	CRITICAL

Total Adverse Events (Risk ratios less than one favor Pycnogenol) (follow up: 3 months)											
2	randomised trials [2,3]	serious [e]	not serious	serious [c]	serious [f]	none	2/69 (2.9%)	4/68 (5.9%)	RR 0.50 (0.10 to 2.61)	29 fewer per 1,000 (from 53 fewer to 95 more)	⨁□□□ VERY LOW	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Pycnogenol) (follow up: 3 months)											
2	randomised trials [2,3]	serious [e]	not serious	serious [c]	serious [f]	none	1/69 (1.4%)	1/68 (1.5%)	RR 1.00 (0.06 to 15.55)	0 fewer per 1,000 (from 14 fewer to 214 more)	⨁□□□ VERY LOW	IMPORTANT
Serious Adverse Events (Risk ratios less than one favor Pycnogenol) (follow up: 3 months)											
1	randomised trial [3]	serious [e]	not assessable	serious [c]	serious [d]	none	0/50 (0.0%)	0/50 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⨁□□□ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Study received High risk of bias rating due to potential reporting bias; inadequate reporting of critical efficacy outcomes and no Adverse Event reporting, despite description of collection in study methods.
b. A standardised mean difference could not be provided due to excessive heterogeneity in efficacy reporting between two eligible studies.
c. All patients in all trials have Knee Osteoarthritis.
d. Sample size≤50 in each study arm.
e. One study received a High risk of bias rating, as well as more than one Unclear risk of bias rating, in important categories.
f. 95% CI crosses null.

PICO 2.4: NUTRACEUTICALS
PICO 2.4.1 (hip): What are the benefits and harms of glucosamine in the management of patients with hip OA?

SUMMARY
There is only one RCT on the effect on hip OA, which failed to demonstrate a benefit. There are concerns in the literature of publication bias, effects being driven by small industry-sponsored trials, and the overall poor quality of the positive trials. Larger publicly funded trials generally demonstrate no effect over placebo. High-quality trial data suggest no effect.

OVERALL QUALITY OF EVIDENCE: ⊕⊕⊕□ MODERATE

Bibliography: 1.Rozendaal, et al. Ann Intern Med. 2008 Feb 19; 148(4): 268-77.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Glucosamine	Placebo	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain [short term] (scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 3 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	N=111	N=111	MD 0.71 lower (5.38 lower to 3.96 higher)		⊕⊕⊕□ MODERATE	CRITICAL
WOMAC Pain [moderate term] (scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	N=111	N=111	MD 0.35 higher (5.35 lower to 6.05 higher)		⊕⊕⊕□ MODERATE	CRITICAL
WOMAC Pain [long term] (scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 24 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	N=111	N=111	MD 2.35 lower (8.59 lower to 3.89 higher)		⊕⊕⊕□ MODERATE	CRITICAL
WOMAC Function [short term] (scale range 0 to 100, with higher scores indicating poorer functional outcome) (follow up: 3 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	N=111	N=111	MD 2.21 lower (5.85 lower to 1.43 higher)		⊕⊕⊕□ MODERATE	CRITICAL
WOMAC Function [moderate term] (scale range 0 to 100, with higher scores indicating poorer functional outcome) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	N=111	N=111	MD 0.1 lower (4.39 lower to 4.19 higher)		⊕⊕⊕□ MODERATE	CRITICAL
WOMAC Function [long term] (scale range 0 to 100, with higher scores indicating poorer functional outcome) (follow up: 24 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	N=111	N=111	MD 2.76 lower (7.86 lower to 2.34 higher)		⊕⊕⊕□ MODERATE	CRITICAL

Lateral Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: 24 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	N=111	N=111	MD 0.02 lower (0.11 lower to 0.07 higher)		⊕⊕⊕◻ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Glucosamine) (follow up: 24 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	4/111 (3.6%)	4/111 (3.6%)	RR 1.00 (0.26 to 3.90)	0 fewer per 1,000 (from 27 fewer to 105 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Glucosamine) (follow up: 24 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	57/111 (51.4%)	59/111 (53.2%)	RR 0.97 (0.75 to 1.24)	16 fewer per 1,000 (from 128 more to 133 fewer)	⊕⊕⊕◻ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Glucosamine) (follow up: 24 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	4/111 (3.6%)	2/111 (1.8%)	RR 2.00 (0.37 to 10.70)	18 more per 1,000 (from 11 fewer to 175 more)	⊕⊕⊕◻ MODERATE	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Glucosamine) (follow up: 24 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [a]	none	58/111 (52.3%)	46/111 (41.4%)	RR 1.26 (0.95 to 1.67)	108 more per 1,000 (from 21 fewer to 278 more)	⊕⊕⊕◻ MODERATE	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. 95% CI crosses null.

PICO 2.4.2 (hip): What are the benefits and harms of chondroitin in the management of patients with hip OA?

SUMMARY
There are a large number of trials on the use of chondroitin where at least seven are industry sponsored. When all 16 studies are pooled, there is a clinically and statistically significant effect on pain and function in the short-term (up to three months), which lessens to clinically not significant by six to 12 months, and no effect is demonstrated at 24 months. However, when the analysis is restricted to studies of higher quality or free of industry sponsorship, no benefit is demonstrated. There are some moderate-term to long-term (12–24 months) benefits on joint space narrowing, but these are not clinically meaningful. The studies are all on participants with knee OA, so extrapolation to OA of hip or other joints requires further caution.
There are concerns in the literature of publication bias, effects being driven by small industry-sponsored trials, and the overall poor quality of the positive trials. High-quality trial data suggest no effect.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1.Bourgeois, et al. Osteoarthritis Cartilage. 1998 May; 6 Suppl A: 25-30; 2. Mazieres, et al. J Rheumatol. 2001 Jan; 28(1): 173-81; 3. Pavelká, et al. Litera Rheumatologica. 1999; 24: 21-30; 4. Zegels, et al. Osteoarthritis Cartilage. 2013 Jan; 21(1): 22-7; 5. Reginster, et al. Ann Rheum Dis. 2017 May 22. pii: annrheumdis-2016-210860; 6. Bucsi, L. and Poór, G. Osteoarthritis Cartilage. 1998 May; 6 Suppl A: 31-6 ; 7. Clegg, et al. N Engl J Med. 2006 Feb 23; 354(8): 795-808 (GAIT); 8. Fransen, et al. Ann Rheum Dis. 2015 May; 74(5): 851-8; 9. Kahan, et al. Arthritis Rheum. 2009 Feb; 60(2): 524-33; 10. Railhac, et al. Clin Rheumatol. 2012 Sep; 31(9): 1347-57; 11. Uebelhart, et al. Osteoarthritis Cartilage. 1998 May; 6 Suppl A: 39-46; 12. Uebelhart, et al. Osteoarthritis Cartilage. 2004 Apr; 12(4): 269-76 ; 13. Wildi, et al. Ann Rheum Dis. 2011 Jun; 70(6): 982-9; 14. Michel, et al. Arthritis Rheum. 2005 Mar; 52(3): 779-86; 15. Sawitzke, et al. Annals of the rheumatic diseases. 2010 Aug 1; 69(8): 1459-64 (GAIT); 16. Sawitzke, et al. Arthritis & Rheumatology. 2008 Oct 1; 58(10): 3183-91.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Chondroitin**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain [short term] (Higher scores indicate higher pain severity) (follow up: 3 months)												
5	randomised trials [1,2,3,4,5]	serious [a]	serious [b]	serious [c]	not serious	none	N=666	N=451	**SMD 0.63 lower (0.91 lower to 0.36 lower)**		⊕□□□ VERY LOW	CRITICAL
Pain [moderate term] (Higher scores indicate higher pain severity) (follow up: range 6 months to 12 months)												
9	randomised trials [5,6,7,8,9,10,11,12,13]	not serious	very serious [d]	serious [c]	not serious	none	N=1109	N=1127	**SMD 0.28 lower (0.49 lower to 0.06 lower)**		⊕□□□ VERY LOW	CRITICAL
Pain [long term] (Higher scores indicate higher pain severity) (follow up: 24 months)												
4	randomised trials [8,9,14,15]	not serious	not serious	serious [c]	serious [e]	none	N=736	N=745	SMD 0.03 lower (0.13 lower to 0.07 higher)		⊕⊕□□ LOW	CRITICAL
Function [short term] (Higher scores indicate poorer functional outcome) (follow up: 3 months)												
5	randomised trials [1,2,3,4,5]	serious [a]	not serious	serious [c]	not serious	none	N=666	N=451	**SMD 0.55 lower (0.78 lower to 0.33 lower)**		⊕⊕□□ LOW	CRITICAL

Function [moderate term] (Higher scores indicate poorer functional outcome) (follow up: range 6 months to 12 months)											
6	randomised trials [5,6,7,8,10,11]	not serious	very serious [f]	serious [c]	not serious	none	N=711	N=724	**SMD 0.33 lower (0.62 lower to 0.04 lower)**	⊕□□□ VERY LOW	CRITICAL
Function [long term] (Higher scores indicate poorer functional outcome) (follow up: 24 months)											
3	randomised trials [8,14,15]	not serious	not serious	serious [c]	serious [e]	none	N=427	N=432	SMD 0.04 lower (0.17 lower to 0.1 higher)	⊕⊕□□ LOW	CRITICAL
Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: range 12 months to 2 years)											
6	randomised trials [8,9,11,12,14,16]	not serious	serious [g]	serious [c]	not serious	none	N=635	N=646	**MD 0.16 mm higher (0.03 mm higher to 0.28 mm higher)**	⊕⊕□□ LOW	IMPORTANT
SF-12 Physical Component Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 2 years)											
1	randomised trials [8]	not serious	not assessable	serious [c]	serious [e]	none	N=151	N=151	MD 1 higher (1.12 lower to 3.12 higher)	⊕⊕□□ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Chondroitin) (follow up: range 3 months to 24 months)											
13	randomised trials [1,2,4,5,6,7,8,9,10,11,12,13]	not serious	not serious	serious [c]	serious [e]	none	85/1691 (5.0%)	67/1547 (4.3%)	RR 1.16 (0.85 to 1.59)	7 more per 1,000 (from 6 fewer to 26 more)	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Chondroitin) (follow up: range 3 months to 24 months)											
5	randomised trials [1,2,3,8,10]	not serious	not serious	serious [c]	serious [e]	none	74/322 (23.0%)	57/286 (19.9%)	RR 1.21 (0.90 to 1.61)	42 more per 1,000 (from 20 fewer to 122 more)	⊕⊕□□ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Chondroitin) (follow up: range 3 months to 24 months)											
6	randomised trials [2,4,11,12,13,15]	serious [h]	not serious	serious [c]	serious [e]	none	10/543 (1.8%)	6/434 (1.4%)	RR 1.32 (0.45 to 3.87)	4 more per 1,000 (from 8 fewer to 40 more)	⊕□□□ VERY LOW	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Chondroitin) (follow up: range 3 months to 24 months)											

10	randomised trials [1,3,5,6,8,9,10,12,13,14]	not serious	not serious	serious [c]	not serious	none	62/1156 (5.4%)	79/1062 (7.4%)	RR 0.72 (0.52 to 0.99)	21 fewer per 1,000 (from 1 fewer to 36 fewer)	⊕⊕⊕◻ MODERATE	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. Two of four studies received High risk of bias ratings. One study did not provide sufficient information to adequately assess bias.
b. I^2= 65%; moderate heterogeneity.
c. All patients in all trials have Knee Osteoarthritis.
d. I^2= 77%
e. 95% CI crosses null.
f. I^2= 86%
g. I^2= 68%; moderate heterogeneity.
h. Five of six studies received at least one High risk of bias rating.

PICO 2.4.3 (hip): What are the benefits and harms of glucosamine and chondroitin in compound form in the management of patients with hip OA?

SUMMARY
With pooling (where possible) of results from the nine available RCTs, no benefit for pain, function or joint space narrowing was demonstrated. Participants in all trials had knee OA, so extrapolation to hip OA needs additional caution.There are concerns in the literature of publication bias, effects being driven by small industry-sponsored trials, and the overall poor quality of the positive trials. High-quality trial data suggest no effect.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Clegg, et al. N Engl J Med. 2006 Feb 23; 354(8): 795-808 (GAIT); 2. Fransen, et al. Ann Rheum Dis. 2015 May; 74(5): 851-8; 3. Kanzaki, et al. J Sci Food Agric. 2012 Mar 15; 92(4): 862-9; 4. Lugo, et al. Nutr J. 2016 Jan 29; 15: 14; 5. Messier, et al. Osteoarthritis Cartilage. 2007 Nov; 15(11): 1256-66; 6. Roman-Blas, et al. Arthritis Rheumatol. 2017 Jan; 69(1): 77-85; 7. Tsuji, et al. Aging Clin Exp Res. 2016 Apr; 28(2): 197-205; 8. Sawitzke, et al. Annals of the rheumatic diseases. 2010 Aug 1; 69(8): 1459-64 (GAIT); 9. Sawitzke, et al. Arthritis & Rheumatology. 2008 Oct 1; 58(10): 3183-91.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Chondroitin + Glucosamine	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 16 weeks to 12 months)												
7	randomised trials [1,2,3,4,5,6,7]	not serious	very serious [a]	serious [b]	serious [c]	none	N=694	N=682	SMD 0.28 higher (0.4 lower to 0.97 higher)		⊕□□□ VERY LOW	CRITICAL
Pain [long term] (Higher scores indicate higher pain severity) (follow up: 24 months)												
2	randomised trials [2,8]	serious [d]	not serious	serious [b]	serious [c]	none	N=280	N=282	SMD 0.02 higher (0.15 lower to 0.19 higher)		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 6 months to 12 months)												
6	randomised trials [1,2,4,5,6,7]	not serious	very serious [e]	serious [b]	serious [c]	none	N=675	N=664	SMD 0.26 higher (0.43 lower to 0.96 higher)		⊕□□□ VERY LOW	CRITICAL
Function [long term] (Higher scores indicate poorer functional outcome) (follow up: 24 months)												
2	randomised trials [2,8]	serious [d]	not serious	serious [b]	serious [c]	none	N=280	N=282	SMD 0.03 higher (0.19 lower to 0.24 higher)		⊕□□□ VERY LOW	CRITICAL
Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: 24 months)												

2	randomised trials [2,9]	serious [d]	serious [f]	not serious	serious [c]	none	N=180	N=191	MD 0.04 lower (0.14 lower to 0.22 higher)		⊕□□□ VERY LOW	IMPORTANT
SF-12 Physical Component Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 24 months)												
1	randomised trial [2]	not serious	not assessable	serious [b]	serious [c]	none	N=151	N=151	MD 0.7 lower (2.92 lower to 1.52 higher)		⊕⊕□□ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Glucosamine + Chondroitin) (follow up: range 16 weeks to 24 months)												
5	randomised trials [1,2,3,4,5]	not serious	not serious	serious [b]	serious [c]	none	28/633 (4.4%)	23/620 (3.7%)	RR 1.18 (0.68 to 2.04)	7 more per 1,000 (from 12 fewer to 39 more)	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Glucosamine + Chondroitin) (follow up: 24 months)												
3	randomised trials [2,4,5]	not serious	serious [g]	serious [b]	serious [c]	none	82/296 (27.7%)	56/287 (19.5%)	RR 1.45 (0.77 to 2.73)	88 more per 1,000 (from 45 fewer to 338 more)	⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Glucosamine + Chondroitin) (follow up: range 16 weeks to 24 months)												
4	randomised trials [3,4,5,8]	not serious	not serious	serious [b]	serious [c]	none	6/294 (2.0%)	5/287 (1.7%)	RR 1.16 (0.36 to 3.75)	3 more per 1,000 (from 11 fewer to 48 more)	⊕⊕□□ LOW	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Glucosamine + Chondroitin) (follow up: 24 months)												
2	randomised trials [2,4]	not serious	serious [h]	serious [b]	serious [c]	none	11/216 (5.1%)	9/209 (4.3%)	RR 1.00 (0.21 to 4.81)	0 fewer per 1,000 (from 34 fewer to 164 more)	⊕□□□ VERY LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. I^2= 97%; removal of Roman-Blas 2017 reduces I^2 to 50% with an effect size (SMD and 95% CI) of -0.14 (-0.33, 0.05).
b. All patients in all trials have Knee Osteoarthritis.
c. 95% CI crosses null.
d. One study was a 2 year follow up study of a subset of patients who underwent a "departure from randomization." Randomization for structural outcomes was adequate. However, for all outcomes, maintenance of blinding from 24 weeks to 24 months is not adequately described. Not enough information is supplied to adequately assess other dimensions of the risk of bias tool.
e. I^2= 97%; removal of Roman-Blas 2017 reduces I^2 to 29% with an effect size (SMD and 95% CI) of -0.16 (-0.31, 0.00).
f. I^2= 56%; moderate heterogeneity.
g. I^2= 69%; moderate heterogeneity.
h. I^2= 61%; moderate heterogeneity.

PICO 2.4.4 (hip): What are the benefits and harms of Vitamin D supplementation in the management of patients with hip OA?

SUMMARY
There were four RCTs (one to three years' duration), all without serious risk of bias. However, there was very serious inconsistent results, leading to low-quality evidence. When combined, there is a suggestion of a favourable effect, statistically significant but not reaching what was regarded as clinical significance (standardised mean difference [SMD] >0.4). The results for function were similar to the effect estimates. There was no evidence of favourable structural effects on cartilage volume or joint space narrowing. Notably, one study in participants who were vitamin D deficient also failed to show clinically meaningful beneficial effects. Participants in all studies had knee OA, so extrapolation to OA of hip or other joints requires additional caution.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Arden, et al. Osteoarthritis Cartilage. 2016 Nov; 24(11): 1858-1866; 2. Jin, et al. JAMA. 2016 Mar 8; 315(10): 1005-13; 3. McAlindon, et al. JAMA. 2013 Jan 9; 309(2): 155-62; 4. Sanghi, et al. Clin Orthop Relat Res. 2013 Nov; 471(11): 3556-62.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Vitamin D	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 12 months to 3 years)												
4	randomised trials [1,2,3,4]	not serious	very serious [a]	serious [b]	not serious	none	N=571	N=565	**SMD 0.36 lower (0.7 lower to 0.02 lower)**		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 12 months to 3 years)												
4	randomised trials [1,2,3,4]	not serious	very serious [c]	serious [b]	not serious	none	N=571	N=565	**SMD 0.34 lower (0.62 lower to 0.07 lower)**		⊕□□□ VERY LOW	CRITICAL
Tibial cartilage volume (mm^3) (Higher values indicate better structural outcome) (follow up: 24 months)												
2	randomised trials [2,3]	not serious	not serious	serious [b]	serious [d]	none	N=282	N=277	MD 35.44 mm^3 higher (13.66 mm^3 lower to 84.54 mm^3 higher)		⊕⊕□□ LOW	IMPORTANT
Radiographic Progression [JSN >0.5 mm] (Risk ratios less than one favor Vitamin D) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	serious [b]	serious [d]	none	92/237 (38.8%)	88/237 (37.1%)	RR 1.05 (0.83 to 1.32)	19 more per 1,000 (from 63 fewer to 119 more)	⊕⊕□□ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Vitamin D) (follow up: range 12 months to 3 years)												

4	randomised trials [1,2,3,4]	not serious	not serious	serious [b]	serious [d]	none	14/572 (2.4%)	14/567 (2.5%)	RR 0.99 (0.48 to 2.05)	0 fewer per 1,000 (from 13 fewer to 26 more)	⊕⊕◯◯ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Vitamin D) (follow up: 24 months)												
2	randomised trials 2,3	not serious	serious [e]	serious [b]	serious [d]	none	103/282 (36.5%)	83/277 (30.0%)	RR 1.20 (0.82 to 1.77)	60 more per 1,000 (from 54 fewer to 231 more)	⊕◯◯◯ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Vitamin D) (follow up: range 24 months to 3 years)												
3	randomised trials [1,2,3]	not serious	not serious	serious [b]	serious [d]	none	86/519 (16.6%)	87/514 (16.9%)	RR 0.97 (0.75 to 1.27)	5 fewer per 1,000 (from 42 fewer to 46 more)	⊕⊕◯◯ LOW	CRITICAL
Hypercalcemia (Risk ratios less than one favor Vitamin D) (follow up: range 24 months to 3 years)												
3	randomised trials [1,2,3]	not serious	not serious	serious [b]	serious [d]	none	7/519 (1.3%)	7/514 (1.4%)	RR 0.99 (0.32 to 3.10)	0 fewer per 1,000 (from 9 fewer to 29 more)	⊕⊕◯◯ LOW	IMPORTANT
Hypercalciuria (Risk ratios less than one favor Vitamin D) (follow up: range 24 months to 3 years)												
2	randomised trials [1,3]	not serious	not serious	serious [b]	serious [d]	none	52/310 (16.8%)	38/310 (12.3%)	RR 1.37 (0.93 to 2.01)	45 more per 1,000 (from 9 fewer to 124 more)	⊕⊕◯◯ LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. I^2= 86%; with the exclusion of Sanghi, 2013, I^2 drops to 0%. Sanghi specifically selected osteoarthritis patients with Vitamin D deficiency at baseline, which may have contributed to discordant results (Clin Orthop Relat Res. 2013 Nov;471(11):3714-5)
b. All patients in all trials have Knee Osteoarthritis.
c. I^2= 78%; with the exclusion of Sanghi, 2013, I^2 drops to 15%. Sanghi specifically selected osteoarthritis patients with Vitamin D deficiency at baseline, which may have contributed to discordant results (Clin Orthop Relat Res. 2013 Nov;471(11):3714-5)
d. 95% CI crosses null.
e. I^2= 68%; moderate heterogeneity.

PICO 2.4.5 (hip): What are the benefits and harms of (omega-3/6) poly-unsaturated fatty acids in the management of patients with hip OA?

SUMMARY
Pooled data from five RCTs (15–26 weeks) demonstrated no benefits on pain and function in people with hip and knee OA. Three studies received at least one high risk of bias. Most trials used marine oil from whole fish, but some used cod liver oil and mussel extracts. High heterogeneity was expected from pooling different sources of omega-3 fatty acids, and measures within each outcome. The optimal type of omega-3 fatty acids could not be established in OA because only a few trials included marine oil from sources other than whole fish. There are high variations in doses of eicosapentaenoic acid (EPA; 0.01–1.7 g/day), and doses of docosahexaenoic acid (DHA; 0.01–1.10 g/day). A controlled trial that was not included found no additional benefit of high dose fish oil (4.5 g/day), compared with low dose fishoil (0.45 g/day).

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Gruenwald, et al. Adv Ther. 2009 Sep; 26(9): 858-71; 2. Lau, et al. Progress in Nutrition. 2004; 3. Stammers, et al. Lancet. 1989 Aug 26; 2(8661): 503; 4. Stammers, et al. Ann Rheum Dis. 1992 Jan; 51(1): 128-9; 5. Stebbings, et al. Annals of the Rheumatic Diseases. 2014 Jun 1; 73(Suppl 2): 755.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	(omega-3/6) poly-unsaturated fatty acids	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 15 weeks to 26 weeks)												
5	randomised trials [1,2,3,4,5]	serious [a]	very serious [b]	serious [c]	serious [d]	none	N=201	N=207	SMD 0.16 lower (0.57 lower to 0.24 higher)		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 15 weeks to 26 weeks)												
5	randomised trials [1,2,3,4,5]	serious [a]	very serious [b]	serious [c]	serious [d]	none	N=201	N=207	SMD 0.11 higher (0.13 lower to 0.35 higher)		⊕□□□ VERY LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Fish oil) (follow up: range 24 weeks to 26 weeks)												
3	randomised trials [1,2,4]	serious [a]	not serious	serious [c]	serious [d]	none	10/174 (5.7%)	7/170 (4.1%)	RR 1.33 (0.52 to 3.39)	14 more per 1,000 (from 20 fewer to 98 more)	⊕□□□ VERY LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Fish oil) (follow up: range 24 weeks to 26 weeks)												
3	randomised trials [1,2,4]	serious [a]	not serious	serious [c]	serious [d]	none	29/174 (16.7%)	21/170 (12.4%)	RR 1.31 (0.79 to 2.18)	38 more per 1,000 (from 26 fewer to 146 more)	⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Fish oil) (follow up: range 15 weeks to 26 weeks)												

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	(omega-3/6) poly-unsaturated fatty acids	Placebo	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials [1,5]	serious [a]	not serious	serious [c]	not serious	none	0/129 (0.0%)	0/129 (0.0%)	Due to zero events in both groups, an absolute risk reduction was not estimable.		⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. 3 of 5 studies received at least one High risk of bias rating; 1 of 5 studies received all Unclear risk of bias ratings due to insufficient information.
b. I^2 >75%
c. Mixed populations of Hip, Knee, Hip/Knee Osteoarthritis patients.
d. 95% CI crosses null.

PICO 2.4.6 (hip): What are the benefits and harms of collagen preparations in the management of patients with hip OA?

SUMMARY
Pooled results from six studies found short-term (13–26 weeks) clinical benefits in pain; however, there have been very serious inconsistent results across the studies. Available data from four studies found no effect in function. All of the studies were conducted in knee OA, so extrapolation to hip or other OA requires additional caution. There are concerns in the literature of publication bias, effects being mostly driven by industrysponsored trials, and the overall poor quality of the positive trials.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Schauss, et al. J Agric Food Chem. 2012 Apr 25; 60(16): 4096-101.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Collagen**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 70 days)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	very serious [c,d]	none	N=40	N=40	MD 0.7 lower (2 lower to 0.6 higher)		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 70 days)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [d]	none	N=40	N=40	**MD 7.4 lower (11.36 lower to 3.44 lower)**		⊕□□□ VERY LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Collagen) (follow up: 70 days)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	very serious [c,d]	none	3/35 (8.6%)	3/33 (9.1%)	RR 0.94 (0.20 to 4.35)	5 fewer per 1,000 (from 73 fewer to 305 more)	⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Collagen) (follow up: 70 days)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [d]	none	0/40 (0.0%)	0/40 (0.0%)	Due to zero events in both groups, an absolute risk reduction was not estimable.		⊕□□□ VERY LOW	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Collagen) (follow up: 70 days)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	very serious [c,d]	none	0/35 (0.0%)	2/33 (6.1%)	RR 0.19 (0.01 to 3.79)	49 fewer per 1,000 (from 60 fewer to 169 more)	⊕□□□ VERY LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Study received a High risk of bias rating due to potential reporting bias.
b. Study involves mixed Hip, Knee, and Hip+Knee Osteoarthritis population. The proportions of each are not described.
c. 95% CI crosses null.
d. Sample size in each study arm <50.

PICO 2.4.7 (hip): What are the benefits and harms of methylsulfonylmethane in the management of patients with hip OA?

SUMMARY
There are three trials with short study durations (12–13 weeks), and pooled data found statistically and clinically significant benefits in pain. Even larger effects were found in function, but with very serious inconsistent results and high heterogeneity across studies. One trial had a high risk of bias because of inappropriate randomisation technique; while the other had potential reporting bias. The doses in the trials ranged from 1.5–6 g/day for 12 weeks. All of the studies were conducted in knee OA, so extrapolation to hip or other OA sites requires additional caution. There are concerns in the literature of publication bias, effects being mostly driven by industrysponsored trials, and the overall poor quality of the positive trials.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1. Debbi, et al. BMC Complement Altern Med. 2011 Jun 27; 11: 50; 2. Kim, et al. Osteoarthritis Cartilage. 2006 Mar; 14(3): 286-94; 3. Usha, P.R., Naidu, M.U.R. Clin Drug Investig. 2004; 24(6): 353-63.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Methylsulfonyl-methane	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 12 weeks to 13 weeks)												
3	randomised trials [1,2,3]	serious [a]	not serious	serious [b]	not serious	none	N=76	N=72	**SMD 0.47 lower (0.8 lower to 0.14 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 12 weeks to 13 weeks)												
3	randomised trials [1,2,3]	serious [a]	very serious [c]	serious [b]	not serious	none	N=76	N=72	**SMD 1.1 lower (1.81 lower to 0.38 lower)**		⊕□□□ VERY LOW	CRITICAL
Quality of Life (Higher scores indicate better quality of life) (follow up: range 12 weeks to 13 weeks)												
2	randomised trials [1,2]	serious [d]	not serious	serious [b]	very serious [e,f]	none	N=46	N=44	SMD 0.42 lower (0.86 lower to 0.01 higher)		⊕□□□ VERY LOW	IMPORTANT
Total Adverse Events (Risk ratios less than one favor Methylsulfonylmethane) (follow up: 13 weeks)												
1	randomised trial [2]	not serious	not assessable	serious [b]	very serious [e,f]	none	21/25 (84.0%)	19/25 (76.0%)	RR 1.1 (0.8 to 1.5)	76 more per 1,000 (from 152 fewer to 380 more)	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. Two studies received at least one High risk of bias rating due to inappropriate randomization technique and potential reporting bias, respectively.
b. All patients in all trials have Knee Osteoarthritis.
c. I^2= 76%

d. One study received at least one High risk of bias rating due to inappropriate randomization technique.
e. 95% CI crosses null.
f. Sample size in each study arm <50.

PICO 2.4.8 (hip): What are the benefits and harms of Diacerein in the management of patients with hip OA?

SUMMARY
Five trials were included, with time durations ranging from eight weeks to 12 months, all receiving high risk of bias because of weak allocation concealment and random sequence generation. Very low-quality evidence from four trials indicated a small clinical benefit on pain reduction. Data from five trials indicated statistically significant effects on function, but this did not reach the clinically meaningful threshold. Analysis of one study demonstrated no benefit in reducing joint space narrowing. There are concerns in the literature of publication bias, effects being mostly driven by industrysponsored trials, and the overall poor quality of the positive trials. A search of regulatory websites found a recommendation from the European Medicines Agency (EMA) Pharmacovigilance Risk Assessment Committee (PRAC) that the marketing authorisation of diacerein should be suspended across Europe because of harms (particularly the risk of severe diarrhoea and potentially harmful effects on the liver) outweighing benefits. However, this guidance is not final as the PRAC recommendation will be re-examined. All of the studies were conducted in knee OA, so extrapolation to hip or other OA sites requires additional caution.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Dougados, et al. Arthritis Rheum. 2001 Nov; 44(11): 2539-47; 2. Nguyen, et al. Arthritis Rheum. 1994 Apr; 37(4): 529-36; 3. Lequesne, et al. Rev Prat. 1998 Nov 1; 48(17 Suppl): S31-5

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Diacerein	Placebo	Relative (95% CI)	Absolute (95% CI)		
VAS Pain (scale range 0 to 100 mm, with higher scores indicating higher pain severity) (follow up: range 8 weeks to 3 years)												
2	randomised trials [1,2]	serious [a]	not serious	not serious	not serious	none	N=330	N=323	**MD 4.62 lower (7.83 lower to 1.42 lower)**		⊕⊕⊕□ MODERATE	CRITICAL
Lequesne Index (scale range 1 to 24, with higher scores indicating poorer functional outcome) (follow up: range 8 weeks to 3 years)												
3	randomised trials [1,2,3]	serious [a]	not serious	not serious	serious [b]	none	N=352	N=357	MD 0.01 lower (0.43 lower to 0.41 higher)		⊕⊕□□ LOW	CRITICAL
Mean annual Joint Space Narrowing rate [mm/year] (Higher values indicate a poorer structural outcome) (follow up: 3 years)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [b]	none	N=131	N=138	MD 0.05 lower (0.11 lower to 0.01 higher)		⊕⊕□□ LOW	IMPORTANT
Proportion of Patients with Radiographic Progression [JSN ≥0.5mm] (Risk ratios less than one favor Diacerein)(follow up: 3 years)												
1	randomised trial [1]	serious [a]	not assessable	not serious	not serious	none	112/221 (50.7%)	136/225 (60.4%)	**RR 0.84 (0.71 to 0.99)**	**97 fewer per 1,000 (from 6 fewer to 175 fewer)**	⊕⊕⊕□ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Diacerein) (follow up: range 8 weeks to 3 years)												
3	randomised trials [1,2,3]	serious [a]	not serious	not serious	serious [b]	none	100/427 (23.4%)	67/423 (15.8%)	RR 1.40 (0.68 to 2.91)	63 more per 1,000 (from 51 fewer to 303 more)	⊕⊕□□ LOW	CRITICAL

Total Adverse Events (Risk ratios less than one favor Diacerein) (follow up: range 6 months to 3 years)												
2	randomised trials [1,3]	serious [a]	serious [c]	not serious	serious [b]	none	277/345 (80.3%)	223/345 (64.6%)	RR 1.87 (0.51 to 6.89)	562 more per 1,000 (from 317 fewer to 1,000 more)	⊕□□□ VERY LOW	CRITICAL
Rash/Pruritus (Risk ratios less than one favor Diacerein) (follow up: 3 years)												
1	randomised trial [1]	serious [a]	not assessable	not serious	not serious	none	17/255 (6.7%)	7/252 (2.8%)	RR 2.40 (1.01 to 5.69)	39 more per 1,000 (from 0 fewer to 130 more)	⊕⊕⊕□ MODERATE	IMPORTANT
Diarrhea (Risk ratios less than one favor Diacerein) (follow up: range 8 weeks to 3 years)												
3	randomised trials [1,2,3]	serious [a]	not serious	not serious	not serious	none	183/420 (43.6%)	45/416 (10.8%)	**RR 4.02 (2.64 to 6.11)**	**327 more per 1,000 (from 177 more to 553 more)**	⊕⊕⊕□ MODERATE	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. All studies received at least one High risk of bias rating due to attrition bias or reporting bias
b. 95% CI crosses null
c. I^2=9

PICO 2.5 (hip): What are the benefits and harms of duloxetine in the management of patients with hip OA?

SUMMARY
In the three trials reviewed, significant response and moderate effects in knee pain (standardised mean difference [SMD] 0.43) and function (SMD 0.45) were found over 13–16 weeks at doses of 60/120 mg. However, most study participants were also already using NSAIDs and paracetamol. The use of duloxetine for knee OA adjunctively with NSAIDs, thus reducing the usage of NSAIDs and paracetamol, would be clinically useful to reduce adverse events. (Brown JP, Boulay LJ. Ther Adv Musculoskelet Dis 2013;5(6):291–304). In addition, results differed as to whether significant reduction in depression symptoms was needed for analgesic impact. There is no direct randomised controlled trial (RCT) evidence for hip OA; thus, using knee OA data to extrapolate to hip or other OA requires additional caution. Duloxetine currently does not have an indication via the Therapeutic Goods Administration (TGA) for OA, and should be considered as an investigational medication only.

QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. Abou-Raya, et al. Age Ageing. 2012 Sep; 41(5): 646-52; 2. Chappell, et al. Pain. 2009 Dec; 146(3): 253-60; 3. Chappell, et al. Pain Pract. 2011 Jan-Feb; 11(1): 33-41.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Duloxetine**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow up: range 13 weeks to 16 weeks)												
3	randomised trials [1,2,3]	serious [a]	not serious	serious [b]	not serious	none	N=379	N=389	**SMD 0.43 lower (0.58 lower to 0.29 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 13 weeks to 16 weeks)												
2	randomised trials [1,3]	serious [a]	not serious	serious [b]	not serious	none	N=272	N=272	**SMD 0.45 lower (0.81 lower to 0.08 lower)**		⊕⊕□□ LOW	CRITICAL
EQ-5D UK index (Higher scores indicate better quality of life) (follow up: 13 weeks)												
1	randomised trial [2]	not serious	not assessable	serious [b]	not serious	none	N=103	N=114	**MD 0.1 higher (0.04 higher to 0.16 higher)**		⊕⊕⊕□ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Duloxetine) (follow-up: range 13 weeks to 16 weeks)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	not serious	none	48/383 (12.5%)	20/392 (5.1%)	**RR 2.42 (1.46 to 4.03)**	**72 more per 1,000 (from 23 more to 155 more)**	⊕⊕⊕⊕ HIGH	CRITICAL
Treatment- Related Adverse Events (Risk ratios less than one favor Duloxetine) (follow-up: 13 weeks)												

1	randomised trial [3]	not serious	not assessable	not serious	not serious	none	65/128 (50.8%)	42/128 (32.8%)	**RR 1.55 (1.15 to 2.09)**	**180 more per 1,000 (from 49 more to 358 more)**	⊕⊕⊕⊕ HIGH	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Duloxetine) (follow-up: range 13 weeks to 16 weeks)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	serious [b]	none	4/383 (1.0%)	4/392 (1.0%)	RR 1.04 (0.25 to 4.33)	0 fewer per 1,000 (from 8 fewer to 34 more)	⊕⊕⊕◻ MODERATE	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. Chappell 2011 received a High risk of bias rating due to potential attrition bias
b. All patients in all trials have Knee Osteoarthritis
c.95% CI crosses null

PICO 2.6 (hip): What are the benefits and harms of doxycycline in the management of patients with hip OA?

SUMMARY
Preclinical research and earlier human studies indicated doxycycline might be useful in managing symptomatic knee OA. However, current evidence found that doxycycline did not reduce the mean severity of joint pain, although pain scores in both treatment groups were low at baseline and remained low throughout the trial, which may suggest the presence of a floor effect. Brown JP, Boulay LJ. Ther Adv Musculoskelet Dis 2013;5(6):291–304). Despite the small benefit (SMD 0.15 mm) in joint space narrowing, it is outweighed by medication harms. There is no RCT of doxycycline for hip OA, thus using knee OA data to extrapolate to hip or other OA requires additional caution. Doxycycline currently does not have an indication via the TGA for OA, and should be considered as an investigational medication only.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. da Costa, et al. Cochrane Database Syst Rev. 2012 Nov 14; 11: CD007323.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Doxycycline	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (multiple measures including 50-ft walking pain, WOMAC pain) (Higher scores indicate higher pain severity) (follow up: range 24 weeks to 30 months)												
2	randomised trials [1]	serious [a]	not serious	serious [b]	serious [c]	none	N=256	N=268	SMD 0.05 lower (0.22 lower to 0.13 higher)		⊕□□□ VERY LOW	CRITICAL
WOMAC function (Higher scores indicate poorer functional outcome) (follow up: range 24 weeks to 30 months)												
2	randomised trials [1]	serious [a]	not serious	serious [b]	serious [c]	none	N=252	N=265	SMD 0.07 lower (0.25 lower to 0.1 higher)		⊕□□□ VERY LOW	CRITICAL
Minimum Joint Space Width (Higher values indicate a better structural outcome) (follow up: 30 months)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	not serious	none	N=181	N=180	**SMD 0.15 lower (0.28 lower to 0.02 lower)**		⊕⊕□□ LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Doxycycline) (follow up: range 24 weeks to 30 months)												
2	randomised trials [1]	not serious	serious [d]	serious [b]	not serious	none	57/334 (17.1%)	27/329 (8.2%)	**RR 2.28 (1.06 to 4.90)**	**105 more per 1,000 (from 5 more to 320 more)**	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Doxycycline) (follow up: 24 weeks)												
1	randomised trial [1]	not serious	not assessable	serious [b]	not serious	none	75/116 (64.7%)	55/116 (47.4%)	**RR 1.36 (1.08 to 1.72)**	**171 more per 1,000 (from 38 more to 341 more)**	⊕⊕⊕□ MODERATE	CRITICAL

Serious Adverse Events (Risk ratios less than one favor Doxycycline) (follow up: range 24 weeks to 30 months)												
2	randomised trials [1]	not serious	not serious	serious [b]	serious [c]	none	34/334 (10.2%)	31/329 (9.4%)	RR 1.07 (0.68 to 1.68)	7 more per 1,000 (from 30 fewer to 64 more)	⊕⊕◻◻ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. Both studies received at least one High risk bias rating due to non-ITT analyses and possible selective reporting
b. All patients in both trials have Knee Osteoarthritis.
c. 95% CI crosses null
d. I^2=55%; moderate heterogeneity

PICO 2.7: ANTI-OSTEOPOROSIS DRUGS
PICO 2.7.1 (hip): What are the benefits and harms of bisphosphonates in the management of patients with hip OA?

SUMMARY
There is one very low-quality trial conducted in 42 participants with hip OA, demonstrating no effect over 24 months. Biphosphonates currently do not have an indication via the Therapeutic Goods Administration (TGA) for OA, and should be considered as investigational medications only.

OVERALL QUALITY OF EVIDENCE: ⊕□□□ VERY LOW

Bibliography: 1.Nishii, et al. Clin Rheumatol. 2013 Dec; 32(12): 1759-66.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Alendronate + Calcium	Calcium alone	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (Scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 24 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	N=30	N=12	MD 0.98 lower (2.95 lower to 0.99 higher)		⊕□□□ VERY LOW	CRITICAL
Short-term WOMAC Pain (Scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 12 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	serious [c]	none	N=30	N=12	**MD 1.74 lower (3.28 lower to 0.19 lower)**		⊕⊕□□ LOW	IMPORTANT
Proportion of Patients Experiencing Structural Progression (decrease in joint space widening >0.30 mm) (Risk ratios less than one favor Bisphosphonate) (follow up: 24 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	18/30 (60.0%)	7/12 (58.3%)	RR 1.03 (0.59 to 1.80)	18 more per 1,000 (from 239 fewer to 467 more)	⊕□□□ VERY LOW	IMPORTANT
Proportion of Patients Experiencing Structural Progression (decrease in joint space widening >0.30 mm) (Risk ratios less than one favor Bisphosphonate) (follow up: 12 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	15/30 (50.0%)	4/12 (33.3%)	RR 1.50 (0.62 to 3.60)	167 more per 1,000 (from 127 fewer to 867 more)	⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Bisphosphonate) (follow up: range 60 days to 24 months)												
1	randomised trial [1]	serious [a]	not assessable	not serious	very serious [b,c]	none	2/33 (6.1%)	1/17 (5.9%)	RR 1.03 (0.10 to 10.57)	2 more per 1,000 (from 53 fewer to 563 more)	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Study received High risk of bias ratings due to inadequate blinding and potential attrition bias
b. 95% CI crosses null
c. Sample size <50 in each study arm

PICO 2.7.2 (hip): What are the benefits and harms of calcitonin in the management of patients with hip OA?

SUMMARY
The two phase III studies found no significant effect of salmon calcitonin on total WOMAC, WOMAC subscores and joint space narrowing. There is a potentially small effect on markers of bone and cartilage degradation, CTX-I and CTX-II respectively, and no positive balance between bone formation and bone resorption. There is no randomised controlled trial (RCT) of calcitonin for hip OA; thus, using knee OA data to extrapolate to hip or other OA requires additional caution.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Karsdal, et al.[1] Osteoarthritis and Cartilage 23 (2015) 532-543; 2. Karsdal, et al.[2] Osteoarthritis and Cartilage 23 (2015) 532-543; 3. Manicourt, et al. Arthritis Rheum. 2006 Oct; 54(10):3205-11

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Calcitonin**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0 to 150, with higher scores indicating higher pain severity) (follow up: 12 months)												
2	randomised trials [1,2]	serious [a]	very serious [b]	serious [c]	serious [d]	none	N=803	N=885	MD 6.65 lower (30.15 lower to 16.85 higher)		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (scale range 0 to 150, with higher scores indicating poorer functional outcome) (follow up: 12 months)												
2	randomised trials [1,2]	serious [a]	serious [e]	serious [c]	serious [d]	none	N=791	N=874	MD 18.42 lower (76.86 lower to 40.01 higher)		⊕□□□ VERY LOW	CRITICAL
Joint Space Widening [mm] (Positive values indicate better structural outcome) (follow up: 12 months)												
2	randomised trials [1,2]	serious [a]	not serious	serious [c]	serious [d]	none	N=804	N=885	MD 0.02 higher (0.04 lower to 0.08 higher)		⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Calcitonin) (follow-up: range 85 days to 24 months)												
3	randomised trials [1,2,3]	serious [a]	not serious	not serious	not serious	none	196/1140 (17.2%)	71/1110 (6.4%)	**RR 2.68 (2.07 to 3.47)**	**107 more per 1,000 (from 68 more to 158 more)**	⊕⊕⊕□ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Calcitonin) (follow-up: range 85 days to 24 months)												

3	randomised trials [1,2,3]	not serious	serious [f]	not serious	serious [d]	none	1034/1140 (90.7%)	981/1110 (88.4%)	RR 1.03 (0.98 to 1.09)	27 more per 1,000 (from 18 fewer to 80 more)	⊕⊕◯◯ LOW	CRITICAL
Withdrawal due to Serious Adverse Events (Risk ratios less than one favor Calcitonin) (follow-up: 24 months)												
2	randomised trials [1,2]	not serious	not serious	not serious	serious [d]	none	18/1105 (1.6%)	11/1092 (1.0%)	RR 1.59 (0.75 to 3.40)	6 more per 1,000 (from 3 fewer to 24 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Calcitonin) (follow-up: range 85 days to 24 months)												
3	randomised trials [1,2,3]	not serious	serious [g]	not serious	not serious	none	480/1140 (42.1%)	303/1110 (27.3%)	**RR 1.55 (1.20 to 2.00)**	**150 more per 1,000 (from 55 more to 273 more)**	⊕⊕⊕◯ MODERATE	IMPORTANT
Hot Flush (Risk ratios less than one favor Calcitonin) (follow-up: range 85 days to 24 months)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	not serious	none	199/1140 (17.5%)	47/1110 (4.2%)	**RR 4.11 (3.02 to 5.59)**	**132 more per 1,000 (from 86 more to 194 more)**	⊕⊕⊕⊕ HIGH	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. All trials received at least one High risk of bias rating due to poor attrition without ITT and/or differential dropout
b. I^2=85%
c. All patients in the studies have Knee Osteoarthritis.
d. 95% CI crosses null
e. I^2=72% moderate heterogeneity
f. I^2=59%; moderate heterogeneity
g. I^2=66%; moderate heterogeneity

PICO 2.7.3 (hip): What are the benefits and harms of strontium ranelate in the management of patients with hip OA?

SUMMARY
Data from one moderate-quality trial found no effect of strontium ranelate in altering OA symptoms. However, strontium ranelate reatment had a beneficial effect on joint space widening, with a mean difference (MD) of 0.12 mm over three years. Similarly, the risk ratio of radiographic progression (joint space narrowing ≥0.5 mm) favoured strontium ranelate over three years. As strontium ranelate is not accessible for people in Australia, the working group considered this treatment as unfeasible for use.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. Reginster, et al. Ann Rheum Dis. 2013 Feb; 72(2): 179-86.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Strontium ranelate**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	N=899	N=472	MD 2.14 lower (6.55 lower to 2.27 higher)		⊕⊕□□ LOW	CRITICAL
WOMAC Function (scale range 0 to 100, with higher scores indicating poorer functional outcome) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	N=899	N=472	MD 1.88 lower (4.63 lower to 0.86 higher)		⊕⊕□□ LOW	CRITICAL
Joint Space Widening [mm] (Positive values indicate better structural outcome) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	serious [a]	not serious	none	N=899	N=472	**MD 0.12 higher (0.05 higher to 0.19 higher)**		⊕⊕⊕□ MODERATE	IMPORTANT
Patients experiencing Radiological Progression [JSN ≥0.5 mm] (Risk ratios less than one favor Strontium ranelate) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	serious [a]	not serious	none	215/899 (23.9%)	156/472 (33.1%)	**RR 0.72 (0.61 to 0.86)**	**93 fewer per 1,000 (from 46 fewer to 129 fewer)**	⊕⊕⊕□ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Strontium ranelate) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [b]	none	45/1112 (4.0%)	27/556 (4.9%)	RR 0.83 (0.52 to 1.33)	8 fewer per 1,000 (from 16 more to 23 fewer)	⊕⊕⊕□ MODERATE	CRITICAL

Gastrointestinal Adverse Events (Risk ratios less than one favor Strontium ranelate) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [b]	none	136/1112 (12.2%)	52/556 (9.4%)	RR 1.31 (0.97 to 1.77)	29 more per 1,000 (from 3 fewer to 72 more)	⊕⊕⊕◻ MODERATE	IMPORTANT
Skin and Subcutaneous Disorders (Sum of N patients experiencing "Dermatitis", "Allergic Dermatitis", "Eczema", and "Rash") (Risk ratios less than one favor Strontium ranelate) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [b]	none	44/1112 (4.0%)	20/556 (3.6%)	RR 1.10 (0.65 to 1.85)	4 more per 1,000 (from 13 fewer to 31 more)	⊕⊕⊕◻ MODERATE	IMPORTANT
Venous Thromboembolism Events (Sum of N patients experiencing "Deep venous thrombosis" and "Pulmonary embolism") (Risk ratios less than one favor Strontium ranelate) (follow up: 3 years)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [b]	none	9/1112 (0.8%)	2/556 (0.4%)	RR 2.25 (0.49 to 10.38)	4 more per 1,000 (from 2 fewer to 34 more)	⊕⊕⊕◻ MODERATE	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. All patients in this trial have Knee Osteoarthritis.
b. 95% CI crosses null

PICO 2.8: INVESTIGATIONAL DMOADs
PICO 2.8.1 (hip): What are the benefits and harms of IL-1 inhibitors in the management of patients with hip OA?

SUMMARY
Results from a three-arm trial of a single intra-articular injection of anakinra at a dose of 50 mg (n = 34) and 150 mg (n = 67) were vailable. The mean improvement from baseline at week 12 on the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) score was not statistically different between the anakinra and placebo groups. A placebo-controlled randomised controlled trial (RCT) of AMG-108, (not included in this review) found non-statistically significant improvement on WOMAC pain after subcutaneous administration of AMG-108. (Cohen SB, et al. Arthritis Res Ther 2011;13(4):R125). Due to the limitations in current efficacy, safety,access and costs, it is considered that IL-1 inhibitors are not a feasible nor cost-effective treatment.There are currently no trials that have investigated the benefits and safety of IL-1 inhibitors in people with hip OA. Using knee OA data to extrapolate to hip or other OA requires additional caution
OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. Chevalier, et al. Arthritis Rheum. 2009 Mar 15; 61(3): 344-52

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	IA Anakinra (single 50 mg/150mg dose)	Placebo	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (Scale range 0 to 500, with higher scores indicating higher pain severity) (follow up: 12 weeks)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	N=101	N=69	MD 2.71 lower (33.78 lower to 28.36 higher)		⊕⊕□□ LOW	CRITICAL
WOMAC Function (Scale range 0 to 1,700, with higher scores indicating poorer functional outcome) (follow up: 12 weeks)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	N=101	N=69	MD 21.12 lower (132.53 lower to 90.29 higher)		⊕⊕□□ LOW	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Anakinra) (follow up: 12 weeks)												
1	randomised trial [1]	not serious	not assessable	serious [a]	not serious	none	0/101 (0.0%)	0/69 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕□ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Anakinra) (follow up: 12 weeks)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	53/101 (52.5%)	41/69 (59.4%)	RR 0.88 (0.67 to 1.16)	71 fewer per 1,000 (from 95 more to 196 fewer)	⊕⊕□□ LOW	CRITICAL

Serious Adverse Events (Risk ratios less than one favor Anakinra) (follow up: 12 weeks)											
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	1/101 (1.0%)	1/69 (1.4%)	RR 0.68 (0.04 to 10.74)	5 fewer per 1,000 (from 14 fewer to 141 more)	⊕⊕◻◻ LOW	CRITICAL
Infections (Risk ratios less than one favor Anakinra) (follow up: 12 weeks)											
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	13/101 (12.9%)	4/69 (5.8%)	RR 2.22 (0.76 to 6.53)	71 more per 1,000 (from 14 fewer to 321 more)	⊕⊕◻◻ LOW	IMPORTANT
Injection Site Reactions (Risk ratios less than one favor Anakinra) (follow up: 12 weeks)											
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	4/101 (4.0%)	4/69 (5.8%)	RR 0.68 (0.18 to 2.64)	19 fewer per 1,000 (from 48 fewer to 95 more)	⊕⊕◻◻ LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. All patients in this study have Knee Osteoarthritis.
b. 95% CI crosses null

PICO 2.8.2 (hip): What are the benefits and harms of TNF-alpha inhibitors in the management of patients with hip OA?

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

No RCT data was found for this question.

PICO 2.8.3 (hip): What are the benefits and harms of anti-nerve growth factor (NGF) therapy in the management of patients with hip OA?

SUMMARY
In these hip studies, statistically significant but less clinically relevant effects were found on WOMAC pain and function scores, with pooled SMDs of 0.33 and 0.4 respectively. The study (Reference 3) evaluated fulranumab with two different dosing frequencies 1 and 3 mg every four weeks; 3, 6 and 10 mg every eight weeks), showing a numerical difference from the active control (oxycodone), although no differentiation was seen between either fulranumab dose and placebo in the same study. Anti-NGF requires off-label prescribing and is expensive, which limited its accessibility and affordability.

OVERALL QUALITY OF EVIDENCE: ⊕⊕⊕◯ MODERATE

Bibliography: 1. Ekman, et al.[1] J Rheumatol. 2014 Nov; 41(11): 2249-59; 2. Ekman, et al. [2] J Rheumatol. 2014 Nov; 41(11): 2249-59; 3. Sanga, et al. Pain. 2013 Oct; 154(10): 1910-9; 4. Spierings, et al. Pain. 2013 Sep; 154(9): 1603-12.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Anti-Nerve Growth Factor Therapy**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (Higher scores indicate higher pain severity) (follow-up: range 8 weeks to 16 weeks)												
4	randomised trials [1,2,3,4]	not serious	not serious	serious [a]	not serious	none	N=1531	N=635	**SMD 0.33 lower (0.43 lower to 0.24 lower)**		⊕⊕⊕◯ MODERATE	CRITICAL
WOMAC Function (Higher scores indicate poorer functional outcome) (follow-up: range 8 weeks to 16 weeks)												
3	randomised trials [1,2,4]	not serious	not serious	serious [a]	not serious	none	N=1142	N=556	**SMD 0.4 lower (0.5 lower to 0.3 lower)**		⊕⊕⊕◯ MODERATE	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Anti-Nerve Growth Factor) (follow-up: range 8 weeks to 16 weeks)												
4	randomised trials [1,2,3,4]	not serious	not serious	not serious	serious [b]	none	59/1533 (3.8%)	19/636 (3.0%)	RR 1.36 (0.81 to 2.27)	11 more per 1,000 (from 6 fewer to 38 more)	⊕⊕⊕◯ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Anti-Nerve Growth Factor) (follow-up: range 8 weeks to 16 weeks)												
4	randomised trials [1,2,3,4]	not serious	not serious	not serious	not serious	none	784/1533 (51.1%)	274/636 (43.1%)	**RR 1.16 (1.05 to 1.29)**	**69 more per 1,000 (from 22 more to 125 more)**	⊕⊕⊕⊕ HIGH	CRITICAL
Treatment-related Adverse Events (Risk ratios less than one favor Anti-Nerve Growth Factor) (follow-up: range 8 weeks to 12 weeks)												

2	randomised trials [3,4]	not serious	not serious	not serious	serious [b]	none	271/699 (38.8%)	55/219 (25.1%)	RR 1.20 (0.89 to 1.61)	50 more per 1,000 (from 28 fewer to 153 more)	⊕⊕⊕◻ MODERATE	IMPORTANT
Serious Adverse Events (Risk ratios less than one favor Anti-Nerve Growth Factor) (follow-up: range 8 weeks to 16 weeks)												
4	randomised trials [1,2,3,4]	not serious	not serious	not serious	serious [b]	none	29/1533 (1.9%)	15/636 (2.4%)	RR 0.88 (0.47 to 1.64)	3 fewer per 1,000 (from 13 fewer to 15 more)	⊕⊕⊕◻ MODERATE	CRITICAL

Ekman, 2014; %knee: 100%, %hip: 0% Ekman, 2014; %knee: 80%, hip%: 20% Sanga, 2013; %knee:77%, %hip: 23% Spierings, 2013; %knee: 77%, %hip: 23%

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. All studies involve mixed populations (hip and knee osteoarthritis)
b. 95% CI crosses null

PICO 2.8.4 (hip): What are the benefits and harms of fibroblast growth factor (FGF) therapy in the management of patients with hip OA?

SUMMARY
There is one trial of 190 participants with knee OA evaluating the effects of intra-articular injection of sprifermin as a single treatment and multiple-dose regimen (three doses of either 10, 30 or 100 µg). Results found that all groups had improved WOMAC pain scores, with statistically significantly less improvement at 12 months in participants receiving the 100 µg dose of sprifermin, compared with participants receiving placebo. No statistically significant relationship between treatment group and reduction in central medial femorotibial compartment cartilage thickness was observed. However, sprifermin was associated with statistically significant, dose-dependent reductions in the loss of total and lateral femorotibial cartilage thickness and volume, and in joint space widening in the lateral femorotibial compartment. The reasons for the seemingly preferential effect on the lateral knee compartment in the present and previous studies are not clear. In OA, the status of cartilage differs between the medial and lateral femorotibial compartments, with the medial compartment more commonly severely affected. An anabolic agent acting on cartilage may be less effective in tissue that is severely damaged. Currently, sprifermin is expensive and mainly available in phase II trials. No trial has investigated the benefits and safety of sprifermin in people with hip OA. Using knee OA data to extrapolate to hip or other OA requires additional caution.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Lohmander, et al. Arthritis Rheumatol. 2014 Jul; 66(7): 1820-31; 2. Dahlberg, et al. Clin Exp Rheumatol. 2016 May-Jun; 34(3): 445-50.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Fibroblast Growth Factor**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (Scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	N=122	N=41	**MD 2.14 higher (0.61 higher to 3.67 higher)**		⊕⊕□□ LOW	CRITICAL
WOMAC Function (Scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	very serious [b,c]	none	N=122	N=41	MD 4.23 higher (0.64 lower to 9.1 higher)		⊕□□□ VERY LOW	CRITICAL
Joint Space Widening, mm (medial) (Positive values indicate better structural outcome) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	very serious [b,c]	none	N=120	N=38	MD 0.03 higher (0.3 lower to 0.36 higher)		⊕□□□ VERY LOW	IMPORTANT
Joint Space Widening, mm (lateral) (Positive values indicate better structural outcome) (follow up: 12 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [b]	none	N=120	N=38	**MD 0.31 higher (0.03 higher to 0.59 higher)**		⊕⊕□□ LOW	IMPORTANT
Medial femorotibial compartment cartilage thickness (mm) (Higher values indicate better structural outcome) (follow up: 12 months)												

1	randomised trial [1]	not serious	not assessable	serious [a]	very serious [b,c]	none	N=115	N=37	MD 0.05 higher (0.03 lower to 0.12 higher)		⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Fibroblast Growth Factor) (follow-up: range 24 weeks to 12 months)												
2	randomised trials [1,2]	serious [d]	not serious	serious [a]	serious [c]	none	3/199 (1.5%)	1/66 (1.5%)	RR 1.00 (0.11 to 9.39)	0 fewer per 1,000 (from 13 fewer to 127 more)	⊕□□□ VERY LOW	CRITICAL
Treatment-emergent Adverse Events (Risk ratios less than one favor Fibroblast Growth Factor) (follow-up: range 24 weeks to 12 months)												
2	randomised trials [1,2]	serious [d]	very serious [e]	serious [a]	serious [c]	none	131/199 (65.8%)	42/66 (63.6%)	RR 0.94 (0.48 to 1.82)	38 fewer per 1,000 (from 331 fewer to 522 more)	⊕□□□ VERY LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Fibroblast Growth Factor) (follow-up: 24 weeks)												
1	randomised trial [2]	serious [d]	not assessable	serious [a]	serious [b]	none	5/55 (9.1%)	5/18 (27.8%)	**RR 0.33 (0.11 to 1.00)**	**186 fewer per 1,000 (from 0 fewer to 247 fewer)**	⊕□□□ VERY LOW	CRITICAL
Local Treatment-emergent Adverse Events (Risk ratios less than one favor Fibroblast Growth Factor) (follow-up: range 24 weeks to 12 months)												
2	randomised trials [1,2]	serious [d]	not serious	serious [a]	serious [c]	none	50/199 (25.1%)	12/66 (18.2%)	RR 1.37 (0.78 to 2.41)	67 more per 1,000 (from 40 fewer to 256 more)	⊕□□□ VERY LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. All patients in both trials have Knee Osteoarthritis.
b. Sample size in one study arm <50.
c. 95% CI crosses null
d. Dahlberg, 2016: 25% more patients in the Placebo group were Kellgren Lawrence Grade IV than in the Sprifermin groups at baseline
e. I^2=92%

PICO 2.8.5 (hip): What are the benefits and harms of colchicine in the management of patients with hip OA?

SUMMARY

There is currently a lack of high-quality evidence supporting the use of colchicine for symptomatic relief for people with knee OA. While two small trials (one comparing colchicine to placebo; one comparing the combination of colchicine and an anti-inflammatory medication to the anti-inflammatory medication alone) indicate colchicine may provide symptomatic relief, its efficacy and safety remains unproven. In the trials, participants who received colchicine reported more gastrointestinal adverse effects, and the benefit to risk profile needs to be investigated in larger studies. One randomised placebo-controlled trial for people with knee OA that commenced enrolment of 120 participants in June 2014 in Singapore is reported to have been completed (Identifier: NCT02176460; ClinicalTrials.gov), but the results have not been published. One additional trial was identified in a search of the World Health Organization's (WHO's) International Clinical Trials Registry Platform (ICTRP). This trial is reported to have recruited 81 participants between March and September 2012 in Iran, and was retrospectively registered in September 2015 (IRCT2015071623240N1). These results have also not been published. There are currently no trials investigating the benefits and safety of colchicine in people with OA of the hip. Colchicine currently does not have an indication via the Therapeutic Goods Administration (TGA) for OA, and should be considered as an investigational medication only.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Das, et al. Arthritis Rheum. 2002 Jun 15; 47(3): 280-4 [concomitant nimesulide]; 2. Das, et al. Osteoarthritis Cartilage. 2002 Apr; 10(4): 247-52 [concomitant piroxicam]; 3. Aran, et al. Clin Exp Rheumatol. 2011 May-Jun; 29(3): 513-8 [concomitant OA treatment, various]; 4. Ediz, et al. Journal of Clinical and Analytical Medicine 3, no. 1 (2012): 63-67[concomitant acetaminophen].

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Colchicine**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
VAS Pain (scale range 0 to 15 cm, with higher scores indicating higher pain severity) (follow up: 20 weeks)												
2	randomised trials [1,2]	serious [a]	very serious [b]	serious [c]	very serious [d,e]	none	N=38	N=37	MD 1.24 lower (3.27 lower to 0.79 higher)		⊕□□□ VERY LOW	CRITICAL
Modified HAQ (Quality of Life) (range unclear, unvalidated measure) (Higher scores indicate better quality of life) (follow up: 20 weeks)												
2	randomised trials [1,2]	serious [a]	not serious	serious [c]	serious [d]	none	N=38	N=37	**MD 3.13 lower (4.43 lower to 1.83 lower)**		⊕□□□ VERY LOW	IMPORTANT
Patients' Global Assessment of Disease Severity (scale range 0 to 15 cm, with higher scores indicating higher disease severity) (follow up: range 3 months to 20 weeks)												
2	randomised trials [1,3]	serious [f]	very serious [g]	serious [c]	serious [d]	none	N=48	N=46	**MD 5.76 lower (10.17 lower to 1.35 lower)**		⊕□□□ VERY LOW	IMPORTANT
Physician's Global Assessment of Disease Severity (scale range 0 to 15 cm, with higher scores indicating higher disease severity) (follow up: range 3 months to 20 weeks)												
2	randomised trials [1,3]	serious [f]	very serious [h]	serious [c]	serious [d]	none	N=48	N=46	**MD 4.77 lower (7.32 lower to 2.21 lower)**		⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Colchicine) (follow up: range 3 months to 6 months)												

4	randomised trials [1,2,3,4]	not serious	not serious	not serious	serious [e]	none	2/101 (2.0%)	0/100 (0.0%)	RR 2.89 (0.31 to 26.79)	NA [i]	⊕⊕⊕◯ MODERATE	CRITICAL
Total Adverse Events (Risk ratios less than one favor Colchicine) (follow up: 3 months)												
1	randomised trial [3]	not serious	not assessable	not serious	very serious [d,e]	none	1/30 (3.3%)	0/31 (0.0%)	RR 3.10 (0.13 to 73.16)	NA [i]	⊕⊕◯◯ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Colchicine) (follow up: 20 weeks)												
1	randomised trial [2]	not serious	not assessable	not serious	very serious [d,e]	none	0/19 (0.0%)	0/20 (0.0%)	Due to zero events in both groups, an absolute risk reduction was not estimable.		⊕⊕◯◯ LOW	CRITICAL
Gastrointestinal Adverse Events (Risk ratios less than one favor Colchicine) (follow up: range 20 weeks to 6 months)												
2	randomised trials [2,4]	serious [j]	not serious	not serious	serious [e]	none	19/52 (36.5%)	15/52 (28.8%)	RR 1.26 (0.83 to 1.93)	75 more per 1,000 (from 49 fewer to 268 more)	⊕⊕◯◯ LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Both studies received High risk of bias ratings due to potential attrition bias
b. I^2=79%
c. All patients in all studies have Knee Osteoarthritis.
d. Sample size <50 in each study arm.
e. 95% CI crosses null
f. One study received a High risk of bias rating due to potential attrition bias
g. I^2=93%
h. I^2=77%
i. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.
j. One study received High risk of bias rating due to unblinded design

PICO 2.8.6 (hip): What are the benefits and harms of methotrexate in the management of patients with hip OA?

SUMMARY

There is very low-quality evidence from one small trial of 56 participants who used 7.5 mg of methotrexate weekly versus placebo for painful knee OA, which did not find a reduction in pain at four months. Another open-label study evaluated the effects of methotrexate for pain relief in participants with knee OA. At 24 weeks, 13/30 participants (43%) achieved ≥30% reduction in Visual Analogue Scale (VAS) pain, of whom, seven (23%) had achieved ≥50% reduction. Conversely, four participants (13%) experienced a flare. Thirteen of 30 (43%) participants achieved Osteoarthritis Research Society International's responder criteria.(Wenham CY, et al. Rheumatology, 2013;52(5):888–92). An ongoing pragmatic phase III trial (ISRCTN77854383) has been designed to confirm these inconsistent findings. In terms of cost and access, methotrexate is a relatively cheap and widely available. Currently, there is no direct evidence for hip OA.Methotrexate does not currently have an indication via the TGA for OA, and should be considered as an investigational medication only.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. De Holanda, et al. Revista Brasileira de Reumatologia 47, no. 5 (2007): 334-340.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Methotrexate	Placebo	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 4 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	very serious [b,c]	none	N=28	N=28	MD 0.62 higher (1.28 lower to 2.52 higher)		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 4 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	very serious [b,c]	none	N=28	N=28	MD 3.18 lower (10.1 lower to 3.74 higher)		⊕□□□ VERY LOW	CRITICAL
Lequesne Index (scale range 0 to 24, with higher scores indicating poorer functional outcome) (follow up: 4 months)												
1	randomised trial [1]	not serious	not assessable	serious [a]	very serious [b,c]	none	N=28	N=28	MD 0.1 lower (2.43 lower to 2.23 higher)		⊕□□□ VERY LOW	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Methotrexate) (follow up: 4 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [c]	none	0/29 (0.0%)	0/29 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕□ MODERATE	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Methotrexate) (follow up: 4 months)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [c]	none	0/29 (0.0%)	0/29 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕□ MODERATE	CRITICAL

Gastrointestinal Adverse Events (Risk ratios less than one favor Methotrexate) (follow up: 4 months)												
1	randomised trial [1]	not serious	not assessable	not serious	very serious [b,c]	none	6/29 (20.7%)	6/29 (20.7%)	RR 1.00 (0.37 to 2.74)	0 fewer per 1,000 (from 130 fewer to 360 more)	⨁⨁◯◯ LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. All patients in the study have Knee Osteoarthritis.

b. 95% CI crosses null

c. Sample size <50 in each study arm.

PICO 2.9: INTRA-ARTICULAR INJECTIONS

PICO 2.9.1 (hip): What are the benefits and harms of corticosteroids in the management of patients with hip OA?

SUMMARY
For hip pain, the clinical benefits were demonstrated for up to 12 weeks; however, there is lack of long-term data. In addition, considering the complexity of the hip joint, image guidance would be required, which would further add to the costs.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Atchia, et al. Ann Rheum Dis. 2011 Jan; 70(1): 110-6; 2. Lambert, et al. Arthritis Rheum. 2007 Jul; 56(7): 2278-87; 3. Kullenberg, et al. J Rheumatol. 2004 Nov; 31(11): 2265-8; 4. Qvistgaard, et al. Osteoarthritis Cartilage. 2006 Feb; 14(2): 163-70; 5. Flanagan, et al. Ann R Coll Surg Engl. 1988 May; 70(3): 156-7.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intra-articular Steroid	Placebo	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher Scores indicate higher pain severity) (follow-up: range 3 weeks to 12 weeks)												
4	randomised trials [1,2,3,4]	not serious	very serious [a]	not serious	not serious	none	N=122	N=115	**SMD 1.37 lower (2.72 lower to 0.02 lower)**		⊕⊕□□ LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow-up: range 3 weeks to 12 weeks)												
4	randomised trials [1,2,3,4]	not serious	very serious [b]	not serious	serious [c]	none	N=122	N=115	SMD 1.42 lower (3.06 lower to 0.21 higher)		⊕□□□ VERY LOW	CRITICAL
SF-36 Composite Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow-up: 8 weeks)												
1	randomised trial [2]	not serious	not assessable	not serious	serious [d]	none	N=31	N=21	**MD 4.2 higher (0.23 higher to 8.17 higher)**		⊕⊕⊕□ MODERATE	IMPORTANT
Withdrawals due to Adverse Events (Risk ratios less than one favor Intra-articular corticosteroids) (follow-up: range 8 weeks to 12 weeks)												
2	randomised trials [2,4]	not serious	not serious	not serious	not serious	none	0/65 (0.0%)	0/57 (0.0%)	Due to zero events in both study arms, an absolute risk reduction was not estimable.		⊕⊕⊕⊕ HIGH	CRITICAL
Total Adverse Events (Risk ratios less than one favor Intra-articular corticosteroids) (follow-up: range 8 weeks to 12 weeks)												

1	randomised trial [2]	not serious	not assessable	not serious	very serious [c,d]	none	16/31 (51.6%)	11/21 (52.4%)	RR 0.99 (0.58 to 1.68)	5 fewer per 1,000 (from 220 fewer to 356 more)	⊕⊕◻◻ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Intra-articular corticosteroids) (follow up: range 8 weeks to 12 weeks)												
2	randomised trials [2,4]	not serious	not serious	not serious	serious [c]	none	1/65 (1.5%)* *deep vein thrombosis at 3 months	0/57 (0.0%)	RR 2.06 (0.09 to 48.34)	NA [e]	⊕⊕⊕◻ MODERATE	CRITICAL
Patients Experiencing Worsening of Pain after injection (Risk ratios less than one favor Intra-articular corticosteroids) (follow-up: 8 weeks)												
2	randomised trials [2,5]	not serious	not serious	not serious	very serious [c,d]	none	9/43 (20.9%)	4/33 (12.1%)	RR 2.01 (0.73 to 5.49)	122 more per 1,000 (from 33 fewer to 544 more)	⊕⊕◻◻ LOW	IMPORTANT
Local Reactions (Risk ratios less than one favor Intra-articular corticosteroids) (follow-up: 8 weeks)												
1	randomised trial [1]	not serious	not assessable	not serious	serious [d]	none	0/19 (0.0%)	0/19 (0.0%)	Due to zero events in both study arms, an effect was inestimable		⊕⊕⊕◻ MODERATE	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. I^2=95%
b. I^2=96%
c. 95% CI crosses null
d. Sample size in each study arm <50.
e. Due to zero events in the comparator arm, an absolute risk reduction was inestimable.

PICO 2.9.2 (hip): What are the benefits and harms of viscosupplementation in the management of patients with hip OA?

SUMMARY
The recommendation for hip OA is based on three small randomised controlled trials (RCTs), which were judged to not be at serious risk of bias.The overall quality of evidence was judged to low. No effect on pain nor function was demonstrated, and the risk of total and serious adverse events and local reactions was greater in the viscosupplementation group. In addition, for a hip injection,image guidance would be required, further adding to complexity and cost. The increased risk of total and serious adverse events are of concern.

OVERALL QUALITY OF EVIDENCE: ⊕⊕□□LOW

Bibliography: 1. Atchia, et al. Ann Rheum Dis. 2011 Jan; 70(1): 110-6; 2. Qvistgaard, et al. Osteoarthritis Cartilage. 2006 Feb; 14(2): 163-70; 3. Richette, et al. Arthritis Rheum. 2009 Mar; 60(3): 824-30.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Intra-articular Hyaluronic Acid	Control	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 60 days to 90 days)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	serious [a]	none	N=93	N=97	SMD 0.18 lower (0.50 lower to 0.13 higher)		⊕⊕⊕□ MODERATE	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 60 days to 90 days)												
3	randomised trials [1,2,3]	not serious	not serious	not serious	serious [a]	none	N=93	N=97	SMD 0.18 lower (0.47 lower to 0.10 higher)		⊕⊕⊕□ MODERATE	CRITICAL
Withdrawals due to Adverse Events (Risk ratios less than one favor Intra-articular Hyaluronic Acid) (follow up: 90 days)												
1	randomised trials [2]	not serious	not assessable	not serious	very serious [a,b]	none	2/34 (5.9%)	0/36 (0.0%)	RR 5.29 (0.26 to 106.27)	NA [c]	⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Intra-articular Hyaluronic Acid) (follow up: 90 days)												
1	randomised trials [3]	not serious	not assessable	not serious	very serious [a,b]	none	17/42 (40.5%)	15/43 (34.9%)	RR 1.16 (0.67 to 2.01)	56 more per 1,000 (from 115 fewer to 352 more)	⊕⊕□□ LOW	CRITICAL
Serious Adverse Events (Risk ratios less than one favor Intra-articular Hyaluronic Acid) (follow up: 90 days)												
2	randomised trials [2,3]	not serious	not serious	not serious	serious [a]	none	1/76 (1.3%)	0/79 (0.0%)	RR 3.07 (0.13 to 73.30)	NA [c]	⊕⊕⊕□ MODERATE	CRITICAL
Local Reactions (Risk ratios less than one favor Intra-articular Hyaluronic Acid) (follow up: range 60 days to 90 days)												

2	randomised trials [1,3]	not serious	not serious	not serious	serious [a]	none	9/61 (14.8%)	2/62 (3.2%)	RR 3.44 (0.86 to 13.74)	79 more per 1,000 (from 5 fewer to 411 more)	⊕⊕⊕◻ MODERATE	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

a. 95% CI crosses null.
b. Sample size in each study arm <50.
c. Due to zero events in the comparator arm, an absolute risk reduction was not estimable.

PICO 2.9.3 (hip): What are the benefits and harms of platelet-rich plasma in the management of patients with hip OA?

SUMMARY

The studies upon which the recommendation is based were at serious risk of bias and inconsistency, and were generally small in size. The overall quality of the evidence was judged to be very low. Beneficial effects on both knee pain and Western Ontario and McMaster Universities (WOMAC) function were demonstrated at six months. With the concern of potential reporting bias and low-quality data, the beneficial effects are likely to be overinflated. In addition, there is no consensus on eligible participant selection, number and frequency of injections, preparation technique, or appropriate platelet concentration,(Chang KV, et al. Arch Phys Med Rehabil, 2014;95(3):562–75) leading to large variationsin the design of PRP trials. No RCT was conducted in hip OA. However, during working group discussions, it was suggested that the mechanism of action should be no different in hip OA. Therefore, the findings might be transferrable to hip OA, but with a particular caution in terms of the complexity of the hip joint. The cost of PRP treatment is high, and additional equipment might be required for the preparation and administration.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Patel, et al. Am J Sports Med. 2013 Feb; 41(2): 356-64; 2. Rayegani, et al. Orthop Rev (Pavia). 2014 Sep 18; 6(3): 5405; 3. Smith, Patrick A. Am J Sports Med. 2016 Apr; 44(4): 884-91 4. Görmeli, et al. Knee Surg Sports Traumatol Arthrosc. 2017 Mar;25(3):958-965.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Platelet-rich Plasma**	**Placebo**	**Relative (95% CI)**	**Absolute (95% CI)**		
Pain (Higher scores indicate higher pain severity) (follow-up: 6 months)												
4	randomised trials [1,2,3,4]	serious [a]	very serious [b]	serious [c]	not serious	none	N=231	N=132	**SMD 1.87 lower (2.47 lower to 1.27 lower)**		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow-up: 6 months)												
3	randomised trials [1,2,3]	serious [a]	very serious [d]	serious [c]	not serious	none	N=148	N=92	**MD 16.07 lower (22.76 lower to 9.37 lower)**		⊕□□□ VERY LOW	CRITICAL
SF-36 Physical Component Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow-up: 6 months)												
1	randomised trial [2]	serious [a]	not assessable	serious [c]	serious [e]	none	N=31	N=31	**MD 9.54 higher (0.22 higher to 18.86 higher)**		⊕□□□ VERY LOW	IMPORTANT
SF-36 Mental Component Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow-up: 6 months)												
1	randomised trial [2]	serious [a]	not assessable	serious [c]	very serious [e,f]	none	N=31	N=31	MD 10.17 higher (0.67 lower to 21.01 higher)		⊕□□□ VERY LOW	IMPORTANT
Treatment-related Adverse Events (Risk Ratios less than one favor Platelet-rich Plasma) (follow-up: range 6 months to 12 months)												

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Platelet-rich Plasma	Placebo	Relative (95% CI)	Absolute (95% CI)		
2	randomised trials [1,3]	not serious	not serious	serious [c]	serious [f]	none	17/117 (14.5%)	1/61 (1.6%)	RR 3.06 (0.39 to 23.87)	34 more per 1,000 (from 10 fewer to 375 more)	⊕⊕◻◻ LOW	CRITICAL

CI: Confidence interval; **SMD:** Standardised mean difference; **MD:** Mean difference; **RR:** Risk ratio

a. 2 studies received at least one High risk of bias rating; 1 due to unblinded design and 1 due to potential reporting bias
b. I^2=81%
c. Patients in all three studies have Knee Osteoarthritis; there were no available placebo-controlled RCTs in which patients with Hip Osteoarthritis receive PRP.
d. I^2=77%
e. Sample size in each study arm <50.
f. 95% CI crosses null

PICO 2.9.4 (hip): What are the benefits and harms of stem cell therapy in the management of patients with hip OA?

SUMMARY
The two studies upon which the recommendation is based were at very serious risk of bias and were small in size. The overall quality of the evidence was judged to be low to very low. Beneficial effects on pain and function were demonstrated at up to six months. The between-group differences reported for pain and function appeared to be remarkably good. As they deviate significantly from those of other successful interventions, replication is required in high quality, large RCTs before a more favourable recommendation can be considered. onsistent with a recent position statement from the Australian College of Sports and Exercise Physicians, stem cell administration should be part of a rigorously designed study and the priority for individual health and welfare. (Osborne H, et al. Br J Sports Med 2016;50(20):1237–44).

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1. Varma HS, et al. J Indian Med Assoc 2010;108:583–58; 2. Tan, et al. J Tradit Chin Orthop Traumatol 2013;10:35–8.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**IA Mesenchymal Stem Cells**	**Control**	**Relative (95% CI)**	**Absolute (95% CI)**		
VAS Pain (scale range 0 to1 0, with higher scores indicating higher pain severity) (follow up: 6 months)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [c]	none	N=25	N=25	**MD 3.4 lower (3.94 lower to 2.86 lower)**		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 6 months to 12 months)												
2	randomised trials [1,2]	serious [a]	very serious [d]	serious [b]	not serious	none	N=61	N=61	**SMD 5.05 lower (7.01 lower to 3.1 lower)**		⊕□□□ VERY LOW	CRITICAL
Change in Cartilage Thickness [mm] (Higher values indicate better structural outcome) (follow up: 12 months)												
1	randomised trial [2]	serious [a]	not assessable	serious [b]	serious [c]	none	N=36	N=36	**MD 0.6 higher (0.54 higher to 0.66 higher)**		⊕□□□ VERY LOW	IMPORTANT

CI: Confidence interval; **MD:** Mean difference; **SMD:** Standardised mean difference

a. Both trials received High risk of bias ratings, and/or "Low" quality ratings. Adverse Events are not reported for either study.
b. All patients in both trials have Knee Osteoarthritis
c. Sample size <50 in each study arm.
d. I^2= 83%

PICO 2.9.5 (hip): What are the benefits and harms of dextrose prolotherapy in the management of patients with hip OA?

SUMMARY
The recommendation is based on the evidence of only one small RCT of low quality. The risk of bias in this study was not determined to be serious. No clinically significant effects were found for pain at 24 and 52 weeks follow-up. In terms of function, no clinically significant effects were found for pain at 24 weeks, but a marginally significant effect was recorded at 52 weeks. Furthermore, high-quality RCTs with low risk of bias and specifically for hip OA are required. As prolotherapy is relatively cheap and accessible, it is likely to be injudiciously used.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

Bibliography: 1.Rabago, et al. Ann Fam Med. 2013 May-Jun; 11(3): 229-37.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Dextrose Prolotherapy	Placebo	Relative (95% CI)	Absolute (95% CI)		
WOMAC Pain (Scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 24 weeks)												
1	randomised trial [1]	not serious	not assessable	serious [a]	very serious [b,c]	none	N=28	N=25	MD 9.1 lower (19.07 lower to 0.87 higher)		⊕□□□ VERY LOW	CRITICAL
WOMAC Pain (Scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 52 weeks)												
1	randomised trial [1]	not serious	not assessable	serious [a]	very serious [b,c]	none	N=26	N=25	MD 6.8 lower (16.9 lower to 3.3 higher)		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (Scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 24 weeks)												
1	randomised trial [1]	not serious	not assessable	serious [a]	very serious [b,c]	none	N=28	N=25	MD 9.57 lower (19.43 lower to 0.29 higher)		⊕□□□ VERY LOW	CRITICAL
WOMAC Function (Scale range 0 to 100, with higher scores indicating higher pain severity) (follow up: 52 weeks)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [c]	none	N=26	N=25	**MD 10.79 lower (20.26 lower to 1.32 lower)**		⊕⊕□□ LOW	CRITICAL
Total Adverse Events (Risk ratios less than one favor Dextrose Prolotherapy) (follow up: 52 weeks)												
1	randomised trial [1]	not serious	not assessable	serious [a]	serious [c]	none	0/30 (0.0%)	0/29 (0.0%)	Due to zero events in both study arms, an effect was not estimable.		⊕⊕□□ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. All patients have Knee Osteoarthritis.
b. 95% CI crosses null

c. Sample size <50 in each study arm. (Authors noted the following for adverse event collection: "The study was not large enough to detect uncommon adverse events, such as intolerance to study medication or rare injection-related sequelae")

GRADE tables for hip osteoarthritis

Section 3: Surgical Interventions

PICO 3.1 (hip): What are the benefits and harms of arthroscopic lavage and debridement interventions in the management of patients with hip OA?

No RCT data was found which related to this question.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

PICO 3.2 (hip): What are the benefits and harms of arthroscopic procedures for cartilage repair interventions in the management of patients with hip OA?

No RCT data was found which related to this question.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

GRADE tables for hip osteoarthritis

Section 4: Combination Therapies

PICO 4.1 (hip): What are the benefits and harms of combination weight management and exercise interventions compared to mono-therapy in patients with hip OA?

SUMMARY
There is limited evidence of very low quality that weight loss alone (achieved through diet and exercise) has no significant effect on either pain or function in people with knee OA although benefits appear to be more significant with higher amounts of weight loss – starting at a minimum of 5–7.5% body weight loss. Ideally, it is beneficial to achieve a greater amount of weight loss.(Atukorala I, et al. Arthritis Care Res 2016;68(8):1106–14; Messier SP, et al. JAMA, 2013;310(12):1263–73). There is no randomised controlled trial (RCT) on the effects of bariatric surgery in people with hip and/or knee OA. There are no RCTs investigating weight management specifically in people with hip OA, although the numerous other systemic health benefits of weight loss and maintaining a healthy body weight are most likely transferable to people with hip OA. There are limitations of the available RCT evidence in OA. Clinicians are advised to refer to the National Health and Medical Research Council's (NHMRC's) guidelines for the most effective strategies for managing overweight/obesity in primary care. National Health and Medical Research Council. Clinical practice guidelines for the management of overweight and obesity in adults, adolescents and children in Australia. Melbourne: NHMRC, 2013).

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

WEIGHT MANAGEMENT+EXERCISE VS. EXERCISE ALONE

Bibliography: 1. Messier, et al. Arthritis Rheum. 2004 May; 50(5): 1501-10 (ADAPT); 2. Messier, et al. JAMA. 2013 Sep 25; 310(12): 1263-73 (IDEA); 3. Focht, et al. Arthritis Rheum. 2005 Oct 15; 53(5): 659-65 (ADAPT).

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Weight Management + Exercise**	**Exercise Alone**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 18 months)												
2	randomised trials [1,2]	serious [a]	not serious	serious [b]	not serious	none	N=205	N=208	**MD 1.43 lower (2.09 lower to 0.77 lower)**		⊕⊕□□ LOW	CRITICAL
WOMAC Function (scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 18 months)												
2	randomised trials [1,2]	serious [a]	not serious	serious [b]	not serious	none	N=205	N=208	**MD 4.21 lower (6.46 lower to 1.95 lower)**		⊕⊕□□ LOW	CRITICAL
Walking Self-Efficacy (Patient confidence in walking around a gymnasium twice without stopping; score range 0-100 with higher scores indicating more confidence) (follow up: 18 months)												
1	randomised trial [3]	serious [a]	not assessable	serious [b]	serious [c]	none	N=76	N=80	MD 3.97 higher (5.52 lower to 13.46 higher)		⊕□□□ VERY LOW	IMPORTANT
SF-36 Physical Component Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 18 months)												
1	randomised trial [2]	serious [a]	not assessable	serious [b]	not serious	none	N=129	N=128	**MD 3 higher (0.76 higher to 5.24 higher)**		⊕⊕□□ LOW	IMPORTANT

Percentage Weight Loss (Percentage of weight at baseline lost by follow up time, with more loss indicating positive outcome) (follow up: 18 months)											
2	randomised trials [1,2]	serious [a]	very serious [d]	serious [b]	serious [c]	none	N=197	N=195	MD 5.42% more lost (1.86% less lost to 12.69% more lost)	⊕□□□ VERY LOW	IMPORTANT
Lateral Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: 18 months)											
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [c]	none	N=76	N=80	MD 0.16 mm lower (0.66 mm lower to 0.34 mm higher)	⊕□□□ VERY LOW	IMPORTANT
Medial Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: 18 months)											
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [c]	none	N=76	N=80	MD 0.11 mm lower (0.68 mm lower to 0.46 mm higher)	⊕□□□ VERY LOW	IMPORTANT
Non-Compliance with Regimen [defined as "non-adherence"] (Risk ratios less than one favor Weight Management + Exercise) (follow up: 18 months)											
2	randomised trials [1,2]	serious [a]	not serious	serious [b]	serious [c]	none	91/228 (39.9%)	101/230 (43.9%)	RR 0.91 (0.73 to 1.13)	40 fewer per 1,000 (from 57 more to 119 fewer)	⊕□□□ VERY LOW	IMPORTANT
Withdrawal due to Lack of Interest (Risk ratios less than one favor Weight Management + Exercise) (follow up: 18 months)											
1	randomised trial [2]	serious [a]	not assessable	serious [b]	serious [c]	none	2/152 (1.3%)	2/150 (1.3%)	RR 0.99 (0.14 to 6.92)	0 fewer per 1,000 (from 11 fewer to 79 more)	⊕□□□ VERY LOW	IMPORTANT
Serious Adverse Events (Risk ratios less than one favor Weight Management + Exercise) (follow up: 18 months)											
1	randomised trial [2]	serious [a]	not assessable	serious [b]	serious [c]	none	6/152 (3.9%)	3/150 (2.0%)	RR 1.97 (0.50 to 7.75)	19 more per 1,000 (from 10 fewer to 135 more)	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Both studies received High risk of bias ratings due to single blinded design, and due to potential for attrition bias.
b. All patients in all trials have Knee Osteoarthritis.
c. 95% CI crosses null.
d. I^2= 79%

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

WEIGHT MANAGEMENT+EXERCISE VS. WEIGHT MANAGEMENT ALONE

Bibliography: 1. Messier, et al. Arthritis Rheum. 2004 May; 50(5): 1501-10 (ADAPT); 2. Messier, et al. JAMA. 2013 Sep 25; 310(12): 1263-73 (IDEA); 3. Focht, et al. Arthritis Rheum. 2005 Oct 15; 53(5): 659-65 (ADAPT).

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Weight management + Exercise**	**Weight management alone**	**Relative (95% CI)**	**Absolute (95% CI)**		
WOMAC Pain (scale range 0 to 20, with higher scores indicating higher pain severity) (follow up: 18 months)												
2	randomised trials [1,2]	serious [a]	not serious	serious [b]	not serious	none	N=205	N=206	**MD 1.18 lower (1.85 lower to 0.51 lower)**		⊕⊕□□ LOW	CRITICAL
WOMAC Function (scale range 0 to 68, with higher scores indicating poorer functional outcome) (follow up: 18 months)												
2	randomised trials [1,2]	serious [a]	not serious	serious [b]	not serious	none	N=205	N=206	**MD 2.7 lower (4.99 lower to 0.4 lower)**		⊕⊕□□ LOW	CRITICAL
Walking Self-Efficacy (Patient confidence in walking around a gymnasium twice without stopping; score range 0-100 with higher scores indicating more confidence) (follow up: 18 months)												
1	randomised trial [3]	serious [a]	not assessable	serious [b]	not serious	none	N=76	N=82	**MD 13.4 higher (3.84 higher to 22.96 higher)**		⊕⊕□□ LOW	IMPORTANT
SF-36 Physical Component Score (scale range 0 to 100, with higher scores indicating better quality of life) (follow up: 18 months)												
1	randomised trial [2]	serious [a]	not assessable	serious [b]	serious [c]	none	N=129	N=124	MD 2 higher (0.27 lower to 4.27 higher)		⊕□□□ VERY LOW	IMPORTANT
Percentage Weight Loss (Percentage of weight at baseline lost by follow up time, with more loss indicating positive outcome) (follow up: 18 months)												
2	randomised trials [1,2]	serious [a]	not serious	serious [b]	serious [c]	none	N=197	N=202	MD 1.14% more lost (1.88% less lost to 4.17% more lost)		⊕□□□ VERY LOW	IMPORTANT
Lateral Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: 18 months)												
1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [c]	none	N=76	N=82	MD 0.16 mm lower (0.65 mm lower to 0.33 mm higher)		⊕□□□ VERY LOW	IMPORTANT
Medial Joint Space Narrowing [mm] (Positive values indicate better structural outcome) (follow up: 18 months)												

1	randomised trial [1]	serious [a]	not assessable	serious [b]	serious [c]	none	N=76	N=82	MD 0.1 mm lower (0.62 mm lower to 0.42 mm higher)		⊕□□□ VERY LOW	IMPORTANT
Non-Compliance with Regimen [defined as "non-adherence"] (Risk ratios less than one favor Weight Management + Exercise) (follow up: 18 months)												
2	randomised trials [1,2]	serious [a]	not serious	serious [b]	serious [c]	none	83/228 (36.4%)	82/234 (35.0%)	RR 1.03 (0.80 to 1.34)	11 more per 1,000 (from 70 fewer to 119 more)	⊕□□□ VERY LOW	IMPORTANT
Withdrawal due to Lack of Interest (Risk ratios less than one favor Weight Management + Exercise) (follow up: 18 months)												
1	randomised trial [2]	serious [a]	not assessable	serious [b]	serious [c]	none	2/152 (1.3%)	3/152 (2.0%)	RR 0.67 (0.11 to 3.93)	7 fewer per 1,000 (from 18 fewer to 58 more)	⊕□□□ VERY LOW	IMPORTANT
Serious Adverse Events (Risk ratios less than one favor Weight Management + Exercise) (follow up: 18 months)												
1	randomised trial [2]	serious [a]	not assessable	serious [b]	serious [c]	none	6/152 (3.9%)	1/152 (0.7%)	RR 6.00 (0.73 to 49.24)	33 more per 1,000 (from 2 fewer to 317 more)	⊕□□□ VERY LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

a. Both studies received High risk of bias ratings due to single blinded design, and due to potential for attrition bias.
b. All patients in all trials have Knee Osteoarthritis.
c. 95% CI crosses null.

PICO 4.2 (hip): What are the benefits and harms of combination exercise and cognitive behavioural interventions compared to mono-therapy in patients with hip OA?

SUMMARY
Low-quality evidence in people with knee OA shows that CBT programs can have small benefits for pain, and can also improve self-efficacy, pain coping, depression and anxiety (O'Moore KA, et al. 2017;70(1):61–70). CBT for knee OA is associated with a low risk of adverse events. Very low-quality evidence from a limited number of studies indicates that combining CBT with exercise is more effective than either alone. While there is no evidence of the effects of CBT, specifically in people with hip OA. Benefits seen in people with knee OA and mixed samples of hip/knee OA are likely to be generalisable to those with hip OA.

OVERALL QUALITY OF EVIDENCE: ⊕□□□VERY LOW

COGNITIVE BEHAVIORAL THERAPY + EXERCISE VS. EXERCISE ALONE

Bibliography: 1. Bennell, et al. Arthritis Care Res (Hoboken). 2016 May; 68(5): 590-602; 2. Somers, et al. Pain. 2012 Jun; 153(6): 1199-209.

Quality assessment							№ of events/№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cognitive Behavioral Therapy + Exercise†	Exercise Alone	Relative (95% CI)	Absolute (95% CI)		
Pain (Higher scores indicate higher pain severity) (follow up: range 12 weeks to 24 weeks)												
2	randomised trials [1,2]	very serious [a]	serious [b]	serious [c]	not serious	none	N=130	N=126	**SMD 0.52 lower (1 lower to 0.03 lower)**		⊕□□□ VERY LOW	CRITICAL
Function (Higher scores indicate poorer functional outcome) (follow up: range 12 weeks to 24 weeks)												
2	randomised trials [1,2]	very serious [a]	not serious	serious [c]	not serious	none	N=130	N=126	**SMD 0.43 lower (0.68 lower to 0.18 lower)**		⊕□□□ VERY LOW	CRITICAL
Self-Efficacy (Higher scores indicate higher self-efficacy) (follow up: range 12 weeks to 24 weeks)												
2	randomised trials [1,2]	very serious [a]	not serious	serious [c]	not serious	none	N=130	N=126	**SMD 0.28 higher (0.03 higher to 0.53 higher)**		⊕□□□ VERY LOW	IMPORTANT
Depression (Higher scores indicate more severe depression) (follow up: range 12 weeks to 24 weeks)												

2	randomised trials [1,2]	very serious [a]	not serious	serious [c]	serious [d]	none	N=130	N=126	SMD 0.04 lower (0.29 lower to 0.21 higher)		⊕□□□ VERY LOW	IMPORTANT
Treatment-related Adverse Events (Risk ratios less than one favor Cognitive Behavioral Therapy + Exercise) (follow up: range 12 weeks to 24 weeks)												
2	randomised trials [1,2]	very serious [a]	not serious	serious [c]	serious [d]	none	25/135 (18.5%)	28/134 (20.9%)	RR 0.90 (0.58 to 1.39)	21 fewer per 1,000 (from 81 more to 88 fewer)	⊕□□□ VERY LOW	CRITICAL
No Participation due to Lack of Interest [defined as withdrawal due to "no response" or "dissatisfaction" or "no longer interested"] (Risk ratios less than one favor Cognitive Behavioral Therapy + Exercise) (follow up: range 24 weeks to 52 weeks)												
2	randomised trials [1,2]	very serious [a]	serious [e]	serious [c]	serious [d]	none	4/135 (3.0%)	5/134 (3.7%)	RR 1.02 (0.08 to 12.39)	1 more per 1,000 (from 34 fewer to 425 more)	⊕□□□ VERY LOW	IMPORTANT

CI: Confidence interval; **SMD:** Standardised mean difference; **RR:** Risk ratio

† **The following Cognitive Behavioral Therapeutic Technique was included in the analysis:** Bennell 2016- Pain coping skills training + Strengthening exercise vs. Strengthening exercise; Somers 2012- Pain coping skills training + Diet and Exercise vs. Diet and Exercise

a. All studies received High risk of bias ratings due to single blind study design or inadequate description of blinding and potential attrition bias. Bennell 2016 received an additional High risk of bias for reporting bias due to errors in reporting.
b. I^2= 74%; moderate heterogeneity.
c. All patients in all trials have Knee Osteoarthritis.
d. 95% CI crosses null.
e. I^2= 56%; moderate heterogeneity.

Serious Adverse Event Descriptions	
Non-Pharmacologic Interventions	
1.4 Exercise	
PICO 1.5.1 Muscle Strengthening—Knee	
Ettinger, 1997	Six serious adverse events (death or injury possibly related to participation in the trial and requiring medical attention) occurred during the study. Of the 6 adverse events, 5 occurred in persons participating in the exercise intervention: 2 people in the aerobic group fell while walking (1 fall resulted in a fracture of the distal radius); 2 people fell during participation in the weight-training program; and 1 dropped a dumbbell on her foot resulting in a fracture. One person in the health education group had sudden death while walking from her car to an intervention session
PICO 1.5.2 Walking—Knee	
Ettinger, 1997	Six serious adverse events (death or injury possibly related to participation in the trial and requiring medical attention) occurred during the study. Of the 6 adverse events, 5 occurred in persons participating in the exercise intervention: 2 people in the aerobic group fell while walking (1 fall resulted in a fracture of the distal radius); 2 people fell during participation in the weight-training program; and 1 dropped a dumbbell on her foot resulting in a fracture. One person in the health education group had sudden death while walking from her car to an intervention session
PICO 1.5.4 Tai Chi- Knee/Hip	
Wang, 2009	"One participant in each assignment group reported newly diagnosed cancer (1 breast cancer, 1 colon cancer) during the 12-week intervention period"
PICO 1.5.5 Yoga—Knee	
Cheung, 2017	None
PICO 1.7.2 Manual Therapy (manipulation & mobilization)—Hip	
Abbott, 2013	"We detected no trial-related serious adverse events."
PICO 1.8 Weight Management—Knee/Hip	
Miller, 2006	"There were no adverse, or at least no serious adverse, events attributed to the weight loss intervention during our study."
Morgan, 2009	"There were three serious adverse events over the study period. One subject in the flavocoxid group was hospitalized because of a broken pelvis suffered in a fall. One subject in the placebo group was hospitalized with interstitial pneumonitis after completing the protocol. Another subject in the placebo group was hospitalized with chest pain after completion of the protocol. All of these serious events were deemed as unrelated to the study protocol or to the treatment."
PICO 1.10.3 Realigning Patellofemoral Braces—Knee	
Callaghan, 2015	"One subject had a serious adverse event, bilateral leg swelling, which was felt to be unrelated to treatment (the brace was used on one knee). No other adverse events were reported."
PICO 1.14 Assistive Walking Device—Hip	
Jones, 2012	None.
PICO 1.15.1 Electromagnetic Therapy—Knee	
Thamsborg, 2005	"There were no serious adverse effects"
PICO 1.19 Acupuncture—Knee	
Berman, 2004	"Serious Adverse Events" per Table 4 after 26 weeks. In the paragraph about safety reporting, the events are simply described as "adverse events" "True" Acupuncture: heart disease (N=1), cancer (N=2), "non-study-related injuries" (N=3), "non-arthritis-related surgery" (N=6), stroke (N=1), pneumonia (N=1)

	"Sham" Acupuncture: "non-study-related injuries" (N=1), "exacerbation of knee pain" (N=1), "non-arthritis-related surgery" (N=3)
Scharf, 2006	26 week follow up: "Fifty serious adverse events were reported (23 in 20 patients in the TCA group, 9 in 9 patients in the sham acupuncture group, and 18 in 16 patients in the conservative therapy group). Hematoma was reported more often in the TCA and sham acupuncture groups than in the conservative therapy group." "One 83-year-old woman in the TCA group died of a myocardial infarction during the study. The investigator saw no causal relationship between her death and the treatment." "Other notable serious adverse events were syncope and stroke in 1 patient each in the TCA group, myocardial infarction in 1 patient in the sham acupuncture group; and renal failure, melena, and deep venous thrombosis in 1 patient each in the conservative therapy group."
Witt, 2005	"One patient from the minimal acupuncture group died from myocardial infarction. All cases were admitted to hospital and regarded as unrelated to the study condition or the intervention."
Pharmacologic Interventions	
PICO 2.1.1 Paracetamol—Knee/Hip	
Miceli-Richard, 2004 (Hip/Knee)	Specific SAEs were not reported. None of them were attributable to the treatment.
Pincus, 2004a (Hip/Knee)	"…eight adverse events were classified as serious because they required admission to hospital: one in the celecoxib group—intestinal obstruction and neuropathy; three in the acetaminophen group—one case of anxiety, one of cholelothiasis, and one of cholecystitis; and four in the placebo group—one patient with raised liver function tests, one with urinary tract malformation and rectal disorder, one with accidental fracture, and one with sepsis. Two of the events, the intestinal obstruction that occurred with the patient taking celecoxib and increased liver enzymes, which occurred while a patient was taking placebo, were regarded by the investigators as potentially related to the study drug. The other serious adverse events were regarded as probably not related to the study drug."
Pincus, 2004b (Hip/Knee)	"…four adverse events were classified as serious, because they required admission to hospital: two in the celecoxib group—one case of cholelithiasis and one case of instable angina; one in the acetaminophen group- chest pain, probably musculoskeletal in origin, and one in the placebo group- angina pectoris. All events were considered unrelated to study drug by the investigators."
Herrero-Beaumont, 2007 (Hip/Knee)	"There were 5 serious adverse events in the placebo group (precordial chest pain, apnea, pneumonia, elective surgery, and lumbar pain), 5 in the acetaminophen group (atrial flutter, carpal tunnel syndrome, vertebral fracture, meniscus rupture, and crush injury), and 2 in the glucosamine sulfate group (meniscus rupture, and elective surgery)."
Altman, 2007 (Hip/Knee)	Specific SAEs were not reported. None of them were attributable to the treatment.
Prior, 2014 (Hip/Knee)	"Serious adverse events were reported for eight patients treated with acetaminophen ER and two patients treated with placebo; all serious adverse events were considered to be not related or unlikely to be related to study medication. Serious adverse event in placebo-treated patients included asthma and pneumonia, and atrial fibrillation and congestive heart failure, and in acetaminophen-treated patients included chest pain and hypertension; lower limb fracture; dehydration and urosepsis; atrial fibrillation failure, chest pain, and dyspnea; and angina pectoris."
Williams, 2014 (Hip/Knee)	Hospitalization due to SAEs in the "as-needed" treatment group included acute pancreatitis, severe back pain (discharged with analgesia and physiotherapy), and chest infection, in the "regular" treatment group included asthma attack, abdominal pain (adhesions and ruptured ovaries), Crohn's disease, endometriosis, and severe back pain (prolapsed disc), and in the "placebo" treatment group included cellulitis, mental health reasons, bleeding bowel, pseudo gout, emergency hernia, and metastatic bony disease with a compression fracture of L1.
PICO 2.1.2 Oral NSAIDs—Knee	
Boswell, 2008	"Nonfatal serious adverse events occurred in 14 patients receiving GW406381, 1 patient receiving celecoxib, and 2 patients receiving placebo. Of these events, 5 (cardiac arrest, advanced heart block, congestive cardiac failure, arrhythmia, and the sick sinus syndrome) were reported in 2

	patients receiving GW406381 and were considered by investigators as possibly related to use of the study medication."
Conaghan, 2013	"There were no serious treatment-related AEs reported in any group and there were no deaths."
Fleischmann, 2005	"Occurrences of serious GI events (e.g. perforations, obstructions and bleeding) were monitored, and an independent, blinded GI safety committee." "Lumiracoxib was well tolerated in this study. There were no deaths and only six patients reported SAEs that were suspected to be related to study drug"
Hochberg, 2011A	"The incidence of serious AEs (SAEs) was low and similar between active treatment groups in both studies (Table 4). There was only one possibly treatment-related serious AE (SAE): an anaphylactic reaction experienced by a patient receiving celecoxib in Study 307 who had a history of allergy to certain drugs, including hydrocodone."
Hochberg, 2011B	"There were no treatment-related SAEs in Study 309 and no deaths in either study."
Lehmann, 2005	"One death occurred during the study in the lumiracoxib 100 mg od group. The patient did not display any cardiac or other symptoms prior to death. The case was forwarded to the CV Safety Committee where it was adjudicated as being a probably CV death, but it was not considered by the investigator to be related to study medication."
Schnitzer, 2010	"There were four GI SAEs (two GI hemorrhages in the naproxcinod 750 mg group, one ischemic colitis in the naproxen group, and one colitis in the placebo group); all were considered treatment-related."
Schnitzer, 2011	"Three patients experienced GI SAEs (2 of whom were taking concomitant low-dose aspirin), 2 patients with naproxcinod 750 mg bid (1 event of severe upper GI hemorrhage and 1 of lower GI hemorrhage), and 1 patient with naproxen 500 mg bid (severe lower GI hemorrhage)." "Two patients in the placebo group and 1 patient on active treatment experienced a CV SAE (PT: arteriosclerosis coronary artery in the naproxcinod 750 mg bid group), and this event was not considered to be treatment-related." "SAEs occurred most frequently in the GI, cardiac disorders, and infection and infestation disorders SOCs"
Sheldon, 2005	"There was only 1 death (myocardial infarction in a patient with a history of coronary heart disease, hypertension, and hypercholesterolemia), which occurred in the group that received lumiracoxib 100 mg QD with loading dose and was not considered by the investigator to be related to study medication."
Tannenbaum, 2004	"There were no deaths during the study and the incidence of SAEs was similar in all active treatment groups and the placebo group (2.5% of patients receiving lumiracoxib 200 mg od, 2.9% receiving lumiracoxib 400 mg od, 2.9% receiving celecoxib 200 mg od, and 3.3% receiving placebo; table 4)."
PICO 2.1.2 Oral NSAIDs—Hip	
Baerwald, 2010	Naproxcinod and Naproxen groups are combined "One death occurred during the study: a patient in the naproxen group died following surgery on an aortic intramural hematoma. This event was not considered by the investigator to be related to the study treatment." "Serious GI adverse events were reported in no patients receiving active treatment and in 2 patients treated with placebo (inguinalinguinal hernia and colitis)." "Serious CV adverse events were reported in 4 patients in total (3 in the naproxen group [worsening of hypertension, hypotension,and the above-mentioned death] and 1 patient in the placebo group [obstructive coronary disease])"
Schnitzer, 2011	Lumiracoxib and Celecoxib groups are combined "Two deaths were reported, both in the celecoxib group: one sudden death (CV/cerebrovascular adjudication, probably CV death) and one death resulting from atherosclerotic CV diseasewhich occurred more than 30 days after the last known dose of study medication. Neither of the deaths was considered to be related to the study treatment by the investigators. One other event was adjudicated by the CV safety committee, a lacunar infarction in the placebo group that was confirmed as an ischemic stroke and not considered related to the study treatment. There were three prespecified CV events excluding chest pain (celecoxib [n=1], angina pectoris; placebo [n=2], angina pectoris and syncope). Three GI events were adjudicated by the GI safety committee, one in the celecoxib 200-mg o.d. group (hematemesis from a possible nonulcerative upper GI complication) and two in the placebo group (hematochezia [n=2])."
PICO 2.1.3 Oral Opioids—Knee	
Afilalo, 2010	"During double-blind treatment (titration and/or maintenance periods) and within 30 days of the last dose of study medication, 20 patients

	experienced serious AEs (placebo, n = 6 [1.8%]; tapentadol ER, n = 4 [1.2%]; oxycodone CR, n = 10 [2.9%]). One patient died because of a myocardial infarction 90 days after receiving the first dose of oxycodone CR; this patient had a history of morbid obesity, and the death was considered by the investigator to be unrelated to study medication"
Fleischmann, 2001	"There were no serious adverse events among patients treated with tramadol. Placebo patients experienced 2 serious adverse events: marked arthralgia requiring knee replacement and back pain requiring hospitalization."
NCT00486811	**13/15 patients who experienced an SAE while on an opioid were taking Oxycodone CR** Tapentadol SAEs: 2 patients experienced Upper Abdominal Pain (N=1), Constipation (N=1), Diarrhea (N=1), Vomiting (N=1), and Syncope (N=1) Oxycodone SAEs: 13 patients experienced Atrial Fibrillation (N=3), Tachycardia Paroxysmal (N=1), Ventricular Arrhythmia (N=1), Vertigo (N=1), Inappropriate Antidiuretic Hormone Secretion (N=1), Constipation (N=2), Colonic Polyp (N=1), Nausea (N=1), Lower Respiratory Tract Infection Viral (N=1), Neck injury (N=1), Rectal Cancer (N=1), Dizziness (N=1), Dyspnea (N=1), and Interstitial Lung Disease (N=1) Placebo SAEs: 4 patients experienced Atrial Fibrillation (N=1), Foot Fracture (N=1), Tumor Hemorrhage (N=1), and Uterine Leimyoma (N=1),
PICO 2.1.3 Oral Opioids—Hip	
Friedmann, 2011	"Serious AEs were experienced by 11 patients and included four patients during the open-label titration period and seven patients following randomization (placebo, n 2; Remoxy, n 5). Only two of these serious AEs (orthostatic hypotension [open-label titrationperiod] and fecaloma [double-blind treatment period]) were considered related to treatment with Remoxy. There were no deaths during the study."
Gana, 2006	Serious adverse events reported by more than one subject receiving tramadol ER which were not considered related to study treatment included cholelithiasis (0.4%), chest pain (0.4%), and pancreatitis (0.2%). One serious adverse event- drug withdrawal syndrome in a subject in the tramadol ER 400 mg group was considered to be related to the study treatment.
Katz, 2010	Morphine/Naltrexone: pancreatitis and renal cell carcinoma, malignant lung neoplasm, cholelithiasis, intestinal blockage, viral gastroenteritis, and basal cell carcinoma Placebo: chest pain, abdominal pain, and transient ischemic attack). Only 1 SAE (abdominal pain in a patient taking placebo) was considered by the investigator to be treatment related.
NCT00980798	Hydromorphone HCl SAEs: 4 patients experienced Dyspepsia (N=2), Upper Abdominal Pain (N=1), Diarrhea (N=1), Nausea (N=1), Asthenia (N=1), Skin laceration (N=1), Cerebrovascular accident (N=1), and Hypertensive crisis (N=1) Placebo SAEs: 7 patients experienced Atrial fibrillation (N=1), Myocardial infarction (N=1), Supraventricular tachycardia (N=1), Acute pyelonephritis (N=1), Road traffic accident (N=1), Hyperglycemia (N=1), Cerebrovascular accident (N=1), Allergic dermatitis (N=1), Cardioversion (N=1)
Rauck, 2013	"None of the SAEs for patients who received placebo was considered treatment-related; 2/319 (0.63%) patients, who received OROS hydromorphone ER 8 mg and 4/330 (1.2%), who received 16 mg experienced treatment-related SAEs... No deaths occurred during the study or within 30 days of termination of study drug"
2.2 Topical Analgesics	
PICO 2.2.1 Topical NSAIDs—Knee/Hip	
Baraf, 2010	"No deaths occurred. Six serious AEs occurred, but none were considered related to the study drug."
Barthel, 2009	"No serious AEs were suspected of being related to the study drug. Four patients in the DSG (diclofenac) group and 1 in the vehicle group experienced cardiovascular AEs. A 76-year old man in the DSG group with multiple pre-existing medical problems died of atrial fibrillation. None of the cardiovascular events, including the death, were considered related to treatment."
Conaghan, 2013	"There were no serious treatment-related AEs reported in any group and there were no deaths"
Roth, 2004	None.
Simon, 2009	"There was no serious adverse event in the TDiclo arm....one in DMSO vehicle (acute enteritis)"
PICO 2.2.2 Transdermal Opioids—Knee/Hip	
Langford, 2006 (Hip/Knee)	One patient receiving transdermal fentanyl (TDF) had moderate dyspnea, pruritus, and rash, which were considered to be related to the trial medication, whereas all of the other serious adverse events (heart disorder, hepatitis, duodenal ulcer, urinary tract infection, transient ischemic attack, and prolongation of hospitalization due to arthritis) were considered unrelated to treatment with TDF. One patient in the placebo group died of hepatocellular carcinoma, and another patient receiving placebo had dyspnea.
2.3 Herbal Therapies	
PICO 2.3.3 Curcuma/Curcuminoid—Knee	
Nakagawa, 2014	None.

Panahi, 2014	None.
PICO 2.3.4 Pycnogenol—Knee	
Cisár, 2008	None.
2.4 Nutraceuticals	
PICO 2.4.1 Glucosamine—Knee	
Giordano, 2009	"Serious AEs occurred in 2 patients (6.7%) from the GS group and 1 patient (3.3%) from the placebo group; therefore, they were withdrawn from the study." FYI: "In the GS group, 1 patient (3.3%) withdrew because of heartburn that developed 2 weeks after treatment initiation, and another (3.3%) withdrew because of a diffuse itch that developed in the first week of treatment." "In the placebo group, 1 subject (3.3%) withdrew because of constipation that developed during the first week of treatment..."
Hughes, 2002	"No serious adverse effects of treatment were reported during the trial and there were no differences between the treatment groups in the numbers or severity of adverse events reported."
McAlindon, 2004	"One participant in the placebo group was hospitalized due to dizziness, nausea, and chest pressure." 12 weeks
Sawitzke, 2010	"There were 84 SAEs occurring in 64 patients, of which 5 were felt to be possibly related to the study medications. These included myocardial infarction (in a patient receiving the glucosamine/chondroitin sulfate combination, hereafter referred to as “combination”), coronary angioplasty (in placebo group), and hip arthroplasty, cerebrovascular accident and abdominal wall abscess (all receiving celecoxib). Other SAEs reported regardless of relatedness are one death, which occurred as a completed suicide (placebo group), two myocardial infarctions (one glucosamine, one combination), two cerebrovascular accidents (one glucosamine, one celecoxib), two hypertension (one combination, one placebo), one case with palpitations (combination), and one transient ischemic attack (combination). There were no serious adverse gastrointestinal bleeding events reported."
Usha, 2004	"...no patient had any serious side effect." 12 weeks follow up
PICO 2.4.1 Glucosamine—Hip	
Rozendaal, 2008	"Four patients, 3 of whom were randomly assigned to glucosamine sulfate, had a stroke. Two patients, 1 in each group, reported cancer."
PICO 2.4.2 Chondroitin—Knee/Hip	
Mazieres, 2001	"Two patients in the CS group had a serious adverse event during treatment (dysphagia and surgery for cataract). No serious adverse event was judged to be drug related by the investigator or the sponsor."
Sawitzke, 2010	"There were 84 SAEs occurring in 64 patients, of which 5 were felt to be possibly related to the study medications. These included myocardial infarction (in a patient receiving the glucosamine/chondroitin sulfate combination, hereafter referred to as “combination”), coronary angioplasty (in placebo group), and hip arthroplasty, cerebrovascular accident and abdominal wall abscess (all receiving celecoxib). Other SAEs reported regardless of relatedness are one death, which occurred as a completed suicide (placebo group), two myocardial infarctions (one glucosamine, one combination), two cerebrovascular accidents (one glucosamine, one celecoxib), two hypertension (one combination, one placebo), one case with palpitations (combination), and one transient ischemic attack (combination). There were no serious adverse gastrointestinal bleeding events reported."
Uebelhart, 1998	One patient in the Chondroitin group died within the 12 month period.
Uebelhart, 2004	"Only minor adverse events (AEs) occurred during this 1-year study."
Wildi, 2011	Chondroitin: Atrial fibrillation (N=1) Placebo: Cholecystitis (N=1) Within 6 months.
Zegels, 2012	"Eight serious adverse events, all related to hospitalization, occurred during the study: two in the placebo arm (i.e., endourethral prostate resection, surgery related to frequent angina and snoring), two in the CS 1200 group (i.e., cystitis, radical prostatectomy) and four in the CS 3*400 group (i.e., transient ischaemic attack, acute intermediate syndrome, surgery on lumbar spinal stenosis, myocardial infraction)."
PICO 2.4.3 Glucosamine and Chondroitin—Knee/Hip	

Kanzaki, 2012	"All of the self-reported adverse events were mild or intermediate in intensity and occurred only temporarily, and were judged by the investigator as medically unrelated to the treatment."
Lugo, 2016	"Severe Adverse Events" per Table 10
Roman-Blas, 2017	Patients with at least one SAE within 6 months Table 4
Sawitzke, 2010	"There were 84 SAEs occurring in 64 patients, of which 5 were felt to be possibly related to the study medications. These included myocardial infarction (in a patient receiving the glucosamine/chondroitin sulfate combination, hereafter referred to as "combination"), coronary angioplasty (in placebo group), and hip arthroplasty, cerebrovascular accident and abdominal wall abscess (all receiving celecoxib). Other SAEs reported regardless of relatedness are one death, which occurred as a completed suicide (placebo group), two myocardial infarctions (one glucosamine, one combination), two cerebrovascular accidents (one glucosamine, one celecoxib), two hypertension (one combination, one placebo), one case with palpitations (combination), and one transient ischemic attack (combination). There were no serious adverse gastrointestinal bleeding events reported."
PICO 2.4.4 Vitamin D—Knee	
Arden, 2016	"Only two SAE's were classified as possibly related to treatment (one placebo with pancreatitis and one vitamin D with calculus ureteric), the remaining SAE's were classified as unrelated to treatment."
Jin, 2016	Sum of all serious adverse events; may be slight overestimation if certain patients experienced more than one event Vitamin D: Death N=1, Malignancy N=4, Coronary artery disease N=1, Major depression N=1, Nephrolithiasis N=1, Hospitalization N=3 (Two participants were admitted to the hospital after a fall and 1 was admitted because of severe diarrhea) Placebo: Malignancy N=2, Coronary artery disease N=1, Severe infection N=3, Nephrolithiasis N=1
McAlindon, 2013	"All except 1were considered unrelated, a possibly related hip fracture."
PICO 2.4.6 Collagen—Knee	
Lugo, 2016	"Severe Adverse Events" per Table 10
McAlindon, 2011	"Only one event (appendicitis) was classified as serious and this occurred in the CH group and was considered unrelated to treatment... There were no serious adverse event reports among the placebo group."
PICO 2.4.6 Collagen—Hip	
Schauss, 2011	"None of the sites opened the coded envelopes until the end of the study because there were no serious adverse events."
PICO 2.4.8 Diacerein—Knee	
Chevalier, 2009	ONLY ANAKINRA TRIAL
Pelletier, 2000	No serious or severe AEs regarding the upper GI tract occurred during the study.
PICO 2.5 Duloxetine—Knee	
Abou-Raya, 2012	None.
Chappell, 2009	"A total of 3 (1.3%) patients experienced 6 serious adverse events, including 2 patients in the placebo group (dehydration, gouty arthritis, myocardial infarction) and 1 patient in the duloxetine group (asthma, bronchitis, allergic rhinitis)."
Chappell, 2011	"A total of 5 (2.0%) patients experienced five serious adverse events, including two patients in the placebo group (atrial fibrillation and acute pyelonephritis) and three patients in the duloxetine group (drug intolerance, memory impairment, and supraventricular tachycardia)."
PICO 2.5 Duloxetine—Hip	
Abou-Raya, 2012 (Knee OA)	None.
Chappell, 2009 (Knee OA)	"A total of 3 (1.3%) patients experienced 6 serious adverse events, including 2 patients in the placebo group (dehydration, gouty arthritis, myocardial infarction) and 1 patient in the duloxetine group (asthma, bronchitis, allergic rhinitis)."
Chappell, 2011 (Knee OA)	"A total of 5 (2.0%) patients experienced five serious adverse events, including two patients in the placebo group (atrial fibrillation and acute pyelonephritis) and three patients in the duloxetine group (drug intolerance, memory impairment, and supraventricular tachycardia)."
2.7 Doxycycline	
PICO 2.6 Doxycycline—Knee/Hip	
Brandt, 2005	Specific SAEs were not reported. None of them were attributable to the treatment.

Snijders, 2011	“A total of five SAE occurred during the trial: one traumatic patella fracture (placebo group); two myocardial infarctions (both doxycycline group); one total knee replacement (doxycycline group) and one arthroscopic meniscectomy (placebo group). None of the SAE were likely to have been attributable to doxycycline, therefore no suspected unexpected serious adverse reactions occurred during the trial.”
2.8 Anti-Osteoporosis Drugs	
PICO 2.7.1 Bisphosphonates—Knee	
Laslett, 2012	*6 and 1 Non-elective hospital admissions in the zoledronic acid and placebo groups, respectively, which included removal of lymph node (in a patient with lymphoma), insertion of heart stent, heart valve repair, colonoscopy and cystoscopy (two operations in a patient with bladder cancer), knee pain, fractured pelvis and fractured elbow. eGFR, estimated glomerular filtration rate. Two cases of cancer- bladder cancer and non-Hodgkin’s lymphoma, one in each group "There was a trend to a higher rate of serious adverse events in the ZA group but these were disparate in nature and none were considered causally related to ZA."
Rossini, 2015	Of 10 side effects reported by the Clodronate group, one case involved a severe cough that led to hospitalization for suspected lung cancer during follow-up
Spector, 2005	"Overall, 34 patients reported a total of 53 serious AEs. Investigators considered four serious AEs as possibly related to study treatment; two of these (rash and diarrhea) were in patients treated with placebo and two (anaemia and increased general joint pain) were in patients treated with risedronate at 5 mg."
PICO 2.7.2 Calcitonin—Knee/Hip	
Karsdal, 2015 (2301)	Specific SAEs were not reported.
Kardsal, 2015 (2302)	Specific SAEs were not reported.
2.9 Investigational DMOADs	
PICO 2.8.1 IL-1 inhibitors—Knee/Hip	
Chevalier, 2009	“Four serious AEs were reported in 2 patients, none of which were determined by the investigators to be related to the study drug. One of these patients (in the placebo group) experienced a single serious AE of noncritical coronary artery stenosis. The other 3 serious AEs reported by a second patient (in the anakinra 150 mg group) included intense menstrual bleeding, altered mental status, noncardiac chest pain, and extremity pain. No deaths were reported.”
PICO 2.8.3 Anti-Nerve Growth Factor—Knee	
Brown, 2012	Specific SAEs were not reported.
Brown, 2013	Specific SAEs were not reported.
Lane, 2010	Serious adverse events were reported in 6 patients (2%) receiving tanezumab (appendicitis, bacterial arthritis, cellulitis, spinal stenosis, breast cancer, and syncope) and in 1 patient (1%) receiving placebo (noncardiac chest pain).
Mayorga, 2016	*Serious TEAEs in double-blind Phase: Placebo: chronic obstructive pulmonary disease; Fulranumab: myocardial infarction. †Serious TEAEs in Posttreatment Phase: Placebo: intervertebral disc protrusion and cerebrovascular accident); Fulranumab: bilateral totalknee arthroplasty and right total knee arthroplasty"
Nagashima, 2011	"Three serious AEs (SAEs) were reported: inguinal hernia (tanezumab 10 mg/kg group); atrioventricular block (tanezumab 25 mg/kg group) and contusion (prior to study drug administration). None of the SAEs were considered to be related to the study drug and were resolved or resolving at the end of the study"
Tiseo, 2014	"These serious AEs were osteoarthritis (2), intervertebral disc degeneration (1), pulmonary embolism (1), deep vein thrombosis (1), atrial fibrillation (1), pyelonephritis (1), and squamous cell carcinoma (1). None of these events was characterized as being related to study drug."
PICO 2.8.3 Anti-Nerve Growth Factor—Hip	

Ekman 2014A	Specific SAE's were not described.
Ekman, 2014B	Specific SAE's were not described.
Sanga, 2013	Serious adverse events occurred in 3 patients, but they were not neurologically related and resolved before study completion.
Spierings, 2013	None of the adverse events were considered serious.
PICO 2.8.4 Fibroblast growth factor—Knee/Hip	
Dahlberg, 2016	Placebo: 1 patient died from myocardial infarction. 1 gastrointestinal hemorrhage, 1 gouty arthritis, 1 post-procedural infection, all were classified as TEAEs. 1 2nd degree atrioventricular block (pre-treatment) not classified as TEAE. =5 SAE total FGF: 1 patient died from pulmonary embolism. 1 joint contracture, 1 angioedema, 1 decreased mobility, and 1 spinal OA, not classified as TEAE. =5 SAE total
PICO 2.8.5 Colchicine—Knee/Hip	
Das, 2002	1 patient in colchicine group withdrew due to drug-related diarrhea
PICO 2.8.6 Methotrexate—Knee/Hip	
De Holanda, 2007	Portuguese translate from Google: ""Adverse events reported were few and without gravity." Assuming that "without gravity" means not serious. There is further description of mild AEs which were most common among patients.
2.10 Intra-articular Injections	
PICO 2.9.1 Corticosteroids—Knee	
Henriksen, 201 5	3 SAE and 1 SAE reported for IACS and Placebo, respectively. No safety descriptions given within the publication, but the values correspond to withdrawals due to adverse events. Sought out other publications, but unable to find full details of safety data. Assume Juni 2015 used withdrawals due to AE as a proxy for SAE.
Lyons, 2005	N/A. No SAEs reported or described.
Ozturk, 2006	N/A. No SAEs reported or described.
Petrella, 2015	In the Hydros group, the 3 reported SAEs were colitis, broncho-pneumonia and arthralgia. In the Hydros-TA group, the 2 reported SAEs included a report of a meniscal lesion and a cyst aspiration. (Hydros: control, Hydros-TA: corticosteroid).
Ravaud, 1999	N/A. No SAEs reported or described.
PICO 2.9.1 Corticosteroids—Hip	
Lambert, 2007	None.
Qvistgaard, 2006	None.
PICO 2.9.2 Viscosupplementation—Knee	
Wobig, 1998	None.
Dickson, 2001	"Serious adverse events were reported in four patients during the study. One patient had a suspected gastrointestinal bleed, which was considered to be related to treatment with diclofenac. The other events were considered unrelated to treatment. These were myocardial infarction and transient ischaemic attack in two patients treated with hylan G-F 20 (both patients had had similar episodes prior to study entry) and blurred vision due to a macular hole in the retina of a patient in the control group."
Cubukcu, 2005	None.
Diracoglu, 2009	None.
Karlsson, 2002	"We cannot tell from our data if any of these serious adverse events resulted from the use of concomitant analgesics"
Rolf, 2005	None.
Bragantini, 1987	None.
Grecomoro, 1987	None.
Dougados, 1993	None.
Carrabba, 1995	None.
Formiguera Sala, 1995	None.
Jubb, 2003	"Fourteen patients in the placebo and 27 in the HA group experienced serious AEs (one patient in the HA group died as a result of a myocardial infarction). All serious AEs were considered by the investigators to be the result of primary concomitant disease and not to be drug related."
Sanofi, 2008	"Eight subjects experienced SAEs: 2 subjects in the HYALGAN group and 6 subjects in the PB-saline group; none of the SAEs were considered device related. Four subjects experienced AEs that led to study discontinuation: 2 HYALGAN subjects for moderate arthralgia; 1 HYALGAN subject for myocardial infarction; and 1 PB-Saline subject for knee arthroplasty."

Lundsgaard, 2008	None.
Huang, 2011	"[Serious adverse events] included a forearm fracture, intestinal obstruction, and aggravated urinary incontinence in the sodium hyaluronate group, and upper gastrointestinal bleeding and joint sprain in the placebo group. All were considered to be unrelated to study treatment."
Lohmander, 1996	None.
France [unpublished]	None.
Hizmetli, 1999 [unpublished]	None.
Brandt, 2001	"These [serious] adverse events included diverticulitis, esophagitis, cholecystitis, hyperglycemia, atrial fibrillation, congestive heart failure, deep vein thrombosis, pneumonia, asthma, congenital hernia, prostatic disorder, and carcinoma. Only cholecystitis was reported by more than one patient (n=2). None of the serious adverse events was thought by the investigator to have been related to treatment… No patient died during the study."
Butun, 2002	None.
Neustadt, 2005	"These [serious] adverse events included angina, myocardial infarction, gastrointestinal (GI) hemorrhage, and GI tract cancer None of the serious adverse events were considered by the investigator to be related to Orthovisc® treatment… There were no patient deaths during the study."
Sezgin, 2005	None.
Kul-Panza, 2010	None.
Altman, 2009	"One death occurred in the IA-SA group as a result of a motor vehicle accident, which was considered unrelated to treatment. Eighteen (3%) subjects experienced 22 serious TEAEs, with similar proportions in the IA-BioHA (n=9, 3%) and IA-SA (n=9, 3%) groups. The most common serious TEAEs were pneumonia and transient ischemic attack [each affecting 2 (0.3%) subjects]. However, these events were not notably more frequent than any other serious TEAEs (each with an incidence of 0.2%). None of the serious TEAE were considered related to study treatment"
Ferring, 2012 (NCT00988091)	Bradycardia (N=1 BioHA), Congestive heart failure (N=1 BioHA), Myopericarditis (N=1 BioHA), Angina pectoris (N=1 saline), Atrial fibrillation (N=2 saline), Myocardial infarction (N=1, open label period), urethral intrinsic sphincter deficiency (N=1 BioHA), Chest pain (N=1 BioHA), Death (N=1 BioHA), Cholecystitis (N=1, open label period), Intraspinal abscess (N=1 BioHA), Pneumonia (N=1 BioHA), Diverticulitis (N=1, open label period), Urosepsis (N=1, open label period), Femur fracture (N=1 BioHA), Upper limb fracture (N=1 saline), Joint injury (N=1, open label period), irregular heart rate (N=1 saline), Hyponatraemia (N=1, open label period), Obesity (N=1, open label period), Arthralgia (N=1 BiohA; N=1 open label period), Back pain (N=1 saline), Neoplasms (N=1 saline), Headache (N=1 BioHA), Syncope (N=1, saline), Transient ischaemic attack (N=1, open label period), Psychotic disorder (N=1, open label period), Urinary incontinence (N=1 BioHA), Acute renal failure (N=1, open label period), Pelvic prolapse (N=1 BioHA), Nephrectomy (N=1 BioHA), Spinal fusion surgery (N=1 BioHA), Knee Arthroplasy (N=1, open label period).
Strand, 2012	"Serious adverse events (SAEs) were reported in eight patients, including five cases of cancer. None were judged by investigators to be related to study treatment, although all SAEs occurred in the Gel-200 group, including one death." Others listed included Cardiac arrest, cryptogenic cirrhosis, acute bilateral pulmonary edema, respiratory failure, acute renal failure, and hypokalemia in one patient; transient ischemic attack in one patient, exertional dyspnea, transient blurry vision, and dizziness in one patient; incarcerated right femoral hernia and abdominal pain in one patient; the remaining three patients experienced four cases of cancer.
Blanco, 2008	One SAE in the HA group led to withdrawal; One SAE in the placebo group led to withdrawal and was specifically listed as "Not related" (Figure 2)
Altman, 2004	"There were no treatment-related serious adverse events"
Arden, 2014	Discontinuation due to unrelated serious adverse event (Figure 1)
Kosuwon, 2012	None.
Baltzer, 2009	None.
Petrella, 2006	None.
Van der Weegen, 2014	None.
PICO 2.9.2 Viscosupplementation—Hip	
Qvistgaard, 2006	None.
Richette, 2009	None.
Surgical Interventions	
PICO 3.2 Meniscectomy- Knee	
Sihvoenen, 2013	"The only observed serious adverse reaction was a deep infection of the index knee 4 months after surgery and 1 week after a dental procedure,

	leading to joint irrigation."
Combination Therapies	
PICO 4.1 Combination weight management and exercise—Knee	
Messier, 2013	Diet+Exercise: ALS (N=1), Stroke (N=1), Lung Infection (N=1), Cancer (N=2), Staph Infection (N=1) Diet: Cancer (N=1) All events were classified by investigators as unrelated to the study

RACGP

Royal Australian College *of* General Practitioners

Healthy Profession.
Healthy Australia.